Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications

EUROCOMB'23
Prague, August 28 - September 1, 2023

Editors: Daniel Kráľ, Jaroslav Nešetřil

MUNI PRESS · RSJ Foundation · ELSEVIER · erc DYNASNET · DIMATIA · matfyz · CHARLES UNIVERSITY

MASARYK
UNIVERSITY
PRESS

Editors: Daniel Král' and Jaroslav Nešetřil

Reviewed conference proceedings, anonymous reviewers

Published by Masaryk University, Žerotínovo nám. 617/9, 601 77 Brno, Czech Republic

First electronic edition, 2023

ISBN 978-80-280-0344-9 (online ; pdf)
ISSN 2788-3116

`https://journals.muni.cz/eurocomb`

Contents

Foreword

This is a collection of extended abstracts of the talks presented at the 12th European Conference on Combinatorics, Graph Theory and Applications 2023 (Eurocomb'23), which was held in Prague from August 28 to September 1, 2023. This year's Eurocomb conference was organized by Charles University with support from colleagues from Masaryk University. Continuing the tradition of the previous 11 conferences of the Eurocomb series, Eurocomb'23 covered the full range of combinatorics and graph theory including applications in other areas of mathematics, computer science and engineering.

Combinatorics is witnessing a steady increase in its significance within contemporary research, particularly in the fields of mathematics and computer science. However, its impact is not limited to these areas alone. This is also reflected in the increasing interest in Eurocomb conferences, which became an established forum of cutting-edge research in combinatorics world-wide. This year's Eurocomb attracted a total of 170 submissions. Out of them, 115 were presented as talks at the conference, while additional 17 were presented as posters. In addition to the contributed presentations, the conference program featured invited plenary talks by nine renowned experts - Gwenaël Joret (ULB), Eun Jung Kim (LAMSADE, Paris-Dauphine), Matthew Kwan (IST Austria), Shoham Letzter (University College London), Rose McCarty (Princeton), Dhruv Mubayi (UI Chicago), Oleg Pikhurko (Warwick), Luke Postle (Waterloo) and Martin Tancer (Charles University) - who covered recent advancements across a wide range of topics encompassing extremal, probabilistic, and structural combinatorics, as well as theoretical computer science.

Since Eurocomb'03, which was held in Prague 20 years ago, the European Prize in Combinatorics has been awarded to recognize groundbreaking contributions in combinatorics, discrete mathematics, and their applications by young European researchers not older than 35. The prize is supported by the Centre for Discrete Mathematics, Theoretical Computer Science and Applications (DIMATIA) of Charles University in Prague, the organizers of Eurocomb, Elsevier and additional sponsors.

We would like to thank all participants of Eurocomb'23, particularly the presenters, for contributing to the success of this year's Eurocomb conference. We express our special gratitude to the nine invited speakers for accepting our invitations to deliver plenary talks. We would also like to acknowledge the remarkable dedication and hard work of the pro-

gram committee members who contributed to creating an engaging conference program; their commitment played a vital role in making the program appealing. Last but not least, we would like to thank Zdeněk Dvořák, the chair of the organizing committee, together with all organizing committee members for their tremendous efforts, ensuring the smooth execution of the conference. Their commitment and meticulous planning significantly contributed to the overall success of the event.

Dan Kráľ and Jaroslav Nešetřil

Program committee

Agnes Backhausz (Eötvös Loránd University)
Marthe Bonamy (LABRI, Bordeaux)
Boris Bukh (CMU)
David Conlon (Caltech)
Michael Drmota (TU Wien)
Zdeněk Dvořák (Charles University)
Stefan Felsner (TU Berlin)
Andrzej Grzesik (Jagiellonian University)
Ervin Győri (Alfréd Rényi Institute, Budapest)
Dan Kráľ (Masaryk University, Brno) - co-chair
Bojan Mohar (SFU, Vancouver and University of Ljubljana)
Rob Morris (IMPA, Rio de Janeiro)
Jaroslav Nešetřil (Charles University, Prague) - chair
Marc Noy (UPC, Barcelona)
Patrice Ossona de Mendez (CNRS and EHESS, Paris)
Marco Pellegrini (IIT-CNR, Pisa)
Michał Pilipczuk (University of Warsaw)
Oriol Serra (UPC, Barcelona)
Jozef Širáň (Open University)
Jozef Skokan (LSE, London)
Sophie Spirkl (Waterloo)
Maya Stein (University of Chile, Santiago de Chile)
David Wood (Monash University)
Xuding Zhu (Zhejiang Normal University, Jinhua)

Organizing committee

Martin Balko (Charles University, Prague)
Zdeněk Dvořák (Charles University, Prague) - chair
Frederik Garbe (Masaryk University, Brno)
Jan Hubička (Charles University, Prague)
Dan Kráľ (Masaryk University, Brno) - co-chair
Jan Kratochvíl (Charles University, Prague)
Martin Mareš (Charles University, Prague)
Petra Milštainová (Charles University, Prague)
Jaroslav Nešetřil (Charles University, Prague) - co-chair
Milena Zeithamlová (Action M)

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

On the number of tangencies among 1-intersecting curves

(Extended abstract)

Eyal Ackerman* Balázs Keszegh†

Abstract

Let $\mathcal{C}$ be a set of curves in the plane such that no three curves in $\mathcal{C}$ intersect at a single point and every pair of curves in $\mathcal{C}$ intersect at exactly one point which is either a crossing or a touching point. According to a conjecture of János Pach the number of pairs of curves in $\mathcal{C}$ that touch each other is $O(|\mathcal{C}|)$. We prove this conjecture for x-monotone curves.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-001`

1 Introduction

We study the number of tangencies within a family of 1-intersecting x-monotone planar curves. A planar curve is a *Jordan arc*, that is, the image of an injective continuous function from a closed interval into $\mathbb{R}^2$. If no two points on a curve have the same x-coordinate, then the curve is *x-monotone*. We consider families of curves such that every pair of curves

*Department of Mathematics, Physics and Computer Science, University of Haifa at Oranim, Tivon 36006, Israel.

†Alfréd Rényi Institute of Mathematics and ELTE Eötvös Loránd University, Budapest, Hungary. Research supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, by the National Research, Development and Innovation Office – NKFIH under the grant K 132696 and FK 132060, by the ÚNKP-22-5 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund and by the ERC Advanced Grant "ERMiD". This research has been implemented with the support provided by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, financed under the ELTE TKP 2021-NKTA-62 funding scheme.

intersect at a finite number of points. Such a family is called t-intersecting if every pair of curves intersects at at most t points. An intersection point p of two curves is a *crossing point* if there is a small disk D centered at p which contains no other intersection point of these curves, each curve intersects the boundary of D at exactly two points and in the cyclic order of these four points no two consecutive points belong to the same curve. If two curves intersect at exactly one point which is not a crossing point, then we say that they are *touching* or *tangent* at that point.

The number of *tangencies* is the number of tangent pairs of curves. If more than two curves are allowed to intersect at a common point, then every pair of curves might be tangent, e.g., for the graphs of the functions x^{2i}, $i = 1, 2, \ldots, n$, in the interval $[-1, 1]$. Therefore, we restrict our attention to families of curves in which no three curves intersect at a common point. It is not hard to construct such a family of n (x-monotone) 1-intersecting curves with $\Omega(n^{4/3})$ tangencies based on a famous construction of Erdős (see [14]) of n lines and n points admitting that many point-line incidences. János Pach [13] conjectured that requiring every pair of curves to intersect (either at crossing or a tangency point) leads to significantly less tangencies.

Conjecture 1 ([13])**.** *Let $\mathcal{C}$ be a set of n curves such that no three curves in $\mathcal{C}$ intersect at a single point and every pair of curves in $\mathcal{C}$ intersect at exactly one point which is either a crossing or a tangency point. Then the number tangencies among the curves in $\mathcal{C}$ is $O(n)$.*

Györgyi, Hujter and Kisfaludi-Bak [8] proved Conjecture 1 for the special case where there are constantly many faces in the arrangement of $\mathcal{C}$ that together contain all the endpoints of the curves. In this paper we show that Conjecture 1 also holds for x-monotone curves.

Theorem 2. *Let $\mathcal{C}$ be a set of n x-monotone curves such that no three curves in $\mathcal{C}$ intersect at a single point and every pair of curves in $\mathcal{C}$ intersect at exactly one point which is either a crossing or a tangency point. Then the number tangencies among the curves in $\mathcal{C}$ is $O(n)$.*

We prove Theorem 2 by considering two types of tangencies according to whether a tangency point is between two curves such that their projections on the x-axis are nested or overlapping. In each case we consider the tangencies graph whose vertices represent the curves and whose edges represent tangent pairs of curves. In the latter case we show that it is possible to disregard some ratio of the edges using the pigeonhole principle and the dual of Dilworth's Theorem and then order the remaining edges such that there is no long monotone increasing path with respect to this order. In the first case, we show that after disregarding some ratio of the edges the remaining edges induce a forest. Due to space limitations most of the details of the proof are omitted. The interested reader can find them in [3].

Related Work. It follows from a result of Pach and Sharir [17] that n x-monotone 1-intersecting curves admit $O\left(n^{4/3}(\log n)^{2/3}\right)$ tangencies. Note that this bound almost matches the lower bound mentioned above. It also follows from [17] that for bi-infinite x-monotone 1-intersecting curves the maximum number of tangencies is $\Theta(n \log n)$. Pálvölgyi et nos [2] showed that there are $O(n)$ tangencies among families of n 1-intersecting curves that can be partitioned into two sets such that all the curves within each set are pairwise disjoint. Variations of this bipartite setting were also studied in [1, 10, 19].

Pach, Rubin and Tardos [15, 16] settled a long-standing conjecture of Richter and Thomassen [20] concerning the number of crossing points determined by pairwise intersecting curves. In particular, they showed that in any set of curves admitting linearly many tangencies the number of crossing points is superlinear with respect to the number of tangencies. This implies that for any fixed t every set of n t-intersecting curves admits $o(n^2)$ tangencies. Salazar [22] already pointed that out for such families which are also pairwise intersecting. Better bounds for families of t-intersecting curves were found in [5, 10]. Specifically, it follows from [10] that n 1-intersecting curves determine $O(n^{7/4})$ tangencies.

There are several other problems in combinatorial geometry that can be phrased in terms of bounding the number of tangencies between certain curves, see, e.g., [4]. The most famous of which is the unit distance problem of Erdős [6] which asks for the maximum number of unit distances among n points in the plane. It is easy to see that this problem is equivalent to asking for the maximum number of tangencies among n unit circles.

2 Proof of Theorem 2

Let $\mathcal{C}$ be a set of n x-monotone curves such that no three curves in $\mathcal{C}$ intersect at a single point and every pair of curves in $\mathcal{C}$ intersect at exactly one point which is either a crossing or a tangency point. By slightly extending the curves if needed, we may assume that every intersection point of two curves is an interior point of both of them and that all the endpoints of the curves are distinct.

Let $p = (x_1, y_1)$ and $q = (x_2, y_2)$ be two points. We write $p <_x q$ if $x_1 < x_2$ and we write $p <_y q$ if $y_1 < y_2$. We mainly consider the order of points from left to right, so when we use terms like 'before', 'after' and 'between' they should be understood in this sense. For a curve $c \in \mathcal{C}$ we denote by $L(c)$ and $R(c)$ the left and right endpoints of c, respectively. If $p, q \in c$, then $c(p, q)$ denotes the part of c between these two points. We denote by $c(-, p)$ (resp., $c(p, +)$) the part of c between $L(c)$ (resp., $R(c)$) and p. For another curve $c' \in \mathcal{C}$ we denote by $I(c, c')$ the intersection point of c and c'. We may also write, e.g., $c(c', q)$ instead of $c(I(c, c'), q)$

Suppose that an x-monotone curve c_1 lies *above* another x-monotone curve c_2, that is, the two curves are non-crossing (but might be touching) and there is no vertical line ℓ such that $I(c_1, \ell) <_y I(c_2, \ell)$. Assuming the endpoints of c_1 and c_2 are distinct there are four possible cases: (1) $L(c_1) <_x L(c_2) <_x R(c_2) <_x R(c_1)$; (2) $L(c_2) <_x L(c_1) <_x R(c_1) <_x R(c_2)$; (3) $L(c_1) <_x L(c_2) <_x R(c_1) <_x R(c_2)$; and (4) $L(c_2) <_x L(c_1) <_x R(c_2) <_x R(c_1)$. We denote by $c_2 \prec_i c_1$ the relation that corresponds to case i, for $i = 1, 2, 3, 4$. It is not hard to see that each $\prec_i$ is a partial order.

Proposition 3. *For every $i = 1, 2, 3, 4$ there are no three curves $c_1, c_2, c_3 \in \mathcal{C}$ such that $c_1 \prec_i c_2 \prec_i c_3$.*

We say that the tangency point of two touching curves $c_1, c_2 \in \mathcal{C}$ is of *Type i* if $c_1 \prec_i c_2$. We will count separately tangency points of Types 1 and 2 and tangency points of Types 3 and 4.

Lemma 4. *There are $O(n)$ tangency points of Type 1 or 2.*

Proof. Since all the curves in $\mathcal{C}$ are pairwise intersecting and x-monotone there is a vertical line ℓ that intersects all of them. By slightly shifting ℓ if needed we may assume that no two curves intersect ℓ at the same point. We assume without loss of generality that at least half of all the tangency points of Types 1 and 2 are to the right of ℓ, for otherwise we may reflect all the curves about ℓ. We may further assume that at least half of the tangency points of Types 1 and 2 to the right of ℓ are of Type 2, for otherwise we may reflect all the curves about the x-axis. Henceforth, we consider only Type 2 tangency points to the right of ℓ.

By Proposition 3 a curve cannot touch one curve from above and another curve from below at Type 2 tangency points. Thus, we may partition the curves into *blue* curves and *red* curves such that at every tangency point a blue curve touches a red curve from below (we ignore curves that contain no tangency points among the ones that we consider).

Proposition 5. *Every pair of blue curves cross each other.*

We proceed by marking the rightmost tangency point on every red curve. Clearly, at most n tangency points are marked. Henceforth, we consider only unmarked tangency points. Let G be the (bipartite) *tangencies graph* of the blue and red curves. That is, the vertices of G correspond to the blue and red curves and its edges correspond to pairs of touching blue and red curves (recall that we consider only unmarked tangency points of Type 2 to the right of ℓ). We will show that G is a forest and hence has at most $n-1$ edges.

Suppose that G contains a cycle and let $C = b_0 - r_0 - b_1 - r_1 - \ldots - b_k - r_k - b_0$ be a shortest cycle in G, such that b_i corresponds to a blue curve and r_i corresponds to a red curve, for every $i = 0, 1, \ldots, k$. We may assume without loss of generality that b_1 has the lowest intersection point with ℓ among the blue curves in C and that $I(b_0, \ell) <_y I(b_2, \ell)$.

Proposition 6. *For every $i \geq 1$ the curve r_i intersects ℓ above r_0 and intersects $b_0(-, \ell)$, $r_0(b_0, +)$ and $b_1(b_0, +)$. See Figure 1 for an illustration.*

It follows from Proposition 6 that r_k intersects b_0 to the left of ℓ and therefore (b_0, r_k) cannot be an edge in G. Thus G is a forest and has at most $n-1$ edges. This implies that there are at most $2n-1$ Type 2 tangency points to the right of ℓ and at most $8n-4$ tangency points of Types 1 and 2. □

Lemma 7. *There are $O(n)$ tangency points of Type 3 or 4.*

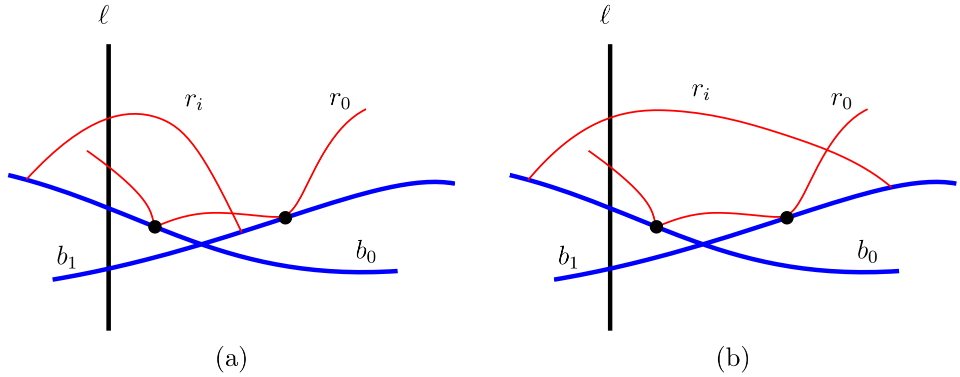


Figure 1: Illustrations for the statement of Proposition 6: r_i intersects ℓ above r_0 and intersects $b_0(-,\ell)$, $r_0(b_0,+)$ and $b_1(b_0,+)$.

Proof. As in the proof of Lemma 4, we may assume that there is a vertical line ℓ that intersects all the curves at distinct points and it is enough to consider only Type 4 tangency points to the right of ℓ.

By Proposition 3 a curve cannot touch one curve from above and another curve from below at Type 4 tangency points. Thus, we may partition the curves into *blue* curves and *red* curves such that at every tangency point a blue curve touches a red curve from below (we ignore curves that contain no tangency points among the ones that we consider).

Clearly, there are no Type 4 tangencies among the blue curves, however, there might be tangencies of other types among them. Next we wish to obtain a subset of the blue curves such that every pair of them are crossing and they together contain a percentage of the tangency points that we consider. It follows from Proposition 3 that the largest chain in the partially ordered set of the blue curves with respect to $\prec_1$ is of length two. Therefore, by Mirsky's Theorem (the dual of Dilworth's Theorem) the blue curves can be partitioned into two antichains with respect to $\prec_1$. The blue curves of one of these antichains contain at least half of the tangency points that we consider. By continuing with this set of blue curves and applying the same argument twice more with respect to $\prec_2$ and $\prec_3$ we obtain a set of pairwise crossing blue curves that together contain at least $1/8$ of the tangency points of Type 4 to the right of ℓ. Henceforth we consider these blue curves and the red curves that touch at least one of them at a Type 4 tangency point to the right of ℓ.

Let $G = (B \cup R, E)$ be the (bipartite) *tangencies graph* of these blue and red curves. That is, B corresponds to the blue curves, R corresponds to the red curves and E corresponds to pairs of touching blue and red curves (at Type 4 tangency points to the right of ℓ). We order the edges of G according to the order of their corresponding tangency points from left to right. We will show that G has linearly many edges using the following fact, attributed to Rödl [21] in [7].

Proposition 8. *Let $G = (V, E)$ be a graph and let $<$ be a total order of its edges. Let k be an integer and suppose that G does not contain a monotone increasing path of k edges, that is, a path $e_1 - e_2 - \ldots - e_k$ such that $e_1 < e_2 < \ldots < e_k$. Then $|E| < \binom{k}{2}|V|$.*

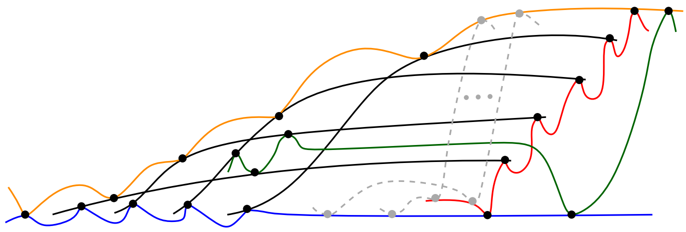


Figure 2: n x-monotone pairwise intersecting 1-intersecting curves might determine $3n-4$ tangencies.

Recall that we order the edges of G according to the order of their corresponding tangency points from left to right. The lemma follows from Proposition 8 and the next claim.

Proposition 9. *G does not contain a monotone increasing path of 7 edges starting at B.*

We conclude from Propositions 8 and 9 that G has at most $28n$ edges. This in turn implies that there are at most $8 \cdot 2 \cdot 2 \cdot 28n = 896n$ tangency points of Types 3 and 4. $\square$

By Lemmata 4 and 7 there are at most $904n - 4$ tangency points among the curves in $\mathcal{C}$. This concludes the proof of Theorem 2.

3 Discussion

We have shown that n x-monotone pairwise intersecting 1-intersecting curves determine $O(n)$ tangencies. The constant hiding in the big-O notation is rather large, since, for simplicity, we did not make much of an effort to get a smaller constant. In particular, our upper bound can be improved by considering more cases. It would be interesting to determine the exact maximum number of tangencies among a set of n x-monotone curves each two of which intersect at exactly one point. The best lower bound we came up with is $3n-4$, see Figure 2.

References

[1] Eyal Ackerman. The maximum number of tangencies among convex regions with a triangle-free intersection graph. In Pach [12], pages 19–30.

[2] Eyal Ackerman, Balázs Keszegh, and Dömötör Pálvölgyi. On tangencies among planar curves with an application to coloring L-shapes. In Jaroslav Nešetřil, Guillem Perarnau, Juanjo Rué, and Oriol Serra, editors, *Extended Abstracts EuroComb 2021*, pages 123–128, Cham, 2021. Springer International Publishing.

[3] Eyal Ackerman and Balázs Keszegh. On the number of tangencies among 1-intersecting curves. *CoRR*, abs/2305.13807, 2023.

[4] Pankaj K Agarwal, Eran Nevo, János Pach, Rom Pinchasi, Micha Sharir, and Shakhar Smorodinsky. Lenses in arrangements of pseudo-circles and their applications. *Journal of the ACM (JACM)*, 51(2):139–186, 2004.

[5] Maya Bechler-Speicher. A crossing lemma for families of jordan curves with a bounded intersection number. *CoRR*, abs/1911.07287, 2019.

[6] P. Erdős. On sets of distances of n points. *The American Mathematical Monthly*, 53(5):248–250, 1946.

[7] Dániel Gerbner, Abhishek Methuku, Dániel T. Nagy, Dömötör Pálvölgyi, Gábor Tardos, and Máté Vizer. Turán problems for edge-ordered graphs. *Journal of Combinatorial Theory, Series B*, 160:66–113, 2023.

[8] Péter Györgyi, Bálint Hujter, and Sándor Kisfaludi-Bak. On the number of touching pairs in a set of planar curves. *Computational Geometry*, 67:29–37, 2018.

[9] Klara Kedem, Ron Livne, János Pach, and Micha Sharir. On the union of jordan regions and collision-free translational motion amidst polygonal obstacles. *Discrete & Computational Geometry*, 1(1):59–71, Mar 1986.

[10] Balázs Keszegh and Dömötör Pálvölgyi. The number of tangencies between two families of curves, 2021.

[11] P. Koebe. Kontaktprobleme der konformen Abbildung. *Ber. Saechs. Akad. Wiss. Leipzig, Math.-Phys. Kl.*, 88:141–164, 1936.

[12] J. Pach, editor. *Thirty Essays on Geometric Graph Theory*. Springer New York, 2012.

[13] János Pach. personal communication.

[14] János Pach and Pankaj K. Agarwal. *Combinatorial Geometry*, chapter 11, pages 177–178. John Wiley and Sons Ltd, 1995.

[15] János Pach, Natan Rubin, and Gábor Tardos. On the Richter-Thomassen conjecture about pairwise intersecting closed curves. *Combinatorics, Probability and Computing*, 25(6):941–958, 2016.

[16] János Pach, Natan Rubin, and Gábor Tardos. A crossing lemma for Jordan curves. *Advances in Mathematics*, 331:908–940, 2018.

[17] János Pach and Micha Sharir. On vertical visibility in arrangements of segments and the queue size in the Bentley-Ottmann line sweeping algorithm. *SIAM J. Comput.*, 20(3):460–470, 1991.

[18] János Pach and Micha Sharir. On the boundary of the union of planar convex sets. *Discrete & Computational Geometry*, 21(3):321–328, 1999.

[19] János Pach, Andrew Suk, and Miroslav Treml. Tangencies between families of disjoint regions in the plane. *Comput. Geom.*, 45(3):131–138, 2012.

[20] R.B. Richter and C. Thomassen. Intersections of curve systems and the crossing number of $C_5 \times C_5$. *Discrete Comput. Geom.*, 13(2):149–159, 1995.

[21] V. Rödl. Master's thesis, Charles University, 1973.

[22] Gelasio Salazar. On the Richter-Thomassen conjecture about pairwise intersecting closed curves. *J. Comb. Theory, Ser. B*, 75(1):56–60, 1999.

[23] Jack Snoeyink and John Hershberger. Sweeping arrangements of curves. In *Proceedings of the fifth annual symposium on Computational geometry*, pages 354–363. ACM, 1989.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Partition universality for hypergraphs of bounded degeneracy and degree

(Extended abstract)

Peter Allen* Julia Böttcher† Domenico Mergoni Cecchelli‡

Abstract

We consider the following question. When is the random k-uniform hypergraph $\Gamma = G^{(k)}(N, p)$ likely to be r-partition universal for k-uniform hypergraphs of bounded degree and degeneracy? That is, for which p can we guarantee asymptotically almost surely that in any r-colouring of $E(\Gamma)$ there exists a colour χ such that in Γ there are χ-monochromatic copies of all k-uniform hypergraphs of maximum vertex degree Δ, degeneracy at most D, and cN vertices for some constant $c = c(D, \Delta) > 0$. We show that if $\mu > 0$ is fixed, then $p \geq N^{-1/D+\mu}$ suffices for a positive answer if N is large. On the other hand, for $p = o(N^{-1/D})$ we show that $G^{(k)}(N, p)$ is likely not to contain some graphs of maximum degree Δ and degeneracy D on cN vertices at all.

This improves the best upper bounds on the minimum number of edges required for a k-uniform hypergraph to be partition universal (even for $k = 2$) and also for the size-Ramsey problem for most k-uniform hypergraphs of bounded degree and degeneracy.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-002`

1 Introduction

Ramsey theory refers to the area of combinatorics that studies unavoidable organised subsystems. At the centre of this area of research is Ramsey's [24] seminal result which states

*London School of Economics, Department of Mathematics, Houghton Street, London WC2A 2AE, UK. E-mail: p.d.allen@lse.ac.uk.

†London School of Economics, Department of Mathematics, Houghton Street, London WC2A 2AE, UK. E-mail: j.boettcher@lse.ac.uk.

‡London School of Economics, Department of Mathematics, Houghton Street, London WC2A 2AE, UK. E-mail: d.mergoni@lse.ac.uk.

that for any graph F there exists N such that in any 2-edge-colouring of K_N there is a monochromatic copy of F. This is often indicated with the notation $K_N \to_2 F$. The study of the minimal value of N for which $K_N \to_2 F$, which is denoted by $R_2(F)$ and called the *Ramsey number* of F, is one of the main topics of Ramsey theory. This question can be easily generalised to a setting with r colours and a different host graph G: we write $G \to_r F$ if every r-colouring of $E(G)$ contains a monochromatic copy of F.

Another line of research that generated a lot of attention was started by Erdős, Faudree, Rousseau and Schelp [13]. The authors analysed other sparser structures for which we have $G \to_2 F$ and in particular, for fixed F and $r \in \mathbb{N}$, they asked to determine what they called the *size-Ramsey number* of the graph F, defined as $\hat{R}_r(F) := \min\{|E(G)| : G \to_r F\}$. We remark that this well defined as $K_{R_r(F)} \to_r F$.

One can also ask for a stronger property of G, namely that it is r-partition-universal for a family $\mathcal{F}$ of graphs. We say that G is r-partition-universal for $\mathcal{F}$ if for every r-colouring of $E(G)$, there exists some colour χ such that a colour χ copy of every $F \in \mathcal{F}$ appears. Observe that while $G \to_r F$ for every $F \in \mathcal{F}$ is certainly necessary for r-partition universality, it is not sufficient: it could be that for an r-colouring of $E(G)$ different members of $\mathcal{F}$ appear in different colours.

These questions have primarily been studied for graphs, but the above definitions naturally extend to k-uniform hypergraphs with $k \geq 3$. We will write *k-graphs* for k-uniform hypergraphs.

2 Previous results

For several classes of 'tree-like' graphs, the size-Ramsey number is known to be linear in the number of vertices. Beck [5] proved that $\hat{R}_r(P_n)$ is linear in n. Trees were dealt with by Friedman and Pippenger [14], and Haxell, Kohayakawa and Łuczak [17] proved it for cycles. Much later, Clemens et al. [7] showed that the same holds for powers of paths with 2 colours (Han et al. [15] extended this to r colours), Berger et al. [6] and Kamčev, Liebenau, Wood, and Yepremyan[18] for graphs of bounded maximum degree and treewidth.

For k-graphs for $k \geq 3$, Han et al. [16] proved a linear bound for 3-uniform tight paths, and Letzter, Pokrovskiy, Yepremyan [22] for all uniformities and more generally powers of bounded degree hypertrees.

However, not all bounded degree graphs have linear size-Ramsey numbers. Rödl and Szemerédi [25] proved that there is a family of graphs of maximum degree 3 whose size-Ramsey numbers grow as $n \log^{1/60} n$. Recently, Tikhomirov [27] improved this to $n \exp\big(\Omega(\sqrt{\log n})\big)$. In k-graphs, Dudek, Fleur, Mubayi and Rödl [12] proved a lower bound similar to that of Rödl and Szemerédi for k-graphs of maximum degree $k+1$. It remains a conjecture that these bounds can be improved to $n^{1+\varepsilon}$ (possibly for larger maximum degrees).

In terms of general upper bounds, Kohayakawa, Rödl, Schacht and Szemerédi [21] proved the first non-trivial upper bound $O\big(n^{2-1/\Delta} \log^{1/\Delta} n\big)$ on the size-Ramsey number of any graph on n vertices and maximum degree Δ; this was recently improved by Draganić and Petrova [11] for $\Delta = 3$.

For k-graphs, the upper bound $O(n^k)$ follows from the linearity of the usual Ramsey number, proved by Cooley, Fountoulakis, Kühn and Osthus [10]. The only improvement over this, to $O(n^{k-\varepsilon})$ for some (very) small $\varepsilon > 0$, was by Allen et al. [2].

All of the mentioned general upper bounds in fact are not only upper bounds on size-Ramsey numbers, but actually give graphs which are r-partition universal for bounded degree k-graphs.

3 Our result

If F is a k-graph, we write $\Delta(F)$ for the *maximum vertex degree* of F, i.e. the maximum number of edges which all contain a given vertex. We say F is D-degenerate if there is an ordering of the vertex set of F such that each vertex v in the ordering is in at most D edges whose other vertices are strictly before v. We prove the following main theorem.

Theorem 1. *For all $r, D, \Delta \in \mathbb{N}$, $k \geq 2$, and $\mu > 0$, there is $c > 0$ such that for all positive integers N and $p = N^{-\frac{1}{D}+\mu}$, the random k-graph $\Gamma = G^{(k)}(N, p)$ asymptotically almost surely has the following property. For every r-edge-colouring there is a colour χ such that, for any D-degenerate k-graph F of maximum degree at most Δ with cN vertices, there is a colour χ-monochromatic copy of F in Γ.*

For $k = 2$ this appears in the preprint [1, Theorem 3] of the first two authors.

Theorem 1 is asymptotically sharp. Indeed, if F is a D-degenerate graph on cN vertices with bounded maximum degree and approximately DcN edges (it is easy to construct such graphs) a first moment method argument shows that $G^{(k)}(N, p)$ with $p = o(N^{\frac{1}{D}})$ is likely not to contain a copy of F (let alone if we adversarily colour the edges of G).

4 Outline of the proof of Theorem 1

We now sketch the proof of Theorem 1. Our argument is intricate. Hence, to simplify the discussion here, we look at the case $k = 2$ first (and thus naturally refer to the lemmas in [1]), and then outline the changes needed for $k \geq 3$. For lack of space, we do not state the lemmas required precisely here, but only give a general idea.

Given a typical $\Gamma = G(N, p)$ and an adversarial r-colouring of its edges, we begin by identifying the colour χ and a subgraph G of colour χ edges in Γ into which we will embed our graph F. This G will be a h_1-partite subgraph of Γ, where h_1 is a reasonably large constant (depending on D, Δ and μ). We require that G has a property 'few unpromising subgraphs' which we will explain shortly. The precise statement is [1, Lemma 19].

We now want to embed a fixed D-degenerate graph F on cN vertices of maximum degree at most Δ to G. A method of Nenadov [23], which we formulate as a lemma in [1, Lemma 21], reduces this problem to that of finding 'robust homomorphisms' from F to G. Roughly speaking, this means we need to find a way of constructing a sequence

$\psi_0, \psi_1, \ldots, \psi_{v(F)}$ of partial homomorphisms from F to G, each embedding one more vertex than the previous, taken in the degeneracy order on G. The key property this sequence must have is: at each step, the number of choices we have to embed the next vertex is comparable to the number of choices we would have to embed it in the random graph Γ. That is, if the next vertex y of F has $d^-(y)$ neighbours before it in the degeneracy order, we should have $\Omega\big(p^{d^-(y)}N\big)$ choices for our embedding of y. The advantage this gives us is that we no longer have to care, when we construct our embedding of F, about avoiding previously used vertices.

We now explain the property 'few unpromising subgraphs'. Before we begin embedding F, we assign its vertices to the parts of G, with the property that vertices of F which are somewhat close together in graph distance will be assigned to different parts of G. When we come to embed y, we have already embedded its $d^-(y)$ neighbours $N^-(y)$ preceding it. In order to construct robust homomorphisms, we need many choices for our embedding of y. That is, we need that these embedded neighbours have many G-common neighbours in the part of G to which y is assigned. So the first idea is the following: we look at copies of $F\big[N^-(y)\big]$ in G, and designate them as *promising* if they have sufficiently many common neighbours in the correct part of G, and *unpromising* otherwise. For technical reasons—to be explained shortly—we actually have to modify this slightly, looking at copies not of $F\big[N^-(y)\big]$ but of a supergraph $H_0'(y)$ which contains some additional vertices of F that are within a (small) distance ℓ_0 of y. The property *few unpromising subgraphs* that we require of G is then the following. For each of the (boundedly many) choices of H_0' and parts to which the vertices of H_0' might be assigned, at most a tiny fraction of the embeddings of H_0' into G are unpromising.

Our embedding procedure will succeed if we are never forced to embed $H_0'(y)$ to an unpromising copy in G. To avoid this, we need to look ahead a (large) constant ℓ_1 number of steps. That is, when we embed a vertex x of F preceding y, we look at the graph distance from x to y. If it exceeds ℓ_1, we will not *cross off for y* any vertices to which we could embed x. If it is at most ℓ_1, we will examine each candidate vertex v for x and determine whether v is *dangerous*, i.e. assigning v to x makes it too hard in the future to avoid embedding $H_0'(y)$ to an unpromising copy. If it does, we will *cross off* v, which means we will not choose to embed x to v. What we need to argue is that we will not need to cross off many vertices at any given step.

To decide whether v is dangerous, we look at all the embeddings of $H_1'(y)$ (all the vertices at distance ℓ_1 or less from y that precede y) that are consistent with our current embedding ψ_{x-1}. Some of these embed $H_0'(y)$ to an unpromising copy, and we call these *unpromising extensions*. If the fraction of unpromising extensions which embed x to v is exceptionally large, then v is dangerous.

We can use a double-counting argument to show that for vertices at distance strictly less than ℓ_1 from y, there can only be few vertices which are dangerous. For vertices at distance exactly ℓ_1 from y, this argument fails. However we can show that the embeddings 'mix rapidly' and there will in fact be no dangerous vertices in this case. This is [1, Lemma 20].

To complete this sketch, we require that before we begin to embed F, there are very few unpromising extensions of $H_1'(y)$. When $H_0'(y)$ is carefully chosen—this is the technical reason mentioned above—this follows from a theorem of Spencer [26] counting rooted copies of graphs in a random graph together with the few unpromising subgraphs property of G.

Putting the pieces together, when we use this embedding strategy we will (inductively) have many candidates for each x preceding y, at most a few of which are crossed off for y (or for any other vertex). When we embed x to a vertex which is not crossed off for v, the fraction of unpromising extensions of $H'_1(y)$ can increase, but since this fraction starts tiny and is increased only by a bounded amount a bounded number of times, it will remain small. In particular, we will end up embedding $H'_0(y)$ to a promising subgraph, which guarantees y has many candidates as we needed.

We briefly mention one place in this sketch where significant technical work is required. This is in the construction of G, [1, Lemma 19]. To make this strategy work, the number of unpromising subgraphs we can allow must be very tiny indeed. In particular, a standard application of the Sparse Regularity Lemma [20] will not give sufficiently good control of the constants, and we need instead to use a strengthened version of the Sparse Regularity Lemma together with a 'cleaning' process. In order to obtain the required counting results, we in addition need the 'Counting KŁR' results of Conlon, Gowers, Samotij and Schacht [8].

We should also note that the 'mixing rapidly' proof of [1, Lemma 20] relies on Spencer's theorem [26] on rooted copies.

We now sketch how this strategy can be modified to work for $k \geq 3$. The object G into which we embed F needs to be a k-complex (hypergraph with edges of size at most k) with its edges of size k selected from edges of the same colour in Γ. In order to show that a suitable G exists, we need to use a strengthened sparse version of the Strong Hypergraph Regularity Lemma, proved by Allen, Parczyk and Pfenninger [4]. We need to develop a hypergraph version of the Counting KŁR results of [8] for this setting of complexes. And, finally, we need to use a rather more involved 'cleaning' process in order that we obtain the required control of our constants for 'few unpromising subgraphs'.

Apart from this, much of our strategy sketched above works in a broadly similar way for hypergraphs. We need to use the polynomial concentration theorem of Kim and Vu [19] replacing Spencer's theorem. We need to view F as a complex by down-closure and hence we need to consider edges of all uniformities up to k, not just of uniformity k, throughout. In particular, although all edges of F of uniformity smaller than k are contained in edges of uniformity k, this property is not preserved for the subgraphs $H'_0(y)$ and $H'_1(y)$; these k-complexes can have edges of smaller uniformity that are not in any k-edge.

References

[1] Peter Allen and Julia Böttcher. Partition universality for graphs of bounded degeneracy and degree. *arXiv preprint arXiv:2211.15819*, 2022.

[2] Peter Allen, Julia Böttcher, Ewan Davies, Eng Keat Hng, and Jozef Skokan. A sparse hypergraph blow-up lemma. In preparation.

[3] Peter Allen, Julia Böttcher, and Domenico Mergoni Cecchelli. Partition universality for hypergraphs of bounded degeneracy and degree. *To appear*, 2023+.

[4] Peter Allen, Olaf Parczyk, and Vincent Pfenninger. Resilience for tight hamiltonicity. *arXiv preprint arXiv:2105.04513*, 2021.

[5] József Beck. On size Ramsey number of paths, trees, and circuits. I. *J. Graph Theory*, 7(1):115–129, 1983.

[6] Sören Berger, Yoshiharu Kohayakawa, Giulia Satiko Maesaka, Taísa Martins, Walner Mendonça, Guilherme Oliveira Mota, and Olaf Parczyk. The size-Ramsey number of powers of bounded degree trees. *J. Lond. Math. Soc. (2)*, 103(4):1314–1332, 2021.

[7] Dennis Clemens, Matthew Jenssen, Yoshiharu Kohayakawa, Natasha Morrison, Guilherme Oliveira Mota, Damian Reding, and Barnaby Roberts. The size-Ramsey number of powers of paths. *J. Graph Theory*, 91(3):290–299, 2019.

[8] D. Conlon, W. T. Gowers, W. Samotij, and M. Schacht. On the KŁR conjecture in random graphs. *Israel J. Math.*, 203(1):535–580, 2014.

[9] David Conlon, Rajko Nenadov, and Miloš Trujić. The size-Ramsey number of cubic graphs. Bulletin of the London Mathematical Society, published online.

[10] Oliver Cooley, Nikolaos Fountoulakis, Daniela Kühn, and Deryk Osthus. Embeddings and Ramsey numbers of sparse k-uniform hypergraphs. *Combinatorica*, 29(3):263–297, 2009.

[11] Nemanja Draganić and Kalina Petrova. Size-Ramsey numbers of graphs with maximum degree three. arXiv:2207.05048.

[12] Andrzej Dudek, Steven La Fleur, Dhruv Mubayi, and Vojtech Rödl. On the size-ramsey number of hypergraphs. *Journal of Graph Theory*, 86(1):104–121, 2017.

[13] Paul Erdős, Ralph J Faudree, Cecil C Rousseau, and Richard H Schelp. The size ramsey number. *Periodica Mathematica Hungarica*, 9(1-2):145–161, 1978.

[14] J. Friedman and N. Pippenger. Expanding graphs contain all small trees. *Combinatorica*, 7(1):71–76, 1987.

[15] Jie Han, Matthew Jenssen, Yoshiharu Kohayakawa, Guilherme Oliveira Mota, and Barnaby Roberts. The multicolour size-Ramsey number of powers of paths. *J. Combin. Theory Ser. B*, 145:359–375, 2020.

[16] Jie Han, Yoshiharu Kohayakawa, Shoham Letzter, Guilherme Oliveira Mota, and Olaf Parczyk. The size-Ramsey number of 3-uniform tight paths. *Adv. Comb.*, pages Paper No. 5, 12, 2021.

[17] P. E. Haxell, Y. Kohayakawa, and T. Ł uczak. The induced size-Ramsey number of cycles. *Combin. Probab. Comput.*, 4(3):217–239, 1995.

[18] Nina Kamcev, Anita Liebenau, David R. Wood, and Liana Yepremyan. The size Ramsey number of graphs with bounded treewidth. *SIAM J. Discrete Math.*, 35(1):281–293, 2021.

[19] Jeong Han Kim and Van H. Vu. Concentration of multivariate polynomials and its applications. *Combinatorica*, 20(3):417–434, 2000.

[20] Y. Kohayakawa. Szemerédi's regularity lemma for sparse graphs. In *Foundations of computational mathematics (Rio de Janeiro, 1997)*, pages 216–230. Springer, Berlin, 1997.

[21] Y. Kohayakawa, V. Rödl, M. Schacht, and E. Szemerédi. Sparse partition universal graphs for graphs of bounded degree. *Advances in Mathematics*, 226(6):5041–5065, 2011.

[22] Shoham Letzter, Alexey Pokrovskiy, and Liana Yepremyan. Size-ramsey numbers of powers of hypergraph trees and long subdivisions. *arXiv preprint arXiv:2103.01942*, 2021.

[23] Rajko Nenadov. *Ramsey and universality properties of random graphs*. PhD thesis, ETH Zürich, 2016.

[24] F. P. Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, s2-30(1):264–286, 1930.

[25] Vojtech Rodl and Endre Szemerédi. On size ramsey numbers of graphs with bounded degree. *Combinatorica*, 20(2):257–262, 2000.

[26] Joel Spencer. Counting extensions. *J. Combin. Theory Ser. A*, 55(2):247–255, 1990.

[27] Konstantin Tikhomirov. On bounded degree graphs with large size-ramsey numbers. arXiv:2210.05818.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Expander graphs, strong blocking sets and minimal codes

(Extended abstract)

Noga Alon* Anurag Bishnoi† Shagnik Das‡ Alessandro Neri§

Abstract

We give a new explicit construction of strong blocking sets in finite projective spaces using expander graphs and asymptotically good linear codes. Using the recently found equivalence between strong blocking sets and linear minimal codes, we give the first explicit construction of $\mathbb{F}_q$-linear minimal codes of length n and dimension k such that n is at most a constant times qk. This solves one of the main open problems on minimal codes.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-003`

1 Introduction

Blocking sets are sets of points in a finite projective or affine space that meet every hyperplane non-trivially. Studying these objects is a classical topic in finite geometry [15,17]. A stronger notion of blocking sets is that of a set of points that meets every hyperplane in a spanning set. For example, in a projective plane, the set of all points on a single line is a blocking set while the set of all points on three non-concurrent lines is a strong blocking set. These special kind of blocking sets have been studied under the names of generating

*Department of Mathematics, Princeton University, United States of America. E-mail: `nalon@math.princeton.edu` . Supported by NSF grant DMS-2154082 and BSF grant 2018267.

†Delft Institute of Applied Mathematics, Delft University of Technology, Netherlands. E-mail: `A.Bishnoi@tudelft.nl`.

‡Department of Mathematics, National Taiwan University, Taiwan. E-mail: `shagnik@ntu.edu.tw`.

§Department of Mathematics: Analysis, Logic and Discrete Mathematics, Ghent University, Belgium. E-mail: `Alessandro.Neri@ugent.be`.

sets [23, 25], cutting blocking sets [1, 12, 16] and strong blocking sets [21, 24]. It is the last terminology that we use in this paper.

Strong blocking sets have recently been shown to be in one-to-one correspondence with the notion of *minimal codes* [1, 32]. Minimal codes are linear subspaces of $\mathbb{F}_q^n$ such that the support of any non-zero vector in the subspace does not contain the support of any other non-zero vector of the subspace as a proper subset. These codes have been studied for their application in decoding algorithms [27] and cryptography [18, 29]. Recently, minimal codes have also been linked to perfect hash families [14], which have important applications in computer science. The main problem is to find minimal codes of dimension k and the shortest possible length n as a function of k and the size of the underlying finite field $\mathbb{F}_q$ [18]. It is known that any strong blocking set in the $(k-1)$-dimensional projective space obtained from $\mathbb{F}_q^k$, denoted by $\mathrm{PG}(k-1, q)$, must have size at least $(q+1)(k-1)$ [3], which implies that any minimal code of length n and dimension k over $\mathbb{F}_q$ must satisfy $n \geq (q+1)(k-1)$. Therefore, we would like to construct minimal codes whose length is at most a constant times qk. It follows from [3, Theorem 2.8] that such a minimal code will also be an asymptotically good error-correcting code, which provides another motivation for the problem. While it is easy to show the existence of such short minimal codes using the probabilistic method (for the best results, see [30] for $q = 2$ and [2, 14] for $q > 2$), it is a challenging and central open problem to give explicit constructions [20]. Many constructions of minimal codes have appeared in the last few years [1, 10, 20, 22, 23], and the current best explicit construction has length $n \sim q^4k/4$ [11, 19].

In this paper, we give a new graph-theoretical construction of strong blocking sets, and thus minimal codes. By using asymptotically good linear codes and constant-degree expander graphs, we obtain an explicit construction of strong blocking sets of size cqk, in the projective space $\mathrm{PG}(k-1, q)$, for an absolute constant c.

A graph parameter known as the (vertex) integrity of a graph plays a crucial role in our construction. We prove a new lower bound on the vertex integrity of d-regular graphs in terms of their eigenvalues. Our lower bound implies that any expander graph of bounded degree on n vertices has vertex integrity at least a constant times n. We combine explicit constructions of such graphs with explicit constructions of asymptotically good linear codes, to get explicit minimal codes.

There is a rich history of using expander graphs to construct asymptotically good linear codes [6, 31, 33]. Our work contributes to this line of research by using these graphs in a novel way to construct (asymptotically good) minimal codes. Our construction is the first of its kind in finite geometry as it uses graphs to pick a subset of lines in a finite projective space whose union has certain intersection properties with hyperplanes. This construction has already led to explicit constructions of small affine blocking sets [14], and we expect that it will lead to many new results in finite geometry.

2 Preliminaries

Definition 2.1. The (Hamming) support of a vector $v \in \mathbb{F}_q^n$ is the set $\sigma(v) := \{i : v_i \neq 0\} \subseteq [n]$. The (Hamming) weight of v is $\mathrm{wt}(v) := |\sigma(v)|$.

Definition 2.2. An $[n, k, d]_q$ code $\mathcal{C}$ is a k-dimensional subspace of $\mathbb{F}_q^n$, with $d := \min\{\mathrm{wt}(v) : v \in \mathcal{C} \setminus \{\vec{0}\}\}$ is called the minimum distance of $\mathcal{C}$. The elements of $\mathcal{C}$ are called codewords. Moreover, a generator matrix for $\mathcal{C}$ is a matrix $G \in \mathbb{F}_q^{k\times n}$ such that $\mathcal{C} = \{uG : u \in \mathbb{F}_q^k\}$.

Definition 2.3. Let $\{n_i\}_{i\geq 1}$ be an increasing sequence of positive numbers and suppose that there exist sequences $\{k_i\}_{i\geq 1}$ and $\{d_i\}_{i\geq 1}$ such that for all $i \geq 1$ there exists an $[n_i, k_i, d_i]_q$ code $\mathcal{C}_i$. Then the sequence $\{\mathcal{C}_i\}_{i\geq 1}$ is called an $(R, \delta)_q$-family of codes, where the rate R of this family is defined as $R = \liminf_{i\to\infty} \frac{k_i}{n_i}$, and the relative distance δ is defined as $\delta = \liminf_{i\to\infty} \frac{d_i}{n_i}$.

One of the central problems on error-correcting codes is to understand the trade-off between the rate and the relative distance of codes. A family of codes for which $R > 0$ and $\delta > 0$, is known as an *asymptotically good code*. An easy probablistic argument known as the Gilbert-Varshamov bound shows the existence of such codes for every $\delta \in [0, 1 - 1/q)$ and $R = 1 - H_q(\delta)$, where $H_q(x) := x \log_q(q - 1) - x \log_q(x) - (1 - x) \log_q(1 - x)$, is the q-ary entropy function, defined on the domain $0 \leq x \leq 1 - 1/q$. The first explicit construction of asymptotically good codes was given by Justesen [28], who showed that for every $0 < R < 1/2$, there is an explicit family of codes with rate R and relative distance $\delta \geq (1 - 2R)H_q^{-1}\left(\frac{1}{2}\right)$. Note that for any prime power q, $H_q^{-1}\left(\frac{1}{2}\right) \geq H_2^{-1}\left(\frac{1}{2}\right) > 0.11$, and thus there are absolute constants $R, \delta > 0$, not depending on q, for which we have an explicit construction of a family of $\mathbb{F}_q$-linear codes with rate R and relative distance δ.

Definition 2.4. Let $\mathcal{C}$ be an $[n, k, d]_q$ code. A nonzero codeword $v \in \mathcal{C}$ is said to be minimal (in $\mathcal{C}$) if $\sigma(v)$ is minimal with respect to the inclusion in the set $\sigma(\mathcal{C}) := \{\sigma(u) : u \in \mathcal{C}\setminus\{\vec{0}\}\}$. The code $\mathcal{C}$ is a minimal linear code if all its nonzero codewords are minimal.

For $k > 1$, the finite projective space of dimension $k - 1$ over the finite field $\mathbb{F}_q$ is defined as $\mathrm{PG}(k - 1, q) := \left(\mathbb{F}_q^k \setminus \{\bar{0}\}\right) / \sim$, where $u \sim v$ if $u = \lambda v$ for some non-zero $\lambda \in \mathbb{F}_q$ (in some circles the same object will be denoted by $\mathbb{P}^{k-1}(\mathbb{F}_q)$). The equivalence class that a non-zero vector v belongs to is denoted by $[v]$. The 1-dimensional, 2-dimensional, ..., $(k-1)$-dimensional vector subspaces of $\mathbb{F}_q^k$ correspond to the points, lines, ..., hyperplanes of $\mathrm{PG}(k-1, q)$. We denote the span of a subset S of points in a projective space by $\langle S\rangle$ and the dimension $\dim(\langle S\rangle)$ is one less than the vector space dimension of the corresponding vector subspace. For example, the span of two distinct points P, Q in a projective space, which we will also denote by $\langle P, Q\rangle$, is a 1-dimensional projective subspace corresponding to a 2-dimensional vector subspace, and we refer to it as the line joining P and Q in $\mathrm{PG}(k - 1, q)$.

Definition 2.5. A projective $[n, k, d]_q$ system is a (multi)set of n points, $\mathcal{M} \subseteq \mathrm{PG}(k-1, q)$, such that $\langle\mathcal{M}\rangle = \mathrm{PG}(k - 1, q)$ and $d = n - \max\{|H \cap \mathcal{M}| : H \text{ is a hyperplane}\}$.

A projective $[n,k,d]_q$ system is simply a dual interpretation of a nondegenerate $[n,k,d]_q$ code, that is, codes with no identically zero entry in all the codewords. If G is the $k \times n$ generator matrix of the code, then the columns of G correspond to a multiset of n points in $\mathrm{PG}(k-1,q)$ with the property that the maximum intersection with a hyperplane of this multiset is equal to $n-d$. This process can clearly be reversed.

Definition 2.6. A set $\mathcal{M} \subseteq \mathrm{PG}(k-1,q)$ is said to be a strong blocking set if $\langle H \cap \mathcal{M} \rangle = H$, for every hyperplane H of $\mathrm{PG}(k-1,q)$.

Theorem 2.7 (see [1], [32])**.** Let $\mathcal{C}$ be a nondegenerate $[n,k,d]_q$ code and let $G = (g_1 \mid \ldots \mid g_n) \in \mathbb{F}_q^{k \times n}$ be any of its generator matrices. The following are equivalent:

1. $\mathcal{C}$ is a minimal code;
2. $\mathcal{M} = \{[g_1], \ldots, [g_n]\}$ is a strong blocking set in $\mathrm{PG}(k-1,q)$.

All known explicit constructions of strong blocking sets are obtained as union of lines in the projective space. This is mainly due to the fact that with such a structure it is easy to control their intersections with subspaces. In particular, the main feature that these constructions possess is the following stronger property than being a strong blocking set.

Definition 2.8. A set $\mathcal{L}$ of lines in a projective space satisfies the avoidance property if there is no codimension-2 space meeting every line $\ell \in \mathcal{L}$.

The relation between these sets of lines and strong blocking sets is the observation of Fancsali and Sziklai [23, Theorem 11] that if a set $\mathcal{L}$ of lines satisfying the avoidance property, then the point-set $\mathcal{B} := \cup_{\ell \in \mathcal{L}} \ell$ is a strong blocking set.

For our explicit construction of strong blocking sets we will need explicit constructions of constant-degree expander graphs. Informally, expander graphs have the property that for any vertex subset which is not too large, its boundary is at least a constant times its size. Expansion in graphs can be measured by their spectral properties (see [26]). For a graph G we denote the eigenvalues of its adjacency matrix by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. If G is d-regular, then $\lambda_1 = d$. Moreover, if it is also connected then $\lambda_2 < d$. A graph G is called an (n,d,λ)-graph if it is a d-regular graph on n vertices with $|\lambda_i| \leq \lambda$ for all $i > 1$. The smaller the value of λ, the larger is the expansion of an (n,d,λ)-graph. Asymptotically, the smallest possible value is close to $2\sqrt{d-1}$, and the graphs achieving that bound are known as Ramanujan graphs. We will use the following result of Alon on explicit constructions of almost Ramanujan graphs.

Theorem 2.9 (see [5, Theorem 1.3])**.** For every positive integer d, and every $\varepsilon > 0$, there is an $n_0(d,\varepsilon)$, such that for all $n \geq n_0(d,\varepsilon)$, with nd even, there is an explicit construction of an (n,d,λ)-graph $G^{\varepsilon}_{n,d}$ with $\lambda \leq 2\sqrt{d-1} + \varepsilon$

3 Integrity of a graph

We will need the following graph parameter, known as the (vertex) *integrity* of a graph, which was introduced in the late 1980s as a measure of the robustness of the connectivity of a network under vertex deletion [7, 9].

Definition 3.1. Let $G = (V, E)$ be a simple connected graph. For any subgraph H, let $\kappa(G)$ denote the largest size of a connected component in H. The integrity of G is the integer

$$\iota(G) := \min_{S \subseteq V} (|S| + \kappa(G - S)) .$$

It is a challenging problem to determine the integrity of graphs precisely, or even asymptotically (see [7] for an old survey and [8, 13] for some recent bounds on different families of graphs). We prove new lower bounds on the vertex integrity of (n, d, λ)-graphs. First, we relate the integrity of a graph to another graph parameter.

Definition 3.2. For a graph G, let $z(G)$ denote the largest integer z so that there are two disjoint sets of vertices in G, each of size z, with no edge connecting them.

Proposition 3.3. For every graph $G = (V, E)$ on n vertices,

$$n - 2z(G) \leq \iota(G) \leq n - z(G).$$

Theorem 3.4. For any (n, d, λ)-graph G, we have $\iota(G) \geq \left(\frac{d-\lambda}{d+\lambda}\right) n$.

Proof. Let $z(G)$ be the maximum integer z such that the vertices of a graph G contains two disjoint parts of size z each with no edge between them. A direct application of the expander mixing lemma implies that

$$z(G) \leq \frac{\lambda n}{d + \lambda}.$$

Applying the lower bound $\iota(G) \geq n - 2z(G)$ from Proposition 3.3, implies $\iota(G) \geq n - 2\frac{\lambda}{d+\lambda}n = \frac{d-\lambda}{d+\lambda}n$. □

4 Constructing Strong Blocking Sets from Graphs

Definition 4.1. Let $\mathcal{M} = \{P_1, \ldots, P_n\}$ be a set of n points in $\mathrm{PG}(k-1, q)$ and let $G = (\mathcal{M}, E)$ be a graph with vertex set equal to $\mathcal{M}$. We define the following sets of lines

$$\mathcal{L}(\mathcal{M}, G) := \{\langle P_i, P_j \rangle : P_i P_j \in E\}$$

and the following set of points

$$\mathcal{B}(\mathcal{M}, G) := \bigcup_{\ell \in \mathcal{L}(\mathcal{M}, G)} \ell,$$

obtained from $\mathcal{M}$ and G.

We make the following crucial observation relating the properties of the graph G and the projective sets defined above.

Proposition 4.2. Let $\mathcal{M} = \{P_1, \dots, P_n\}$ be a set of points in $\mathrm{PG}(k-1, q)$ and let $G = (\mathcal{M}, E)$ be a graph whose set of vertices is $\mathcal{M}$. If for every $S \subseteq \mathcal{M}$ there exists a connected component C in $G - S$ such that $\langle S \cup C \rangle = \mathrm{PG}(k-1, q)$, then the set $\mathcal{L}(\mathcal{M}, G) = \{\langle P_i, P_j \rangle : P_iP_j \in E\}$ satisfies the avoidance property, that is, no codimension-2 subspace of $\mathrm{PG}(k-1, q)$ meets every line of $\mathcal{L}(\mathcal{M}, G)$.

Lemma 4.3. Let $\mathcal{M}$ be a projective $[n, k, d]_q$ system and let $G = (\mathcal{M}, E)$ be a graph such that $\iota(G) \geq n - d + 1$. Then $\mathcal{L}(\mathcal{M}, G)$ satisfies the avoidance property, and thus $\mathcal{B}(\mathcal{M}, G)$ is a strong blocking set in $\mathrm{PG}(k-1, q)$ of size at most $n + (q-1)|E|$.

Proof. Let S be an arbitrary subset of $\mathcal{M}$. Since $\iota(G) \geq n-d+1$, there exists a connected component C in $G - S$ such that $|S| + |C| \geq n - d + 1$. From the definition of projective systems, it follows that every hyperplane meets $\mathcal{M}$ in at most $n - d$ points. Therefore, $S \cup C \subseteq \mathcal{M}$ is not contained in any hyperplane of $\mathrm{PG}(k-1, q)$, thus implying $\langle S \cup C \rangle = \mathrm{PG}(k-1, q)$. From Proposition 4.2, we conclude that $\mathcal{L}(\mathcal{M}, G)$ satisfies the avoidance property, and thus $\mathcal{B}(\mathcal{M}, G)$ is a strong blocking set. Each line in $\mathcal{L}(\mathcal{M}, G)$ contains exactly $q + 1$ points, of which at most $q - 1$ are non-vertices. As there are $|E|$-many lines in this set, we get $|\mathcal{B}(\mathcal{M}, G)| \leq n + (q-1)|E|$. □

Finally, we prove the main result of our paper.

Theorem 4.4. There is an absolute constant c such that for every prime power q, there is an explicit construction of strong blocking sets of size at most cqk_i in $\mathrm{PG}(k_i - 1, q)$, for some increasing infinite sequence $\{k_i\}_{i \in \mathbb{N}}$.

Proof. Let R be any constant satisfying $0 < R < 1/2$ and let $\delta = (1-2R)0.11$. Let $\mathcal{M}_i$ be projective $[n_i, k_i, d_i]_q$ systems given by the Justesen construction [28]. Then $\lim_{i\to\infty} k_i/n_i = R$ and $\lim_{i\to\infty} d_i/n_i \geq (1-2R)H_q^{-1}(1/2) > \delta$. Therefore, there exists an i_0 such that for all $i \geq i_0$, we have $d_i/n_i \geq \delta$ and $k_i/n_i \geq R/2$. For the rest of this proof let i_0 be large enough. Let $\{G_i\}_{i \geq i_0}$ be an explicit family of (n_i, d, λ)-graphs, where d and λ are positive constants for which $(d-\lambda)/(d+\lambda) \geq 1 - \delta + 1/n_i$. From Theorem 2.9, it follows that such an explicit construction of graphs is always possible. By Theorem 3.4, we have $\iota(G_i) \geq (1-\delta)n_i + 1 \geq n_i - d_i + 1$. Therefore, by Lemma 4.3, $\mathcal{B}(\mathcal{M}_i, G_i)$ is a strong blocking set in $\mathrm{PG}(k_i - 1, q)$ of size at most

$$n_i + (q-1)\frac{dn_i}{2} < \frac{d}{2}qn_i \leq \frac{d}{R}qk_i.$$

This concludes the proof with $c = \frac{d}{R}$. □

In the expanded version of this short abstract [4], we obtain the optimal value of the constant c by using algebraic-geometric codes, and in particular, we show that we can take $c = 20$ for large enough q. Moreover, for any fixed $q \geq 7$, and $k \to \infty$, we show that our explicit construction is better than [11].

References

[1] G. N. Alfarano, M. Borello, and A. Neri. A geometric characterization of minimal codes and their asymptotic performance. *Advances in Mathematics of Communications*, 16(1):115–133, 2022.

[2] G. N. Alfarano, M. Borello, and A. Neri. Outer strong blocking sets. *preprint arXiv:2301.09590*, 2023.

[3] G. N. Alfarano, M. Borello, A. Neri, and A. Ravagnani. Three combinatorial perspectives on minimal codes. *SIAM Journal on Discrete Mathematics*, 36(1):461–489, 2022.

[4] N. Alon, A. Bishnoi, S. Das, and A. Neri. Strong blocking sets and minimal codes from expander graphs. *arXiv:2305.15297*, 2023.

[5] Noga Alon. Explicit expanders of every degree and size. *Combinatorica*, pages 1–17, 2021.

[6] Noga Alon, Jehoshua Bruck, Joseph Naor, Moni Naor, and Ron M Roth. Construction of asymptotically good low-rate error-correcting codes through pseudo-random graphs. *IEEE Transactions on Information Theory*, 38(2):509–516, 1992.

[7] K. S. Bagga, L. W. Beineke, W. D. Goddard, M. J. Lipman, and R. E. Pippert. A survey of integrity. *Discrete Applied Mathematics*, 37:13–28, 1992.

[8] József Balogh, Tamás Mészáros, and Adam Zsolt Wagner. Two results about the hypercube. *Discrete Applied Mathematics*, 247:322–326, 2018.

[9] Curtis A Barefoot, Roger Entringer, and Henda Swart. Vulnerability in graphs-a comparative survey. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 1(38):13–22, 1987.

[10] D. Bartoli and M. Bonini. Minimal linear codes in odd characteristic. *IEEE Transactions on Information Theory*, 65(7):4152–4155, 2019.

[11] Daniele Bartoli and Martino Borello. Small strong blocking sets by concatenation. *SIAM Journal on Discrete Mathematics*, 37(1):65–82, 2023.

[12] Daniele Bartoli, Antonio Cossidente, Giuseppe Marino, and Francesco Pavese. On cutting blocking sets and their codes. *Forum Mathematicum*, 34(2):347–368, 2022.

[13] D Benko, C Ernst, and Dominic Lanphier. Asymptotic bounds on the integrity of graphs and separator theorems for graphs. *SIAM Journal on Discrete Mathematics*, 23(1):265–277, 2009.

[14] A. Bishnoi, J. D'haeseleer, D. Gijswijt, and A. Potukuchi. Blocking sets, minimal codes and trifferent codes. *arXiv:2301.09457*, 2023.

[15] A. Blokhuis, P. Sziklai, and T. Szonyi. Blocking sets in projective spaces. *Current research topics in Galois geometry*, pages 61–84, 2011.

[16] Matteo Bonini and Martino Borello. Minimal linear codes arising from blocking sets. *Journal of Algebraic Combinatorics*, 53:327–341, 2021.

[17] A. E Brouwer and A. Schrijver. The blocking number of an affine space. *Journal of Combinatorial Theory, Series A*, 24(2):251–253, 1978.

[18] H. Chabanne, G. Cohen, and A. Patey. Towards secure two-party computation from the wire-tap channel. In *International Conference on Information Security and Cryptology*, pages 34–46. Springer, 2013.

[19] G. Cohen, S. Mesnager, and H. Randriam. Yet another variation on minimal linear codes. *Advances in Mathematics of Communications*, 10(1):53–61, 2016.

[20] G. D. Cohen, S. Mesnager, and A. Patey. On minimal and quasi-minimal linear codes. In *IMA International Conference on Cryptography and Coding*, pages 85–98. Springer, 2013.

[21] A. A. Davydov, M. Giulietti, S. Marcugini, and F. Pambianco. Linear nonbinary covering codes and saturating sets in projective spaces. *Advances in Mathematics of Communications*, 5(1):119–147, 2011.

[22] C. Ding. Linear codes from some 2-designs. *IEEE Transactions on Information Theory*, 61(6):3265–3275, 2015.

[23] S. Fancsali and P. Sziklai. Lines in higgledy-piggledy arrangement. *Electronic Journal of Combinatorics*, 21, 2014.

[24] T. Héger and Z. L. Nagy. Short minimal codes and covering codes via strong blocking sets in projective spaces. *IEEE Transactions on Information Theory*, 68(2):881–890, 2021.

[25] T. Héger, B. Patkós, and M. Takáts. Search problems in vector spaces. *Designs, Codes and Cryptography*, 76(2):207–216, 2015.

[26] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.

[27] Tai-Yang Hwang. Decoding linear block codes for minimizing word error rate (corresp.). *IEEE Transactions on Information Theory*, 25(6):733–737, 1979.

[28] Jørn Justesen. Class of constructive asymptotically good algebraic codes. *IEEE Transactions on Information Theory*, 18(5):652–656, 1972.

[29] J. L. Massey. Minimal codewords and secret sharing. In *Proceedings of the 6th joint Swedish-Russian international workshop on information theory*, pages 276–279, 1993.

[30] D Miklós. Linear binary codes with intersection properties. *Discrete Applied Mathematics*, 9(2):187–196, 1984.

[31] Michael Sipser and Daniel A Spielman. Expander codes. *IEEE Transactions on Information Theory*, 42(6):1710–1722, 1996.

[32] C. Tang, Y. Qiu, Q. Liao, and Z. Zhou. Full characterization of minimal linear codes as cutting blocking sets. *IEEE Transactions on Information Theory*, 67(6):3690–3700, 2021.

[33] R Tanner. A recursive approach to low complexity codes. *IEEE Transactions on Information Theory*, 27(5):533–547, 1981.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Moderate Deviations of Triangle Counts – the lower tail

(Extended abstract)

José D. Alvarado* Gabriel D. do Couto† Simon R. Griffiths‡

Abstract

Two recent papers [11] and [19] study the lower tail of triangle count deviations in random graphs $G(n,m)$ with positive density $t := m/\binom{n}{2} \in (0,1)$. Let us write $D_{\triangle}(G)$ for the deviation of the triangle count from its mean. Results of [11] and [19] determine the order of magnitude of the log probability $\log(\mathbb{P}\left(D_{\triangle}(G(n,m)) < -\tau\binom{n}{3}\right))$ for the ranges $n^{-3/2} \ll \tau \ll n^{-1}$ and $n^{-3/4} \ll \tau \ll 1$ respectively. Furthermore, in [19] it is proved that the log probability is at least $\Omega(\tau^2 n^3)$ in the "missing" range $n^{-1} \ll \tau \ll n^{-3/4}$, and they conjectured that this result gives the correct order of magnitude. Our main contribution is to prove this conjecture.

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-004

1 Introduction

The study of subgraph count deviations, and especially triangle count deviations has been a very active area of research in recent decades. In particular, a great many results have been proved regarding small deviations (of the order of the standard deviation) beginning

*Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brazil. E-mail: josealvarado.mat17@ime.usp.br. Supported by FAPESP (Proc. 2020/10796-0).

†Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brazil E-mail: gabrilord@gmail.com. Supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

‡Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brazil E-mail: simon@mat.puc-rio.br. Supported by CNPq (Proc. 307521/2019-2) and FAPERJ (Proc. E-26/202.713/2018 and Proc. E-26/201.194/2022).

with Ruciński [22], see also [2, 14, 13, 15, 17, 20, 21]. There have also been many results which focus on large deviations (of the order of the mean) including the seminal articles of Vu [23] and Janson and Ruciński [16] in the early 2000s, and continuing with Chatterjee and Varadhan [4] who related these deviations to solutions to variational problems, which were resolved in certain cases by Lubetzky and Zhao [18] and Zhao [24]. The survey of Chatterjee [3] and the references therein give a detailed overview. Further developments related to these techniques may be found in [1, 5, 9]. A major breakthrough by Harel, Mousset and Samotij [12] essentially resolved the large deviation (upper tail) problem for triangles.

There has also been some interest in deviations of intermediate value, which we call moderate deviations. These deviations are considered in the $G(n,p)$ model in [7, 8, 10]. It is argued by the third author, together with Goldschmidt and Scott [11] that for many moderate deviation problems the $G(n,m)$ model is more appropriate as it is possible to study finer causes of deviatons, and that, in any case, one may deduce results for $G(n,p)$ by a simple conditioning argument. See also [6], which extends these results to sparser random graphs.

Let us now consider the model $G_m \sim G(n,m)$, in which G_m is selected uniformly from graphs with n vertices and m edges. Suppose that $t \in (0,1)$ is fixed and that our random graphs have density t, that is $t = m/\binom{n}{2}$. Let $N_\triangle(G)$ be the number of triangles in the graph G and let $D_\triangle(G_m)$ be the deviation of the triangle count in G_m, i.e., $D_\triangle(G_m) := N_\triangle(G_m) - \mathbb{E}\left[N_\triangle(G_m)\right]$.

We also remark that the majority of results previously mentioned have focussed on the upper tail, whereas we shall focus on the lower tail. That is, we consider the question of how likely it is that a random graph has many *fewer* triangles than expected.

By the main results of [11] and Neeman, Radin and Sadun [19] respectively, we have

$$-\log\left[\mathbb{P}\left(D_\triangle(G_m) \,<\, -\tau\binom{n}{3}\right)\right] \,=\, \begin{cases}\Theta(\tau^2 n^3) & n^{-3/2} \ll \tau \ll n^{-1}\\ \Theta(\tau^{2/3} n^2) & n^{-3/4} \ll \tau \ll 1\end{cases}$$

Furthermore, Neeman, Radin and Sadun [19] obtained the bounds

$$\exp(-C\tau^{2/3}n^2) \,\leqslant\, \mathbb{P}\left(D_\triangle(G_m) \,<\, -\tau\binom{n}{3}\right) \,\leqslant\, \exp(-c\tau^2 n^3)$$

in the "missing" range $n^{-1} \ll \tau \ll n^{-3/4}$, for some constants $c, C > 0$. They conjectured that the final quantity is the correct probability of this deviation, up to the choice of the constant C. We prove this conjecture, thus completing the understanding of the order of magnitude of deviations in the lower tail across essentially the entire range of possible deviations.

Theorem 1. *Let $t \in (0,1)$. There exists a constant $c > 0$ such that the following holds. Suppose that n is sufficiently large and $c^{-1}n^{-1} \leqslant \tau \leqslant cn^{-3/4}$ then*

$$\mathbb{P}\left(D_\triangle(G_m) \,<\, -\tau\binom{n}{3}\right) \,\geqslant\, \exp(-c\tau^2 n^3)\,.$$

2 Preliminaries

As we are claiming a lower bound on the deviation probability we must justify that there is a certain reasonably likely "cause" of this deviation. In works that consider the upper tail this is often a fixed subgraph (such as a clique or hub) which occurs with a certain probability. As Neeman, Radin and Sadun [19] discovered, the situation is more subtle for the lower tail. In the range of slightly larger deviations they showed that the likeliest "cause" corresponds to a deviation event of the smallest eigenvalue.

We shall give a quite different "cause" of the triangle deficit. Roughly speaking, we consider running the majority of the process, and then, near the end, we select a certain set of pairs (non-edges) which have *small* codegree. If in the rest of the process we select many more of these pairs than expected then this causes a deficit of triangles in the final graph G_m. We show that this cause has a cost of $\exp\left(-\Theta(\tau^2 n^3)\right)$, thus proving Theorem 1.

In fact, to implement this approach we have to be slightly more careful about the set of pairs of low co-degree, which we will call F_-. It will be useful that F_- is close to regular. We therefore introduce a concept of synergy, which we use instead of codegree when defining F_-.

Notation

We write $d_u(G)$ for the degree of a vertex u in G and $d_{uv}(G)$ for the codegree of the pair u, v.

We now define *synergy*. The synergy of u and v with respect to G is

$$Syn_{uv}(G) := d_{uv}(G) - td_u(G) - td_v(G) + t^2(n-2).$$

The synergy of a pair of vertices can be thought of as how well their neighbourhoods intersect. As we are dealing with well behaved graphs, i.e. graphs with high probability properties, if a pair has positive synergy, then a high proportion of the neighbourhoods of its vertices intersect, and the opposite is true for negative synergy.

As we said, our proof involves revealing G_m into two parts, which we call G_m^0 and G_m^1. The first part will correspond to the first $m_0 := (1-\eta)m$ edges added to G_m, where $\eta \in (0,1)$. We note that $G_m^0 \sim G(n,(1-\eta)m)$, and we shall assume at various points in the proof that G_m^0 has the standard properties which hold with high probability in such random graphs.

The negative deviation of triangles will come with the selection of G_m^1. We note that G_m^1 corresponds to the last $m_1 := \eta m$ edges of the random process.

Let (f_i) be the sequence of non-edges of G_0 with non-decreasing order of synergies. The set of non-edges of low synergy is

$$F_- := \left\{ f_i : i \in \left\{ 1, \ldots, \frac{\binom{n}{2} - m_0}{2} \right\} \right\}$$

and the set of non-edges of high synergy is

$$F_+ := \left\{ f_i : i \in \left\{ \frac{\binom{n}{2} - m_0}{2} + 1, \ldots, \binom{n}{2} - m_0 \right\} \right\}.$$

Discussion of our approach

Let us first ask: What is the expected value of $|F_- \cap E(G_m^1)|$? Since F_- and F_+ have the same size, we have that $\mathbb{E}\left[|F_- \cap E(G_m^1)|\right] = \mathbb{E}\left[|F_+ \cap E(G_m^1)|\right] = \frac{m_1}{2} = \frac{\eta}{2} m$. Moreover, the same equalities holds if we replace the expectation by the conditional expectation (given G_m^0). With this in mind, we define the event

Definition 2 (The Event $\mathcal{E}(\alpha)$)**.** Let $\alpha \in (0,1)$ be a parameter. We denote by $\mathcal{E}(\alpha)$ the event defined by "$|F_- \cap E(G_m^1)| = (1+\alpha)m_1/2$".

Note that this event relies on first revealing G_m^0, as F_- is defined as a function of G_m^0. Since pairs of low synergy tend to have smaller codegree, there ought to be a relation between selecting more edges of G_m^1 in F_- and a deficit of triangles in the final random graph G_m. This idea is central to our approach.

3 Main result

As we will simply provide a proof overview here, some of the details will be left somewhat vague. For example, it is useful to have a graph property $\mathcal{P}_0$ such that $\mathbb{P}_{G_0}(\mathcal{P}_0) = 1 + o(1)$. This graph property consist of various properties which hold with high probability in random graphs. We note that, by monotony and conditioning, and the fact that $\mathcal{E}(\alpha)$ is independent of G_m^0 (and so also $\mathcal{P}_0$) we have

$$\begin{aligned}&\mathbb{P}\left(N_\triangle(G) < \mathbb{E}\left[N_\triangle(G)\right] - a\right) \geqslant \\ &\quad (1+o(1))\,\mathbb{P}\left(\mathcal{E}(\alpha)\right)\,\mathbb{P}\big(N_\triangle(G) < \mathbb{E}\left[N_\triangle(G)\right] - a \,\big|\, \mathcal{E}(\alpha), \mathcal{P}_0\big).\end{aligned}$$

Given this inequality, it suffices to prove the following two lemmas:

Lemma 3. *Let $\alpha := \alpha_n$ with $n^{-1} \ll \alpha_n \ll n^{-1/4}$. Then*

$$\mathbb{P}\left(\mathcal{E}(\alpha)\right) \geqslant \exp\left(-O_{t,\eta}(\alpha^2 n^2)\right). \tag{3.1}$$

Lemma 4. *There exists $C > 0$ such that the following holds. If n is sufficiently large and $\alpha n^{5/2} \geqslant Ca$ then*

$$\mathbb{P}\big(N_\triangle(G) < \mathbb{E}\left[N_\triangle(G)\right] - a \,\big|\, \mathcal{E}(\alpha), \mathcal{P}_0\big) = n^{-O_{t,\eta}(1)} = e^{o(n)}. \tag{3.2}$$

We remark that Lemma 3 follows easily from known bounds on the tail of the hypergeometric distribution. The proof of Lemma 4 is more involved. However, the result clearly follows from the following statement (and Markov's inequality)

$$\mathbb{E}\left[N_{\triangle}(G) \mid \mathcal{E}(\alpha), \mathcal{P}_0\right] \leqslant \mathbb{E}\left[N_{\triangle}(G)\right] - 2a\,. \tag{3.3}$$

To prove (3.3) we consider the various types of triangle which occur in the final graph. We divide the triangle count into four categories: three edges from G_m^0, $\triangle_{(3,0)}$, two edges from G_m^0 and one from G_m^1, $\triangle_{(2,1)}$, one edges from G_m^0 and two from G_m^1, $\triangle_{(1,2)}$ and three edges from G_m^1, $\triangle_{(0,3)}$. The idea is that the number of edges of type $(3,0)$ is predictable, as G_m^0 is a random graph; the number of type $(2,1)$ is significantly *less* than one would expect, because we are conditioning on the event $\mathcal{E}(\alpha)$; and we shall prove that the conditioning does not change the expected number of types $(1,2)$ and $(0,3)$ by very much.

The following result states that conditioning on $\mathcal{E}(\alpha)$ does indeed have the effect of reducing the expected number of triangles of type $(2,1)$. The result mentions μ_- and μ_+ which are defined to be the average codegree of pairs in F_- and F_+ respectively. One of the properties in $\mathcal{P}_0$ is that $\mu_+ - \mu_- = \Omega(n^{1/2})$.

Lemma 5.

$$\mathbb{E}\left[\triangle_{(2,1)} | \mathcal{P}_0, \mathcal{E}_\eta(\alpha)\right] = \mathbb{E}_{G_m^0}[\triangle_{(2,1)}] + \alpha m \eta(\mu_- - \mu_+)$$

The alert reader may question why we chose to define F_- in terms of synergy, rather than simply taking F_- to be those pairs with smaller codegree. Indeed, Lemma 5 would work just as well with this alternative definition. However, the problem arises when trying to control the effect that conditioning (on $\mathcal{E}(\alpha)$) has on the expected number of triangles of types $(1,2)$ and $(0,3)$. We are able to prove sufficiently strong bounds

$$\begin{aligned} \mathbb{E}\left[\triangle_{(1,2)} | \mathcal{P}_0, \mathcal{E}_\eta(\alpha)\right] &= O(\alpha\eta m n^{1/2}) \quad \text{and} \\ \mathbb{E}\left[\triangle_{(0,3)} | \mathcal{P}_0, \mathcal{E}_\eta(\alpha)\right] &= O(\alpha\eta m n^{1/2}) \end{aligned}$$

using the fact that F_- is close to regular. (These bounds are seen to be sufficient by taking the constant η sufficiently small.) We remark that our proofs of these bounds rely on the fact that F_- is close to regular, and so would fail if F_- was defined directly in terms of codegrees.

In order to prove F_- is close to regular we actually prove the following stronger statement about the set of synergies $Syn_{uw}(G_m^0) : w \in V \setminus N_u$ between a fixed vertex u and its non-neighbours w. Let σ denote the standard deviation of $Syn_{uw}(G_m^0)$ (which is of order $n^{1/2}$). The following result states an approximate central limit theorem for the empirical distribution of synergies.

Lemma 6. *There exists a constant $C > 0$ such that, with high probability the following holds simultaneously for all vertices $u \in V(G_m^0)$:*

$$\left|\{w \in V \setminus N_u \,:\, Syn_{uw}(G_m^0) \leqslant a\sigma\}\right| = \left(\Phi(a) \pm Cn^{-1/4}\right)(n - d_u(G_m^0))\,.$$

Using this approximation of the distribution of synergies together with other concentration bounds and tools such as Goodman's theorem it is possible to control the expected number of triangles of type $(1,2)$ and $(0,3)$ and thereby prove (3.3).

4 Remarks

As we said, Neeman, Radin and Sadun [19] showed that their construction for the lower bound holded even for the missing range. Their construction involved partitioning the graph into two parts in which the smaller part (much smaller than the other) have a lower density of edges. Hence, they focused on partitioning the vertex and we, instead, partitioned the edges, which should be more effective for $G(n, m)$ type of graphs since their structure is more rigid.

There are a number of questions which remain open. For example, is it possible to extend these results, and the results of [19] to sparser random graphs, as [6] did with the results of [11]. One may also ask whether stronger bounds may be proved. Perhaps it is possible to determine the log probability asymptotically, rather than up to a multiplicative constant. It seems possible to divide our construction further into finer steps with scaling tendencies towards low synergy pairs to get optimal results.

Finally, it would be interesting to investigate other graphs. We remark that the results we prove here correspond to a regime which simply doesn't exist for odd cycles of length at least 5. Surprisingly [19] showed that the log probability exhibits a large discontinuity when considering odd cycles of length at least 5.

References

[1] F. Augeri. Nonlinear large deviation bounds with applications to traces of Wigner matrices and cycles counts in Erdős-Rényi graphs. *arXiv:*, (1810.01558), 2018.

[2] A.D. Barbour, M. Karoński, and A. Ruciński. A central limit theorem for decomposable random variables with applications to random graphs. *J. Combin. Theory Ser. B*, 47(2):125–145, 1989.

[3] S. Chatterjee. An introduction to large deviations for random graphs. *Bull. Amer. Math. Soc.*, 53(4):617–642, 2016.

[4] S Chatterjee and S.R.S. Varadhan. The large deviation principle for the Erdős-Rényi random graph. *European J. Combin.*, 32(7):1000–1017, 2011.

[5] N. Cook and A. Dembo. Large deviations of subgraph counts for sparse Erdős-Rényi graphs. *Adv. Math.*, 343, 2020.

[6] L. G. de Oliveira. Moderate deviations of triangle counts in sparse random graphs. *Doctoral thesis*, https://www.maxwell.vrac.puc-rio.br/61161/61161.PDF.

[7] H. Döring and P. Eichelsbacher. Moderate deviations in a random graph and for the spectrum of bernoulli random matrices. *Electron. J. Probab.*, 14:2636–2656, 2009.

[8] H. Döring and P. Eichelsbacher. Moderate deviations via cumulants. *J. Theoret. Probab.*, 2:360–385, 2013.

[9] R. Eldan. Gaussian-width gradient complexity, reverse log-sobolev inequalities and nonlinear large deviations. *Geom. Funct. Anal.*, 28:1548–1596, 2018.

[10] V. Féray, P.-L. Méliot, and A. Nikeghbali. Mod-ϕ convergence: Normality zones and precise deviations. *Springer Briefs in Probability and Mathematical Statistics*, 2016.

[11] C. Goldschmidt, S. Griffiths, and A. Scott. Moderate deviations of subgraph counts in the Erdős-Rényi random graphs $g(n, m)$ and $g(n, p)$. *Trans. Amer. Math. Soc.*, 343:5517–5585, 2020.

[12] M. Harel, F. Mousset, and W. Samotij. Upper tails via high moments and entropic stability. *arXiv:*, (1904.08212), 2019.

[13] S. Janso. Orthogonal decompositions and functional limit theorems for random graph statistics. *Mem. Amer. Math. Soc.*, (534), 1994.

[14] S. Janson. A functional limit theorem for random graphs with applications to subgraph count statistics. *Random Structures Algorithms*, 1:15–37, 1990.

[15] S. Janson and K. Nowicki. The asymptotic distributions of generalized U-statistics with applications to random graphs. *Probab. Theory Related Fields*, 90(3):341–375, 1991.

[16] S. Janson and A. Ruciński. The infamous upper tail. *Random Structures Algorithms*, 20(3):317–342, 2002.

[17] K. Krokowski, A. Reichenbachs, and C. Thäle. Discrete Malliavin-Stein method: Berry-Esseen bounds for random graphs and percolation. *Ann. Probab.*, 47(2):1071–1109, 2017.

[18] E. Lubetzky and Y. Zhao. On the variational problem for upper tails in sparse random graphs. *Random Structures Algorithms*, 50:420–436, 2017.

[19] J. Neeman, C. Radin, and L. Sadun. Moderate deviations in cycle count. *arXiv:*, (2101.08249), 2021.

[20] G. Reinert and A. Röllin. Random subgraph counts and U-statistics: multivariate normal approximation via exchangeable pairs and embedding. *J. Appl. Probab.*, 47(2):378–393, 2010.

[21] A. Röllin. Kolmogorov bounds for the Normal approximation of the number of triangles in the Erdős-Rényi random graph. *arXiv:*, 1704.00410, 2017.

[22] A. Ruciński. When are small subgraphs of a random graph normally distributed? *Probab. Theory Related Fields*, 78(1):1–10, 1988.

[23] V. Vu. A large deviation result on the number of small subgraphs of a random graph. *Combin. Probab. Comput.*, 10(1):79–94, 2001.

[24] Y. Zhao. On the lower tail variational problem for random graphs. *Combin. Probab. Comput.*, 26(2):301–320, 2017.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Constructing Hamilton cycles and perfect matchings efficiently

(Extended abstract)

Michael Anastos *

Abstract

Starting with the empty graph on $[n]$, at each round, a set of $K = K(n)$ edges is presented chosen uniformly at random from the ones that have not been presented yet. We are then asked to choose at most one of the presented edges and add it to the current graph. Our goal is to construct a Hamiltonian graph with $(1 + o(1))n$ edges within as few rounds as possible.

We show that in this process, one can build a Hamiltonian graph of size $(1+o(1))n$ in $(1+o(1))(1+(\log n)/2K)n$ rounds w.h.p. The case $K = 1$ implies that w.h.p. one can build a Hamiltonian graph by choosing $(1 + o(1))n$ edges in an online fashion as they appear along the first $(0.5 + o(1))n \log n$ rounds of the random graph process. This answers a question of Frieze, Krivelevich and Michaeli. Observe that the number of rounds is asymptotically optimal as the first $0.5n \log n$ edges do not span a Hamilton cycle w.h.p. The case $K = \Theta(\log n)$ implies that the Hamiltonicity threshold of the corresponding Achlioptas process is at most $(1+o(1))(1+(\log n)/2K)n$. This matches the $(1-o(1))(1+(\log n)/2K)n$ lower bound due to Krivelevich, Lubetzky and Sudakov and resolves the problem of determining the Hamiltonicity threshold of the Achlioptas process with $K = \Theta(\log n)$.

We also show that in the above process one can construct a graph G that spans a matching of size $\lfloor V(G)/2)\rfloor$ and $(0.5+o(1))n$ edges within $(1+o(1))(0.5+(\log n)/2K)n$ rounds w.h.p.

Our proof relies on a robust Hamiltonicity property of the strong 4-core of the binomial random graph which we use as a black-box. This property allows it to absorb

*Institute of Science and Technology Austria, Klosterneuburg 3400, Austria. E-mail: michael.anastos@ist.ac.at. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101034413.

paths covering vertices outside the strong 4-core into a cycle.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-005`

1 Introduction

Let $G_0, G_1, ..., G_N$, $N = \binom{n}{2}$ be the random graph process. That is, G_0 is the empty graph on $[n]$ and G_{i+1} is formed by adding to G_i an edge chosen uniformly at random from the non-present ones, for $0 \leq i < N$. Equivalently let $e_1, e_2, ..., e_N$ be a permutation of the edges of the complete graph K_n chosen uniformly at random and set $G_i = ([n], \{e_1, ..., e_i\})$, $0 \leq i < N$. Let τ_2 be the minimum i such that G_i has minimum degree 2 and τ_H be the minimum i such that G_i is Hamiltonian. Building upon work of Pósa [13] and Korshunov [11], Bollobás [6] and independently Ajtai, Komlós and Szemerédi [1] proved that $\tau_2 = \tau_H = 0.5n(\log n + (1 + o(1)) \log\log n)$ w.h.p.[1] Thus, to achieve Hamiltonicy, one has to wait until the minimum degree becomes 2. Unfortunately, this necessary condition is satisfied w.h.p. only by graphs of the random graphs process that have at least $0.5n \log n$ edges, while a Hamilton cycle uses only n of them. This raises the following question. Can one built a Hamiltonian subgraph of G_t that spans $(1 + o(1))n$ edges in an online fashion for some $t = (1 + o(1))\tau_2$?

Frieze, Krivelevich and Michaeli studied a generalization of this question in the following setting [9]. Once again let $e_1, e_2, ..., e_N$ be a permutation of $E(K_n)$ chosen uniformly at random. The sequence $e_1, e_2, ..., e_N$ is revealed, one edge at a time. Starting with the empty graph on $[n]$, as soon as an edge is revealed we must decide, immediately and irrevocably, whether to choose and add it to our graph. Let B_i be the graph constructed after the ith edge has been revealed. Let $\mathcal{B}'_{HAM}$ be the set of pairs (t, b) for which there exists an algorithm that builds a Hamiltonian graph of size at most b within the first t rounds of the above process w.h.p. Clearly, as $B_i \subseteq G_i$ for all i and $\tau_2 > 0.5n \log n$ w.h.p., a necessary condition for $(t, b) \in \mathcal{B}'_{HAM}$ is that $t \geq 0.5n \log n$ and $b \geq n$. Frieze, Krivelevich and Michaeli proved that for every $\epsilon > 0$ there exists $C > 0$ such that if $t \geq (0.5+\epsilon)n \log n$ and $b \geq 9n$ or $t \geq Cn \log n$ and $b \geq (1+\epsilon)n$ then $(t, b) \in \mathcal{B}'_{HAM}$. They also asked whether there exist $\epsilon > 0$ and a pair t, b such that $t \leq (0.5 + \epsilon)n \log n$, $b \leq (1 + \epsilon)n$ and $(t, b) \notin \mathcal{B}'_{HAM}$. Theorem 1.1 answers this question.

A second way to generalize our question is within the framework of the Achlioptas processes. Inspired by the "power of two choices" paradigm Achiloptas proposed the following process. Starting with the empty graph on $[n]$, at each round, a set of $K = K(n)$ edges is presented chosen uniformly at random from the ones that have not been presented yet (or from all $\binom{n}{2}$ possible ones). We are then asked to choose one of them to add to the current graph, immediately and irrevocably. The aim of the Achlioptas process is to accelerate or delay a given graph property. For example, Bohman and Frieze proved that there exist $\epsilon > 0$ and a strategy that w.h.p. ensure that one can construct a graph with no component

[1]We say that a sequence of events $\{\mathcal{E}_n\}_{n\geq 1}$ holds *with high probability* if $\lim_{n\to\infty} \mathbf{Pr}(\mathcal{E}_n) = 1 - o(1)$.

of size $\Omega(n)$ after $(1+\epsilon)n/2$ rounds, thus delaying the appearance of the giant [4]. Krivelevich, Lubetzky and Sudakov studied $\tau_H(K)'$, the minimum number of rounds needed to construct a Hamiltonian graph in the above process [12]. They proved that w.h.p.

$$(1+o(1))\left(1+\frac{\log n}{2K}\right)n \leq \tau_H(K) \leq (1+o(1))\left(3+\frac{\log n}{K}\right)n. \tag{1}$$

To obtain the upper bound, they constructed a random 3-out graph which is known to be Hamiltonian [5]. For the lower bound they proved that for any algorithm $\mathcal{A}$ and any $\epsilon > 0$, after $(1-\epsilon)(1+0.5\log n/K)n$ rounds there exist $n^{\epsilon/2}$ vertices of degree smaller than 2 w.h.p. Their argument goes as follows. After $0.5(1-\epsilon)n$ rounds, the graph constructed so far by $\mathcal{A}$ contains at least ϵn vertices of degree smaller than 2, deterministically. From those vertices, at least $n^{\epsilon/2}$ will not be incident to any edge that will be presented in the next $0.5(1-\epsilon)n(\log n)/K$ rounds w.h.p. Any such vertices have degree at most 1 in the graph constructed so far.

Krivelevich, Lubetzky and Sudakov also proved that the lower bound in (1) is the correct one, in the sense that it is equal to $(1+o(1))\tau_H(K)$ w.h.p., in the regimes $K = o(\log n)$ and $K = \omega(\log n)$. In these regimes the lower bound reduces to $(1+o(1))(n\log n)/2K$ and $(1+o(1))n$ respectively. Theorem 1.1 implies that the lower bound in (1) is always the correct one. The problem of improving the bounds in (1) is also stated as Problem 43 in Frieze's bibliography on Hamilton cycles in random graphs [7].

Formally the process that we consider is the following one. Starting with the empty graph on $[n]$, at each round, a set of $K = K(n)$ edges is presented chosen uniformly at random from the ones that have not been presented yet. We are then asked to choose *at most* one of them to add to the current graph immediately and irrevocably. We let B_i be the graph constructed after i rounds. We let $\mathcal{B}_{HAM} = \mathcal{B}_{HAM}(K)$ be the set of pairs $(t,b) = (t(K), b(K))$ for which there exists an algorithm that builds a Hamiltonian graph of size at most b within the first t rounds of the above process w.h.p. Similarly, we let $\mathcal{B}_{PM} = \mathcal{B}_{PM}(K)$ be the set of pairs (t,b) for which there exists an algorithm that builds a graph of size at most b that spans a matching of size $\lfloor n/2 \rfloor$ within the first t rounds of the above process w.h.p.

Theorem 1.1. *Let $K = K(n) = O(\log n)$. Then,*

$$\left(\left(1+\frac{250}{\log\log n}\right)\left(1+\frac{\log n}{2K}\right)n, \left(1+\frac{11}{\log\log n}\right)n\right) \in \mathcal{B}_{HAM}.$$

The case $K = \omega(\log n)$ of the above theorem follows from Theorem 1.2 of [12]. Once again, as G_t has minimum degree 0 for $t \leq 0.5n\log n$ w.h.p., one has that $(t,b) \in \mathcal{B}_{PM}$ only if $t \geq 0.5n\log n$ and $b \geq n/2$.

Theorem 1.2.

$$\left(\left(1+\frac{250}{\log\log n}\right)\left(0.5+\frac{\log n}{2K}\right)n, \left(0.5+\frac{11}{\log\log n}\right)n\right) \in \mathcal{B}_{PM}.$$

Ramark 1.3. *Frieze, Krivelevich and Michaeli gave an alternative proof to Theorem 1.2 for the case $K = 1$ (See Theorem 4 of [9]).*

Ramark 1.4. *One may consider the variations of the process where at every round, the K edges that are presented are chosen uniformly at random from all $\binom{n}{2}$ possible edges or from the ones that are missing from the graph that is constructed so far. Theorems 1.1 and 1.2 as stated also hold for these variations.*

In this note we sketch the proof of Theorem 1.1. Theorem 1.2 can be proven in a similar manner. Both proofs are based on structural properties of the strong 4-core of a random graph which we describe in the next section.

2 The strong k-core

For a graph G we define the *strong k-core* of G to be the maximal subset S of $V(G)$ with the property that every vertex in $S \cup N(S)$ has at least k neighbors in S. By $N(S)$ we denote the set of vertices in $V(G) \setminus S$ that are adjacent to S. Observe that if the sets $S_1, S_2 \subset V(G)$ have this property, then so does the set $S_1 \cup S_2$. Thus the strong k-core of a graph is well-defined. It also naturally partitions the vertex set of a graph G into 3 sets which we denote by $V_{k,black}(G), V_{k,blue}(G)$ and $V_{k,red}(G)$ where $V_{k,black}(G)$ is the strong k-core of G, $V_{k,blue}(G)$ is its neighborhood and $V_{k,red}(G)$ is the rest i.e. $V_{k,red}(G) = V(G) \setminus (V_{k,black}(G) \cup N(V_{k,black}(G))$. In our knowledge, the strong 3-core was first used in [3] for finding the longest cycle in sparse random graphs while the concept of the strong k-core was first formalized in [2]. There it was observed that the strong 4-core of $G(n, c/n)$ is robustly Hamiltonian for $c \geq 20$ as described below. For a graph G and $U \subseteq V(G)$ denote by $G[U]$ the subgraph of G induced by U. By $G(n, p)$ we denote the binomial random graph i.e., the random graph on $[n]$ where every edge appears independently with probability p.

Theorem 2.1 (Theorem 3.3 of [2])**.** *Let $G \sim G(n, c/n)$, $c \geq 20$. Let G' be the subgraph of G induced by $V_{4,black}(G) \cup V_{4,blue}(G)$. Then for every $U \subseteq V_{4,blue}(G)$ and matching M on $V_{4,blue} \setminus U$ w.h.p. we have that $G'[V(G') \setminus U] \cup M$ has a Hamilton cycle that spans M.*

Theorem 2.1 enable us to use the strong 4-core of $G(n, 20/n)$ as an absorber for finding large cycles. Indeed, assume that a graph G contains $G' \sim G(n, 20/n)$ as a subgraph. In addition assume that there exists a set of vertex disjoint paths $\mathcal{P}$ that do not intersect $V_{4,black}(G') \cup V_{4,blue}(G')$ internally and whose endpoints lie in $V_{4,blue}(G')$. Then, given G' and $\mathcal{P}$, one can contract each path of $\mathcal{P}$ into an edge. This results to a matching M on $V_{4,blue}(G')$. Theorem 2.1 then gives that $G' \cup M$ spans a Hamilton cycle which spans all the edges in M. Replacing the edges in M with the corresponding paths in $\mathcal{P}$ gives a cycle of G whose vertex set consists of $V_{4,black}(G')$, $V_{4,black}(G')$ and the set of vertices spanned by the paths in $\mathcal{P}$. This will be our main strategy in proving Theorem 2.1.

The next lemma will also be used in the proof of Theorem 1.1. For its proof see Lemma 3.3 of [2].

Lemma 2.2. *Let $G \sim G(n, c/n)$, $c \geq 20$. Then $|V_{4,blue}(G)| \geq 0.1 \cdot (2c)^3 e^{-2c} n$ w.h.p.*

3 Constructing a Hamilton cycle online, efficiently

We now sketch the proof of Theorem 1.1. To simplify its description we only consider the case $K = 1$. Thus at round i we are presented with an edge e_i chosen uniformly at random from the ones that have not been presented yet, for $i \in [N]$. For its proof we describe an algorithm $\mathcal{A}$ that chooses $(1 + 11/\log\log n)n$ edges within the first $(1 + 250/\log\log n)(1 + \log n/2)n$ rounds and constructs a Hamiltonian graph w.h.p. Let

$$n' = \frac{n}{\log\log n}, \quad t_\epsilon = \left(\frac{50}{\log\log n}\right)\left(1 + \frac{\log n}{2}\right)n,$$

$t_0 = 0$, $t_1 = t_\epsilon$, $t_2 = t_1 + t_\epsilon + n$, $t_3 = t_2 + t_\epsilon$, $t_4 = t_3 + t_\epsilon + n(\log n/2)$ and $t_5 = t_4 + t_\epsilon$. $\mathcal{A}$ consists of 5 phases. Its ith phase starts when $e_{t_{i-1}+1}$ is presented and ends once $\mathcal{A}$ decides whether to keep the edge e_{t_i}.

During its first phase, $\mathcal{A}$ picks the first $10n'$ edges that are spanned by $[n']$. Let G_1 be the graph $\mathcal{A}$ constructed during Phase 1 of $\mathcal{A}$, $U = V_{4,black}(G) \cup V_{4,blue}(G')$, $W = V_{4,blue}(G')$ and $Z = [n] \setminus U$. Lemma 2.2 implies that $|U| = \Omega(n')$ w.h.p. The rest of the phases of $\mathcal{A}$ aim to cover the vertices in Z by a set $\mathcal{P}'$ of vertex disjoint paths with endpoints in W that do not internally intersect U. To do so, during its second phase, $\mathcal{A}$ greedily covers Z with at most $n/(\log\log n)^2$ vertex disjoint paths, each of length at most $\log n$. Here we allow paths of length 0 which correspond to single vertices. Let $\mathcal{P}$ be the set of these paths. Then, during Phase 3, $\mathcal{A}$ greedily matches the endpoints of the paths in $\mathcal{P}$ to W, each path $P \in \mathcal{P}$ is therefore potentially extended to a path with a pair of unique endpoints in W. Let $End(\mathcal{P})$ be the set of endpoints of paths in $\mathcal{P}$ that lie in Z (are left unmatched). During Phase 4, $\mathcal{A}$ attempts to match the vertices in $End(\mathcal{P})$ to $\log^{0.8} n$ many vertices in the interior of distinct paths in $\mathcal{P}$. This is possible as $t_4 - t_3 = t_\epsilon + 0.5n\log n$, which implies that each vertex in $End(\mathcal{P})$ is incident to $\omega(\log^{0.8} n)$ edges in $\{e_{t_3+1}, ..., e_{t_4}\}$ whose other endpoint lies in $[n] \setminus U$. Finally, during Phase 5, using the edges selected during Phase 4, $\mathcal{A}$ reroutes the paths in $\mathcal{P}$ with an endpoint in $End(\mathcal{P})$ through the rest of the paths. Such a rerouting may look as follows. Let $Q = v_1, v_2, ..., v_k$ and $P = u_1, u_2, ..., u_r$ be vertex disjoint paths with $v_1, v_k, u_r \in W$ and $u_1 \in Z$. In such a case, adding the edges u_1v_i and $v_{i+1}v$ with $v \in W$, $1 \le i \le k-1$ (selected during phases 4 and 5 respectively) and removing the edge v_iv_{i+1} from $E(P) \cup E(Q)$ results to 2 vertex disjoint paths that cover $V(P) \cup V(Q)$ and have their endpoints in W.

One may show that the set of edges selected during the last 4 phases span a set $\mathcal{P}'$ of vertex disjoint paths with endpoints in W that do not internally intersect U w.h.p. Given G_1 and $\mathcal{P}$, one may appeal to Theorem 2.1, as discussed in the previous section, to show the existence of a Hamilton cycle spanned by the constructed graph. Finally note that the edges selected during phases 2 and 3 span a set of paths, thus there are at most n. Therefore, in total, $\mathcal{A}$ selects $10n' + n + |End|(\log^{0.8} n + 2)$ which is equal to $(1 + o(1))n$ in the high probability event that $|End(\mathcal{P})| = o(n/\log n)$.

References

[1] Miklós Ajtai, János Komlós, and Endre Szemerédi. First occurrence of Hamilton cycles in random graphs. *North-Holland Mathematics Studies*, 115(C):173–178, 1985.

[2] Michael Anastos. A note on long cycles in sparse random graphs. *The Electronic Journal of Combinatorics 30.2*, 2023.

[3] Michael Anastos and Alan Frieze. A scaling limit for the length of the longest cycle in a sparse random graph. *Journal of Combinatorial Theory, Series B*, 148:184–208, 2021.

[4] Tom Bohman and Alan Frieze. Avoiding a giant component. *Random Structures & Algorithms*, 19(1):75–85, 2001.

[5] Tom Bohman and Alan Frieze. Hamilton cycles in 3-out. *Random Structures & Algorithms*, 35(4):393–417, 2009.

[6] Béla Bollobás. The evolution of sparse graphs, graph theory and combinatorics, 1984. *MR*, 777163(2):35–57, 1984.

[7] Alan Frieze. Hamilton cycles in random graphs: a bibliography. *arXiv preprint arXiv:1901.07139*, 2019.

[8] Alan Frieze and Michal Karoński. *Introduction to random graphs*. Cambridge University Press, 2016.

[9] Alan Frieze, Michael Krivelevich, and Peleg Michaeli. Fast construction on a restricted budget. *arXiv preprint arXiv:2207.07251*, 2022.

[10] János Komlós and Endre Szemerédi. Limit distribution for the existence of Hamiltonian cycles in a random graph. *Discrete mathematics*, 43(1):55–63, 1983.

[11] Aleksei Dmitrievich Korshunov. Solution of a problem of Erdős and Rényi on Hamiltonian cycles in nonoriented graphs. In *Doklady Akademii Nauk*, volume 228, pages 529–532. Russian Academy of Sciences, 1976.

[12] Michael Krivelevich, Eyal Lubetzky, and Benny Sudakov. Hamiltonicity thresholds in achlioptas processes. *Random Structures & Algorithms*, 37(1):1–24, 2010.

[13] Lajos Pósa. Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14(4):359–364, 1976.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Algorithms for subgraph complementation to some classes of graphs

(Extended abstract)

Dhanyamol Antony* Sagartanu Pal† R. B. Sandeep†

Abstract

For a class $\mathcal{G}$ of graphs, the objective of Subgraph Complementation to $\mathcal{G}$ is to find whether there exists a subset S of vertices of the input graph G such that modifying G by complementing the subgraph induced by S results in a graph in $\mathcal{G}$. We obtain a polynomial-time algorithm for the problem when $\mathcal{G}$ is the class of graphs with minimum degree at least k, for a constant k, answering an open problem by Fomin et al. (Algorithmica, 2020). When $\mathcal{G}$ is the class of graphs without any induced copies of the star graph on $t+1$ vertices (for any constant $t \geq 3$) and diamond, we obtain a polynomial-time algorithm for the problem. This is in contrast with a result by Antony et al. (Algorithmica, 2022) that the problem is NP-complete and cannot be solved in subexponential-time (assuming the Exponential Time Hypothesis) when $\mathcal{G}$ is the class of graphs without any induced copies of the star graph on $t+1$ vertices, for every constant $t \geq 5$.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-006`

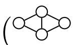
*Department of Computer Science and Automation, Indian Institute of Science, Bangalore, India. E-mail: dhanyamola@iisc.ac.in. Partially supported by SERB Core Research Grant CRG/2022/006770: "Bridging Quantum Physics with Theoretical Computer Science and Graph Theory" under Prof. Sunil Chandran Leela while Dhanyamol Antony was a Postdoc at IISC Bangalore.

†Department of Computer Science and Engineering, Indian Institute of Technology Dharwad, India. E-mail: {183061001,sandeeprb}@iitdh.ac.in. Partially supported by SERB MATRICS Grant MTR/2022/000692: "Algorithmic study on hereditary graph properties".

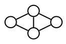
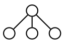
1 Introduction

Complementation is a very fundamental graph operation and modifying a graph by complementing an induced subgraph to satisfy certain properties is a natural algorithmic problem on graphs. The operation of complementing an induced subgraph, known as subgraph complementation, is introduced by Kamiński et al. [1] in connection with clique-width of graphs. For a class $\mathcal{G}$ of graphs, the objective of SUBGRAPH COMPLEMENTATION TO $\mathcal{G}$ is to find whether there exists a subset S of the vertices of the input graph G such that complementing the subgraph induced by S in G results in a graph in $\mathcal{G}$. Fomin et al. [2] studied this problem on various classes $\mathcal{G}$ of graphs. They obtained that the problem can be solved in polynomial-time when $\mathcal{G}$ is bipartite, d-degenerate, or co-graphs. In addition to this, they proved that the problem is NP-complete when $\mathcal{G}$ is the class of all regular graphs. Antony et al. [3] studied this problem when $\mathcal{G}$ is the class of H-free graphs (graphs without any induced copies of H). They proved that the problem is polynomial-time solvable when H is a complete graph on t vertices. They also proved that the problem is NP-complete when H is a star graph on at least 6 vertices or a path or a cycle on at least 7 vertices. Later Antony et al. [4] proved that the problem is polynomial-time solvable when H is paw, and NP-complete when H is a tree, except for 41 trees of at most 13 vertices. It has been proved [3,4] that none of these hard problems admit subexponential-time algorithms (algorithms running in time $2^{o(n)}$), assuming the Exponential Time Hypothesis.

Fomin et al. [2] proved that the problem is polynomial-time solvable not only when $\mathcal{G}$ is the class of d-degenerate graphs but also when $\mathcal{G}$ is any subclass of d-degenerate graphs recognizable in polynomial-time. This implies that the problem is polynomial-time solvable when $\mathcal{G}$ is the class of r-regular graphs or the class of graphs with maximum degree at most r (for any constant r). They asked whether the problem can be solved in polynomial-time when $\mathcal{G}$ is the class of graphs with minimum degree at least r, for a constant r. We resolve this positively and obtain a stronger result - a simple quadratic kernel for the following parameterized problem: Given a graph G and an integer k, find whether G can be transformed into a graph with minimum degree at least k by subgraph complementation (here the parameter is k). The result follows from an observation that if G has more than $2k^2 - 2$ vertices, then it is a yes-instance of the problem.

When $\mathcal{G}$ is the class of graphs without any induced copies of the star graph on $t + 1$ vertices (for any fixed $t \geq 3$) and the diamond (), we obtain a polynomial-time algorithm. When $t = 3$ this graph class is known as linear domino and is the class of line graphs of triangle-free graphs. Cygan et al. [5] have studied the polynomial kernelization of edge deletion problem for this target graph class. When $t = 4$, the graph class is the line graphs of linear hypergraphs of rank 3. The technique that we use is similar to that given in [3] and [4] for obtaining polynomial-time algorithms when $\mathcal{G}$ is H-free, for H being a complete graph on t vertices or a paw. Our result is in contrast with the result by Antony et al. [3] that the problem is NP-complete and cannot be solved in subexponential-time (assuming the Exponential Time Hypothesis) when H is a star graph on $t+1$ vertices, for every constant $t \geq 5$.

Preliminaries

A diamond is the graph  , and a star graph on $t+1$ vertices, denoted by $K_{1,t}$, is the tree with t degree-1 vertices and one degree-t vertex. The degree-t vertex of a star is known as the center of the star. For example, $K_{1,3}$, also known as a claw, is the graph  . A complete graph on t vertices is denoted by K_t. By $\overline{G}$ we denote the complement graph of G. The open neighborhood and closed neighborhood of a vertex v are denoted by $N(v)$ and $N[v]$ respectively. The underlying graph will be evident from the context. For a subset S of vertices of G, by $G[S]$ we denote the graph induced by S in G. For a given graph G and a set $S \subseteq V(G)$, we define the graph $G \oplus S$ as the graph obtained from G by complementing the subgraph induced by S, i.e., an edge uv is in $G \oplus S$ if and only if uv is a nonedge in G and $u, v \in S$, or uv is an edge in G and $\{u, v\} \setminus S \neq \emptyset$. The operation is called subgraph complementation. Let $\mathcal{H}$ be a set of graphs. We say that a graph G is $\mathcal{H}$-free if G does not have any induced copies of any of the graphs in $\mathcal{H}$. If $\mathcal{H} = \{H\}$, then we say that G is H-free. The general definition of the problem that we deal with is given below.

SC-TO-$\mathcal{G}$: Given a graph G, find whether there is a set $S \subseteq V(G)$ such that $G \oplus S \in \mathcal{G}$.

In a parameterized problem, apart from the usual input, there is an additional integer input known as the parameter. A graph problem is fixed-parameter tractable (FPT) if it can be solved in time $f(k)n^{O(1)}$, where n is the number of vertices and $f(k)$ is any computable function. A parameterized problem admits a kernel if there is a polynomial-time algorithm which takes as input an instance (I', k') of the problem and outputs an instance (I, k) of the same problem so that $|I|, k \leq f(k)$ for some computable function $f(k)$, and (I', k') is a yes-instance if and only if (I, k) is a yes-instance (here, k' and k are the parameters). A kernel is a polynomial kernel if $f(k)$ is a polynomial function. It is known that a problem admits an FPT algorithm if and only if it admits a kernel. An FPT algorithm implies that there is a polynomial-time algorithm to solve the problem when the parameter is a constant. We refer to the book [6] for further exposition on these topics.

2 Algorithms

We obtain our results in this section. Let $\mathcal{G}_k$ be the class of graphs with minimum degree at least k. We prove that a no-instance of SC-TO-$\mathcal{G}_k$ cannot be very large.

Lemma 2.1. *Let G be a graph with more than $2k^2 - 2$ vertices. Then G is a yes-instance of* SC-TO-$\mathcal{G}_k$.

Proof. Let M be the set of vertices in G with degree less than k. Clearly, $M \subseteq S$ for every solution S (i.e., $G \oplus S \in \mathcal{G}_k$). Let $|M| = m$. Let M' be the set of vertices in $V(G) \setminus M$ adjacent to at least one vertex in M. As each vertex in M has degree at most $k-1$, we obtain that $|M'| \leq m(k-1)$.

Let $M'' = V(G) \setminus (M \cup M')$. Let X be the set of vertices in M'' having degree at least $2k - m - 1$ in G. If $|X| \geq k$, then $G \oplus (M \cup X') \in \mathcal{G}_k$, where X' is any subset of k vertices of X - note that degree of every vertex in X' is at least $(2k-m-1)+m-(k-1)=k$, in $G \oplus (M \cup X')$. Therefore, assume that $|X| \leq k-1$. Every vertex in $M'' \setminus X$ has degree at most $2k-m-2$ in G. Then, every maximal independent set in $M'' \setminus X$ has size at least $|M'' \setminus X|/(2k-m-1)$. Therefore, if $|M'' \setminus X| \geq k(2k-m-1)$, then for any maximal independent set I of $M'' \setminus X$, $G \oplus (M \cup I) \in \mathcal{G}_k$. Hence assume that $|M'' \setminus X| \leq k(2k-m-1)-1$. Therefore, if G is a no-instance of SC-TO-$\mathcal{G}_k$, then the number of vertices in G is at most $|M|+|M'|+|X|+|M'' \setminus X| \leq m+m(k-1)+(k-1)+k(2k-m-1)-1 = 2k^2-2$. □

Lemma 2.1 gives a polynomial-time algorithm for the problem: If G has more than $2k^2-2$ vertices, then return YES, and do an exhaustive search for a solution otherwise. Lemma 2.1 also gives a simple quadratic kernel for the problem parameterized by k: For an input (G,k) if G has more than $2k^2-2$ vertices, then return a trivial yes-instance, and return the same instance otherwise. By a result from [3], SC-TO-$\mathcal{G}$ and SC-TO-$\overline{\mathcal{G}}$ are polynomially equivalent. Therefore, we obtain a polynomial-time algorithm for SC-TO-$\mathcal{G}$ when $\mathcal{G}$ is the class of graphs with maximum degree at most $n-k$, for a constant k. It also implies a quadratic kernel for the problem parameterized by k. It remains open whether the following problem is NP-complete: Given a graph G and an integer k, find whether G can be subgraph complemented to a graph with minimum degree at least k. We note that, the problem is NP-complete if the objective is to make the input graph k-regular [2].

Destroying stars and diamonds

Let $\mathcal{G}$ be the class of $\{K_{1,t}, \text{diamond}\}$-free graphs, for any fixed $t \geq 3$. We give a polynomial-time algorithm for SC-TO-$\mathcal{G}$. The concept of (p,q)-split graphs was introduced by Gyárfás [7]. For $p \geq 1$, and $q \geq 1$, if the vertices of a graph G can be partitioned into two sets P and Q in such a way that the clique number of $G[P]$ and the independence number of $G[Q]$ are at most p and q respectively (i.e., $G[P]$ is K_{p+1}-free and $G[Q]$ is $(q+1)K_1$-free), then G is called a (p,q)-split graph and (P,Q) is a (p,q)-split partition of G.

Proposition 2.2 ([3, 8, 9]). *For any fixed constants $p \geq 1$ and $q \geq 1$, recognizing a (p,q)-split graph and obtaining all (p,q)-split partitions of a (p,q)-split graph can be done in polynomial-time.*

Algorithm for SC-TO-$\mathcal{G}$, where $\mathcal{G}$ is $\{K_{1,t}, \textit{diamond}\}$-free graphs, for any constant $t \geq 3$.
Input: A graph G.
Output: If G is a yes-instance of SC-TO-$\mathcal{G}$, then returns YES; otherwise returns NO.

Step 1 : Let S be the set of all degree-2 vertices of all the induced diamonds in G. If $G \oplus S \in \mathcal{G}$, then return YES.

Step 2 : Let r be the center of any induced $K_{1,t}$ in G and let I be the set of isolated vertices in the subgraph induced by $N(r)$ in G. For every subset $S \subseteq I$ such that $|S| \geq |I|-t+2$, if $G \oplus S \in \mathcal{G}$, then return YES.

Step 3 : For every edge uv in G, do the following:

1. If $N(u) \setminus N[v]$ or $N(v) \setminus N[u]$ does not induce a $(t-1, t-1)$-split graph, then continue with Step 3.
2. Compute $L(u\overline{v})$, the list of all $(t-1, t-1)$-split partitions of the graph induced by $N(u) \setminus N[v]$.
3. Compute $L(\overline{u}v)$, the list of all $(t-1, t-1)$-split partitions of the graph induced by $N(v) \setminus N[u]$.
4. Compute $L(uv)$, the list of all partitions of the graph induced by $N(u) \cap N(v)$ into an independent set of size at most $t-1$ and the rest.
5. For every $(S_1, T_1) \in L(u\overline{v})$, for every $(S_2, T_2) \in L(\overline{u}v)$, for every $(S_3, T_3) \in L(uv)$, do the following:
   (a) Let $S = S_1 \cup S_2 \cup S_3 \cup \{u, v\}$. If $G \oplus S \in \mathcal{G}$, return YES.
   (b) For every vertex $w \in \overline{N[u]} \cap \overline{N[v]}$, let $S = S_1 \cup S_2 \cup S_3 \cup \{u, v, w\}$. If $G \oplus S \in \mathcal{G}$, return YES.
   (c) For every edge xy in the graph induced by $\overline{N[u]} \cap \overline{N[v]}$, if the graph induced by $J = N[x] \cap N[y] \cap \overline{N[u]} \cap \overline{N[v]}$ is not a split graph then continue with the current step. Otherwise, for every split partition (S_4, T_4) of the graph induced by J, let $S = S_1 \cup S_2 \cup S_3 \cup S_4 \cup \{u, v\}$. If $G \oplus S \in \mathcal{G}$, then return YES.

Step 4 : Return NO.

Lemma 2.3 and 2.4 deals with the case when G is a yes-instance having a solution which is an independent set, the case handled in Step 1 and 2 of the algorithm.

Lemma 2.3. *Assume that G is not diamond-free. Let $S \subseteq V(G)$ such that $G \oplus S \in \mathcal{G}$ and S is an independent set. Then S is the set of all degree-2 vertices of all the induced diamonds in G.*

Proof. Since S is an independent set and $G \oplus S \in \mathcal{G}$, both the degree-2 vertices of every induced diamond in G must be in S. Assume for a contradiction that S has a vertex v which is not a degree-2 vertex of any of the induced diamonds in G. Let $D = \{d_1, d_2, d_3, d_4\}$ induces a diamond in G, where d_1 and d_2 are the degree-2 vertices of the diamond. Clearly, $S \cap D = \{d_1, d_2\}$. We know that $v \neq d_1$ and $v \neq d_2$. If v is not adjacent to d_3 in G, then $\{v, d_1, d_2, d_3\}$ induces a diamond in $G \oplus S$, which is a contradiction. Therefore, v is adjacent to d_3. Similarly, v is adjacent to d_4. Then $\{v, d_1, d_3, d_4\}$ induced a diamond in G, where v and d_1 are the degree-2 vertices, which is a contradiction. □

Lemma 2.4. *Assume that G has no induced diamond but has at least one induced $K_{1,t}$. Let $S \subseteq V(G)$ such that $G \oplus S \in \mathcal{G}$ and S is an independent set. Let r be the center of any induced $K_{1,t}$ in G. Let I be the set of isolated vertices in the subgraph induced by $N(r)$ in G. Then $S \subseteq I$ and $|S| \geq |I| - t + 2$.*

Proof. If $r \in S$, then none of the vertices in $N(r)$ is in S - recall that S is an independent set. But then, none of the induced $K_{1,t}$ centered at r is destroyed in $G \oplus S$. Therefore, $r \notin S$. Since G is diamond-free, $N(r)$ induces a cluster (graph with no induced path of length 3) J in G. Since r is the center of an induced $K_{1,t}$ in G, there are at least t cliques in J. Since $G \oplus S$ is $K_{1,t}$-free, S must contain all vertices of at least two cliques in J. Since S is an independent set, S contains at least two isolated vertices, say s_1 and s_2, in J. First we prove that $S \subseteq N(r)$. For a contradiction, assume that there is a vertex $v \in S$ such that v is not adjacent to r. Then $\{v, s_1, s_2, r\}$ induces a diamond in $G \oplus S$, which is a contradiction. Therefore, $S \subseteq N(r)$. Next we prove that $S \subseteq I$. For a contradiction, assume that there is a vertex $v \in S \setminus I$. Then v is part of a clique J' of size at least 2 in J. Let v' be any other vertex in J'. Since S is an independent set, $v' \notin S$. Then $\{v, v', s_1, r\}$ induces a diamond in $G \oplus S$, which is a contradiction. Therefore, $S \subseteq I$. If $|S| < |I| - t + 2$, then there is a $K_{1,t}$ centered at r in $G \oplus S$, which is a contradiction. □

Let G be a yes-instance of SC-TO-$\mathcal{G}$. Let $S \subseteq V(G)$ be such that $|S| \geq 2$, $G \oplus S \in \mathcal{G}$, and S be not an independent set. Let u, and v be two adjacent vertices in S. Then with respect to S, u, v, we can partition the vertices in $V(G) \setminus \{u, v\}$ into eight sets as given below, and shown in Figure 1.

(i) $N_S(uv) = S \cap N(u) \cap N(v)$

(ii) $N_S(\bar{u}\bar{v}) = S \cap \overline{N[u]} \cap \overline{N[v]}$

(iii) $N_S(u\bar{v}) = S \cap (N(u) \setminus N[v])$

(iv) $N_S(\bar{u}v) = S \cap (N(v) \setminus N[u])$

(v) $N_T(uv) = (N(u) \cap N(v)) \setminus S$

(vi) $N_T(\bar{u}\bar{v}) = (\overline{N[u]} \cap \overline{N[v]}) \setminus S$

(vii) $N_T(u\bar{v}) = (N(u) \setminus N[v]) \setminus S$

(viii) $N_T(\bar{u}v) = (N(v) \setminus N[u]) \setminus S$

We notice that $S = N_S(uv) \cup N_S(\bar{u}\bar{v}) \cup N_S(u\bar{v}) \cup N_S(\bar{u}v) \cup \{u, v\}$.

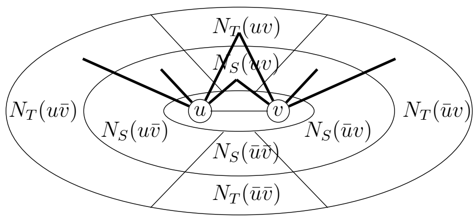


Figure 1: Partitioning of vertices of G based on S and two adjacent vertices $u, v \in S$. The bold lines represent the adjacency of vertices u and v [3].

Observation 2.5. *Then the following statements are true.*

(i) $N(u) \setminus N[v]$ *induces a* $(t-1, t-1)$*-split graph with a* $(t-1, t-1)$*-split partition of* $(N_S(u\bar{v}), N_T(u\bar{v}))$.

(ii) $N(v) \setminus N[u]$ *induces a* $(t-1, t-1)$*-split graph with a* $(t-1, t-1)$*-split partition of* $(N_S(v\bar{u}), N_T(v\bar{u}))$.

(iii) *$N_T(uv)$ induces an independent set with at most $(t-1)$ vertices.*

(iv) *$N_S(\bar{u}\bar{v})$ induces a clique. If xy is an edge of the clique, then $N[x]\cap N[y]$ in $\overline{N[u]}\cap\overline{N[v]}$ induces a split graph with one split partition being $(N_S(\bar{u}\bar{v}), (N[x]\cap N[y]\cap \overline{N[u]}\cap \overline{N[v]})\setminus (N_S(\bar{u}\bar{v})))$.*

Proof. If $N_S(u\overline{v})$ has a K_t, then v along with the vertices of the K_t induce a $K_{1,t}$ in $G\oplus S$. If $N_T(u\overline{v})$ has an independent set of size t, then u along with the vertices of the independent set induce a $K_{1,t}$ in $G\oplus S$. Therefore, (i) holds true. Similarly we can prove the correctness of (ii). If there are two adjacent vertices x and y in $N_T(uv)$, then $\{x,y,u,v\}$ induces a diamond in $G\oplus S$. Therefore, $N_T(uv)$ is an independent set. If it has at least t vertices then there is an induced $K_{1,t}$ formed by those vertices and u in $G\oplus S$. Therefore, (iii) holds true. If there are two nonadjacent vertices x and y in $N_S(\bar{u}\bar{v})$, then there is a diamond induced by $\{x,y,u,v\}$ in $G\oplus S$. Therefore, $N_S(\bar{u}\bar{v})$ is a clique. Assume that $x,y\in N_S(\bar{u}\bar{v})$. If x and y have two adjacent common neighbors x' and y' in $N_T(\bar{u}\bar{v})$, then $\{x,y,x',y'\}$ induces a diamond in $G\oplus S$. Therefore, $N[x]\cap N[y]\cap \overline{N[u]}\cap \overline{N[v]}$ is a split graph with one split partition being $(N_S(\bar{u}\bar{v}), (N[x]\cap N[y]\cap \overline{N[u]}\cap \overline{N[v]})\setminus (N_S(\bar{u}\bar{v})))$. □

Lemma 2.6. *G is a yes-instance of* SC-TO-$\mathcal{G}$ *if and only if the algorithm returns YES.*

Proof. Since the algorithm returns YES only when a solution is found, the backward direction of the statement is true. For the forward direction, let G be a yes-instance. Assume that there exists a solution S which is an independent set. Further, assume that G has an induced diamond. Then by Lemma 2.3, S is the set of all degree-2 vertices of the induced diamonds in G. Then Step 1 returns YES. Assume that G is diamond-free. Then by Lemma 2.4, $S\subseteq I$, where I is the set of isolated vertices in the graph induced by the neighbors of r, for a center r of an induced $K_{1,t}$ in G. Further $|S|\geq |I|-t+2$. Then Step 2 returns YES. Let S be a solution which is not an independent set. Let uv be an edge in the graph induced by S. The algorithm will discover uv in one iteration of Step 3. By Observation 2.5, we know that the graph induced by $N(u)\setminus N[v]$ is a $(t-1,t-1)$-split graph with a $(t-1,t-1)$-split partition $(N_S(u\overline{v}), N_T(u\overline{v}))$. Similarly, the graph induced by $N(v)\setminus N[v]$ is a $(t-1,t-1)$-split graph with a $(t-1,t-1)$-split partition $(N_S(\overline{u}v), N_T(\overline{u}v))$. Further, $N_T(uv)$ is an independent set of size at most $t-1$. Therefore, in one iteration of Step 3.5, we obtain $S_1=N_S(u\overline{v}), S_2=N_S(\overline{u}v)$, and $S_3=N_S(uv)$. If $N_S(\bar{u}\bar{v})$ is empty, then Step 3.5(a) returns YES. If $N_S(\bar{u}\bar{v})$ is a singleton set, then Step 3.5(b) returns YES. Assume that $|N_S(\bar{u}\bar{v})|\geq 2$. By Observation 2.5, $N_S(\bar{u}\bar{v})$ is a clique and for every edge xy in it, the common neighborhood of x and y in $\overline{N[u]}\cap\overline{N[v]}$ is a split graph with a partition being $N_S(\bar{u}\bar{v})$ and the rest. The algorithm will discover such an edge xy in one of the iterations of Step 3.5(c) and $N_S(\bar{u}\bar{v})$ will be discovered as S_4. Then YES is returned at Step 3.5(c). □

By Proposition 2.2, $(t-1,t-1)$-split graphs can be recognized in polynomial-time and all $(t-1,t-1)$-split partitions of a $(t-1,t-1)$-split graph can be found in polynomial-time. Therefore, each step in the algorithm runs in polynomial-time. Then we obtain Theorem 2.7 from Lemma 2.6.

Theorem 2.7. *Let $\mathcal{G}$ be the class of $\{K_{1,t}, diamond\}$-free graphs for any constant $t \geq 3$. Then* SC-TO-$\mathcal{G}$ *can be solved in polynomial-time.*

It remains open whether the problem is polynomial-time solvable when $\mathcal{G}$ is H-free for an $H \in \{K_{1,3}, K_{1,4}, \text{diamond}\}$.

References

[1] Marcin Kamiński, Vadim V. Lozin, and Martin Milanič. Recent developments on graphs of bounded clique-width. Discret. Appl. Math., 157(12):2747–2761, 2009.

[2] Fedor V. Fomin, Petr A. Golovach, Torstein J. F. Strømme, and Dimitrios M. Thilikos. Subgraph complementation. Algorithmica, 82(7):1859–1880, 2020.

[3] Dhanyamol Antony, Jay Garchar, Sagartanu Pal, R. B. Sandeep, Sagnik Sen, and R. Subashini. On subgraph complementation to H-free graphs. Algorithmica, 84(10):2842–2870, 2022.

[4] Dhanyamol Antony, Sagartanu Pal, R. B. Sandeep, and R. Subashini. Cutting a tree with subgraph complementation is hard, except for some small trees. In Armando Castañeda and Francisco Rodríguez-Henríquez, editors, LATIN 2022: Theoretical Informatics - 15th Latin American Symposium, Guanajuato, Mexico, November 7-11, 2022, Proceedings, volume 13568 of Lecture Notes in Computer Science, pages 3–19. Springer, 2022.

[5] Marek Cygan, Marcin Pilipczuk, Michal Pilipczuk, Erik Jan van Leeuwen, and Marcin Wrochna. Polynomial kernelization for removing induced claws and diamonds. Theory Comput. Syst., 60(4):615–636, 2017.

[6] Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. Parameterized Algorithms. Springer, 2015.

[7] András Gyárfás. Generalized split graphs and ramsey numbers. J. Comb. Theory, Ser. A, 81(2):255–261, 1998.

[8] Sudeshna Kolay and Fahad Panolan. Parameterized algorithms for deletion to (r, ℓ)-graphs. In Prahladh Harsha and G. Ramalingam, editors, 35th IARCS Annual Conference on Foundation of Software Technology and Theoretical Computer Science, FSTTCS 2015, volume 45 of LIPIcs, pages 420–433. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2015.

[9] Sudeshna Kolay and Fahad Panolan. Parameterized algorithms for deletion to (r, l)-graphs. arXiv preprint arXiv:1504.08120, 2015.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

A LOWER BOUND FOR SET-COLOURING RAMSEY NUMBERS

(EXTENDED ABSTRACT)

Lucas Aragão* Maurício Collares† João Pedro Marciano‡
Taísa Martins§ Robert Morris¶

Abstract

The set-colouring Ramsey number $R_{r,s}(k)$ is defined to be the minimum n such that if each edge of the complete graph K_n is assigned a set of s colours from $\{1,\dots,r\}$, then one of the colours contains a monochromatic clique of size k. The case $s=1$ is the usual r-colour Ramsey number, and the case $s=r-1$ was studied by Erdős, Hajnal and Rado in 1965, and by Erdős and Szemerédi in 1972.

The first significant results for general s were obtained only recently, by Conlon, Fox, He, Mubayi, Suk and Verstraëte, who showed that $R_{r,s}(k)=2^{\Theta(kr)}$ if s/r is bounded away from 0 and 1. In the range $s=r-o(r)$, however, their upper and lower bounds diverge significantly. In this note we introduce a new (random) colouring, and use it to determine $R_{r,s}(k)$ up to polylogarithmic factors in the exponent for essentially all r, s and k.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-007`

*IMPA, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ, Brazil.

†Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria. This research was funded in whole, or in part, by the Austrian Science Fund (FWF) P36161.

‡IMPA, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ, Brazil.

§Instituto de Matemática, Universidade Federal Fluminense, Niterói, Brazil. Supported by CNPq (Proc. 406248/2021-4).

¶IMPA, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ, Brazil. Supported by FAPERJ (Proc. E-26/200.977/2021) and CNPq (Proc. 303681/2020-9).
This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES), Finance Code 001.

1 Introduction

The r-colour Ramsey number $R_r(k)$ is defined to be the minimum $n \in \mathbb{N}$ such that every r-colouring $\chi\colon E(K_n) \to \{1, \dots, r\}$ of the edges of the complete graph on n vertices contains a monochromatic clique of size k. These numbers (and their extensions to general graphs, hypergraphs, etc.) are among the most important and extensively-studied objects in combinatorics, see for example the beautiful survey article [4].

In this paper we will study the following generalisation of the r-colour Ramsey numbers.

Definition 1.1. The set-colouring Ramsey number $R_{r,s}(k)$ is the least $n \in \mathbb{N}$ such that every colouring $\chi\colon E(K_n) \to \binom{[r]}{s}$ contains a monochromatic clique of size k, that is, a set $S \subset V(K_n)$ with $|S| = k$ and a colour $i \in [r]$ such that $i \in \chi(e)$ for every $e \in \binom{S}{2}$.

That is, we assign a set $\chi(e) \subset [r] = \{1, \dots, r\}$ of s colours to each edge of the complete graph, and say that a clique is monochromatic if there exists a colour $i \in [r]$ that is assigned to every edge of the clique. Note that when $s = 1$ this is simply the usual r-colour Ramsey number. The study of set-colouring Ramsey numbers was initiated in the 1960s by Erdős, Hajnal and Rado [5], who conjectured that $R_{r,r-1}(k) \le 2^{\delta(r)k}$ for some function $\delta(r) \to 0$ as $r \to \infty$. This conjecture was proved by Erdős and Szemerédi [6] in 1972, who showed that

$$2^{\Omega(k/r)} \le R_{r,r-1}(k) \le r^{O(k/r)}.$$

For more general values of s, the first significant progress was made only recently, by Conlon, Fox, He, Mubayi, Suk and Verstraëte [2], who showed that

$$\exp\left(\frac{c'k(r-s)^3}{r^2}\right) \le R_{r,s}(k) \le \exp\left(\frac{ck(r-s)^2}{r}\log\frac{r}{\min\{s, r-s\}}\right) \tag{1}$$

for absolute constants $c, c' > 0$. While the exponents in the lower and upper bounds differ by only a factor of $\log r$ when $r - s = \Omega(r)$, they diverge much more significantly when $(r-s)/r \to 0$. We remark that the range $s = r - o(r)$ was of particular interest to the authors of [2], who were motivated by an application to hypergraph Ramsey numbers [3].

The main result of this paper is the following improved lower bound, which allows us to determine $R_{r,s}(k)$ up to a poly-logarithmic factor in the exponent for essentially all r, s, k.

Theorem 1.2. *There exist constants $C > 0$ and $\delta > 0$ such that the following holds. If $r, s \in \mathbb{N}$ with $s \le r - C\log r$, then*

$$R_{r,s}(k) \ge \exp\left(\frac{\delta k(r-s)^2}{r}\right) \tag{2}$$

for every $k \ge (C/\varepsilon)\log r$, where $\varepsilon = (r-s)/r$.

Note that the bound (2) matches the upper bound (1) on $R_{r,s}(k)$, proved in [2], up to a factor of $O(\log r)$ in the exponent for all $s \le r - C\log r$. When $s \ge r - C\log r$ our method does not provide a construction, but in this case the bounds from [2] only differ by a factor of order $(\log r)^2$ in the exponent, the lower bound coming from a simple random colouring.

The lower bound on k in Theorem 1.2 is also not far from best possible, since if $k \le 1/\varepsilon$ then the most common colour has density at least $1 - 1/k$, and therefore $R_{r,s}(k) \le k^2$, by Turán's theorem. A simpler version of the construction described in this paper (taking complete $(k-1)$-partite graphs instead of blow-ups of random graphs) extends Theorem 1.2 to a wider range of s and k, as stated in the following corollary. We omit its proof for space reasons.

Corollary 1.3. *Let $r > s \ge 1$ and $\delta > 0$, and set $\varepsilon = (r-s)/r$. We have $R_{r,s}(k) = 2^{\tilde{\Theta}(\varepsilon^2 rk)}$ for every $k \ge (1+\delta)/\varepsilon + 1$.*

2 The construction

In this section we will define the (random) colouring that we use to prove Theorem 1.2, and prove that it has the desired properties with high probability. The idea behind our construction, to let each colour be a random copy of some pseudorandom graph, was introduced in the groundbreaking work of Alon and Rödl [1] on multicolour Ramsey numbers, and has been used in several recent papers in the area [7, 8, 10, 9]. However, our approach differs from that used in these previous works in several important ways; in particular, we will not count independent sets, and it will be important that our colour classes are chosen (almost) independently at random.

Fix a sufficiently small[1] constant $\delta > 0$, and set $C = 1/\delta^3$. Recall that $r - s = \varepsilon r$, and let

$$m = 2^{\delta^2 \varepsilon k} \qquad \text{and} \qquad n = 2^{\delta^4 \varepsilon^2 rk}.$$

Note that $\varepsilon\sqrt{m} \ge k$, since $k \ge (C/\varepsilon)\log r$ and $\varepsilon \ge 1/r$, and by our choice of C.

Set $p = 1 - 5\delta\varepsilon$, and for each colour $i \in [r]$, let

- H_i be an independently chosen copy of the random graph $G(m,p)$, and
- $\phi_i \colon [n] \to [m]$ be an independently and uniformly chosen random function.

Now define G_i to be the (random) graph with vertex set $[n]$ and edge set

$$E(G_i) = \big\{ uv : \{\phi_i(u), \phi_i(v)\} \in E(H_i) \big\},$$

that is, a random blow-up of H_i, with parts given by ϕ_i. Define a colouring χ' of K_n by $\chi'(e) = \big\{ i \in [r] : e \in E(G_i) \big\}$, and define the set of *bad* edges to be

$$B = \big\{ e \in E(K_n) : |\chi'(e)| < s \big\}. \tag{3}$$

We will also say that an edge $e = uv \in E(K_n)$ is *i-crossing* if $\phi_i(u) \ne \phi_i(v)$, and define

$$\kappa(e) = \big\{ i \in [r] : e \text{ is } i\text{-crossing} \big\}.$$

We can now define the colouring that we will use to prove Theorem 1.2.

[1] In fact taking $\delta = 2^{-5}$ would suffice, but we will not make any attempt to optimise the value of δ.

Definition 2.1. For each $e \in E(K_n)$, we define the set of colours $\chi(e) \subset [r]$ by

$$\chi(e) = \begin{cases} \chi'(e) & \text{if} \quad e \notin B, \\ \kappa(e) & \text{if} \quad e \in B. \end{cases}$$

Our task is to show that with high probability $|\chi(e)| \geq s$ for every $e \in E(K_n)$, and moreover that χ contains no monochromatic copy of K_k. We start with the former.

Lemma 2.2. *With high probability, $|\chi(e)| \geq s$ for every $e \in E(K_n)$.*

Proof. Note that for each $i \in [r]$ we have $\Pr(i \notin \kappa(e)) = 1/m$ all independently, by the definition of the functions ϕ_i. By the union bound over the set of $r - s = \varepsilon r$ missed colours,

$$\Pr\big(|\chi(e)| < s\big) \leq \Pr\big(|\kappa(e)| < s\big) \leq \binom{r}{\varepsilon r}\left(\frac{1}{m}\right)^{\varepsilon r} \leq \left(\frac{e}{\varepsilon m}\right)^{\varepsilon r} \leq 2^{-\delta^3\varepsilon^2 rk} \leq \frac{1}{n^3}$$

since $k \geq (C/\varepsilon)\log r$ and $C = \delta^{-3}$ imply that $\varepsilon m \geq \sqrt{m} = 2^{\delta^2\varepsilon k/2}$. Applying Markov's inequality and taking an union bound over edges then proves the lemma. □

To prove that χ contains no monochromatic copy of K_k, we split into two cases, the easier case being the following. Let $t = \delta\varepsilon k^2$.

Lemma 2.3. *With high probability, the colouring χ contains no monochromatic k-clique with at most t bad edges.*

Proof. Suppose χ contains a monochromatic clique $S = \{v_1, \ldots, v_k\}$ of colour $i \in [r]$ such that at most t of the edges $e \in \binom{S}{2}$ are bad. For each $j \in [k]$, let $w_j = \phi_i(v_j) \in V(H_i)$, and observe that the set $W = \{w_1, \ldots, w_k\}$ has size k, since by Definition 2.1, and noting that $\chi'(e) \subset \kappa(e)$, every edge $e \in E(K_n)$ such that $i \in \chi(e)$ is i-crossing.

Now, if $e = v_j v_\ell \in \binom{S}{2}$ is not a bad edge, then $i \in \chi(e) = \chi'(e)$, and hence $w_j w_\ell \in E(H_i)$. Since there are at most t bad edges in $\binom{S}{2}$, it follows that $e\big(H_i[W]\big) \geq \binom{k}{2} - t > p\binom{k}{2} + \delta\varepsilon k^2$, since $p = 1 - 5\delta\varepsilon$ and $t = \delta\varepsilon k^2$. Since $H_i[W] \sim G(k,p)$, it follows from Chernoff's inequality that this event has probability at most $e^{-\delta^2\varepsilon k^2}$. By the union bound, the probability that χ contains a monochromatic clique with at most t bad edges is at most

$$r\binom{m}{k}e^{-\delta^2\varepsilon k^2} \leq r\big(2^{\delta^2\varepsilon k} \cdot e^{-\delta^2\varepsilon k}\big)^k.$$

Since $\delta^3\varepsilon k \geq \log r$, the right-hand side tends to zero as $k \to \infty$, as required. □

Our remaining task is to show that, with high probability, no graph F of the family

$$\mathcal{F} = \big\{F \subset K_n \,:\, v(F) = k \text{ and } e(F) = t\big\}$$

is such that $F \subset B$.[2] We will not be able to prove this using a simple first-moment argument, summing over all graphs $F \in \mathcal{F}$, since the probability of the event $\{F \subset B\}$ is not always sufficiently small. Instead, we will identify a 'bottleneck event' for each $F \in \mathcal{F}$.

[2]Here, and below, we abuse notation slightly by treating the set of bad edges B as a graph.

To do so, choose a total ordering $\prec_F$ on the vertices of F such that $u \prec_F v$ implies $d_F(u) \geq d_F(v)$. In other words, we order the vertices according to their degrees in F, breaking ties arbitrarily. Now, define

$$Q_i(F) = \big\{ v \in V(F) \,:\, \exists\, u \in V(F) \text{ with } u \prec_F v \text{ such that } \phi_i(u) = \phi_i(v) \big\}$$

to be the set of vertices which share a part of ϕ_i with another vertex of F that comes earlier in the order $\prec_F$. We remark that if $u \prec_F v$, then $u \neq v$. In the following two lemmas, we will bound the probability in two different ways, depending on the size of

$$X_F = \sum_{i=1}^{r} \sum_{v \in Q_i(F)} d_F(v).$$

Lemma 2.4. *With high probability, there does not exist $F \in \mathcal{F}$ with $X_F \leq \varepsilon rt/2$ and $F \subset B$.*

Proof. We first reveal the random functions $\phi_1, \ldots, \phi_r$, and therefore the sets $Q_i(F)$ (and hence also the random variable X_F) for each $F \in \mathcal{F}$. To prove the lemma we will only use the randomness in the choice of $H_1, \ldots, H_r$. More precisely, we will consider the set

$$Y = \big\{ (uv, i) \in E(F) \times [r] : u, v \notin Q_i(F) \big\}$$

of pairs $(e, i) \in E(F) \times [r]$ such that neither endpoint of e is contained in $Q_i(F)$, and

$$Z = \sum_{(e,i) \in Y} \mathbb{1}\big[e \notin E(G_i) \big],$$

the number of such pairs for which $i \notin \chi'(e)$. Note that, for each $i \in [r]$, the graph $\{e : (e,i) \in Y\}$ is contained in a clique with at most one vertex in each part of ϕ_i. The events $\{e \in E(G_i)\}$ for $(e,i) \in Y$ are therefore independent, and hence $Z \sim \mathrm{Bin}(|Y|, 1-p)$. Since $|Y| \leq rt$, Z is dominated by a binomial random variable with expectation $(1-p)rt = 5\delta\varepsilon rt$.

If $F \subset B$, then for each edge $e \in E(F)$, there are at least εr colours $i \in [r]$ such that $e \notin E(G_i)$. Thus

$$\sum_{i=1}^{r} \sum_{e \in E(F)} \mathbb{1}\big[e \notin E(G_i) \big] \geq \varepsilon rt.$$

Therefore, if $X_F \leq \varepsilon rt/2$, then $Z \geq \varepsilon rt/2$, since for each vertex $v \in Q_i(F)$ we remove at most $d_F(v)$ edges from Y. By Chernoff's inequality, it follows that for a fixed $F \in \mathcal{F}$ we have $\Pr(X_F \leq \varepsilon rt/2 \text{ and } F \subset B) \leq e^{-\delta\varepsilon rt}$. Taking a union bound and recalling that $t = \delta\varepsilon k^2$, it follows that the probability that there exists $F \in \mathcal{F}$ with $X_F \leq \varepsilon rt/2$ and $F \subset B$ is at most

$$\binom{n}{k} \binom{\binom{k}{2}}{t} e^{-\delta\varepsilon rt} \leq \left(\frac{en}{k} \left(\frac{e}{\delta\varepsilon} \right)^{\delta\varepsilon k} e^{-\delta^2\varepsilon^2 rk} \right)^{k} \to 0,$$

as claimed, where in the final step we used our choice of $n = 2^{\delta^4\varepsilon^2 rk}$, the bound $\varepsilon \geq (\log r)/r$, which holds by our assumption that $s \leq r - C\log r$, and our choice of $C = 1/\delta^3$. $\square$

Finally, we will use the randomness in $\phi_1, \ldots, \phi_r$ to show that X_F is always small.

Lemma 2.5. *With high probability, $X_F \le \varepsilon rt/2$ for every $F \in \mathcal{F}$.*

Proof. For each graph $F \in \mathcal{F}$, and each $j \in \big\{1, \ldots, \lceil \log_2 k \rceil \big\}$, define

$$A_j(F) = \Big\{ v \in V(F) : 2^{-j}k \le d_F(v) < 2^{-j+1}k \Big\} \quad \text{and} \quad s_j(F) = \sum_{i=1}^{r} |A_j(F) \cap Q_i(F)|.$$

Note that the random functions $\phi_1, \ldots, \phi_r$ determine $Q_1(F), \ldots, Q_r(F)$, and hence $s_j(F)$. The key step is the following claim, which provides us with our bottleneck event.

Claim. *If $X_F \ge \varepsilon rt/2$, then there exists $\ell \in \big\{1, \ldots, \lceil \log_2 k \rceil \big\}$ satisfying the inequality $s_\ell(F) > \delta \varepsilon r \sum_{j=1}^{\ell} |A_j(F)|$.*

Proof of claim. Observe that

$$\frac{\varepsilon rt}{2} \le X_F = \sum_{i=1}^{r} \sum_{v \in Q_i(F)} d_F(v) \le \sum_{i=1}^{r} \sum_{j=1}^{\lceil \log_2 k \rceil} \frac{k}{2^{j-1}} \cdot |A_j(F) \cap Q_i(F)| = \sum_{j=1}^{\lceil \log_2 k \rceil} \frac{k}{2^{j-1}} \cdot s_j(F).$$

Thus, if the conclusion of the claim fails to hold for every $\ell \in \big\{1, \ldots, \lceil \log_2 k \rceil \big\}$ then we have

$$\begin{aligned} \frac{t}{4\delta k} &\le \frac{1}{\delta \varepsilon r} \sum_{\ell=1}^{\lceil \log_2 k \rceil} \frac{s_\ell(F)}{2^\ell} \le \sum_{\ell=1}^{\lceil \log_2 k \rceil} \frac{1}{2^\ell} \sum_{j=1}^{\ell} |A_j(F)| \\ &= \sum_{j=1}^{\lceil \log_2 k \rceil} |A_j(F)| \sum_{\ell=j}^{\lceil \log_2 k \rceil} \frac{1}{2^\ell} \le \sum_{j=1}^{\lceil \log_2 k \rceil} \frac{|A_j(F)|}{2^{j-1}} \le \frac{2}{k} \sum_{v \in V(F)} d_F(v) = \frac{4t}{k}. \end{aligned}$$

Since $\delta < 2^{-4}$, this is a contradiction, and so the claim follows. $\square$

Fix $\ell \in \big\{1, \ldots, \lceil \log_2 k \rceil \big\}$ such that the conclusion of the claim holds, and set $A := \cup_{j=1}^{\ell} A_j(F)$ and $a := |A|$. Now, if we reveal ϕ_i for the vertices of F one vertex at a time using the order $\prec_F$, then for each vertex $v \in Q_i(F)$ we must choose $\phi_i(v)$ to be one of the (at most k) previously selected elements of $[m]$. The expected number of sets A such that the conclusion of the claim holds is thus at most

$$\sum_{a=1}^{k} n^a \binom{ar}{\delta \varepsilon ar} \left(\frac{k}{m}\right)^{\delta \varepsilon ar} \le \sum_{a=1}^{k} \left(n \cdot \left(\frac{e}{\delta \varepsilon} \cdot \frac{k}{m} \right)^{\delta \varepsilon r} \right)^{a} \to 0$$

as $k \to \infty$, as required, since $n = 2^{\delta^4 \varepsilon^2 rk}$ and $\varepsilon m/k \ge \sqrt{m} = 2^{\delta^2 \varepsilon k/2}$. $\square$

Proof of Theorem 1.2. Combining Lemmas 2.2, 2.3, 2.4 and 2.5, we see that, with high probability, the random colouring χ satisfies $|\chi(e)| \ge s$ for every $e \in E(K_n)$ and contains no monochromatic K_k. Therefore $R_{r,s}(k) > n = 2^{\delta^4 \varepsilon^2 rk}$, as desired. $\square$

References

[1] N. Alon and V. Rödl, Sharp bounds for some multicolor Ramsey numbers, *Combinatorica*, **25** (2005), 125–141.

[2] D. Conlon, J. Fox, X. He, D. Mubayi, A. Suk and J. Verstraëte, Set-coloring Ramsey numbers via codes, arXiv:2206.11371

[3] D. Conlon, J. Fox, X. He, D. Mubayi, A. Suk and J. Verstraëte, Hypergraph Ramsey numbers of cliques versus stars, *Random Structures Algorithms*, to appear, arXiv:2210.03545

[4] D. Conlon, J. Fox and B. Sudakov, Recent developments in graph Ramsey theory, Surveys in Combinatorics, **424** (2015), 49–118.

[5] P. Erdős, A. Hajnal and R. Rado, Partition relations for cardinal numbers, *Acta Math. Acad. Sci.*, **16** (1965), 93–196.

[6] P. Erdős and A. Szemerédi, On a Ramsey type theorem, *Period Math. Acad.*, **2** (1972), 295–299.

[7] X. He and Y. Wigderson, Multicolor Ramsey numbers via pseudorandom graphs, *Electron. J. Combin.*, **27** (2020), Article No. P1.32.

[8] D. Mubayi and J. Verstraëte, A note on pseudorandom Ramsey graphs, *J. Eur. Math. Soc.*, to appear, arXiv:1909.01461.

[9] W. Sawin, An improved lower bound for multicolor Ramsey numbers and a problem of Erdős, *J. Combin. Theory Ser. A*, **188** (2022), Paper No. 105579, 11 pp.

[10] Y. Wigderson, An improved lower bound on multicolor Ramsey numbers, *Proc. Amer. Math. Soc.*, **149** (2021), 2371–2374.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Type-respecting amalgamation and big Ramsey degrees

(Extended abstract)

Andrés Aranda* Samuel Braunfeld* David Chodounský*
Jan Hubička* Matěj Konečný* Jaroslav Nešetřil† Andy Zucker‡

Abstract

We give an infinitary extension of the Nešetřil–Rödl theorem for category of relational structures with special type-respecting embeddings.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-008`

1 Introduction

We use the standard model-theoretic notion of structures allowing functions to be partial. Let L be a *language* with relational symbols $R \in L$ and functional symbols $f \in L$ each having its *arity*. An L-*structure* $\mathbf{A}$ on A is a structure with *vertex set* A, relations $R_{\mathbf{A}} \subseteq A^r$ for every relation symbol $R \in L$ of arity r and partial functions $f_{\mathbf{A}}\colon A^s \to A$ for every function symbol $f \in L$ of arity s. If the set A is finite say that $\mathbf{A}$ is *finite* (it may still have infinitely many relations if L is infinite). We consider only L-structures with finitely many or countably infinitely many vertices. Language L is *relational* if it contains no function symbols. We say that $\mathbf{A}$ is a *substructure* of $\mathbf{B}$ and write $\mathbf{A} \subseteq \mathbf{B}$ if the identity map is an embedding $\mathbf{A} \to \mathbf{B}$. Let $\mathcal{K}$ be a class of L-structures. We say that $\mathcal{K}$ is *hereditary* if it is closed for substructures. We say that L-structure $\mathbf{U} \in \mathcal{K}$ is $\mathcal{K}$-*universal* if every L-structure in $\mathcal{K}$ embeds to $\mathbf{U}$.

*Department of Applied Mathematics (KAM), Charles University, Prague, Czech Republic. E-mail: `{aranda|braunfeld|chodounsky|hubicka|matej}@kam.mff.cuni.cz`

†Computer Science Institute (IUUK), Charles University, Prague, Czech Republic. E-mail: `nesetril@iuuk.mff.cuni.cz`

‡Department of Pure Mathematics, University of Waterloo, Canada. E-mail: `a3zucker@uwaterloo.ca`

Given L-structures $\mathbf{A}$ and $\mathbf{B}$, we denote by $\binom{\mathbf{B}}{\mathbf{A}}$ the set of all embeddings from $\mathbf{A}$ to $\mathbf{B}$. We write $\mathbf{C} \longrightarrow (\mathbf{B})^{\mathbf{A}}_{k,l}$ to denote the following statement: for every colouring χ of $\binom{\mathbf{C}}{\mathbf{A}}$ with k colours, there exists an embedding $f\colon \mathbf{B} \to \mathbf{C}$ such that χ does not take more than l values on $\binom{f(\mathbf{B})}{\mathbf{A}}$. For a countably infinite L-structure $\mathbf{B}$ and its finite substructure $\mathbf{A}$, the *big Ramsey degree* of $\mathbf{A}$ in $\mathbf{B}$ is the least number $D \in \mathbb{N} \cup \{\infty\}$ such that $\mathbf{B} \longrightarrow (\mathbf{B})^{\mathbf{A}}_{k,D}$ for every $k \in \mathbb{N}$. We say that L-structure $\mathbf{B}$ has *finite big Ramsey degrees* if the big Ramsey degree of every finite substructure $\mathbf{A}$ of $\mathbf{B}$ is finite. In general, we are interested in the following question: Given a hereditary class of L-structures $\mathcal{K}$, do $\mathcal{K}$-universal L-structures in $\mathcal{K}$ have finite big Ramsey degrees? Notice that if one $\mathcal{K}$-universal L-structure in $\mathcal{K}$ has finite big Ramsey degrees then all of them do. The study of big Ramsey degrees originated in 1960's Laver's unpublished proof that the big Ramsey degrees of the order of rationals are finite. This result was refined and a precise formula was obtained by Devlin [8]. This area has recently been revitalized with a rapid progress regarding big Ramsey degrees of structures in finite binary languages (see e.g. recent survey [10]).

We call an L-structure $\mathbf{A}$ *irreducible* if for every pair of vertices $u, v \in A$ there exist a relational symbol $R \in L$ and a tuple $\vec{t} \in R_{\mathbf{A}}$ such that $u, v \in \vec{t}$. Given set of L-structures $\mathcal{F}$, L-structure $\mathbf{A}$ is $\mathcal{F}$*-free* if there there is no $\mathbf{F} \in \mathcal{F}$ with an embedding $\mathbf{F} \to \mathbf{A}$. The class of all (finite and countably infinite) $\mathcal{F}$-free L-structures is denoted by $\mathrm{Forb_e}(\mathcal{F})$. With these definitions we can state a recent result:

Theorem 1.1 (Zucker [21])**.** *Let L be a finite binary relational language, and $\mathcal{F}$ a finite set of finite irreducible L-structures. Then every $\mathrm{Forb_e}(\mathcal{F})$-universal L-structure has finite big Ramsey degrees. (In other words, for every finite substructure $\mathbf{A}$ of $\mathbf{U}$ there exists finite $D = D(\mathbf{A})$ such that $\mathbf{U} \longrightarrow (\mathbf{U})^{\mathbf{A}}_{k,D}$ for every $k > 1$.)*

This result can be seen as an infinitary variant of well known Nešetřil–Rödl theorem (one of the fundamental results of structural Ramsey theory) which can be stated as follows:

Theorem 1.2 (Nešetřil–Rödl theorem [17, 18])**.** *Let L be a relational language, $\mathcal{F}$ a set of finite irreducible L-structures. Then for every finite $\mathbf{A} \in \mathrm{Forb_e}(\mathcal{F})$ there exists a finite integer $d = d(\mathbf{A})$ such that for every finite $\mathbf{B} \in \mathrm{Forb_e}(\mathcal{F})$ and finite $k > 0$ there exists a finite $\mathbf{C} \in \mathrm{Forb_e}(\mathcal{F})$ satisfying $\mathbf{C} \longrightarrow (\mathbf{B})^{\mathbf{A}}_{k,d}$.*

To see the correspondence of Theorems 1.1 and 1.2 choose $\mathcal{F}$ as in Theorem 1.1 and a finite $\mathcal{F}$-free L-structure $\mathbf{A}$. By Theorem 1.1 there is a finite $D = D(\mathbf{A})$ such that every $\mathrm{Forb_e}(\mathcal{F})$-universal L-structure $\mathbf{U}$ satisfies $\mathbf{U} \longrightarrow (\mathbf{U})^{\mathbf{A}}_{k,D}$ for every $k > 0$. By $\mathrm{Forb_e}(\mathcal{F})$-universality of $\mathbf{U}$ for every $\mathcal{F}$-free L-structure $\mathbf{B}$ we have $\mathbf{U} \longrightarrow (\mathbf{B})^{\mathbf{A}}_{k,D}$ and by compactness there exists a finite substructure $\mathbf{C}$ of $\mathbf{U}$ such that $\mathbf{C} \longrightarrow (\mathbf{B})^{\mathbf{A}}_{k,D}$. In general, $D(\mathbf{A})$ (characterised precisely in [1]) differs from $d(\mathbf{A})$, the number of linear orderings of A. However, the proof of Theorem 1.1 can be used to recover precise bounds on $d(\mathbf{A})$.

Comparing Theorems 1.1 and 1.2, it is natural to ask whether the assumptions about finiteness of $\mathcal{F}$, finiteness of the language L, and relations being only binary can be dropped from Theorem 1.1. It is known that the first condition can not be omitted: Sauer [19] has shown that there exist infinite families $\mathcal{F}$ of finite irreducible L-structures where $\mathrm{Forb_e}(\mathcal{F})$-universal structures have infinite big Ramsey degrees of vertices. This is true even for language L containing only one binary relation (digraphs). The latter two conditions remain open.

Until recently, most bounds on big Ramsey degrees were for L-structures in binary languages only. Techniques to give bounds on big Ramsey degrees of 3-uniform hypergraphs have been announced in Eurocomb 2019 [4] and published recently [5]; they were later extended to languages with arbitrary relational symbols [6]. Extending links between the Hales–Jewett theorem [12], Carlson–Simpson theorem [7] and big Ramsey degrees established in [15], a Ramsey-type theorem for trees with successor operations has been introduced [3] which extends to all known big Ramsey results on L-structures. However, the following problem remains open:

Problem 1.3. *Let $L = \{E, H\}$ be a language with one binary relation E and one ternary relation H. Let $\mathbf{F}$ be the L-structure where $F = \{0, 1, 2, 3\}$, $R^F = \{(1, 0), (1, 2), (1, 3)\}$, $H = \{(0, 2, 3)\}$. Denote by $\mathcal{K}$ the class of all L-structures $\mathbf{A}$ such that there is no monomorphism $\mathbf{F} \to \mathbf{A}$. Do $\mathcal{K}$-universal L-structures have finite big Ramsey degrees?*

Thise problem demonstrates unforeseen obstacles on giving a natural infinitary generalization of the Nešetřil–Rödl theorem. We give a new approach which avoids this issue and which suggests perhaps the proper setting for big Ramsey degrees.

Finite structural Ramsey results are most often proved by refinements of the Nešetřil–Rödl partite construction [18]. This technique does not generalize to infinite structures due to essential use of backward induction. Upper bounds on big Ramsey degrees are based on Ramsey-type theorems on trees (e.g. the Halper–Läuchli theorem [13], Milliken's tree theorem [16], the Carlson–Simpson theorem [7], and their various refinements [21, 9, 11]). This proof structure may seem unexpected at first glance but is justified by the existence of unavoidable colourings (based on idea of Sierpiński) which are constructed by assigning colors according to subtrees of the tree of 1-types (see e.g. [2, 10]). The exact characterisations of big Ramsey degrees can then be understood as an argument that this proof structure is in a very specific sense the only possible: the trees used to give upper bounds are also encoded in the precise characterisations of big Ramsey degrees.

We briefly review the construction of tree of 1-types. Recall that a (model-theoretic) *tree* is a partial order $(T, \leq)$ where the down-set of every $x \in T$ is a finite chain. An *enumerated L-structure* is simply an L-structure $\mathbf{U}$ whose underlying set is the ordinal $|\mathbf{U}|$. Fix a countably infinite enumerated L-structure $\mathbf{U}$. Given vertices u, v and an integer n satisfying $\min(u, v) \geq n \geq 0$, we write $u \sim^{\mathbf{U}}_n v$, and say that *u and v are of the same (quantifier-free) type over* $0, 1, \ldots, n-1$, if the L-structure induced by $\mathbf{U}$ on $\{0, 1, \ldots, n-1, u\}$ is identical to the L-structure induced by $\mathbf{U}$ on $\{0, 1, \ldots, n-1, v\}$ after renaming vertex v to u. We write $[u]^{\mathbf{U}}_n$ for the $\sim^{\mathbf{U}}_n$-equivalence class of vertex u.

Definition 1.1 (Tree of 1-types)**.** Let $\mathbf{U}$ be an infinite (relational) enumerated L-structure. Given $n < \omega$, write $\mathbb{T}_{\mathbf{U}}(n) = \omega/\sim^{\mathbf{U}}_n$. A (quantifier-free) *1-type* is any member of the disjoint union $\mathbb{T}_{\mathbf{U}} := \bigsqcup_{n<\omega} \mathbb{T}_{\mathbf{U}}(n)$. We turn $\mathbb{T}_{\mathbf{U}}$ into a tree as follows. Given $x \in \mathbb{T}_{\mathbf{U}}(m)$ and $y \in \mathbb{T}_{\mathbf{U}}(n)$, we declare that $x \leq^{\mathbb{T}}_{\mathbf{U}} y$ if and only if $m \leq n$ and $x \supseteq y$.

One can associate every vertex of $v \in \mathbf{U}$ with its corresponding equivalence class in $\simeq^{\mathbf{U}}_{v}$. This way every substructure $\mathbf{A} \subseteq \mathbf{U}$ corresponds to a subset of nodes of the tree $\mathbb{T}_{\mathbf{U}}$. Sierpiński-like colourings can be then constructed by considering shapes of the meet closures of nodes corresponding to each given copy. Every type in $x \in \omega/\sim^{\mathbf{U}}_{n}$ can be described as an L-structure $\mathbf{T}$ with vertex set $T = \{0, 1, \dots n-1, t\}$ such that for every $v \in x$ it holds that L-structure induced by $\mathbf{U}$ on $\{0, 1, \dots n-1, v\}$ is $\mathbf{T}$ after renaming t to v. This is very useful in the setting where types originating from multiple enumerated L-structures are considered (see for instance [15, 2]).

The concept of the tree of 1-types was implicit in early proofs (such as in Devlin's thesis) and became explicit later. The tree of 1-types itself is, however, not sufficient to give upper bounds on big Ramsey degrees for L-structures in languages containing symbols of arity 3 and more. Upper bounds in [5] and [6] are based on the product form of the Milliken tree theorem which in turn suggests the following notion of a weak type.

For the rest of this note, fix a relational language L containing a binary symbol $\leq$. For all L-structures, $\leq$ will always be a linear order on vertices which is either finite or of order-type ω. This will describe the enumeration. All embeddings will be monotone.

Definition 1.2 (Weak type)**.** We denote by L^f the language L extended by unary function symbol f. An L^f-structure $\mathbf{T}$ is a *weak type* of *level* ℓ if

1. $T = \{0, 1, \dots, \ell-1, t_0, t_1, \dots\}$ where vertices t_i are called *type vertices.*
2. For every $R \in L$ and $\vec{t} \in R_{\mathbf{T}}$ it holds that $\vec{t} \cap \{t_0, t_1, \dots\}$ is a (possibly empty) initial segment of type vertices (i.e. set of the form $\{t_i : i \in k\}$ for some $k \in \omega$) and $\vec{t} \cap \{0, 1, \dots, \ell-1\} \neq \emptyset$.
3. For every $i > 0$ we put $f_{\mathbf{T}}(t_i) = t_{i-1}$, $f_{\mathbf{T}}(t_0) = t_0$, and $f_{\mathbf{T}}$ is undefined otherwise.

Weak types thus give less information than standard model-theoretic k-types [14]. Function f is added to type vertices to distinguish them from normal vertices. This will be useful in later constructions. Notice that while technically weak type has infinitely many types vertices, thanks to condition 2 of Definition 1.2, if the language L contains no relations of arity $r+1$ or more, vertices $t_{r-1}, t_r, \dots$ will be isolated. In particular:

Observation 1.4. *If L contains only unary and binary symbols then there is one-to-one correspondence between 1-types and weak types because only type vertex t_0 carries interesting structure.*

1-types describes one vertex extensions of an initial part of the enumerated L-structure. The weak-type equivalent of this is the following:

Definition 1.3 (Weak type of a tuple)**.** Let $\mathbf{A}$ be an enumerated L-structure, $\mathbf{T}$ a weak type of level $\ell \in A \subseteq \omega$ and $\vec{a} = (a_0, a_1, \dots, a_{k-1})$ an increasing tuple of vertices from $A \setminus \ell$. We say that $\vec{a}$ *has type* $\mathbf{T}$ on level ℓ if the function $h \colon T \to A$ given by:

$$h(x) = \begin{cases} x & \text{if } x \in \ell, \\ a_i & \text{if } x = t_i \text{ for some } i < k \end{cases}$$

has the property that for every $R \in L$ and $\vec{b}$ a tuple of vertices in $\{0, 1, \ldots, \ell{-}1, t_0, t_1, \ldots, t_{k-1}\}$ such that $\vec{b} \cap \{t_0, t_1, \ldots\}$ is an initial segment of type vertices and $\vec{b} \cap \{0, 1, \ldots, \ell - 1\} \neq \emptyset$ it holds that $\vec{b} \in R_{\mathbf{T}} \iff h(\vec{b}) \in R_{\mathbf{A}}$.

Definition 1.4 (Tree of weak types)**.** Given an enumerated L-structure $\mathbf{U}$, its *tree of weak types* consists of all L^f-structures $\mathbf{T}$ that are weak types of some tuple of $\mathbf{U}$ on some level $\ell \in U$ ordered by $\subseteq$.

Given an enumerated L-structure $\mathbf{A}$ and a weak type $\mathbf{T}$, we say that $\mathbf{T}$ *extends* $\mathbf{A}$ if $\mathbf{T} \setminus \{t_0, t_1, \ldots\} = \mathbf{A}$. Given two types $\mathbf{T}$ and $\mathbf{T}'$ that extend $\mathbf{A}$, and $n \geq 0$, we say that $\mathbf{T}$ and $\mathbf{T}'$ agree as n-types if $\mathbf{T} \restriction (A \cup \{t_0, t_1, \ldots t_{n-1}\}) = \mathbf{T}' \restriction (A \cup \{t_0, t_1, \ldots t_{n-1}\})$.

A standard technique for proving infinite Ramsey-type theorems is to work with finite approximations of the embeddings considered. See e.g. Todorcevic's axiomatization of Ramsey spaces [20]. Initial approximations of our embeddings will be described as follows:

Definition 1.5 (Structure with types)**.** Given a finite enumerated L-structure $\mathbf{A}$, $\mathbf{A}^+$ denotes the L-structure created from the disjoint union of all weak types extending $\mathbf{A}$ by

1. identifying all copies of $\mathbf{A}$, and,
2. identifying the copy of vertex t_i of weak type $\mathbf{T}$ and with the copy of t_i of weak type $\mathbf{T}'$ whenever $\mathbf{T}$ and $\mathbf{T}'$ agree as $i + 1$ types.

Observe that thanks to the function f added to weak types, for any two L-structures with types $\mathbf{A}^+$ and $\mathbf{B}^+$, every embedding $h\colon \mathbf{A}^+ \to \mathbf{B}^+$ is also a map from weak types of $\mathbf{A}$ on level $|A|$ to weak types of $\mathbf{B}$ of level $|B|$.

Given an L-structure $\mathbf{A}$ and a vertex v, we denote by $\mathbf{A}({<}v)$ the L-structure induced by $\mathbf{A}$ on $\{a \in A; a < v\}$ and call it the *initial segment* of $\mathbf{A}$. The key notion for our approach is to restrict attention to embedding which behave well with respect to weak types. That is, for every initial segment of the L-structure, the rest of the embedding can be summarized via embedding of weak types extending the initial segments.

Definition 1.6 (Type-respecting embeddings of L-structures)**.** Given enumerated L-structures $\mathbf{A}$ and $\mathbf{B}$ and an embedding $h\colon \mathbf{A} \to \mathbf{B}$, we say that h is *type-respecting* if for every $v \in A$ there exists an embedding $h^v\colon \mathbf{A}({<}v)^+ \to \mathbf{B}({<}h(v))^+$ such that the weak types of tuples in $\mathbf{B}$ on level $h(v)$ consisting only of vertices of $h[A]$ are all in the image $h^v[\mathbf{A}]$.

Definition 1.7 ($\mathcal{K}$-type-respecting embeddings of initial segments)**.** Let $\mathbf{A}$ and $\mathbf{B}$ be two finite enumerated L-structures. Embedding $h\colon \mathbf{A}^+ \to \mathbf{B}^+$ is *type-respecting* if for every (possibly infinite) L-structure $\mathbf{A}'$ with initial segment $\mathbf{A}$ there exists an L-structure $\mathbf{B}'$ with initial segment $\mathbf{B}$ and a type-respecting embedding $g\colon \mathbf{A} \to \mathbf{B}$ finitely approximated by h. That is $g \restriction A = h \restriction A$ and every weak type in $\mathbf{B}'$ of a tuple consisting of vertices of $g[A]$ of level $g(\max A)$ is in $h[A^+]$.

Given class $\mathcal{K}$ of L-structures we say that $h\colon \mathbf{A}^+ \to \mathbf{B}'^+$ is $\mathcal{K}$-*type-respecting* if for every L-structure $\mathbf{A}' \in \mathcal{K}$ with initial segment $\mathbf{A}$ there exists an structure $\mathbf{B}' \in \mathcal{K}$ with initial segment $\mathbf{B}$ and a type-respecting embedding $g\colon \mathbf{A} \to \mathbf{B}$ finitely approximated by h.

Definition 1.8 (Type-respecting amalgamation property)**.** Let $\mathcal{K}$ be a hereditary class of enumerated L-structures. We say that $\mathcal{K}$ has *type-respecting amalgamation property* if given three finite enumerated L-structures $\mathbf{A}$, $\mathbf{B}$, $\mathbf{B}' \in \mathcal{K}$ such that $B' \setminus B = \{\max B'\}$ and $\mathbf{B}' \restriction \mathbf{B} = \mathbf{B}$, two $\mathcal{K}$-type-respecting embeddings $f\colon \mathbf{A}^+ \to \mathbf{B}^+$, $f'\colon \mathbf{A}^+ \to \mathbf{B}'^+$ and a type-respecting (but not necessarily $\mathcal{K}$-type-respecting) embedding $g\colon \mathbf{B}^+ \to \mathbf{B}'^+$ such that $g \restriction B$ is the identity and $g \circ f = f'$, there exists a $\mathcal{K}$-type-respecting embedding $g'\colon \mathbf{B}^+ \to \mathbf{B}'^+$ such that $g' \circ f = f'$ and $g' \restriction B = \mathrm{Id}$.

Given a class of L-structures $\mathcal{K}$, finite $\mathbf{A} \in \mathcal{K}$ and $\mathbf{B} \in \mathcal{K}$, we denote by $\binom{\mathbf{B}}{\mathbf{A}}^{\mathcal{K}}$ the set of all $\mathcal{K}$-type-respecting embeddings $\mathbf{A}^+ \to \mathbf{B}'^+$ for $\mathbf{B}'$ an initial segment of $\mathbf{B}$. We write $\mathbf{C} \longrightarrow^{\mathcal{K}} (\mathbf{B})^{\mathbf{A}}_{k,l}$ to denote the following statement: for every colouring χ of $\binom{\mathbf{C}}{\mathbf{A}}^{\mathcal{K}}$ with k colours, there exists a type-respecting embedding $f\colon \mathbf{B} \to \mathbf{C}$ such that χ does not take more than l values on $\binom{f(\mathbf{B})}{\mathbf{A}}^{\mathcal{K}}$. For a countably infinite L-structure $\mathbf{B}$ and its finite suborder $\mathbf{A}$, the *big Ramsey degree of* $\mathbf{A}$ *in* $\mathcal{K}$*-type-respecting embeddings* of $\mathbf{A}$ in $\mathbf{B}$ is the least number $D \in \mathbb{N} \cup \{\infty\}$ such that $\mathbf{B} \longrightarrow^{\mathcal{K}} (\mathbf{B})^{\mathbf{A}}_{k,D}$ for every $k \in \mathbb{N}$.

For type-respecting embeddings we can prove the Ramsey property in the full generality (showing that, in this situation, Problem 1.3 is not a problem).

Theorem 1.5. *Let L be a finite relational language. Let $\mathcal{F}$ be a finite family of finite irreducible enumerated L-structures. Denote by $\mathcal{K}_{\mathcal{F}}$ the class of all finite or countably-infinite enumerated L-structures $\mathbf{A}$ where $\leq_{\mathbf{A}}$ is either finite or of order-type ω such that for every $\mathbf{F} \in \mathcal{F}$ there no embedding $\mathbf{F} \to \mathbf{A}$. Assume that $\mathcal{K}_{\mathcal{F}}$ has the type-respecting amalgamation property. Then for every universal L-structure $\mathbf{U} \in \mathcal{K}_{\mathcal{F}}$ and every finite $\mathbf{A} \in \mathcal{K}_{\mathcal{F}}$ there is a finite $D = D(\mathbf{A})$ such that $\mathbf{U} \longrightarrow^{\mathcal{K}} (\mathbf{U})^{\mathbf{A}}_{k,D}$ for every $k \in \mathbb{N}$.*

We show the following:

Proposition 1.6. *Let L be a finite language consisting of binary and unary relational symbols only. Let $\mathcal{F}$ be a finite family of enumerated irreducible L-structures. Then $\mathcal{K}_{\mathcal{F}}$ has the type-respecting amalgamation property. Moreover, Theorem 1.5 implies Theorem 1.1.*

Proof. Fix L, $\mathcal{F}$ and $\mathcal{K}_{\mathcal{F}}$. Let $\mathbf{A}$, $\mathbf{B}$, $\mathbf{B}' \in \mathcal{K}_{\mathcal{F}}$, $f\colon \mathbf{A}^+ \to \mathbf{B}^+$, $f'\colon \mathbf{A}^+ \to \mathbf{B}'^+$ and $g\colon \mathbf{B}^+ \to \mathbf{B}'^+$ be as in Definition 1.8. By Observation 1.4, in order to specify g', it is only necessary to give, for every weak type $\mathbf{T}$ extending $\mathbf{A}$, an image of its type vertex $t_0 \in T$. Let $t' \in B^+$ be a vertex corresponding to t_0. We consider two cases. (1) If $t' \in f[A^+]$ then we put $g'(t') = g(t')$. (2) If $t' \notin f[A^+]$ we put $g'(t') = t''$ where t'' is the only possible image of t' such that there is no relational symbol $R \in L$ such that $R_{\mathbf{B}^+}$ contains a tuple with both t'' and $\max B'$.

To verify that g' is $\mathcal{K}_{\mathcal{F}}$-type-respecting, choose $\mathbf{A}' \in \mathcal{K}_{\mathcal{F}}$ with initial segment $\mathbf{B}$. Construct $\mathbf{A}''$ from $\mathbf{A}$ by inserting a new vertex v after $\max B$ and extending $\leq_{\mathbf{A}''}$. Add the needed tuples to relations to make $\mathbf{B}'$ the initial segment of $\mathbf{A}''$. Finally, for every $R \in L$ and $u \in A'$ with $u > v$, put $(u, v) \in R_{\mathbf{A}''}$ if and only if $(g'(t), u) \in R_{\mathbf{B}^+}$ where is t is the type vertex of $\mathbf{B}$ corresponding to the type of u in $\mathbf{A}'$. Add tuples $(v, u) \in R_{\mathbf{A}''}$ analogously.

To verify that $\mathbf{A}'' \in \mathcal{K}_{\mathcal{F}}$, assume to the contrary that there is $\mathbf{F} \in \mathcal{F}$ and embedding $e\colon \mathbf{F} \to \mathbf{A}''$. Because $\mathbf{A}' \in \mathcal{K}_{\mathcal{F}}$, clearly $v \in e[F]$. Because $\mathbf{B}' \in \mathcal{K}_{\mathcal{F}}$ we also know that

$e[F]$ contains vertices of $\mathbf{A}'' \setminus \mathbf{B}'$. Since $\mathbf{F}$ is irreducible, all such vertices must have types created by condition (1) above. This contradicts that f' is $\mathcal{K}_{\mathcal{F}}$-type-respecting.

To see the moreover part we have to construct a universal $\mathbf{U}$ which is a substructure of some $\mathbf{U}' \in \mathcal{K}_{\mathcal{F}}$ with the property that for every $n \in \mathbb{N}$ there exists $N \in \mathbb{N}$ such that for every $\mathbf{A} \in \mathcal{K}_{\mathcal{F}}$ with n vertices and every embedding $e\colon \mathbf{A} \to \mathbf{U}'$ there exist a structure $\mathbf{E} \in \mathcal{K}_{\mathcal{F}}$ (called an *envelope*) with at most N vertices and a $\mathcal{K}_{\mathcal{F}}$-type-respecting embedding $h\colon \mathbf{E} \to \mathbf{U}$ such that $e[A] \subseteq h[E]$. This follows from Section 4 of [21], because $\mathcal{K}$-type-respecting embeddings in this setup are precisely the aged embeddings from [21]. □

Proposition 1.7. *Let $L' = \{E, H, \leq\}$ and let $\mathbf{F}'$ be an L'-structure created by expanding the L-structure $\mathbf{F}$ from Problem 1.3 by the natural order of vertices. Denote by $\mathcal{K}$ the class of all enumerated L'-structures $\mathbf{A}$ for which there is no monomorphism $\mathbf{F} \to \mathbf{A}$. The class $\mathcal{K}$ has no type-respecting amalgamation property.*

Proof. We give an explicit failure of type-respecting amalgamation showing that the use of Observation 1.4 in the previous proof is essential. Let $\mathbf{A}$ be the empty L'-structure, $\mathbf{B}$ be L'-structure with $B = \{0\}, E_{\mathbf{B}} = H_{\mathbf{B}} = \emptyset$ and $\mathbf{B}'$ be L'-structure with $B' = \{0, 1\}$, $E_{\mathbf{B}'} = \{(0, 1)\}$, $H_{\mathbf{B}'} = \emptyset$. Let $\mathbf{T}_{\mathbf{A}}$ be the unique weak type extending $\mathbf{A}$. Let $\mathbf{T}_{\mathbf{B}}$ be weak type extending $\mathbf{B}$ with $E_{\mathbf{T}_{\mathbf{B}}} = H_{\mathbf{T}_{\mathbf{B}}} = \emptyset$ and $\mathbf{T}'_{\mathbf{B}}$ weak type extending $\mathbf{B}$ with $E_{\mathbf{T}'_{\mathbf{B}}} = \emptyset$ and $H_{\mathbf{T}'_{\mathbf{B}}} = \{(0, t_0, t_1)\}$. Notice that $\mathbf{T}_{\mathbf{B}}$ and $\mathbf{T}'_{\mathbf{B}}$ agree as 1-types and thus in $\mathbf{B}^+$ their vertices t_0 are identified. Finally, let $\mathbf{T}_{\mathbf{B}'}$ and $\mathbf{T}'_{\mathbf{B}'}$ be weak types extending $\mathbf{B}'$ with $E_{\mathbf{T}_{\mathbf{B}'}} = H_{\mathbf{T}'_{\mathbf{B}'}} = \{(0, 1), (1, t_0)\}$, $H_{\mathbf{T}_{\mathbf{B}'}} = \emptyset$, $H_{\mathbf{T}'_{\mathbf{B}'}} = \{(0, t_0, t_1)\}$. Again $\mathbf{T}_{\mathbf{B}}$ and $\mathbf{T}_{\mathbf{B}'}$ agree as 1-types. Now let $f\colon \mathbf{A}^+ \to \mathbf{B}^+$ map $\mathbf{T}_{\mathbf{A}}$ to $\mathbf{T}_{\mathbf{B}}$ and $f^+\colon \mathbf{A}^+ \to \mathbf{B}^+$ map $\mathbf{T}_{\mathbf{A}}$ to $\mathbf{T}_{\mathbf{B}'}$. It is easy to check that these are $\mathcal{K}$-type-respecting. $g\colon \mathbf{B}^+ \to \mathbf{B}'^+$ can be constructed to be type-respecting by mapping type $\mathbf{T}_{\mathbf{B}}$ to $\mathbf{T}_{\mathbf{B}'}$ and $\mathbf{T}'_{\mathbf{B}}$ to $\mathbf{T}'_{\mathbf{B}'}$. However there is no $\mathcal{K}$-type-respecting $g'\colon \mathbf{B}^+ \to \mathbf{B}'^+$. To see that, observe that any image of $\mathbf{T}'_{\mathbf{B}}$ must agree as 1-type with $\mathbf{T}_{\mathbf{B}'}$ and consider $\mathbf{A}'$ with $A' = \{0, 1, 2\}$ and $H_{\mathbf{A}'} = \{(0, 1, 2)\}$. $\mathbf{A}$ is an initial segment of $\mathbf{A}'$ and there is no way to extend g' to a $\mathcal{K}$-type-respecting embedding of $\mathbf{A}'$ to some L-structure in $\mathcal{K}$ since it will always add vertex v after vertex 0 of $\mathbf{A}$ in a way that there is a monomorphism from $\mathbf{F}$ to $\{0, v, 1, 2\}$. □

We conjecture that the answer to Problem 1.3 is in fact negative. It is possible that by concentrating on type-respecting embeddings, the study of big Ramsey degrees can find a proper setting.

Acknowledgement First six authors are supported by the project 21-10775S of the Czech Science Foundation (GAČR). This article is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 810115).

References

[1] Martin Balko, David Chodounský, Natasha Dobrinen, Jan Hubička, Matěj Konečný, Lluis Vena, and Andy Zucker. Exact big Ramsey degrees via coding trees. arXiv:2110.08409, 2021.

[2] Martin Balko, David Chodounský, Natasha Dobrinen, Jan Hubička, Matěj Konečný, Lluis Vena, and Andy Zucker. Characterisation of the big Ramsey degrees of the generic partial order. arXiv:2303.10088, 2023.

[3] Martin Balko, David Chodounský, Natasha Dobrinen, Jan Hubička, Jaroslav Nešetřil, Matěj Konečný, and Andy Zucker. Ramsey theorem for trees with successor operation. In preparation, 2023.

[4] Martin Balko, David Chodounský, Jan Hubička, Matěj Konečný, and Lluis Vena. Big Ramsey degrees of 3-uniform hypergraphs. *Acta Mathematica Universitatis Comenianae*, 88(3):415–422, 2019. Extended abstract for Eurocomb 2021.

[5] Martin Balko, David Chodounský, Jan Hubička, Matěj Konečný, and Lluis Vena. Big Ramsey degrees of 3-uniform hypergraphs are finite. *Combinatorica*, 42(2):659–672, 2022.

[6] Samuel Braunfeld, David Chodounský, Noé de Rancourt, Jan Hubička, Jamal Kawach, and Matěj Konečný. Big Ramsey degrees and infinite languages. arXiv:2301.13116, 2023.

[7] Timothy J. Carlson and Stephen G. Simpson. A dual form of Ramsey's theorem. *Advances in Mathematics*, 53(3):265–290, 1984.

[8] Denis Devlin. *Some partition theorems and ultrafilters on ω*. PhD thesis, Dartmouth College, 1979.

[9] Natasha Dobrinen. The Ramsey theory of the universal homogeneous triangle-free graph. *Journal of Mathematical Logic*, 20(02):2050012, 2020.

[10] Natasha Dobrinen. Ramsey theory of homogeneous structures: current trends and open problems. *arXiv preprint arXiv:2110.00655; Survey for ICM'22*, 2021.

[11] Natasha Dobrinen. The Ramsey theory of Henson graphs. *Journal of Mathematical Logic*, 0(0):2250018, 2023.

[12] Alfred W. Hales and Robert I. Jewett. Regularity and positional games. *Transactions of the American Mathematical Society*, 106:222–229, 1963.

[13] James D. Halpern and Hans Läuchli. A partition theorem. *Transactions of the American Mathematical Society*, 124(2):360–367, 1966.

[14] Wilfrid Hodges. *Model theory*, volume 42. Cambridge University Press, 1993.

[15] Jan Hubička. Big Ramsey degrees using parameter spaces. *arXiv:2009.00967*, 2020.

[16] Keith R. Milliken. A Ramsey theorem for trees. *Journal of Combinatorial Theory, Series A*, 26(3):215–237, 1979.

[17] Jaroslav Nešetřil and Vojtěch Rödl. A structural generalization of the Ramsey theorem. *Bulletin of the American Mathematical Society*, 83(1):127–128, 1977.

[18] Jaroslav Nešetřil and Vojtěch Rödl. The partite construction and Ramsey set systems. *Discrete Mathematics*, 75(1):327–334, 1989.

[19] Norbert Sauer. A ramsey theorem for countable homogeneous directed graphs. *Discrete Mathematics*, 253:45–61, 06 2002.

[20] Stevo Todorcevic. *Introduction to Ramsey spaces*, volume 174. Princeton University Press, 2010.

[21] Andy Zucker. On big Ramsey degrees for binary free amalgamation classes. *Advances in Mathematics*, 408:108585, 2022.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Cycles of every length and orientation in randomly perturbed digraphs

(Extended abstract)

Igor Araujo[*] József Balogh[†] Robert A. Krueger[‡] Simón Piga[§]
Andrew Treglown[§]

Abstract

In 2003, Bohman, Frieze, and Martin initiated the study of randomly perturbed graphs and digraphs. For digraphs, they showed that for every $\alpha > 0$, there exists a constant C such that for every n-vertex digraph of minimum semi-degree at least αn, if one adds Cn random edges then asymptotically almost surely the resulting digraph contains a consistently oriented Hamilton cycle. We generalize their result, showing that the hypothesis of this theorem actually asymptotically almost surely ensures the existence of every orientation of a cycle of every possible length, simultaneously. Moreover, we prove that we can relax the minimum semi-degree condition to a minimum total degree condition when considering orientations of a cycle that do not contain a large number of vertices of indegree 1.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-009`

[*]Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. Email: `igoraa2@illinois.edu`. Supported by UIUC Campus Research Board RB 22000.

[†]Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. Email: `jobal@illinois.edu`. Supported by NSF Grant DMS-1764123, Arnold O. Beckman Research Award (UIUC Campus Research Board RB 22000), the Langan Scholar Fund (UIUC), and NSF RTG Grant DMS-1937241.

[‡]Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. Email: `rak5@illinois.edu`. Supported by the NSF Graduate Research Fellowship Program Grant No. DGE 21-4675.

[§]University of Birmingham, United Kingdom. Email `s.piga@bham.ac.uk`, `a.c.treglown@bham.ac.uk`. Supported by EPSRC grant EP/V002279/1.

1 Introduction

Hamilton cycles are one of the most studied objects in graph theory, and several classical results measure how 'dense' a graph needs to be to force a Hamilton cycle. In particular, in 1952 Dirac [9] proved that every n-vertex graph with minimum degree $\delta(G) \geq n/2$ contains a Hamilton cycle; the minimum degree condition here is best possible.

The Hamiltonicity of directed graphs has also been extensively investigated since the 1960s. A *directed graph*, or *digraph*, is a set of vertices together with a set of ordered pairs of distinct vertices. We think of a digraph as a loop-free multigraph, where every edge is given an orientation from one endpoint to another, and there is at most one edge oriented in each of the two directions between a pair of vertices. An *oriented graph* is a digraph with at most one directed edge between every pair of vertices. An edge from vertex u to vertex v is represented as $\overrightarrow{uv}$ or $\overleftarrow{vu}$. In the digraph setting, there is more than one natural analog of the minimum degree of a graph. The *minimum semi-degree* $\delta^0(D)$ of a digraph D is the minimum of all the in- and outdegrees of the vertices in D; the *minimum total degree* $\delta(D)$ is the minimum number of edges incident to a vertex in D. Ghouila-Houri [14] proved that every strongly connected n-vertex digraph D with minimum total degree $\delta(D) \geq n$ contains a *consistently oriented* Hamilton cycle, that is, a cycle $(v_1, v_2, \ldots, v_n, v_{n+1} = v_1)$ with edges $\overrightarrow{v_iv_{i+1}}$ for all $i \in [n]$. Note that there are n-vertex digraphs D with $\delta(D) = 3n/2 - 2$ that do not contain a consistently oriented Hamilton cycle, so the strongly connected condition in Ghouila-Houri's theorem is necessary.

An immediate consequence of Ghouila-Houri's theorem is that having minimum semi-degree $\delta^0(D) \geq n/2$ forces a consistently oriented Hamilton cycle, and this is best possible. After earlier partial results [15, 16], DeBiasio, Kühn, Molla, Osthus, and Taylor [7] proved that this minimum semi-degree condition in fact forces all possible orientations of a Hamilton cycle, except for the *anti-directed* Hamilton cycle, that is, a cycle $(v_1, v_2, \ldots, v_n, v_{n+1} = v_1)$ with edges $\overrightarrow{v_iv_{i+1}}$ for all odd $i \in [n]$ and $\overleftarrow{v_iv_{i+1}}$ for all even $i \in [n]$, where n is even. Earlier, DeBiasio and Molla [8] showed that the minimum semi-degree threshold for forcing the anti-directed Hamilton cycle is in fact $\delta^0(D) \geq n/2 + 1$.

There has also been interest in Hamilton cycles in random digraphs: the *binomial random digraph* $D(n, p)$ is the digraph with vertex set $[n]$, where each of the $n(n-1)$ possible directed edges is present with probability p, independently of all other edges. Recently, Montgomery [25] determined the sharp threshold for the appearance of any fixed orientation of a Hamilton cycle H in $D(n, p)$, thereby answering a conjecture of Ferber and Long [12] in a strong form. Depending on the orientation of H, the threshold here can vary from $p = \log n/2n$ to $p = \log n/n$.

In this extended abstract, we consider arbitrary orientations of Hamilton cycles in the randomly perturbed digraph model. Introduced in both the undirected and directed setting by Bohman, Frieze, and Martin [3], this model starts with a dense (di)graph and then adds m random edges to it. The overarching question now is how many random edges are required to ensure that the resulting (di)graph *asymptotically almost surely (a.a.s.)* satisfies a given property, that is, with probability tending to 1 as the number of vertices n tends to infinity. For example, Bohman, Frieze, and Martin [3] proved that for every

$\alpha > 0$, there is a $C = C(\alpha)$ such that if we start with an arbitrary n-vertex graph G of minimum degree $\delta(G) \geq \alpha n$ and add Cn random edges to it, then a.a.s. the resulting graph is Hamiltonian. Furthermore, given a constant $0 < \alpha < 1/2$, in a complete bipartite graph with part sizes αn and $(1-\alpha)n$, a linear number of random edges are needed to ensure Hamiltonicity. Thus their result is best possible up to the dependence of C on α. Subsequently, there has been a significant effort to improve our understanding of randomly perturbed graphs. See, e.g., [17, Section 1.3] and the references within for a snapshot of some of the results in the area.

Bohman, Frieze, and Martin [3] also proved the analogous result for consistently oriented Hamilton cycles in the randomly perturbed digraph model. Their result is also best possible up to the dependence of C on α, for similar reasons as the undirected setting.

Theorem 1.1 (Bohman, Frieze, and Martin [3])**.** *For every $\alpha > 0$, there is a $C = C(\alpha)$ such that if D_0 is an n-vertex digraph of minimum semi-degree $\delta^0(D_0) \geq \alpha n$, then $D_0 \cup D(n, C/n)$ a.a.s. contains a consistently oriented Hamilton cycle.*

A notion closely related to Hamiltonicity is *pancyclicity*, which is when a (di)graph contains cycles of every possible length. Bondy [4] generalized Dirac's theorem, showing that if $\delta(G) \geq n/2$ then G is pancyclic or $K_{n/2,n/2}$. Shortly after, Bondy [5] proposed his famous meta-conjecture that any 'non-trivial' sufficient condition for Hamiltonicity should be a sufficient condition for pancyclicity, up to a small number of exceptional graphs. Krivelevich, Kwan, and Sudakov [20] generalized Theorem 1.1 in this way, showing that the same conditions as in Theorem 1.1 imply that the randomly perturbed digraph contains consistently oriented cycles of every length.

Theorem 1.2 (Krivelevich, Kwan, and Sudakov [20])**.** *For every $\alpha > 0$, there is a $C = C(\alpha)$ such that if D_0 is an n-vertex digraph of minimum semi-degree $\delta^0(D_0) \geq \alpha n$, then $D_0 \cup D(n, C/n)$ a.a.s. contains a consistently oriented cycle of every length between 2 and n.*

The original rotation-extension-type proofs of Theorems 1.1 and 1.2 only guarantee consistently oriented cycles. Our main result is a generalization of Theorem 1.2 to allow for all orientations of a cycle of every possible length. Moreover, we find all these cycles simultaneously, i.e., $D_0 \cup D(n, C/n)$ a.a.s. contains all of them. This last property is an example of *universality*, a notion both well-studied in the random graph (e.g., [10, 25]) and randomly perturbed (e.g., [6, 27]) settings.

Theorem 1.3. *For every $\alpha > 0$, there is a $C = C(\alpha)$ such that if D_0 is an n-vertex digraph of minimum semi-degree $\delta^0(D_0) \geq \alpha n$, then $D_0 \cup D(n, C/n)$ a.a.s. contains every orientation of a cycle of every length between 2 and n.*

Theorem 1.3 is best possible in the sense that one really needs to add a linear number of random edges to D_0. Indeed, similarly as before, let D be the complete bipartite digraph with part sizes αn and $(1-\alpha)n$ (where $0 < \alpha < 1/2$). Then one needs to add a linear number of edges to D to ensure a Hamilton cycle of *any* orientation.

It is also natural to try and generalize Theorem 1.1 in another direction, by relaxing the minimum semi-degree condition to a total degree. Unfortunately, this cannot be true for a Hamilton cycle H in which all but $o(n)$ vertices have in- and outdegree 1. Indeed, given $0 < \alpha < 1/2$, let D be the n-vertex digraph which consists of vertex classes S and T of sizes αn and $(1-\alpha)n$ respectively, and whose edge set consists of all possible edges with their startpoint in S and their endpoint in T. Then whilst $\delta(D) = \alpha n$, given any constant C, with probability bounded away from 0, $D \cup D(n, C/n)$ contains a linear number of vertices with outdegree 0 and a linear number of vertices with indegree 0, so it will not contain H.

On the other hand, we show that this type of orientation of a Hamilton cycle is the only one we cannot guarantee. That is, our desired relaxation is possible for all orientations of a Hamilton cycle that contain a linear number of vertices of in- or outdegree 2.

Theorem 1.4. *For every $\alpha, \eta > 0$, there is a $C = C(\alpha, \eta)$ such that if D_0 is an n-vertex digraph of minimum total degree $\delta(D_0) \geq 2\alpha n$, then $D_0 \cup D(n, C/n)$ a.a.s. contains every orientation of a cycle of every length between* 2 *and* n *that contains at most* $(1-\eta)n$ *vertices of indegree* 1*.*

The proof of Theorem 1.4 has the same core ideas as the proof of Theorem 1.3, but there are additional complications and technicalities that come with the weakened degree condition. We prove these two theorems in [1]. In the next section we highlight some of the ideas from the proof of Theorem 1.3.

Notation. We write $\overleftrightarrow{uv}$ if $\overrightarrow{uv}$ and $\overleftarrow{uv}$ are edges and call $\overleftrightarrow{uv}$ a *bidirected edge.* A *bidirected path* is a digraph obtained from an undirected path by replacing each edge uv with a bidirected edge $\overleftrightarrow{uv}$. An *oriented path* is a digraph obtained from an undirected path by replacing each edge uv with a single directed edge; either $\overrightarrow{uv}$ or $\overleftarrow{uv}$. Given an oriented or bidirected path $P = (u_1, \ldots, u_k)$ we call u_1 its *startpoint* and u_k its *endpoint*, distinguishing it from the path $(u_k, \ldots, u_1)$.

2 Some ideas in the proof of Theorem 1.3

Our goal is to show that for a given orientation $\mathcal{C}$ of a cycle, $D_0 \cup D(n, C/n)$ contains $\mathcal{C}$ with probability at least $1 - e^{-n}$. Theorem 1.3 then follows from a union bound over all choices of $\mathcal{C}$, of which there are trivially at most $n2^n$. For the rest of this section we consider only spanning $\mathcal{C}$, as the non-spanning cycle case follows easily from the machinery we set up to deal with arbitrary orientations of a Hamilton cycle.

Let $D^*(n,p)$ denote the random digraph with vertex set $[n]$ where each possible pair of edges $\overrightarrow{uv}$ and $\overleftarrow{uv}$ are included together, independently of other edges, with probability p. In this way $D^*(n,p)$ is the same as the binomial random graph $G(n,p)$ where we replace every undirected edge with a bidirected edge. Via a coupling argument from [22, 25], to prove that $D_0 \cup D(n, C/n)$ contains $\mathcal{C}$ with probability at least $1 - e^{-n}$, it suffices to show that $D_0 \cup D^*(n, C/n)$ contains $\mathcal{C}$ with probability at least $1 - e^{-n}$. This latter goal will be achievable as we only need to access the randomness in $D^*(n, C/n)$ through a simple pseudorandom property that is easily shown to hold with probability at least $1 - e^{-n}$.

Our argument applies the absorbing method, a technique that was introduced systematically by Rödl, Ruciński, and Szemerédi [28], but that has roots in earlier work (see, e.g., [19]).

2.1 Global absorbers

For our problem, a 'global absorber' in $D_0 \cup D^*(n, C/n)$ is a structure A on a small (but linear size) vertex set with the property that for every sufficiently small set of vertices R, $A \cup R$ contains an oriented path on $|V(A) \cup R|$ vertices with prescribed startpoint and endpoint in R, and so that crucially, this oriented path is a segment of our desired orientation of a Hamilton cycle $\mathcal{C}$. If we can obtain such a structure A, then we can proceed as follows: by applying the pseudorandom property of $D^*(n, C/n)$ we find a bidirected path Q in $D^*(n, C/n)$ disjoint from A that covers almost all of the vertices not in A. Let R be the set of vertices consisting of the startpoint x and endpoint y of Q, together with all those vertices not in Q or A. Using the absorbing property of A we ensure that there is an oriented path Q_R on $V(A) \cup R$ with startpoint y and endpoint x, so that Q_R is a segment of $\mathcal{C}$. Joining the startpoints and endpoints of Q and Q_R, we obtain our desired orientation of a Hamilton cycle $\mathcal{C}$.

2.2 Montgomery's absorbing method

Montgomery [24, 23] introduced an approach to absorbing that has already found a number of applications, for example, to spanning trees in random graphs [24], decompositions of Steiner triple systems [11], and tilings in randomly perturbed graphs [17]. The basic idea of the method is to build a global absorber using a special graph H_m as a framework. The bipartite graph H_m has a bounded maximum degree with vertex classes $X \cup Y$ and Z, and has the property that if one deletes *any* set of vertices of a given size from X, then the resulting graph contains a perfect matching.

Roughly speaking, a global absorber is usually built from H_m as follows: every edge xy in H_m is 'replaced' with a 'local absorber' A_{xy} in such a way that all such absorbers A_{xy} are vertex-disjoint. Here a local absorber A_{xy} is some small gadget that can absorb a certain (constant size) set of vertices S_{xy} associated with x and y.

A reason why this approach has found many applications is that, in some sense, it allows one to construct a global absorber in the case when one can only find 'few' local absorbers, where what is meant by 'few' here depends on the precise setting.

In the proofs of Theorems 1.3 and 1.4 in [1] we use H_m again as a framework to build a global absorber. The reason we use H_m, however, is different from most applications of the method (although morally the reason is similar to why Montgomery used this method in [24]). In our case we will replace *every* edge in H_m incident to $z \in Z$ with the *same* local absorbing gadget A_z. Here A_z is not designed to absorb a fixed set of vertices like before; rather, it has some local flexibility about what vertices it will absorb. The idea is that constructing the global absorber in this way gives us the flexibility to know in advance precisely which (constant size) set of vertices of $\mathcal{C}$ an absorbed vertex w can play the role of. Having this 'advanced warning' about what vertices along $\mathcal{C}$ w can play the role of turns out to be a crucial property of our global absorber; see [1, Section 2] for more details.

Acknowledgments

Much of the research related to this extended abstract was carried out during a visit by the fourth and fifth authors to the University of Illinois at Urbana-Champaign. The authors are grateful to the BRIDGE strategic alliance between the University of Birmingham and the University of Illinois at Urbana-Champaign, which partially funded this visit.
Data availability statement. A full paper containing the proofs of our results can be found on arXiv [1].

References

[1] I. Araujo, J. Balogh, R. A. Krueger, S. Piga, and A. Treglown. On oriented cycles in randomly perturbed digraphs. *arXiv:2212.10112*, 2022.

[2] I. Ben-Eliezer, M. Krivelevich, and B. Sudakov. The size Ramsey number of a directed path. *Journal of Combinatorial Theory, Series B*, 102:743–755, 2012.

[3] T. Bohman, A. Frieze, and R. Martin. How many edges make a dense graph Hamiltonian? *Random Structures & Algorithms*, 22:33–42, 2003.

[4] J. A. Bondy. Pancyclic graphs I. *Journal of Combinatorial Theory, Series B*, 11:80–84, 1971.

[5] J. A. Bondy. Pancyclic graphs: recent results, infinite and finite sets. *Colloq. Math. Soc. János Bolyai, Keszthely*, pages 181–187, 1973.

[6] J. Böttcher, J. Han, Y. Kohayakawa, R. Montgomery, O. Parczyk, and Y. Person. Universality for bounded degree spanning trees in randomly perturbed graphs. *Random Structures & Algorithms*, 55:854–864, 2019.

[7] L. DeBiasio, D. Kühn, T. Molla, D. Osthus, and A. Taylor. Arbitrary orientations of Hamilton cycles in digraphs. *SIAM Journal on Discrete Mathematics*, 29:1553–1584, 2015.

[8] L. DeBiasio and T. Molla. Semi-degree threshold for anti-directed Hamilton cycles. *Electronic Journal of Combinatorics*, 22:P4.34, 2015.

[9] G. A. Dirac. Some theorems on abstract graphs. *Proceedings of the London Mathematical Society*, 2:69–81, 1952.

[10] A. Ferber, G. Kronenberg, and K. Luh. Optimal threshold for a random graph to be 2-universal. *Transactions of the American mathematical Society*, 372:4239–4262, 2019.

[11] A. Ferber and M. Kwan. Almost all Steiner triple systems are almost resolvable. *Forum of Mathematics, Sigma*, 8:e39, 2020.

[12] A. Ferber and E. Long. Packing and counting arbitrary Hamilton cycles in random digraphs. *Random Structures & Algorithms*, 54:499–514, 2019.

[13] A. M. Frieze. On large matchings and cycles in sparse random graphs. *Discrete Mathematics*, 59:243–256, 1986.

[14] A. Ghouila-Houri. Une condition suffisante d'existence d'un circuit hamiltonien. *C.R. Acad. Sci. Paris*, 25:495–497, 1960.

[15] D. Grant. Antidirected Hamilton cycles in digraphs. *Ars Combinatoria*, 10:205–209, 1980.

[16] R. Häggkvist and A. Thomason. Oriented Hamilton cycles in digraphs. *Journal of Graph Theory*, 20:471–479, 1995.

[17] J. Han, P. Morris, and A. Treglown. Tilings in randomly perturbed graphs: bridging the gap between Hajnal-Szemerédi and Johansson-Kahn-Vu. *Random Structures & Algorithms*, 58:480–516, 2021.

[18] F. Joos and J. Kim. Spanning trees in randomly perturbed graphs. *Random Struct. Alg*, 56:169–219, 2020.

[19] M. Krivelevich. Triangle factors in random graphs. *Combinatorics, Probability and Computing*, 6:337–347, 1997.

[20] M. Krivelevich, M. Kwan, and B. Sudakov. Cycles and matchings in randomly perturbed digraphs and hypergraphs. *Combinatorics, Probability and Computing*, 25:909–927, 2016.

[21] M. Krivelevich, M. Kwan, and B. Sudakov. Bounded-degree spanning trees in randomly perturbed graphs. *SIAM Journal on Discrete Math*, 31:155–171, 2017.

[22] C. McDiarmid. General first-passage percolation. *Advances in Applied Probability*, 15:149–161, 1983.

[23] R. Montgomery. Embedding bounded degree spanning trees in random graphs. *arXiv:1405.6559*, 2014.

[24] R. Montgomery. Spanning trees in random graphs. *Advances in Mathematics*, 356:106793, 2019.

[25] R. Montgomery. Spanning cycles in random directed graphs. *arXiv:2103.06751*, 2021.

[26] R. Nenadov and Y. Pehova. On a Ramsey-Turán variant of the Hajnal-Szemerédi theorem. *SIAM Journal on Discrete Mathematics*, 34:1001–1010, 2020.

[27] O. Parczyk. 2-universality in randomly perturbed graphs. *European Journal of Combinatorics*, 87:103–118, 2020.

[28] V. Rödl, A. Ruciński, and A E. Szemerédi. Dirac-type theorem for 3-uniform hypergraphs. *Combinatorics, Probability and Computing*, 15:229–251, 2006.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Tiling problems in edge-ordered graphs

(Extended abstract)

Igor Araujo* Simón Piga† Andrew Treglown‡ Zimu Xiang§

Abstract

Given graphs F and G, a perfect F-tiling in G is a collection of vertex-disjoint copies of F in G that together cover all the vertices in G. The study of the minimum degree threshold forcing a perfect F-tiling in a graph G has a long history, culminating in the Kühn–Osthus theorem [Combinatorica 2009] which resolves this problem, up to an additive constant, for all graphs F. We initiate the study of the analogous question for edge-ordered graphs. In particular, we characterize for which edge-ordered graphs F this problem is well-defined. We also apply the absorbing method to asymptotically determine the minimum degree threshold for forcing a perfect P-tiling in an edge-ordered graph, where P is any fixed monotone path.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-010`

1 Introduction

1.1 Monotone paths in edge-ordered graphs

An *edge-ordered graph* G is a graph equipped with a total order $\leq$ of its edge set $E(G)$. Usually we think of a total order of $E(G)$ as a labeling of the edges with labels from $\mathbb{R}$,

*Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. Email: `igoraa2@illinois.edu`. Research partially supported by UIUC Campus Research Board RB 22000.

†University of Birmingham, United Kingdom. Email: `s.piga@bham.ac.uk`. Research supported by EPSRC grant EP/V002279/1.

‡University of Birmingham, United Kingdom. Email: `a.c.treglown@bham.ac.uk`. Research supported by EPSRC grant EP/V002279/1.

§Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. Email: `zimux2@illinois.edu`

where the labels inherit the total order of $\mathbb{R}$ and where edges are assigned distinct labels. A path P in G is *monotone* if the consecutive edges of P form a monotone sequence with respect to $\leq$. We write $P_k^{\leqslant}$ for the monotone path of length k (i.e., on k edges).

The study of monotone paths in edge-ordered graphs dates back to the 1970s. Chvátal and Komlós [7] raised the following question: what is the largest integer $f(K_n)$ such that every edge-ordering of K_n contains a copy of the monotone path $P_{f(K_n)}^{\leqslant}$ of length $f(K_n)$? Over the years there have been several papers on this topic [4, 5, 6, 11, 17, 19]. In a recent breakthrough, Bucić, Kwan, Pokrovskiy, Sudakov, Tran, and Wagner [4] proved that $f(K_n) \geq n^{1-o(1)}$. The best known upper bound on $f(K_n)$ is due to Calderbank, Chung, and Sturtevant [6] who proved that $f(K_n) \leq (1/2 + o(1))n$. There have also been numerous papers on the wider question of the largest integer $f(G)$ such that every edge-ordering of a graph G contains a copy of a monotone path of length $f(G)$. See the introduction of [4] for a detailed overview of the related literature.

A classical result of Rödl [19] yields a Turán-type result for monotone paths: every edge-ordered graph with n vertices and with at least $k(k+1)n/2$ edges contains a copy of $P_k^{\leqslant}$. More recently, Gerbner, Methuku, Nagy, Pálvölgyi, Tardos, and Vizer [10] initiated the systematic study of the Turán problem for edge-ordered graphs.

It is also natural to seek conditions that force an edge-ordered graph G to contain a collection of vertex-disjoint monotone paths $P_k^{\leqslant}$ that cover all the vertices in G, that is, a *perfect $P_k^{\leqslant}$-tiling* in G. Our first result asymptotically determines the minimum degree threshold that forces a perfect $P_k^{\leqslant}$-tiling.

Theorem 1.1. *Given any $k \in \mathbb{N}$ and $\eta > 0$, there exists an $n_0 \in \mathbb{N}$ such that if $n \geq n_0$ where $(k+1)|n$ then the following holds: if G is an n-vertex edge-ordered graph with minimum degree*

$$\delta(G) \geq (1/2 + \eta)n$$

then G contains a perfect $P_k^{\leqslant}$-tiling. Moreover, for all $n \in \mathbb{N}$ with $(k+1)|n$, there is an n-vertex edge-ordered graph G_0 with $\delta(G_0) \geq \lfloor n/2 \rfloor - 2$ that does not contain a perfect $P_k^{\leqslant}$-tiling.

For the edge-ordered graph G_0 in Theorem 1.1, one can take any edge-ordering of the n-vertex graph consisting of two disjoint cliques whose sizes are as equal as possible under the constraint that neither has size divisible by $k+1$. Our proof of Theorem 1.1 in [2] provides the first application of the so-called *absorbing method* in the setting of edge-ordered graphs.

1.2 The general problem

Let F and G be edge-ordered graphs. We say that G *contains* F if F is isomorphic to a subgraph F' of G; here, crucially, the total order of $E(F)$ must be the same as the total order of $E(F')$ that is inherited from the total order of $E(G)$. In this case we say F' is a *copy of F in G*. For example, if G contains a path F' of length 3 with consecutive edges labeled 5, 17 and 4 then F' is a copy of the path F of length 3 with consecutive edges labeled 2, 3 and 1.

Given edge-ordered graphs F and G, an *F-tiling* in G is a collection of vertex-disjoint copies of F in G; an F-tiling in G is *perfect* if it covers all the vertices in G. In light of Theorem 1.1 we raise the following general question.

Question 1.2. *Let F be a fixed edge-ordered graph on $f \in \mathbb{N}$ vertices and let $n \in \mathbb{N}$ be divisible by f. What is the smallest integer $f(n, F)$ such that every edge-ordered graph on n vertices and of minimum degree at least $f(n, F)$ contains a perfect F-tiling?*

Theorem 1.1 implies that $f(n, P_k^{\leqslant}) = (1/2+o(1))n$ for all $k \in \mathbb{N}$. Note that the *unordered* version of Question 1.2 had been well-studied since the 1960s (see, e.g., [1, 8, 12, 14, 15]) and forty-five years later a complete solution, up to an additive constant term, was obtained via a theorem of Kühn and Osthus [15]. Very recently, the *vertex-ordered graph* version of this problem has been asymptotically resolved [3, 9].

Question 1.2 has a rather different flavor to its graph and vertex-ordered graph counterparts. In particular, there are edge-ordered graphs F for which, given *any* $n \in \mathbb{N}$, there exists an edge-ordering $\leq$ of the complete graph K_n that does not contain a single copy of F. Thus, for such F, Question 1.2 is trivial in the sense that clearly there is no minimum degree threshold $f(n, F)$ for forcing a perfect F-tiling. This motivates Definitions 1.3 and 1.4 below.

Definition 1.3 (Turánable)**.** An edge-ordered graph F is *Turánable* if there exists a $t \in \mathbb{N}$ such every edge-ordering of the graph K_t contains a copy of F.

Definition 1.4 (Tileable)**.** An edge-ordered graph F on f vertices is *tileable* if there exists a $t \in \mathbb{N}$ divisible by f such that every edge-ordering of the graph K_t contains a perfect F-tiling.

The following Ramsey-type result, attributed to Leeb (see [10, 18]), says that in every sufficiently large edge-ordered complete graph we must always find a subgraph which is *canonically ordered.* For $n \geq 5$ there are four non-isomorphic canonical edge-orderings of K_n. We omit the definitions of the canonical edge-orderings in this extended abstract, but they can be found in [10, Section 2.1].

Proposition 1.5. *For every $k \in \mathbb{N}$ there is an $m \in \mathbb{N}$ such that every edge-ordered complete graph K_m contains a copy of K_k that is canonically edge-ordered.*

In [10] it was observed that Proposition 1.5 yields the following full characterization of Turánable graphs.

Theorem 1.6 (Turánable characterization)**.** *An edge-ordered graph F on n vertices is Turánable if and only if all four canonical edge-orderings of K_n contain a copy of F.*

In [2], we prove a result analogous to Theorem 1.6 for tileable graphs. More precisely, we provide a full characterization of all n-vertex tileable graphs with respect to twenty fixed edge-orderings of the complete graph K_n. We call those orderings $\star$-*canonical orderings* of K_n. The full definition of the $\star$-canonical orderings is a little involved and we omit the details here, but the precise description can be found in [2].

Theorem 1.7 (Tileable characterization)**.** *An edge-ordered graph F on n vertices is tileable if and only if all twenty $\star$-canonical orderings of K_n contain a copy of F.*

In [2] we study several consequences of Theorems 1.6 and 1.7. In particular, we prove that the notions of Turánable and tileable are genuinely different. More precisely, we show that there are (infinitely many) edge-ordered graphs that are Turánable but not tileable.

In [10] it is proven that no edge-ordering of K_4 is Turánable and consequently, any edge-ordered graph containing a copy of K_4 is not Turánable and therefore not tileable. Thus, for an edge-ordered graph to be tileable it cannot be too 'dense'. We show in [2] that no edge-ordering of K_4^- is tileable.[1]

A graph H is *universally tileable* if for any given ordering of $E(H)$, the resulting edge-ordered graph is tileable. Similarly, we say that H is *universally Turánable* if given any edge-ordering of H, the resulting edge-ordered graph is Turánable. Using [10, Theorem 2.18], in [2] we characterize all those graphs H that are universally tileable.

Theorem 1.8. *Let H be a graph. The following are equivalent:*
(a) *H is universally tileable;*
(b) *H is universally Turánable;*
(c) (i) *H is a star forest (possibly with isolated vertices),*[2] *or*
  (ii) *H is a path on three edges together with a set (possibly empty) of isolated vertices, or*
  (iii) *H is a copy of K_3 together with a (possibly empty) collection of isolated vertices.*

Moreover, in [2] we determine the asymptotic value of $f(n, F)$ in Question 1.2 for all connected universally tileable edge-ordered graphs F.

The characterization of tileable edge-ordered graphs given in Theorem 1.7 lays the ground for the systematic study of Question 1.2. The second and third authors will investigate this problem further in a forthcoming paper. Already though we can say something about this question. Indeed, an almost immediate consequence of the Hajnal–Szemerédi theorem [12] is the following result.

Theorem 1.9. *Let F be a tileable edge-ordered graph and let $T(F)$ be the smallest possible choice of $t \in \mathbb{N}$ in Definition 1.4 for F. Given any integer $n \geq T(F)$ divisible by $|F|$,*

$$f(n, F) \leq \left(1 - \frac{1}{T(F)}\right) n.$$

The proofs of Theorems 1.1, 1.7, 1.8, and 1.9 can be found in [2]. In the next section we outline the main ideas in the proof of Theorem 1.1.

[1] Recall that K_t^- denotes the graph obtained from K_t by removing an edge.
[2] A *star forest* is a graph whose components are all stars.

2 Outline of the proof of Theorem 1.1

As mentioned above, the proof of Theorem 1.1 applies the *absorbing method*. This approach reduces the problem of finding a perfect $P_k^{\leqslant}$-tiling into two sub-tasks: (i) obtain an 'absorbing structure' A in the host graph G, and (ii) find a $P_k^{\leqslant}$-tiling covering almost all of the vertices in $G \setminus A$.

This latter task is achieved via a relatively straightforward application of a result of Komlós [13]. The main task is therefore constructing the absorbing structure.

Roughly speaking, for an edge-ordered graph G, we say that a set of vertices $A \subseteq V(G)$ is a $P_k^{\leqslant}$*-absorber* if, for every sufficiently small set of vertices $W \subseteq V(G) \setminus A$ whose size is divisible by $k+1$, we have that $G[W \cup A]$ contains a perfect $P_k^{\leqslant}$-tiling.

We apply (an edge-ordered version of) a lemma by Lo and Markström [16, Lemma 1.1] which implies that in order to construct such a $P_k^{\leqslant}$-absorber we only need to find many so-called 'local absorbers' for every pair of vertices $x, y \in V(G)$. More precisely, a *local absorber for x and y* is a small set $L \subseteq V(G)$ with the property that both $G[L \cup \{x\}]$ and $G[L \cup \{y\}]$ contain perfect $P_k^{\leqslant}$-tilings.

To build up such local absorbers L for x and y, we prove a supersaturated version of the aforementioned result of Rödl: every edge-ordered graph with linear average degree contains 'many' copies of $P_k^{\leqslant}$. In particular, as our edge-ordered graph G has $\delta(G) \geq (1/2 + o(1))n$ this allows us to find many copies of $P_{k-1}^{\leqslant}$ in the neighborhood $N_G(v)$ of any vertex $v \in V(G)$. In fact, with some care, one can show the following stronger property: for every two different vertices $x, y \in V(G)$ there are many vertices $w \in V(G)$ so that (a) G contains many copies P_{xw} of $P_{k-1}^{\leqslant}$ for which x (resp. w) can be added to the start or end of P_{xw} to form a copy of $P_k^{\leqslant}$ in G, and (b) G contains many copies P_{yw} of $P_{k-1}^{\leqslant}$ for which y (resp. w) can be added to the start or end of P_{yw} to form a copy of $P_k^{\leqslant}$ in G.

This now gives us the structure we need to construct the local absorbers L for x and y. Indeed, for every choice of w, P_{xw} and P_{yw} above, we define a local absorber $L := V(P_{xw}) \cup V(P_{yw}) \cup \{w\}$. Properties (a) and (b) ensure each such L is indeed a local absorber for x and y, as desired.

Note that from the outline above it may not seem clear why our proof is specific to monotone paths, rather than other edge-orderings of paths. However, the details of the proof very much rely on our paths being monotone. For example, one crucial property we exploit is that if $P = u_1 \cdots u_{k+1}$ is a monotone path, then $u_1 \cdots u_k$ is isomorphic to $u_2 \cdots u_{k+1}$. In other words, the path obtained by dropping the last vertex is isomorphic to the one obtained by dropping the first one. It is not hard to see that this property is satisfied only by monotone paths.

3 Almost perfect tilings and open problems

As part of the proof of Theorem 1.1 in [2], we establish the minimum degree threshold that forces an 'almost perfect' $P_k^{\leqslant}$-tiling in an edge-ordered graph. It is also natural to consider this problem more generally. This motivates the following definition.

Definition 3.1 (Almost tileable)**.** *An edge-ordered graph F is* almost tileable *if for every $0 < \varepsilon < 1$ there exists a $t \in \mathbb{N}$ such every edge-ordering of the graph K_t contains an F-tiling covering all but at most εt vertices of K_t.*

It is easy to see that this notion is equivalent to being Turánable.

Proposition 3.2. *An edge-ordered graph F is almost tileable if and only if F is Turánable.*

Proof. The forward direction is immediate. For the reverse direction, consider any F that is Turánable. Given any $0 < \varepsilon < 1$ define $t := \lceil T(F)/\varepsilon \rceil$. (Recall $T(F)$ is defined in the statement of Theorem 1.9.) Then given any edge-ordering of K_t, by definition of $T(F)$ we may repeatedly find vertex-disjoint copies of F in K_t until we have covered all but fewer than $T(F)$ vertices in K_t. That is, we have an F-tiling covering all but at most εt vertices of K_t, as desired. □

In light of Proposition 3.2 we propose the following question.

Question 3.3. *Let F be a fixed Turánable edge-ordered graph. What is the minimum degree threshold for forcing an almost perfect F-tiling in an edge-ordered graph on n vertices? More precisely, given any $\varepsilon > 0$, what is the minimum degree required in an n-vertex edge-ordered graph G to force an F-tiling in G covering all but at most εn vertices?*

Finally, we know that every Turánable (and therefore tileable) edge-ordered graph F does not contain a copy of K_4. However, we are unaware of any result that forbids F from having large chromatic number.

Question 3.4. *Is it true that for every $k \in \mathbb{N}$ there is a Turánable edge-ordered graph F whose underlying graph has chromatic number at least k?*

Acknowledgments

Much of the research discussed in this extended abstract was carried out during a visit by the second and third authors to the University of Illinois at Urbana-Champaign. The authors are grateful to the BRIDGE strategic alliance between the University of Birmingham and the University of Illinois at Urbana-Champaign, which partially funded this visit. The authors are also grateful to József Balogh for helpful discussions.

Data availability statement. The proofs of our results can be found in [2].

References

[1] Noga Alon and Raphael Yuster. H-factors in dense graphs. *Journal of Combinatorial Theory, Series B*, 66(2):269–282, 1996.

[2] Igor Araujo, Simón Piga, Andrew Treglown, and Zimu Xiang. Tiling edge-ordered graphs with monotone paths and other structures. *arXiv:2305.07294*, 2023.

[3] József Balogh, Lina Li, and Andrew Treglown. Tilings in vertex ordered graphs. *Journal of Combinatorial Theory, Series B*, 155:171–201, 2022.

[4] Matija Bucić, Matthew Kwan, Alexey Pokrovskiy, Benny Sudakov, Tuan Tran, and Adam Zsolt Wagner. Nearly-linear monotone paths in edge-ordered graphs. *Israel Journal of Mathematics*, 238(2):663–685, 2020.

[5] Alewyn P. Burger, Ernest J. Cockayne, and Christina M. Mynhardt. Altitude of small complete and complete bipartite graphs. *Australasian Journal of Combinatorics*, 31:167–177, 2005.

[6] Robert A. Calderbank, Fan R.K. Chung, and Dean G. Sturtevant. Increasing sequences with nonzero block sums and increasing paths in edge-ordered graphs. *Discrete mathematics*, 50:15–28, 1984.

[7] Václav Chvátal and János Komlós. Some combinatorial theorems on monotonicity. *Canadian Mathematical Bulletin*, 14(2):151–157, 1971.

[8] Keresztély Corradi and András Hajnal. On the maximal number of independent circuits in a graph. *Acta Mathematica Hungarica*, 14(3-4):423–439, 1963.

[9] Andrea Freschi and Andrew Treglown. Dirac-type results for tilings and coverings in ordered graphs. *Forum of Mathematics, Sigma*, 10:e104, 2022.

[10] Dániel Gerbner, Abhishek Methuku, Dániel T. Nagy, Dömötör Pálvölgyi, Gábor Tardos, and Máté Vizer. Turán problems for edge-ordered graphs. *Journal of Combinatorial Theory, Series B*, 160:66–113, 2023.

[11] Ronald L. Graham and Daniel J. Kleitman. Increasing paths in edge ordered graphs. *Periodica Mathematica Hungarica*, 3(1-2):141–148, 1973.

[12] András Hajnal and Endre Szemerédi. Proof of a conjecture of P. Erdős. *Combinatorial Theory and Its Applications*, 2, 1970.

[13] János Komlós. Tiling Turán theorems. *Combinatorica*, 20:203–218, 2000.

[14] János Komlós, Gábor Sárközy, and Endre Szemerédi. Proof of the Alon-Yuster conjecture. *Discrete Mathematics*, 235(1-3):255–269, 2001.

[15] Daniela Kühn and Deryk Osthus. The minimum degree threshold for perfect graph packings. *Combinatorica*, 29(1):65–107, 2009.

[16] Allan Lo and Klas Markström. F-factors in hypergraphs via absorption. *Graphs and Combinatorics*, 31:679–712, 2015.

[17] Kevin G Milans. Monotone paths in dense edge-ordered graphs. *Journal of Combinatorics*, 8(3):423–437, 2017.

[18] Jaroslav Nešetřil and Vojtěch Rödl. Statistics of orderings. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 87(2):421–433, 2017.

[19] Vojtěch Rödl. Master's thesis. *Charles University*, 1973.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Graphs without a rainbow path of length 3*

(Extended abstract)

Sebastian Babiński† Andrzej Grzesik‡

Abstract

In 1959 Erdős and Gallai proved the asymptotically optimal bound for the maximum number of edges in graphs not containing a path of a fixed length. We investigate a rainbow version of the theorem, in which one considers $k \geq 1$ graphs on a common set of vertices not creating a path having edges from different graphs and asks for the maximum number of edges in each graph. We prove the asymptotically optimal bound in the case of a path on three edges and any $k \geq 1$.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-011`

1 Introduction

A classical problem in graph theory is to determine the Turán number of a graph F, i.e., the maximum possible number of edges in graphs not containing a particular forbidden structure F as a subgraph. The notable results are exact solutions for a triangle by Mantel [16] and for a complete graph by Turán [17], and an asymptotically optimal bound for any non-bipartite graph by Erdős and Stone [6]. Not much is known for bipartite graphs, but the case of a path was solved asymptotically by Erdős and Gallai [5] in 1951, while in 1975 Faudree and Schelp [7] provided an exact solution.

*This work was supported by the National Science Centre grant 2021/42/E/ST1/00193.

†Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. E-mail: `Sebastian.Babinski@alumni.uj.edu.pl`.

‡Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. E-mail: `Andrzej.Grzesik@uj.edu.pl`.

There are many possible ways to define a rainbow version of the problem. In our work we concentrate on a rainbow version without any additional assumptions and when the number of edges in each color is maximized. Formally, for a graph F and a positive integer k we consider k graphs $G_1, G_2, \dots, G_k$ on the same set of vertices and ask for the maximum possible number of edges in each graph avoiding appearance of a copy of F having at most one edge from each graph. In other words, for every i we color edges of G_i in color i (in particular it means that two vertices can be connected by edges in many colors) and forbid all copies of F having non-repeated colors, so-called rainbow copies.

When the forbidden rainbow graph F is a triangle, it follows from a result of Keevash, Saks, Sudakov and Verstraëte [13] that for $k \geq 4$ colors the best possible number of edges in each color without having a rainbow triangle is equal to $\frac{1}{4}n^2$. This is achieved in the balanced complete bipartite graph (the same in each color) as in Mantel's theorem. Surprisingly, Magnant [15] provided a construction showing that for 3 colors the answer is different. Later, Aharoni, DeVos, de la Maza, Montejano and Šámal [1], answering a question of Diwan and Mubayi [4], proved that in this case the asymptotically optimal bound is $\left(\frac{26-2\sqrt{7}}{81}\right) n^2 \approx 0.2557n^2$. They also asked for similar theorems for bigger cliques, other graphs and different color patterns (in this setting some results were proven in [3] and [14]). Recently, Falgas-Ravry, Markström and Räty [8] completely determined the triples of the asymptotic number of edges in each color that force an existence of a rainbow triangle. Similar problems, but where one maximizes other functions of the number of edges (instead of the number of edges in each color), were considered e.g. in [2, 9, 11, 12].

2 The main result

In our work we consider an arbitrary fixed number of colors $k \geq 1$ and we aim to maximize the number of edges in each color avoiding a rainbow path of length 3. The bound obtained is asymptotically tight.

Theorem 1. *For every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that for every $n \geq n_0$, $k \geq 1$ and graphs $G_1, G_2, \dots, G_k$ on a common set of n vertices, each graph having at least $(f(k)+\varepsilon)\frac{n^2}{2}$ edges, where*

$$f(k) = \begin{cases} \lceil \frac{k}{2} \rceil^{-2} & \text{for } k \leq 6, \\ \frac{1}{2k-1} & \text{for } k \geq 7, \end{cases}$$

there exists a rainbow path with 3 edges. Moreover, the above bound on the number of edges is asymptotically optimal for each $k \geq 1$.

We note that in [12] the same forbidden structure is considered, i.e., a rainbow path of length 3, however when one aims to maximize the product of the number of edges in each color and there are only 3 or 4 colors. While the result for 4 colors provided there implies our result for 4 colors, the results for 3 colors are independent of each other.

In order to avoid struggling with the lower-order error terms and to obtain a structure easier to handle, we rewrite Theorem 1 to a bit different setting. Assuming that Theorem 1 does not hold we obtain an arbitrarily large counterexample with at least $(f(k)+\varepsilon)\frac{n^2}{2}$ edges in each color and without a rainbow path with 3 edges. Using colored graph removal lemma [10] we remove all rainbow walks with 3 edges by deleting at most $\frac{1}{4}\varepsilon n^2$ edges in each color. Then, we add all possible edges without creating rainbow walks with 3 edges. Finally, we group all the vertices into clusters based on the colors on the incident edges. Note that if there is an edge between two clusters (or inside one), then all the vertices between these clusters (or inside this cluster) can be connected by edges in the same color without creating a rainbow walk of length 3. Thus between clusters (and inside them) in each color we have all possible edges or none. Additionally, notice that vertices in a cluster incident to only one or two colors can be all connected by edges in those colors, while vertices incident to more than 2 colors need to form an independent set. Therefore, Theorem 1 can be stated in an equivalent form for such kind of clustered graphs.

Definition 2. For any integer $k \geq 1$ a *clustered graph for k colors* is an edge-colored weighted graph on $\binom{k}{2}+k+1$ vertices with vertex weights $b_{ij}=b_{ji}$ for $1 \leq i < j \leq k$, a_i for $i \in [k]$ and x, in which

- $x \geq 0$, $a_i \geq 0$ for $i \in [k]$ and $b_{ij} \geq 0$ for every $1 \leq i < j \leq k$,
- $\displaystyle\sum_{1\leq i<j\leq k} b_{ij} + \sum_{1\leq i\leq k} a_i + x = 1,$
- for every $i \in [k]$ the vertex of weight a_i is connected in color i with itself, the vertex of weight x and all the vertices of weights b_{ip} for $p \neq i$,
- for every $1 \leq i < j \leq k$ each vertex of weight b_{ij} is connected in colors i and j with itself,
- there are no other edges.

The vertex of weight b_{ij} represents the cluster of $b_{ij}n$ vertices incident to edges colored i and j, the vertex of weight a_i – the cluster of $a_i n$ vertices incident only to edges colored i, and x represents the remaining vertices. Clusters for b_{ij} and a_i are cliques in appropriate colors, while cluster for x is an independent set. This is depicted for $k=3$ in Figure 1.

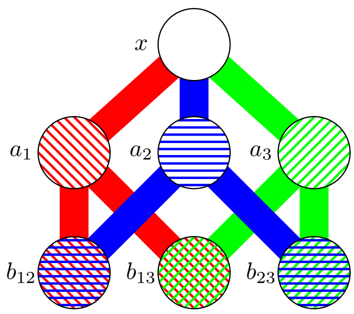


Figure 1: Representation of clusters for $k=3$.

From the definition of a clustered graph it follows that the *density of edges* in color $i \in [k]$ in a clustered graph G is the number $d_i(G) \in [0,1]$ equal to

$$d_i(G) = a_i^2 + \left(\sum_{j \in [k] \setminus \{i\}} b_{ij}^2 \right) + \left(2 \sum_{j \in [k] \setminus \{i\}} a_i b_{ij} \right) + 2a_i x.$$

The equivalent version of Theorem 1 for clustered graphs is the following.

Theorem 3. *For every integer $k \geq 1$, if G is a clustered graph for k colors, then*

$$\min_{i \in [k]} d_i(G) \leq f(k), \text{ where } f(k) = \begin{cases} \lceil \frac{k}{2} \rceil^{-2} & \text{for } k \leq 6, \\ \frac{1}{2k-1} & \text{for } k \geq 7. \end{cases}$$

Theorem 1 follows from Theorem 3, because a possible counterexample leads to a graph with density of edges in each color at least $(f(k) + \frac{1}{2}\varepsilon)\frac{n^2}{2}$ and clusters of vertices behaving as weighted vertices of a related clustered graph. Dividing each cluster size by n we obtain a clustered graph with density of edges in each color at least $f(k) + \frac{1}{2}\varepsilon$, which contradicts Theorem 3. Note that also Theorem 1 implies Theorem 3 as any clustered graph G contradicting Theorem 3 having $d_i(G) \geq f(k) + 2\varepsilon$ for each $i \in [k]$ and some $\varepsilon > 0$ leads for any appropriately large n to a graph on n vertices with at last $(f(k) + \varepsilon)\frac{n^2}{2}$ edges in each color and no rainbow path with 3 edges, which contradicts Theorem 1.

The bound provided in Theorem 3 is tight for every integer $k \geq 1$, because it is possible to construct a clustered graph for k colors G such that $\min_{i \in [k]} d_i(G) = f(k)$:

— for $k = 1$ let $a_1 = 1$;
— for $k = 2$ let $b_{12} = 1$;
— for $k = 3$ let $b_{12} = b_{13} = \frac{1}{2}$;
— for $k = 4$ let $b_{12} = b_{34} = \frac{1}{2}$;
— for $k = 5$ let $b_{12} = b_{34} = b_{15} = \frac{1}{3}$;
— for $k = 6$ let $b_{12} = b_{34} = b_{56} = \frac{1}{3}$;
— for $k = 5$ or $k \geq 7$ let $a_i = \frac{1}{2k-1}$ for each $i \in [k]$, $x = \frac{k-1}{2k-1}$.

In each case the remaining weights are equal to 0.

For $k = 5$ there are two different types of constructions. They are depicted in Figure 2.

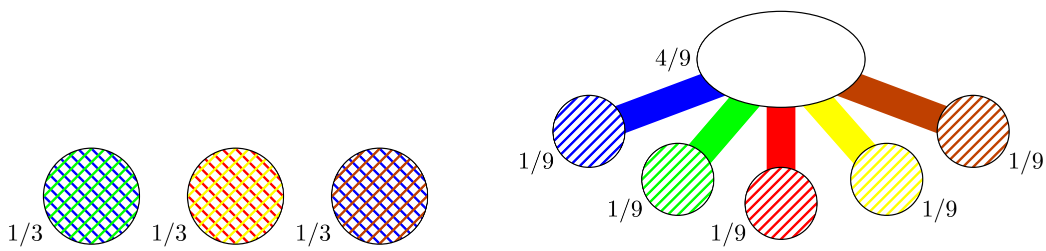


Figure 2: Two possible types of extremal constructions for $k = 5$.

3 Outline of the proof

Theorem 3 is proven by induction. The theorem is trivial for $k \in \{1, 2\}$ as then $f(k) = 1$. Let us fix the smallest $k \geq 3$ for which the theorem does not hold. Take a clustered graph for k colors G maximizing the value of $\min_{i\in[k]} d_i(G)$ and, among such, maximizing the density of edges in any color. The assumption that Theorem 3 does not hold implies that $d_i(G) > f(k)$ for every $i \in [k]$. This and the maximality of G enable to show claims on the weights of the vertices of G, which will lead to a contradiction for each value of $k \geq 3$.

To find many useful bounds on the weights, we introduce an operation of removing and adding weights in a clustered graph for k colors. Intuitively, we remove tiny weights from some of the vertices of positive weight and add them to different vertices. From the maximality of G, such operation cannot enlarge the density of edges in each color, so the density of edges in at least one color needs to drop down (or the densities of edges in every color remain the same). Due to different extremal constructions, we need to consider three main cases: $k \in \{3, 4\}$, $k \in \{5, 6\}$ and $k \geq 7$. Let us denote $d_i = d_i(G)$, $b_i = \sum_{j\in[k],j\neq i} b_{ij}$, and $c_i = a_i + b_i + x$.

In the case of $k = 3$ our conjectured clustered graph G satisfies $\min_{i\in[3]} d_i > \frac{1}{4}$, which implies a simple lower bound on c_i. We prove that for some $i \in [3]$ we have $a_i = 0$ (without loss of generality $a_3 = 0$). By contradiction, if every $a_i > 0$, we can remove appropriate weight from each vertex of weight a_i and add the removed weights to all vertices. It implies, using the maximality of G, a better lower bound for some c_i, say c_3. Now we consider two cases: $x \neq 0$ and $x = 0$. In the former one we find a lower bound on $\sum_{i\in[3]} a_i$ by removing suitable weight from the vertex of weight x and adding weights to each vertex of weight a_i. Together with bounds on $c_i, i \in [3]$ it gives a contradiction. While in the latter case, we show first that $b_{12} > 0$ and removing appropriate weight from the vertex of weight b_{12} and adding weights to each vertex of the graph leads to a contradiction. Once we know that $a_3 = 0$, using the technique of removing and adding weights, we show that $x = 0$ and that $d_3 \leq \min\{d_1, d_2\}$. Then by removing suitable weight from b_{12} and adding weights to a_1 and a_2, we obtain a contradiction which finishes the proof of Theorem 3 for $k = 3$. The case $k = 4$ is a simple corollary of the theorem for $k = 3$ since $f(4) = f(3)$.

The proof of Theorem 3 for $k = 5$ relies on similar techniques. However, as there are two types of constructions, it requires more careful estimations and thus additional bounds and considering more cases. In particular, we prove a different lower bound for b_{ij} when it is positive and a bound on b_i when $a_i = 0$. Having this, depending on the number of $i \in [5]$ such that $a_i = 0$, we bound the sum of all c_i and obtain a contradiction in each case. The proof for $k = 6$ is a simple consequence of the result for $k = 5$.

In the case of $k \geq 7$ we first show that x must be positive. Then we separately prove cases $k = 7$, $k = 8$ and $k \geq 9$. In the first one we sum up lower bounds for c_i, $i \in [7]$ and $\sum_{i\in[7]} a_i$, which implies a lower bound on x. Then, by removing and adding weights between some vertices of weights a_i, a_j and b_{ij}, we get an upper bound on x, which gives a contradiction. The proofs for $k = 8$ and $k \geq 9$ are based on analogous ideas. The main differences come from the fact that the aforementioned bounds on c_i and x are derived using induction and the values of $f(k-2)$ and $f(k)$, which are distinct for each $k \geq 7$.

References

[1] R. Aharoni, M. DeVos, S.G.H. de la Maza, A. Montejano, R. Šámal, A rainbow version of Mantel's Theorem, *Advances in Combinatorics* (2020), 12pp.

[2] D. Chakarborti, J. Kim, H. Lee, H. Liu, J. Seo, On a rainbow extremal problem for color-critical graphs, arXiv: 2204.02575 (2022).

[3] M. DeVos, J. McDonald, A. Montejano, Non-monochromatic triangles in a 2-edge-coloured graph, *Electronic Journal of Combinatoris* (2019), 3–8.

[4] A. Diwan, D. Mubayi, Turán's theorem with colors, preprint, `http://www.math.cmu.edu/~mubayi/papers/webturan.pdf`, 2007.

[5] P. Erdős, T. Gallai, On maximal paths and circuits of graphs, *Acta Math. Hungar.* 10 (1959), 337–356.

[6] P. Erdős, A. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* 52 (1946), 1087–1091.

[7] R.J. Faudree, R.H. Schelp, Path Ramsey numbers in multicolorings, *J. Combin. Theory Ser. B* 19 (1975), 150–160.

[8] V. Falgas-Ravry, K. Markström, E. Räty, Rainbow variations on a theme by Mantel: extremal problems for Gallai colouring templates, arXiv: 2212.07180 (2022).

[9] V. Falgas-Ravry, K. Markström, E. Räty, Minimum degree conditions for rainbow triangles, arXiv: 2305.12772 (2023).

[10] J. Fox, A new proof of the graph removal lemma, *Annals of Mathematics* (2011), 561–579.

[11] P. Frankl, Graphs without rainbow triangles, arXiv: 2203.07768 (2022).

[12] P. Frankl, E. Győri, Z. He, Z. Lv, N. Salia, C. Tompkins, K. Varga, X. Zhu, Extremal results for graphs avoiding a rainbow subgraph, arXiv: 2204.07567 (2022).

[13] P. Keevash, M. Saks, B. Sudakov, J. Verstraëte, Multicolour Turán problems, *Advances in Applied Mathematics* 33 (2004), 238–262.

[14] A. Lamaison, A. Müyesser, M. Tait, On a colored Turán problem of Diwan and Mubayi, *Discrete Mathematics* 345,10 (2022), 113003.

[15] C. Magnant, Density of Gallai Multigraphs, *Electron. J. Comb.* 22 (2015), P1.28.

[16] W. Mantel, Problem 28, *Wiskundige Opgaven* 10 (1907), 60–61.

[17] P. Turán, On an extremal problem in graph theory (in Hungarian), *Matematikai és Fizikai Lapok* 48 (1941), 436–452.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Directed graphs without rainbow triangles*

(Extended abstract)

Sebastian Babiński[†] Andrzej Grzesik[‡] Magdalena Prorok[§]

Abstract

One of the most fundamental questions in graph theory is Mantel's theorem which determines the maximum number of edges in a triangle-free graph of a given order. Recently a colorful variant of this problem has been solved. In such a variant we consider c graphs on a common vertex set, thinking of each graph as edges in a distinct color, and want to determine the smallest number of edges in each color which guarantees the existence of a rainbow triangle. Here, we solve the analogous problem for directed graphs without rainbow triangles, either directed or transitive, for any number of colors. The constructions and proofs essentially differ for $c = 3$ and $c \geq 4$ and the type of the forbidden triangle.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-012`

1 Introduction

A cornerstone of the extremal graph theory is Mantel's theorem from 1907 which determines the maximum possible number of edges in a triangle-free graph of a given order. Its natural generalization, known as a Turán problem, is to determine the maximum possible number

*This work was supported by the National Science Centre grant 2021/42/E/ST1/00193.

[†]Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. E-mail: `Sebastian.Babinski@im.uj.edu.pl`.

[‡]Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. E-mail: `Andrzej.Grzesik@uj.edu.pl`.

[§]AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Krakow, Poland. E-mail: `prorok@agh.edu.pl`.

of edges in an n-vertex graph not containing a given graph F as a subgraph. This is an often studied concept in graph theory with many important results, open problems, and generalizations to various other discrete settings.

In a rainbow version of the Turán problem, for a graph F and an integer c we consider c graphs $G_1, G_2, \dots, G_c$ on the same set of vertices and ask for the maximum possible number of edges in each graph avoiding the appearance of a copy of F having at most one edge from each graph. In other words, for every i we color edges of G_i in color i and forbid all copies of F having non-repeated colors, so-called rainbow copies. Note that if all G_i are exactly the same, then the existence of a rainbow copy of F is equivalent to the existence of a non-colored copy of F, therefore any bound for the rainbow version gives also a bound for the Turán problem.

When the forbidden graph F is a triangle, Keevash, Saks, Sudakov and Verstraëte [10] showed that for $c \geq 4$ colors the best possible number of edges in each color without having a rainbow triangle is equal to $\frac{1}{4}n^2$. This is achieved in the balanced complete bipartite graph (the same in each color) as in Mantel's theorem. Surprisingly, Magnant [11] provided a construction showing that for 3 colors the answer is different. Recently, Aharoni, DeVos, González, Montejano and Šámal [1], answering a question of Diwan and Mubayi [4], proved that for 3 colors the optimal bound is $\left(\frac{26-2\sqrt{7}}{81}\right)n^2 \approx 0.2557n^2$.

Later, Frankl [6] made a conjecture on the optimal bound for the product of the number of edges in each of the 3 colors without having a rainbow triangle. This was disproved by Frankl, Győri, Hel, Lv, Salia, Tomkins, Varga and Zhu [7]. Finally, Falgas-Ravry, Markström and Räty [5] completely determined the triples of the asymptotic number of edges in each color that force an existence of a rainbow triangle. Similar problems were also considered for other rainbow structures than triangles, for instance for paths [2], color-critical graphs [3], or spanning subgraphs [8, 9].

Here, we consider the problem in the setting of directed graphs and solve it for any number of colors when a transitive triangle or a directed triangle is the forbidden rainbow graph. It occurs that for both kinds of triangles and at least 4 colors, the maximum number of edges in each graph is attained when each of them is the same graph maximizing the number of edges without creating the forbidden triangle. While for 3 colors, the behavior is completely different. In case of a transitive triangle the construction is as in the result of Aharoni et al. [1] with all edges replaced by arcs in both ways. While for a directed triangle the construction is again significantly different.

In the next section we introduce the used notation, while in Section 3 we state our theorems and sketch their proofs. If a rainbow directed triangle is forbidden and there are at least 4 colors we show that the optimal asymptotic value of the maximal number of edges in each color follows from the bound on the total number of colored edges (Theorem 1), while for 3 colors we show that the optimal bound follows from the bound on the sum of the number of edges in any two colors (Theorem 3). Using similar case distinction and generalizations, we solve the rainbow Turán problem for a transitive triangle and at least 4 colors by Theorem 5 and finally using Theorem 7 we prove the optimal bound in the 3 colors case.

2 Notation

A *directed graph* G is a pair $(V(G), E(G))$, where $V(G)$ is a set of vertices and $E(G)$ is a set of pairs of distinct vertices. In particular, G does not contain loops or multiple edges. For shortening, we denote $e(G)$ as $|E(G)|$. For two vertices $u, v \in V(G)$, we write uv to denote the edge $(u, v) \in E(G)$. We refer jointly to edges uv and vu as edges *between* u and v. If $uv, vu \in E(G)$ we say that u and v are connected with a *double edge*. A *directed triangle* is a directed graph on the vertex set $\{u, v, w\}$ with edges uv, vw and wu, while a *transitive triangle* is a directed graph on the vertex set $\{u, v, w\}$ with edges uv, vw and uw. We denote by $[c]$ the set of positive integers $\{1, 2, \dots, c\}$.

Having an ordered set of directed graphs $G := (G_1, G_2, \dots, G_c)$ on a common vertex set $V(G)$, we consider the edge set of each graph G_i, for $i \in [c]$, as *edges* of G in color i. For a directed graph F, we say that G contains a rainbow copy of F if $V(F) \subset V(G)$ and there is a coloring φ of the edges of F into distinct colors such that $e \in E(G_{\varphi(e)})$ for every edge $e \in E(F)$.

3 Our results

We start with forbidding a directed triangle in the setting with at least 4 colors and prove the following theorem with its immediate corollary.

Theorem 1. *Let $c \geq 4$ and $G_1, G_2, \dots, G_c$ be directed graphs on a common set of n vertices. If $\sum_{i=1}^{c} e(G_i) > c\lfloor \frac{n^2}{2} \rfloor$, then there exists a rainbow directed triangle.*

Corollary 2. *Let $c \geq 4$ and $G_1, G_2, \dots, G_c$ be directed graphs on a common set of n vertices. If $\min_{i\in[c]} e(G_i) > \lfloor \frac{n^2}{2} \rfloor$, then there exists a rainbow directed triangle.*

The bounds provided above are the best possible. To observe this, consider each graph G_i for $i \in [c]$ as the same directed graph having $\lfloor \frac{n^2}{2} \rfloor$ edges constructed by replacing each edge of a complete bipartite graphs $K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil}$ by a double edge.

Sketch of the proof of Theorem 1. Suppose (for a contradiction) that Theorem 1 is false and choose a counterexample with the smallest number of vertices n. Firstly, we prove that there do not exist two vertices u and v connected with a double edge in two colors. This is obtained by considering how many edges can be between any vertex and vertices u and v, and using the assumed minimality. It implies that between any two vertices, there are at most $c + 1$ edges, where only one of them can be a double edge. Then, using a similar approach and careful counting, we show that the vertices connected with $c + 1$ edges can create only directed paths. Consequently, there are fewer edges than in our extremal construction, which leads to a contradiction. □

In case of 3 colors, Theorem 1 cannot hold, because if G_1 and G_2 are graphs with double edges between each pair of vertices and G_3 is an empty graph, then $\sum_{i=1}^{3} e(G_i) = 2n(n-1) > 3\lfloor \frac{n^2}{2} \rfloor$ and there is no rainbow directed triangle. In this case we prove the following theorem and its corollary.

Theorem 3. *Let G_1, G_2, G_3 be three directed graphs on a common set of n vertices. If $e(G_i) + e(G_j) > \frac{10}{9}n^2 + 2n$ for $1 \le i < j \le 3$, then there exists a rainbow directed triangle.*

Corollary 4. *Let G_1, G_2, G_3 be three directed graphs on a common set of n vertices. If $\min_{i\in[c]} e(G_i) > \frac{5}{9}n^2 + n$, then there exists a rainbow directed triangle.*

To see that the bound in Theorem 3 is asymptotically the best possible, consider three sets of $n/3$ vertices A_1, A_2 and A_3, and for each $i \in [3]$ let $E(G_i)$ contain all double edges inside A_j for $j \neq i$, as well as edges from A_1 to A_2, from A_1 to A_3, and from A_2 to A_3. This is depicted in Figure 1. This construction has approximately $\frac{5}{9}n^2$ edges in each color and does not contain a rainbow directed triangle.

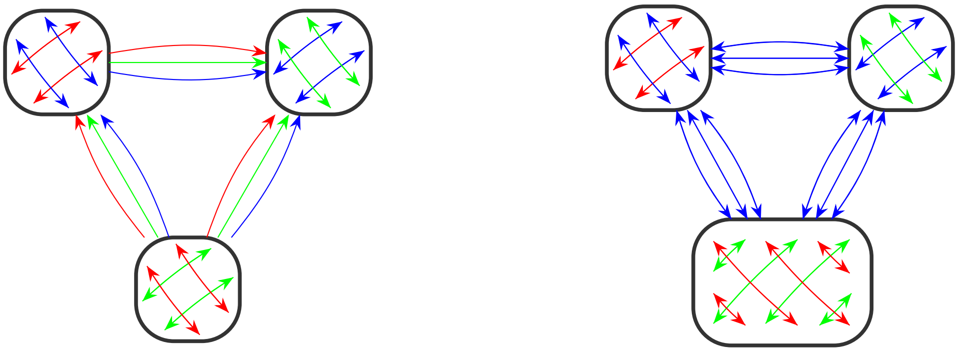


Figure 1: The extremal construction for 3 colors and forbidden rainbow directed triangle (on the left) and rainbow transitive triangle (on the right).

Sketch of the proof of Theorem 3. Consider a counterexample $G = (G_1, G_2, G_3)$ with the smallest number of vertices. Firstly, using the assumed minimality, we prove that there is no pair of vertices connected with double edges in all three colors. Using this, we split the vertices of G into disjoint sets. Consider a set X of vertices forming a maximal matching of double edges in two colors (there might be a single edge in the third color). Then, on the vertices $V(G)\backslash X$ consider a set Y forming a maximal matching consisting of pairs of vertices connected with 4 edges. Next, on the vertices $V(G) \setminus (X \cup Y)$ consider a set Z forming a maximal matching consisting of pairs of vertices connected by a double edge in one color and an edge in a different color. Finally, let T be a set of vertices forming a maximal matching on the vertices $V(G) \setminus (X \cup Y \cup Z)$ consisting of pairs of vertices connected with 3 edges. From the maximality of Z pairs in T are connected by a single edge in each color. From the maximality of T, all pairs of vertices in $D = V(G) \setminus (X \cup Y \cup Z \cup T)$ are connected by at most 2 edges.

Having such partitioning, we bound the total number of edges and the number of edges in each pair of colors in terms of the sizes of the defined sets. It is possible, as the number of edges between the vertices in respective sets are either limited by the definition of the sets or by the possibility of creating a rainbow directed triangle. Moreover, to include the fact that having many edges between set D and sets X, Y and Z is limiting the number of edges inside those sets, we use additional Turán-type bounds on an auxiliary graph.

Altogether, we obtain an optimization problem on 4 variables with linear or quadratic functions, which has a unique solution giving exactly the structure depicted in Figure 1. $\square$

We continue with forbidding a transitive triangle in the setting with at least four colors and prove the following theorem.

Theorem 5. *Let $c \geq 4$ and $G_1, G_2, \dots, G_c$ be directed graphs on a common set of n vertices. If $\sum_{i=1}^{c} e(G_i) > c\lfloor \frac{n^2}{2} \rfloor$, then there exists a rainbow transitive triangle.*

Similarly to the case of a forbidden directed triangle, this theorem easily implies the following corollary.

Corollary 6. *Let $c \geq 4$ and $G_1, G_2, \dots, G_c$ be directed graphs on a common set of n vertices. If $\min_{i \in [c]} e(G_i) > \lfloor \frac{n^2}{2} \rfloor$, then there exists a rainbow transitive triangle.*

The bounds provided in Theorem 5 and Corollary 6 are the best possible. To observe this, consider, like in the case of a forbidden rainbow directed triangle, each graph G_i for $i \in [c]$ as the same directed graph constructed by replacing each edge of a complete bipartite graphs $K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil}$ by a double edge. In this construction, there is no directed triangle and $e(G_i) = \lfloor \frac{n^2}{2} \rfloor$ for each $i \in [c]$.

Sketch of the proof of Theorem 5. Consider a hypothetical counterexample with the smallest number of vertices. Firstly, similarly as in the proof of Theorem 1, we prove that it does not contain a double edge in two colors. Having this limitation, we prove that it cannot contain any double edge at all. This immediately leads to a contradiction with the assumed number of edges. $\square$

Similarly as in the case of the forbidden rainbow directed triangle, in case of 3 colors Theorem 5 cannot hold. In this case we prove the following theorem and its corollary.

Theorem 7. *Let G_1, G_2, G_3 be three directed graphs on a common set of n vertices. If for every $1 \leq i < j \leq 3$ it holds $e(G_i) + e(G_j) > \left(\frac{104-8\sqrt{7}}{81}\right) n^2 + 2n$, then there exists a rainbow transitive triangle.*

Corollary 8. *Let G_1, G_2, G_3 be three directed graphs on a common set of n vertices. If $\min_{i \in [c]} e(G_i) > \left(\frac{52-4\sqrt{7}}{81}\right) n^2 + n$, then there exists a rainbow transitive triangle.*

The bound in Theorem 7 and Corollary 8 is asymptotically optimal in the sense that it is not possible to prove analogous statements with lower constants by the n^2 term. This is a consequence of the construction obtained from the optimal construction of Aharoni et al. [1] for the forbidden rainbow triangle by replacing all edges by double edges, as depicted in Figure 1.

Sketch of the proof of Theorem 7. We follow the idea of the proof from [1] for the forbidden rainbow triangle. It is not straightforward that those bounds translate to the directed setting as here it is possible to contain rainbow triangles as long as they are directed. Nevertheless, we prove bounds that give the same optimization problem (multiplied by a factor of 2). Thus, [1] implies the desired bound. $\square$

References

[1] R. Aharoni, M. DeVos, S. González, A. Montejano, R. Šámal, A rainbow version of Mantel's Theorem, *Advances in Combinatorics* (2020), 12043.

[2] S. Babiński, A. Grzesik, Graphs without a rainbow path of length 3, arXiv: 2211.02308 (2022).

[3] D. Chakarborti, J. Kim, H. Lee, H. Liu, J. Seo, On a rainbow extremal problem for color-critical graphs, arXiv: 2204.02575 (2022).

[4] A. Diwan, D. Mubayi, Turán's theorem with colors, preprint, `http://www.math.cmu.edu/~mubayi/papers/webturan.pdf`, 2007.

[5] V. Falgas-Ravry, K. Markström, E. Räty, Rainbow variations on a theme by Mantel: extremal problems for Gallai colouring templates, arXiv: 2212.07180 (2022).

[6] P. Frankl, Graphs without rainbow triangles, arXiv: 2203.07768 (2022).

[7] P. Frankl, E. Győri, Z. He, Z. Lv, N. Salia, C. Tompkins, K. Varga, X. Zhu, Some remarks on graphs without rainbow triangles, arXiv: 2204.07567 (2022).

[8] P. Gupta, F. Hamann, A. Müyesser, O. Parczyk, A. Sgueglia, A general approach to transversal versions of Dirac-type theorems, arXiv: 2209.09289 (2022).

[9] F. Joos, J. Kim, On a rainbow version of Dirac's theorem, *Bull. Lond. Math. Soc.* 52.3 (2020), 498–504.

[10] P. Keevash, M. Saks, B. Sudakov, J. Verstraëte, Multicolour Turán problems, *Advances in Applied Mathematics* 33(2) (2004), 238–262.

[11] C. Magnant, Density of Gallai Multigraphs, *Electron. J. Comb.* 22 (2015), P1.28.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

On ordered Ramsey numbers of matchings versus triangles

(Extended abstract)

Martin Balko* Marian Poljak†

Abstract

For graphs $G^<$ and $H^<$ with linearly ordered vertex sets, the *ordered Ramsey number* $r_<(G^<, H^<)$ is the smallest $N \in \mathbb{N}$ such that any red-blue coloring of the edges of the complete ordered graph $K_N^<$ on N vertices contains either a blue copy of $G^<$ or a red copy of $H^<$. Motivated by a problem of Conlon, Fox, Lee, and Sudakov (2017), we study the numbers $r_<(M^<, K_3^<)$ where $M^<$ is an n-vertex ordered matching.

We prove that almost all n-vertex ordered matchings $M^<$ with interval chromatic number 2 satisfy $r_<(M^<, K_3^<) \in \Omega((n/\log n)^{5/4})$ and $r_<(M^<, K_3^<) \in O(n^{7/4})$, improving a recent result by Rohatgi (2019). We also show that there are n-vertex ordered matchings $M^<$ with interval chromatic number at least 3 satisfying $r_<(M^<, K_3^<) \in \Omega((n/\log n)^{4/3})$, which asymptotically matches the best known lower bound on these ordered Ramsey numbers for general n-vertex ordered matchings.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-013`

1 Introduction

For graphs G and H, their *Ramsey number* $r(G, H)$ is the smallest positive integer N such that any coloring of the edges of K_N with colors red and blue contains either G as a

*Department of Applied Mathematics, Faculty of Mathematics and Physics, Charles University, Czech Republic. Email: `balko@kam.mff.cuni.cz`. Supported by the grant no. 23-04949X of the Czech Science Foundation (GAČR) and by the Center for Foundations of Modern Computer Science (Charles University project UNCE/SCI/004). This article is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 810115).

†Department of Applied Mathematics, Faculty of Mathematics and Physics, Charles University, Czech Republic. Email: `marian@kam.mff.cuni.cz`.

subgraph with all edges colored red or H as a subgraph with all edges colored blue. The existence of these numbers was proved by Ramsey [18] and later independently by Erdős and Szekeres [13]. The study of the growth rate of these numbers with respect to the number of vertices of G and H is a classical part of combinatorics and plays a central role in Ramsey theory.

Motivated by several applications from discrete geometry and extremal combinatorics, various researchers [3, 4, 6, 10, 16, 19] started investigating an ordered variant of Ramsey numbers. An *ordered graph* $G^<$ is a graph G together with a linear ordering of its vertex set. An ordered graph $G^<$ is an *ordered subgraph* of another ordered graph $H^<$ if G is a subgraph of H and the vertex ordering of G is a suborder of the vertex ordering of H. The *ordered Ramsey number* $r_<(G^<, H^<)$ of ordered graphs $G^<$ and $H^<$ is the smallest positive integer N such that any coloring of the edges of the complete ordered graph $K_N^<$ on N vertices with colors red and blue contains either $G^<$ as an ordered subgraph with all edges red or $H^<$ as an ordered subgraph with all edges blue.

It is known that the ordered Ramsey numbers always exist and that they can behave very differently from the unordered Ramsey numbers. For example, there are ordered *matchings* $M^<$ (that is, 1-regular ordered graphs) on n vertices for which $r_<(M^<, M^<)$ grows superpolynomially in n, in particular, we have $r_<(M^<, M^<) \in 2^{\Omega(\log^2 n/\log\log n)}$ [3, 10] while $r(G, G)$ is linear for all graphs G with bounded maximum degrees [7]. The superpolynomial bound obtained for ordered matchings is almost tight for sparse graphs as, for every fixed $d \in \mathbb{N}$, every d-degenerate ordered graph $G^<$ on n vertices satisfies $r_<(G^<, G^<) \in 2^{O(\log^2 n)}$ [10].

One of the most interesting cases for ordered Ramsey numbers is the study of the growth rate of $r_<(M^<, K_3^<)$ where $M^<$ is an ordered matching on n vertices as this is one of the first non-trivial cases where the exact asymptotics is not known. Conlon, Fox, Lee, and Sudakov [10] observed that the classical bound $r(K_n, K_3) \in O(n^2/\log n)$ immediately gives $r_<(M^<, K_3^<) \in O(n^2/\log n)$. On the other hand, they showed that there exists a positive constant c such that, for all even positive integers n, there is an ordered matching $M^<$ on n vertices with

$$r_<(M^<, K_3^<) \geq c\left(\frac{n}{\log n}\right)^{4/3}. \tag{1}$$

Conlon, Fox, Lee, and Sudakov expect that the upper bound $r_<(M^<, K_3^<) \leq O(n^2/\log n)$ is far from optimal and posed the following open problem [10], which is also mentioned in a survey on recent developments in graph Ramsey theory [11].

Problem 1 ([10])**.** *Does there exist an $\epsilon > 0$ such that for any ordered matching $M^<$ on n vertices $r_<(M^<, K_3^<) \in O(n^{2-\varepsilon})$?*

Problem 1 is one of the most important questions in the theory of ordered Ramsey numbers as in order to get a subquadratic upper bound on $r_<(M^<, K_3^<)$ one has to be able to employ the sparsity of $M^<$ since the bound $r(K_n, K_3) \in O(n^2/\log n)$ is asymptotically tight by a famous result of Kim [15]. Being able to use the sparsity of $M^<$ and thus distinguish $M^<$ from $K_n^<$ could help in numerous problems on ordered Ramsey numbers. However, this is difficult as some ordered matchings $M^<$ can be used to approximate the behavior of complete graphs, which is the reason why the numbers $r_<(M^<, M^<)$ can grow superpolynomially.

Some partial progress on Problem 1 was recently made by Rohatgi [19] who considered ordered matchings with bounded interval chromatic number. The *interval chromatic number* $\chi_<(G^<)$ of an ordered graph $G^<$ is the minimum number of intervals the vertex set of $G^<$ can be partitioned into so that there is no edge of $G^<$ with both vertices in the same interval. The interval chromatic number can be understood as an analogue of the chromatic number for ordered graphs as, for example, there is a variant of the Erdős–Stone–Simonovits theorem for ordered graphs [17] that is expressed in terms of the interval chromatic number.

Rohatgi [19] showed that the subquadratic bound on $r_<(M^<, K_3^<)$ holds for almost all ordered matchings with interval chromatic number 2 by proving the following result.

Theorem 1 ([19])**.** *There is a constant c such that for every even positive integer n, if an ordered matching $M^<$ on n vertices with $\chi_<(M^<) = 2$ is picked uniformly at random, then with high probability*

$$r_<(M^<, K_3^<) \leq cn^{24/13}.$$

Motivated by Problem 1, we study the numbers $r_<(M^<, K_3^<)$ for ordered matchings with bounded interval chromatic number. We strengthen some bounds by Rohatgi [19] and by Conlon, Fox, Lee, and Sudakov [10], obtaining a new partial progress on Problem 1.

From now on, we omit floor and ceiling signs whenever they are not essential. For $n \in \mathbb{N}$, we use $[n]$ to denote the set $\{1, \ldots, n\}$. All logarithms in this paper are base 2.

2 Our results

We try to tackle the first non-trivial instance of Problem 1 by considering the typical behavior of the numbers $r_<(M^<, K_3^<)$ for ordered matchings with interval chromatic number 2. As far as we know, there is no non-trivial lower bound in this case. In his paper, Rohatgi [19] mentions the problem of obtaining lower bounds that would come closer to the upper bound from Theorem 1. As our first result, we prove the first superlinear lower bound for this case.

Theorem 2. *There exists a positive constant c such that, for all even positive integers n, if an ordered matching $M^<$ on n vertices with $\chi_<(M^<) = 2$ is picked uniformly at random, then with high probability*

$$r_<(M^<, K_3^<) \geq c\left(\frac{n}{\log n}\right)^{5/4}.$$

We also show that this lower bound can be improved for ordered matchings $M^<$ with $\chi_<(M^<) > 2$.

Theorem 3. *For every integer $k \geq 3$, there exists a positive constant $c = c(k)$ such that, for all even positive integers n, there exists an ordered matching $M^<$ on n vertices with $\chi_<(M^<) = k$ satisfying*

$$r_<(M^<, K_3^<) \geq c\left(\frac{n}{\log n}\right)^{4/3}.$$

Note that the lower bound from Theorem 3 asymptotically matches the bound (1) by Conlon, Fox, Lee, and Sudakov [10]. Thus, the best known lower bound on $r_<(M^<, K_3^<)$ for general ordered matchings $M^<$ can be obtained also for ordered matchings with bounded interval chromatic number as long as this number is at least 3. The proofs of Theorems 2 and 3 are probabilistic and are based on ideas used by Conlon, Fox, Lee, and Sudakov [10].

Rohatgi [19] was also interested in determining how far from the truth the exponent 24/13 from Theorem 1 is. We narrow the gap there by providing the following upper bound that strengthens Theorem 1.

Theorem 4. *There is a constant c such that for every even positive integer n, if an ordered matching $M^<$ on n vertices with $\chi_<(M^<) = 2$ is picked uniformly at random, then with high probability*

$$r_<(M^<, K_3^<) \le cn^{7/4}.$$

Note that the difference between the exponent in the lower bound from Theorem 2 and the exponent in the upper bound from Theorem 4 is exactly 1/2. The sketch of the proof of Theorem 4 is in Section 4. All proofs can be found in the full version of this paper [5].

3 Open problems

Problem 1 still remains wide open, but there are many interesting intermediate questions that one could try to tackle. The following interesting conjecture was posed by Rohatgi [19].

Conjecture 1 ([19]). *For every integer $k \ge 2$, there is a constant $\varepsilon = \varepsilon(k) > 0$ such that*

$$r_<(M^<, K_3^<) \in O(n^{2-\varepsilon})$$

for almost every ordered matching $M^<$ on n vertices with $\chi_<(M^<) = k$.

It follows from Theorem 4 that $\varepsilon(2) \ge 1/4$. The conjecture is open for all cases with $k \ge 3$. Our results suggest that $\varepsilon(2) > \varepsilon(3)$ might hold.

Concerning the ordered matchings $M^<$ with interval chromatic number 2, even in this case the growth rate of $r_<(M^<, K_3^<)$ is not understood, so we pose the following weaker version of Problem 1.

Conjecture 2. *There exists an $\epsilon > 0$ such that for any ordered matching $M^<$ on n vertices with $\chi_<(M^<) = 2$ we have $r_<(M^<, K_3^<) \in O(n^{2-\varepsilon})$.*

In this paper, we considered the variant of this problem for random ordered matchings with interval chromatic number 2, but there is still a gap between our bounds. It would be very interesting to close it.

Problem 2. *What is the growth rate of $r_<(M^<, K_3^<)$ for uniform random ordered matchings $M^<$ on n vertices with $\chi_<(M^<) = 2$?*

It follows from our results that the answer to Problem 2 lies somewhere between $\Omega((n/\log n)^{5/4})$ and $O(n^{7/4})$. We do not know which of these bounds is closer to the truth.

4 Sketch of the proof of Theorem 4

To prove Theorem 4, we use a multi-thread scanning procedure whose variants were recently used by Cibulka and Kynčl [9], He and Kwan [14], and Rohatgi [19].

First, we associate an ordered matching $M^<$ on $[2n]$ with the permutation $\pi_{M^<}$ on $[n]$ that maps i to $j-n$ for every edge $\{i,j\}$ of $M^<$. Let χ be a red-blue coloring of the edges of $K_{2N}^<$ for some $N \in \mathbb{N}$. Let A be an $N \times N$ matrix where an entry on position $(i,j) \in [N] \times [N]$ is the color of the edge $\{i, N+j\}$ in χ.

We now describe a procedure that we use to find a red copy of $M^<$ in χ. It suffices to find an $n \times n$ submatrix of A with red entries on positions $(i, \phi_{M^<}(i))$ for $i = 1, \dots, n$. Let $T \in \mathbb{N}$. Consider the rows $t+1, \dots, t+n$ of A for every $t \in \{0, 1, \dots, T-1\}$. First, we scan through the row $\pi_{M^<}(1) + t$ of A from left to right until we find a red entry in some position $(\pi_{M^<}(1) + t, j_1)$. For every $i \in \{2, \dots, n\}$, after we have finished scanning through rows $\pi_{M^<}(1) + t, \dots, \pi_{M^<}(i-1) + t$, we scan through the row $\pi_{M^<}(i) + t$ of A, starting from column $j_{i-1} + 1$, until we find a red entry in some position $(\pi_{M^<}(i) + t, j_i)$.

We call this *multi-thread scanning* for $M^<$ and we call the set $Th(t)$ of entries of A that are revealed in step t a *thread*. Note that $Th(t)$ finds a red copy of $M^<$ if and only if some red copy of $M^<$ lies in the rows $t+1, \dots, t+n$ of A.

For a permutation π on $[n]$, we say that a subset $C \subseteq [n]$ with $|C| = k$ is a *shift* of another subset $D \subseteq [n]$ in π if there is a positive integer Δ such that $\pi(c_i) = \pi(d_i) + \Delta$ for each $i \in [k]$ where $c_1 < \cdots < c_k$ and $d_1 < \cdots < d_k$ are the elements of C and D, respectively. Let $L(\pi)$ be the largest positive integer k for which there are sets $C, D \subseteq [n]$, each of size k, such that C is a shift of D.

The multi-thread scanning procedure yields the following result, which gives asymptotically stronger bounds than a similar result obtained by Rohatgi [19].

Theorem 5. *For $n \in \mathbb{N}$, let $M^<$ be an ordered matching on $2n$ vertices with $\chi_<(M^<) = 2$ and $L(\pi_{M^<}) \leq \ell$. If $N \geq 4n(\sqrt{n\ell} + 1)$, then every red-blue coloring of the edges of $K_{2N}^<$ on $[2N]$ satisfies at least one of the following three claims:*

(a) χ contains a blue copy of $K_3^<$,

(b) χ contains a red copy of $K_{2n}^<$, or

(c) χ contains a red copy of $M^<$ between $[N]$ and $\{N+1, \dots, 2N\}$.

For every $\varepsilon > 0$, Theorem 5 immediately implies that $r_<(M^<, K_3^<) \in O(n^{2-\varepsilon})$ for every ordered matching with $\chi_<(M^<) = 2$ and $L(\pi_{M^<}) \leq n^{1-2\varepsilon}$. It suffices to show that this is the case for uniform random ordered matchings with interval chromatic number 2. We do so by using the following result of He and Kwan [14].

Lemma 6 ([14])**.** *A uniform random permutation π on $[n]$ satisfies $L(\pi) \leq 3\sqrt{n}$ with high probability.*

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi. A note on Ramsey numbers. *J. Combin. Theory Ser. A*, 29(3):354–360, 1980.

[2] N. Alon and J. H. Spencer. *The probabilistic method.* Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, fourth edition, 2016.

[3] M. Balko, J. Cibulka, K. Král, and J. Kynčl. Ramsey numbers of ordered graphs. *Electron. J. Combin.*, 27(1), 2020.

[4] M. Balko, V. Jelínek, and P. Valtr. On ordered Ramsey numbers of bounded-degree graphs. *J. Combin. Theory Ser. B*, 134:179–202, 2019.

[5] M. Balko and M. Poljak. On ordered Ramsey numbers of matchings versus triangles. https://arxiv.org/abs/2305.17933, 2023.

[6] S. A. Choudum and B. Ponnusamy. Ordered Ramsey numbers. *Discrete Math.*, 247(1-3):79–92, 2002.

[7] V. Chvátal, V. Rödl, E. Szemerédi, and W. T. Trotter, Jr. The Ramsey number of a graph with bounded maximum degree. *J. Combin. Theory Ser. B*, 34(3):239–243, 1983.

[8] V. Chvatál, V. Rödl, E. Szemerédi, and W. T. Trotter, Jr. The Ramsey number of a graph with bounded maximum degree. *J. Combin. Theory Ser. B*, 34(3):239–243, 1983.

[9] J. Cibulka and J. Kynčl. Better upper bounds on the Füredi-Hajnal limits of permutations. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2280–2293. SIAM, Philadelphia, PA, 2017.

[10] D. Conlon, J. Fox, C. Lee, and B. Sudakov. Ordered Ramsey numbers. *J. Combin. Theory Ser. B*, 122:353–383, 2017.

[11] D. Conlon, J. Fox, and B. Sudakov. Recent developments in graph Ramsey theory. In *Surveys in combinatorics 2015*, volume 424 of *London Math. Soc. Lecture Note Ser.*, pages 49–118. Cambridge Univ. Press, Cambridge, 2015.

[12] P. Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.*, 53:292–294, 1947.

[13] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.

[14] X. He and M. Kwan. Universality of random permutations. *Bull. Lond. Math. Soc.*, 52(3):515–529, 2020.

[15] J. H. Kim. The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$. *Random Structures Algorithms*, 7(3):173–207, 1995.

[16] K. G. Milans and D. B. Stolee, D.and West. Ordered Ramsey theory and track representations of graphs. *J. Comb.*, 6(4):445–456, 2015.

[17] J. Pach and G. Tardos. Forbidden paths and cycles in ordered graphs and matrices. *Israel J. Math.*, 155:359–380, 2006.

[18] F. P. Ramsey. On a Problem of Formal Logic. *Proc. London Math. Soc. (2)*, 30(4):264–286, 1929.

[19] D. Rohatgi. Off-diagonal ordered Ramsey numbers of matchings. *Electron. J. Combin.*, 26(2):Paper No. 2.21, 18, 2019.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

On the sizes of t-intersecting k-chain-free families

(Extended abstract)

József Balogh *   William B. Linz †   Balázs Patkós ‡

Abstract

A set system $\mathcal{F}$ is *t-intersecting*, if the size of the intersection of every pair of its elements has size at least t. A set system $\mathcal{F}$ is *k-Sperner*, if it does not contain a chain of length $k+1$.

Our main result is the following: Suppose that k and t are fixed positive integers, where $n+t$ is even and n is large enough. If $\mathcal{F} \subseteq 2^{[n]}$ is a t-intersecting k-Sperner family, then $\mathcal{F}$ has size at most the size of the sum of k layers, of sizes $(n+t)/2, \ldots, (n+t)/2+k-1$. This bound is best possible. The case when $n+t$ is odd remains open.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-014`

1 Introduction

1.1 Definitions and Notation

For a positive integer n, we write $[n] := \{1, 2, \ldots, n\}$ and $2^{[n]}$ for the power set of $[n]$. For a set S, we denote by $\binom{S}{i}$ the family of all i element subsets of S.

*University of Illinois at Urbana-Champaign. E-mail: `jobal@illinois.edu` Partially supported by NSF grants DMS-1764123 and RTG DMS-1937241, the Arnold O. Beckman Research Award (UIUC Campus Research Board RB 18132), the Langan Scholar Fund (UIUC), and the Simons Fellowship.

†University of South Carolina E-mail: `WLINZ@mailbox.sc.edu` Partially supported by RTG DMS-1937241.

‡Alfréd Rényi Institute of Mathematics E-mail: `patkos@renyi.hu` Partially supported by NKFIH grants SNN 129364 and FK 132060.

For a family of sets $\mathcal{F} \subseteq 2^{[n]}$, we define $\mathcal{F}_i := \{F \in \mathcal{F} : |F| = i\}$ and $f_i := |\mathcal{F}_i|$. We use Δ_i and ∇_i to denote the i-shadow and i-shade of $\mathcal{F}$, respectively, so that $\Delta_i \mathcal{F} := \{A : |A| = i,\ A \subset F \text{ for some } F \in \mathcal{F}\}$ and $\nabla_i \mathcal{F} := \{A : |A| = i,\ A \supset F \text{ for some } F \in \mathcal{F}\}$. If the subscript i is unspecified, then assuming $\mathcal{F}$ is r-uniform, $\Delta \mathcal{F} = \Delta_{r-1}\mathcal{F}$ and similarly $\nabla \mathcal{F} = \nabla_{r+1}\mathcal{F}$.

Definition 1.1. [k-Sperner family]
A $(k+1)$-*chain* is a collection of $k+1$ sets $A_0, A_1, \ldots, A_k$ such that $A_0 \subset A_1 \subset \ldots \subset A_k$. A family of sets $\mathcal{F} \subseteq 2^{[n]}$ is a *k-Sperner family* if there is no $(k+1)$-chain in $\mathcal{F}$. If $k = 1$, then $\mathcal{F}$ is simply called a *Sperner family* or an *antichain.*

Definition 1.2. [t-intersecting family]
A family of sets $\mathcal{F} \subseteq 2^{[n]}$ is *t-intersecting* if for every pair of sets $A, B \in \mathcal{F}$, we have $|A \cap B| \geq t$. If $t = 1$, then we write that $\mathcal{F}$ is *intersecting.*

1.2 History

The maximum size of an antichain in $2^{[n]}$ was determined by Sperner [9].

Theorem 1.3 (Sperner)**.** *Let $\mathcal{F} \subseteq 2^{[n]}$ be an antichain. Then,*

$$|\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

Furthermore, equality holds only if $\mathcal{F}$ is one of the largest layers in the Boolean lattice $2^{[n]}$.

Sperner's theorem was extended to k-Sperner families by Erdős [2].

Theorem 1.4 (Erdős)**.** *The maximum-size k-Sperner family $\mathcal{F} \subseteq 2^{[n]}$ is the union of the largest k layers in the Boolean lattice $2^{[n]}$.*

A different extension of Sperner's theorem was given by Milner [8]. Milner additionally required the family $\mathcal{F}$ to be t-intersecting.

Theorem 1.5 (Milner)**.** *If $\mathcal{F} \subseteq 2^{[n]}$ is a t-intersecting antichain, then*

$$|\mathcal{F}| \leq \binom{n}{\lfloor \frac{n+t+1}{2} \rfloor}.$$

In a different direction, Frankl [3] determined the maximum size of an intersecting k-Sperner family. Different proofs were given by Gerbner [5] and by Gerbner, Methuku and Tompkins [6].

Theorem 1.6 (Frankl)**.** *Let $\mathcal{F} \subseteq 2^{[n]}$ be an intersecting, k-Sperner family. Then,*

$$|\mathcal{F}| \leq \begin{cases} \sum_{i=\frac{n+1}{2}}^{\frac{n+1}{2}+k-1} \binom{n}{i}, & \textit{if } n \textit{ is odd,} \\ \binom{n-1}{\frac{n}{2}-1} + \sum_{i=\frac{n}{2}+1}^{\frac{n}{2}+k-1} \binom{n}{i} + \binom{n-1}{\frac{n}{2}+k}, & \textit{if } n \textit{ is even.} \end{cases}$$

Furthermore, if n is odd, equality holds only if

$$\mathcal{F} = \binom{[n]}{\lfloor \frac{n}{2} \rfloor + 1} \cup \binom{[n]}{\lfloor \frac{n}{2} \rfloor + 2} \cup \ldots \cup \binom{[n]}{\lfloor \frac{n}{2} \rfloor + k},$$

while if n is even and $k > 1$, equality holds only if for some $x \in [n]$,

$$\mathcal{F} = \left\{ F \in \binom{[n]}{\frac{n}{2}} : x \in F \right\} \cup \binom{[n]}{\frac{n}{2} + 1} \cup \ldots \cup \binom{[n]}{\frac{n}{2} + k - 1} \cup \left\{ F \in \binom{[n]}{\frac{n}{2} + k} : x \notin F \right\}.$$

A common generalization of the theorems of Milner and Frankl would be to determine the maximum size of a t-intersecting, k-Sperner family.

Frankl [4] proposed conjectures on the maximum size of a t-intersecting k-Sperner family $\mathcal{F} \subset 2^{[n]}$ and made some progress towards proving these conjectures. The conjectured extremal family depends on the parity of $n + t$.

In the case when $n + t$ is even, the conjectured maximum size of a t-intersecting, k-Sperner family is very easy to describe.

Conjecture 1.7 (Frankl). *If $n+t$ is even, $n > t$, and $\mathcal{F} \subseteq 2^{[n]}$ is a t-intersecting, k-Sperner family, then*

$$|\mathcal{F}| \leq \sum_{i=0}^{k-1} \binom{n}{\frac{n+t}{2} + i}.$$

Conjecture 1.7 is clearly tight if true, as evidenced by the family $\bigcup_{i=0}^{k-1} \binom{[n]}{\frac{n+t}{2}+i}$.

The conjectured extremal families do not have such a simple structure when $n + t$ is odd. We construct two plausible candidates for the maximum size t-intersecting, k-Sperner family:

$$\mathcal{A}(t,k) = \left\{ F \in \binom{[n]}{\frac{n+t-1}{2}} : n \notin F \right\} \cup \left\{ A : \frac{n+t-1}{2} + 1 \leq |A| \leq \frac{n+t-1}{2} + (k-1) \right\}.$$

$$\begin{aligned}\mathcal{B}(t,k) = &\left\{ F \in \binom{[n]}{\frac{n+t-1}{2}} : [1,t] \in F \right\} \cup \left\{ A : \frac{n+t-1}{2} + 1 \leq |A| \leq \frac{n+t-1}{2} + (k-1) \right\} \\ &\cup \left(\left\{ B : |B| = \frac{n+t-1}{2} + k \right\} \setminus \left\{ B : |B| = \frac{n+t-1}{2} + k, [1,t] \in B \right\} \right).\end{aligned}$$

It is not hard to show that $|\mathcal{B}(t,k)| \gg |\mathcal{A}(t,k)|$ for n sufficiently large (in terms of k and t). However, it may be checked by computer that $\mathcal{A}(t,k)$ is optimal for small values of n and specific choices of t and k, for example $t = 2$ and $k = 2$. We conjecture that $\mathcal{B}(t,k)$ is the largest such family when n is sufficiently large.

Conjecture 1.8. *There exists a positive integer $n_0 = n_0(k,t)$ such that if $n+t$ is odd, $n > n_0$, and $\mathcal{F} \subseteq 2^{[n]}$ is a t-intersecting, k-Sperner family, then*

$$|\mathcal{F}| \leq |\mathcal{B}(t,k)| = \binom{n-t}{\frac{n-t-1}{2}} + \sum_{i=1}^{k} \binom{n}{\frac{n+t-1}{2}+i} - \binom{n-t}{\frac{n-t-1}{2}+k}.$$

Frankl [4] more modestly conjectures the following (Frankl's conjecture is formulated for s-union families rather than t-intersecting families, but our formulation is equivalent to Frankl's after taking complements).

Conjecture 1.9 (Frankl)**.** *Let $g(n,t,k) := \max\{|\mathcal{G}| - |\Delta_{\frac{n-t+1}{2}-k}(\mathcal{G})| : \mathcal{G} \subset \binom{[n]}{\frac{n-t+1}{2}}$ is intersecting$\}$. Then, if $n+t$ is odd and $\mathcal{F}$ is a t-intersecting, k-Sperner family, then*

$$|\mathcal{F}| \leq g(n,t,k) + \sum_{i=1}^{k} \binom{n}{\frac{n+t-1}{2}+i}.$$

Note that Conjecture 1.8 can be interpreted as a strengthening of Conjecture 1.9, in that additionally there is a conjecture for the value of the function $g(n,t,k)$ for sufficiently large n. The connection may be made more apparent by noting that, after taking complements, we may equivalently define $g(n,t,k) := \max\{|\mathcal{G}| - |\nabla_{\frac{n+t-1}{2}+k}(\mathcal{G})| : \mathcal{G} \subset \binom{[n]}{\frac{n+t-1}{2}}$ is t-intersecting$\}$.

1.3 New Results

Let us mention that Frankl proved Conjecture 1.7 when $t \geq n - O(\sqrt{n})$. We settle Conjecture 1.7 if t is fixed and n is sufficiently large.

Theorem 1.10. *Let t and k be positive integers, and suppose that $n+t$ is even with $t \leq n$, and n is large enough. If $\mathcal{F} \subseteq 2^{[n]}$ is a t-intersecting k-Sperner family, then*

$$|\mathcal{F}| \leq \binom{[n]}{\frac{n+t}{2}} + \ldots + \binom{[n]}{\frac{n+t}{2}+k-1}.$$

2 Sketch of the Proof of Theorem 1.10

The proof of Theorem 1.10 consists of three parts. The first part is a so-called "push-to-the-middle" argument. By proving this part, one obtains that there exists a maximum size t-intersecting k-Sperner family that contains sets only of cardinality between $\frac{n+t}{2} - (k-1)$ and $\frac{n+t}{2} + 2(k-1)$. This is achieved in two steps, none of which uses the assumption $n+t$ even, so this part of the proof can be applied in arguments for the $n+t$ odd case. The two steps are formulated in the following two lemmas, the first of which applies Katona's shadow t-intersection theorem [7].

Lemma 2.1. *Let $\mathcal{F} \subseteq 2^{[n]}$ be a t-intersecting and k-Sperner family, where $n+t$ is even. Then there exists a t-intersecting k-Sperner family $\mathcal{G} \subseteq 2^{[n]}$ with $|\mathcal{G}| \geq |\mathcal{F}|$ and $\min\{|G| : G \in \mathcal{G}\} \geq \frac{n+t}{2} - (k-1)$.*

Lemma 2.2. *If $\mathcal{F} \subseteq 2^{[n]}$ is a t-intersecting k-Sperner family with $\min\{|F| : F \in \mathcal{F}\} = \frac{n+t}{2} - c$, then there exists a t-intersecting k-Sperner family $\mathcal{F}' \subseteq 2^{[n]}$ with $|\mathcal{F}| \leq |\mathcal{F}'|$, and*

$$\min\{|F| : F \in \mathcal{F}\} = \min\{|F'| : F' \in \mathcal{F}'\} \quad \text{and} \quad \max\{|F'| : F' \in \mathcal{F}'\} \leq \frac{n+t}{2} + c + k - 1.$$

The other two parts of the proof applies Katona's cycle method. We need some definitions. Let σ be a cyclic permutation of $[n]$ and $\mathcal{F}_\sigma$ be the subfamily of those sets in $\mathcal{F}$ that form an interval in σ. Note that there are $(n-1)!$ choices for σ. For a set G, let $w(G) = \binom{n}{|G|}$ and $w(\mathcal{G}) = \sum_{G \in \mathcal{G}} w(G)$. We define m as $m := \frac{n+t}{2} - \min\{|F| : F \in \mathcal{F}''\}$. By the above discussions, we have $0 \leq m \leq k-1$. If $m = 0$ then $\mathcal{F}$ has the required structure, hence we assume $m > 0$. The second and most delicate part of the proof is the following lemma that determines the maximum weight of a t-intersecting k-Sperner family on the cycle.

Lemma 2.3. *Suppose $n+t$ is even with $t \leq n$ and n is large enough. For every cyclic permutation σ and t-intersecting k-Sperner family $\mathcal{F} \subseteq \bigcup_{i=\frac{n+t}{2}-m}^{\frac{n+t}{2}+k-1+m} \binom{[n]}{i}$, we have $w(\mathcal{F}_\sigma) \leq n \sum_{i=0}^{k-1} \binom{n}{\frac{n+t}{2}+i}$.*

Before giving more insight on the proof of Lemma 2.3, let us show how Lemma 2.3 implies Theorem 1.10 that is the part of the proof.

Proof of Theorem 1.10 using Lemma 2.3. As mentioned in the last paragraph of the previous subsection, by Theorem 2.1 and Lemma 2.2, we can assume that $\mathcal{F} \subseteq \bigcup_{i=\frac{n+t}{2}-m}^{\frac{n+t}{2}+k-1+m} \binom{[n]}{i}$ holds. Then using Lemma 2.3 we have:

$$\sum_{\sigma} \sum_{F \in \mathcal{F}_\sigma} w(F) \leq (n-1)! \cdot n \sum_{i=0}^{k-1} \binom{n}{\frac{n+t}{2}+i} = n! \cdot \sum_{i=0}^{k-1} \binom{n}{\frac{n+t}{2}+i}.$$

From the other side,

$$\sum_{\sigma} \sum_{F \in \mathcal{F}_\sigma} w(F) = \sum_{F \in \mathcal{F}} |F|!(n-|F|)! \binom{n}{|F|} = n!|\mathcal{F}|,$$

which implies the required upper bound on $|\mathcal{F}|$. □

Let us return to the proof of Lemma 2.3. The extremal family of the lemma consists of all kn intervals with size between $\frac{n+t}{2}$ and $\frac{n+t}{2} + k - 1$. Also, any k-Sperner family on the cycle may contain at most kn intervals, but those that have cardinality smaller than $\frac{n+t}{2}$ have larger weight than the intervals of the extremal family. To compensate for these, we need to show that if such intervals exist in a t-intersecting K-Sperner family $\mathcal{G}$, then $\mathcal{G}$ must contain intervals that are too large, larger than $\frac{n+t}{2}+k-1$ and therefore have smaller weight than all intervals of the extremal family. (And we have to make sure that the loss of weight is more than the gain.) This is done based on the following simple observation. For a cyclic permutation σ and an interval G define $\overline{G}^t$ as the complement of G together with the (counterclockwise) leftmost $\lfloor \frac{t}{2} \rfloor$ and rightmost $\lceil \frac{t}{2} \rceil$ elements of G with respect to σ. For a family $\mathcal{G}$ of intervals, let $\overline{\mathcal{G}}^t = \{\overline{G}^t : G \in \mathcal{G}\}$. Observe that $|G \cap \overline{G}^t| = t$ and the two endpoints of $\overline{G}^t$ belong to $G \cap \overline{G}^t$. Therefore, if $\mathcal{G}$ is t-intersecting, then for any $G \in \mathcal{G}$ no proper subinterval H of $\overline{G}^t$ belongs to $\mathcal{G}$. $|G| = \frac{n+t}{2} - c$ implies $|\overline{G}^t| = n - (\frac{n+t}{2} - c) + t = \frac{n+t}{2} + c$, so if such a G belongs to $\mathcal{G}$, then many intervals of size close $\frac{n+t}{2}$ cannot belong to $\mathcal{G}$.

The details how this is used to derive Lemma 2.3 can be found in [1].

References

[1] József Balogh, William B Linz, and Balázs Patkós. On the sizes of t-intersecting k-chain-free families. *arXiv preprint arXiv:2209.01656*, 2022.

[2] Paul Erdős. On a lemma of Littlewood and Offord. *Bull. Amer. Math. Soc.*, 51:898–902, 1945.

[3] Peter Frankl. Canonical antichains on the circle and applications. *SIAM Journal on Discrete Mathematics*, 3(3):355–363, 1990.

[4] Peter Frankl. Analogues of Milner's theorem for families without long chains and of vector spaces. *European Journal of Combinatorics*, 93:103279, 2021.

[5] Dániel Gerbner. Profile polytopes of some classes of families. *Combinatorica*, 33:199–216, 2013.

[6] Dániel Gerbner, Abhishek Methuku, and Casey Tompkins. Intersecting P-free families. *Journal of Combinatorial Theory, Series A*, 151:61–83, 2017.

[7] Gyula Katona. Intersection theorems for systems of finite sets. *Acta Mathematica Academiae Scientiarum Hungaricae*, 15(3-4):329–337, 1964.

[8] Eric C Milner. A combinatorial theorem on systems of sets. *Journal of the London Mathematical Society*, 1(1):204–206, 1968.

[9] Emanuel Sperner. Ein satz über untermengen einer endlichen menge. *Mathematische Zeitschrift*, 27(1):544–548, 1928.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Isoperimetric stability in lattices

(Extended abstract)

Ben Barber[*] Joshua Erde[†] Peter Keevash[‡] Alexander Roberts[§]

Abstract

We obtain isoperimetric stability theorems for general Cayley digraphs on $\mathbb{Z}^d$. For any fixed B that generates $\mathbb{Z}^d$ over $\mathbb{Z}$, we characterise the approximate structure of large sets A that are approximately isoperimetric in the Cayley digraph of B: we show that A must be close to a set of the form $kZ \cap \mathbb{Z}^d$, where for the vertex boundary Z is the conical hull of B, and for the edge boundary Z is the zonotope generated by B.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-015`

1 Introduction

An important theme at the interface of Geometry, Analysis and Combinatorics is understanding the structure of approximate minimisers to isoperimetric problems. These problems take the form of minimising surface area of sets with a fixed volume, for various meanings of 'area' and 'volume'. The usual meanings give the Euclidean Isoperimetric Problem considered since the ancient Greek mathematicians, where balls are the measurable subsets of $\mathbb{R}^d$ with a given volume which minimize the surface area. There is a

[*]University of Manchester and Heilbronn Institute for Mathematical Research. E-mail `ben.barber@manchester.ac.uk`

[†]Graz University of Technology, Institute of Discrete Mathematics, Steyrergasse 30, 8010 Graz, Austria. E-mail: `erde@tugraz.at`. This research was funded in part by the Austrian Science Fund (FWF) P36131.

[‡]Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, United Kingdom. E-mail: `keevash@maths.ox.ac.uk`. Research supported in part by ERC Consolidator Grant 647678.

[§]Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, United Kingdom. E-mail: `robertsa@maths.ox.ac.uk`.

large literature on its stability, i.e. understanding the structure of approximate minimisers, culminating in the sharp quantitative isoperimetric inequality of Fusco, Maggi and Pratelli [8].

In the discrete setting, isoperimetric problems form a broad area that is widely studied within Combinatorics (see the surveys [2, 14]) and as part of the Concentration of Measure phenomenon (see [15, 26]). Certain particular settings have been intensively studied due to their applications; for example, there has been considerable recent progress (see [12, 11, 13, 23]) on isoperimetric stability in the discrete cube $\{0,1\}^n$, which is intimately connected to the Analysis of Boolean Functions (see [20]) and the Kahn–Kalai Conjecture (see [10]) on thresholds for monotone properties, which has recently been solved [7, 21]. This paper concerns the setting of integer lattices, which is widely studied in Additive Combinatorics, where the Polynomial Freiman–Ruzsa Conjecture (see [9]) predicts the structure of sets with small doubling.

For an isoperimetric problem on a digraph (directed graph) G, we measure the 'volume' of $A \subseteq V(G)$ by its size $|A|$, and its 'surface area' either by the *edge boundary* $\partial_{e,G}(A)$, which is the number of edges $\overrightarrow{xy} \in E(G)$ with $x \in A$ and $y \in V(G) \setminus A$, or by the *vertex boundary* $\partial_{v,G}(A)$, which is the number of vertices $y \in V(G) \setminus A$ such that $\overrightarrow{xy} \in E(G)$ for some $x \in A$. Here we consider Cayley digraphs: given a generating set B of $\mathbb{Z}^d$, we write G_B for the digraph on $\mathbb{Z}^d$ with edges $E(G_B) = \{\overrightarrow{uv} : v - u \in B\}$.

It is an open problem to determine the minimum possible value of $\partial_{v,G_B}(A)$ or $\partial_{e,G_B}(A)$ for $A \subseteq \mathbb{Z}^d$ of given size, let alone any structural properties of (approximate) minimisers; exact results are only known for a few instances of B (see [3, 4, 27, 24]). It is therefore natural to seek asymptotics. For ease of reference we collect here our notation for the various sets involved in stating the following results.

$C(B) \subseteq \mathbb{R}^d$ The *conical hull* $C(B)$ of B is the convex hull of $B \cup \{0\}$.

$B_n \subseteq \mathbb{Z}^d$ The sets $kC(B) \cap \mathbb{Z}^d$ are increasing as a function of $k > 0$. Write B_n for the smallest of these sets with at least n elements.

$[B] \subseteq \mathbb{Z}^d$ Write $[B] = \left\{ \sum_{b \in B'} b : B' \subseteq B \right\}$ for the set of all sums of subsets of B. Thus $|[B]| \leq 2^{|B|}$, where the bound is strict if multiple subsets of B have equal sums.

$Z(B) \subseteq \mathbb{R}^d$ The *zonotope* generated by B is $\left\{ \sum_{b \in B} x_b b : x \in [0,1]^B \right\}$. Equivalently, $Z(B)$ is the convex (or conical, as $[B]$ contains 0) hull of $[B]$.

For $A \subseteq \mathbb{Z}^d$ of size $n \to \infty$, Ruzsa [25] showed that the minimum value of the vertex boundary $\partial_{v,G_B}(A)$ is asymptotic to that achieved by a set of the form $kC(B) \cap \mathbb{Z}^d$. A corresponding result for the edge boundary was obtained in [1]: the minimum value of $\partial_{e,G_B}(A)$ is asymptotic to that achieved by a set of the form $kZ(B) \cap \mathbb{Z}^d$.

We will prove stability versions of both these results, describing the approximate structure of asymptotic minimisers for both the vertex and edge isoperimetric problems in G_B. We use μ to denote Lebesgue measure.

Theorem 1.1. *Let $d \geq 2$. For every generating set B of $\mathbb{Z}^d$, there is a $K \in \mathbb{N}$ such that whenever*

- $A \subseteq \mathbb{Z}^d$ *with* $|A| = n \geq K$,
- $Kn^{-1/2d} < \varepsilon < K^{-1}$ *and*
- $\partial_{v,G_B}(A) \leq d\mu(C(B))^{1/d} n^{1-1/d}(1+\varepsilon)$,

there is a $v \in \mathbb{Z}^d$ with $|A \triangle (v + B_n)| < Kn\sqrt{\varepsilon}$.

Theorem 1.2. *Let $d \geq 2$. For every generating set B of $\mathbb{Z}^d$ and $\delta > 0$, there are $K \in \mathbb{N}$ and $\epsilon > 0$ such that whenever*

- $A \subseteq \mathbb{Z}^d$ *with* $|A| = n \geq K$ *and*
- $\partial_{e,G_B}(A) \leq d\mu(Z(B))^{1/d} n^{1-1/d}(1+\varepsilon)$,

there is a $v \in \mathbb{Z}^d$ with $|A \triangle (v + [B]_n)| < \delta n$.

The square root dependence in Theorem 1.1 is tight, as may be seen from an example where B consists of the corners of a cube and A is an appropriate cuboid.

Besides drawing on the methods of [25] (particularly Plünnecke's inequality for sumsets) and [1] (a probabilistic reduction to [25]), the most significant new contribution of our paper is a technique for transforming discrete problems to a continuous setting where one can apply results from Geometric Measure Theory. We will employ the sharp estimate on asymmetric index in terms of anisotropic perimeter with respect to any convex set K due to Figalli, Maggi and Pratelli [6] (building on the case when K is a ball, established in [8]).

2 Proof strategy

This section contains an overview of the proof of our tight quantitative stability result for the vertex isoperimetric inequality in general Cayley digraphs. Using ideas from [1] one can deduce from this also a stability result for the edge isoperimetric inequality.

We start with a summary of Ruzsa's approach in [25], during which we record some key lemmas on sumsets and fundamental domains of lattices that we will also use in our proof.

2.1 Ruzsa's approach

The sumset of $A, B \in \mathbb{Z}^d$ is defined by $A + B := \{a + b : a \in A, b \in B\}$. The vertex isoperimetric problem in the Cayley digraph G_B is equivalent to finding the minimum of $|A+B|$ over all sets A of given size. The following result of Ruzsa [25, Theorem 2] implies an asymptotic for this minimum.

Theorem 2.1. *Let B be a generating set of $\mathbb{Z}^d$ with $d \geq 2$. Then for any $A \subseteq \mathbb{Z}^d$ with $|A| = n$ large we have $|A + B| \geq d\mu(C(B))^{1/d} n^{1-1/d}\big(1 - O(n^{-1/2d})\big)$.*

Ruzsa aims to deduce this inequality from the Brunn–Minkowski inequality (in the form due to Lusternik [16]) $\mu(U+V)^{1/d} \geq \mu(U)^{1/d} + \mu(V)^{1/d}$, which is tight when U and V are closed, convex and homothetic (that is, agree up to scaling and translation).

Passing from a discrete inequality to a continuous one can be achieved by adding a *fundamental set* Q to each side; that is, a measurable Q such that any $x \in \mathbb{R}^d$ has a unique representation as $x = z + q$ with $z \in \mathbb{Z}^d$ and $q \in Q$. This ensures that $\mu(X+Q) = |X|$ for any $X \subseteq \mathbb{Z}^d$. One example of a fundamental set is the half-open unit cube $[0,1)^d$, but we will prefer a fundamental set tailored to B rather than to the standard coordinate axes.

Typically $B+Q$ will be far from convex, so a naive application of Brunn–Minkowski gives poor results. Ruzsa smooths out B by using a version of Plünnecke's inequality [22] to replace B by its sumset. We write $\Sigma_k(A)$ for the k-fold sumset of A rather than the commonly used kA, which in this paper denotes the dilate of A by factor k.

Theorem 2.2 (see [25, Statement 6.2])**.** *Let $k \in \mathbb{N}$ and $A, B \subseteq \mathbb{Z}^d$ with $|A| = n$ and $|A+B| = \alpha n$. Then there is a non-empty subset $A' \subseteq A$ with $|A' + \Sigma_k(B)| \leq \alpha^k |A'|$.*

To return to a bound on to discrete sets Ruzsa uses the following lemma. By *nice* we mean that a set is a finite union of bounded convex polytopes.

Lemma 2.3 ([25, Lemma 11.2])**.** *Let B be a generating set of $\mathbb{Z}^d$ with $d \geq 2$ and $0 \in B$. Then there are $p \in \mathbb{N}$, $z \in \mathbb{Z}^d$ and a nice fundamental set $Q \subseteq Z(B)$ such that $kC(B) + Q + z \subseteq \Sigma_{k+p}(B) + Q$ for any $k \in \mathbb{N}$.*

The fact that Q may be chosen to be nice and such that $Q \subseteq Z(B)$ is not stated in [25], but it can be read out of the proof. With a little care Q can be taken to be a parallelepiped, but we make no use of this observation.

Chaining together the inequalities in this section and optimising over k proves Theorem 2.1. A similar process, taking notice of the stability of our application of the Brunn–Minkowski inequality, will prove Theorem 1.1.

2.2 Some Geometric Measure Theory

The next element of our proof incorporates a recent quantitative isoperimetric stability result of Figalli, Maggi and Pratelli [6]. We adopt simplified definitions that suffice for sets that are nice, as defined in the previous subsection; see [17, 18] for the general setting of sets of finite perimeter.

For a closed convex polytope $K \subseteq \mathbb{R}^d$ and a union E of disjoint (possibly non-convex) closed polytopes, the perimeter of E with respect to K is given by

$$\mathrm{Per}_K(E) = \lim_{\varepsilon \to 0^+} \frac{\mu(E + \varepsilon K) - \mu(E)}{\varepsilon}. \tag{1}$$

In our setting, given a nice set A, for all $r \geq 0$ the measure of $A + rK$ and its closure $\overline{A + rK}$ are the same; that is $\mu(A + rK) = \mu(\overline{A + rK})$. Thus for all $r \geq 0$, (1) gives

$$\mathrm{Per}_K(\overline{A + rK}) = \lim_{\varepsilon \to 0^+} \frac{\mu(A + (r+\varepsilon)K) - \mu(A + rK)}{\varepsilon}. \tag{2}$$

The anisotropic isoperimetric problem was posed in 1901 by Wulff [28], who conjectured that minimisers of Per_K up to null sets are homothetic copies of K, giving $\mathrm{Per}_K(E) \geq d\mu(K)^{1/d}\mu(E)^{1-1/d}$. This was established for sets E with continuous boundary by Dinghas [5] and for general sets E of finite perimeter by Gromov [19]. It is equivalent to non-negativity of the *isoperimetric deficit* $\delta_K(E)$ of E with respect to K, defined by

$$\delta_K(E) := \frac{\mathrm{Per}_K(E)}{d\mu(K)^{1/d}\mu(E)^{1-1/d}} - 1.$$

We quantify the structural similarity between K and E via the *asymmetric index* (also known as Fraenkel asymmetry) of E with respect to K, which is given by

$$\mathcal{A}_K(E) = \inf\left\{\frac{\mu(E \triangle (x_0 + rK))}{\mu(E)} : x_0 \in \mathbb{R}^d \text{ and } r^d\mu(K) = \mu(E)\right\}.$$

Theorem 2.4 ([6, Theorem 1.1])**.** *For any $d \in \mathbb{N}$ there exists $D = D(d)$ such that for any bounded convex open set $K \subseteq \mathbb{R}^d$ and $E \subseteq \mathbb{R}^d$ of finite perimeter we have*

$$\mathcal{A}_K(E) \leq D\sqrt{\delta_K(E)}.$$

2.3 Stability

Given these ingredients, let us indicate briefly how Theorem 1.1 follows.

Given a set A which is close to optimal in terms of Theorem 1.2, using Ruzsa's interpretation of the problem in terms of sumsets, we can apply Lemma 2.2 to find a subset $A' \subseteq A$ which is close to optimal in the lattice generated by $\Sigma_{k+p}(B)$. In particular, this leads to a lower bound on the size of A' in terms of $|A' + \Sigma_{k+p}(B)|$. By taking a continuous approximation of this sumset and applying the Brunn-Minkowski inequality we can conclude that $|A'|$ is approximately $|A|$, and so it suffices to show that A' is structurally close to to an appropriate B_n.

Using Lemma 2.3 we can approximate $\Sigma_{k+p}(B)$ by a homothetic copy of $C(B)$, after thickening by an appropraite fundamental set, and hence relate the boundary in this new lattice to the isoperimetric deficit of A' with respect to $C(B)$. In particular, by Theorem 2.4 we can use this to bound the asymmetric index of A' with respect to $C(B)$, and hence by another discrete approximation, to bound the symmetric difference between A' and some B_n.

References

[1] B. Barber and J. Erde. Isoperimetry in integer lattices. *Discrete Anal.*, 7, 2018.

[2] S. Bezrukov. Isoperimetric problems in discrete spaces. *Extremal Problems for Finite Sets, Bolyai Soc. Math. Stud. 3*, pages 59–91, 1994.

[3] B. Bollobás and I. Leader. Compressions and Isoperimetric Inequalities. *J. Comb. Theory Ser. A*, 56:47–62, 1991.

[4] B. Bollobás and I. Leader. Edge-isoperimetric inequalities in the grid. *Combinatorica*, 11:299–314, 1991.

[5] A. Dinghas. Über einen geometrischen Satz von Wulff für die Gleichgewichtsform von Kristallen. *Z. Kristallogr., Mineral. Petrogr.*, 105, 1944.

[6] A. Figalli, F. Maggi, and A. Pratelli. A mass transportation approach to quantitative isoperimetric inequalities. *Invent. Math.*, 182(1):167–211, 2010.

[7] K. Frankston, J. Kahn, B. Narayanan, and J. Park. Thresholds versus fractional expectation-thresholds. *Ann. of Math.*, 194(2):475–495, 2021.

[8] N. Fusco, F. Maggi, and A. Pratelli. The sharp quantitative isoperimetric inequality. *Ann. of Math.*, 168:941–980, 2008.

[9] B. J. Green. The Polynomial Freiman-Ruzsa Conjecture. *Terence Tao's blog*, 2007.

[10] J. Kahn and G. Kalai. Thresholds and expectation thresholds. *Combin. Probab. Comput.*, 16:495–502, 2007.

[11] P. Keevash and E. Long. A stability result for the cube edge isoperimetric inequality. *J. Combin. Theory Ser. A*, 155:360–375, 2018.

[12] P. Keevash and E. Long. Stability for vertex isoperimetry in the cube. *J. Comb. Theory, Ser. B*, 145:113–144, 2020.

[13] N. Keller and N. Lifshitz. Approximation of biased Boolean functions of small total influence by DNF's. *Bull. Lond. Math. Soc*, 50(4):667–679, 2018.

[14] I. Leader. Discrete isoperimetric inequalities. *Probabilistic Combinatorics and its Applications*, 1991.

[15] M. Ledoux. The concentration of measure phenomenon. *Math. Surveys Monogr*, 89, 2005.

[16] L. Lusternik. Die Brunn-Minkowskische Ungleichung für beliebige messbare Mengen. *C. R. Acad. Sci. URSS*, 8:55–58, 1935.

[17] F. Maggi. Some methods for studying stability in isoperimetric type problems. *Bull. Amer. Math. Soc.*, 45:367–408, 2008.

[18] F. Maggi. *Sets of finite perimeter and geometric variational problems: an introduction to Geometric Measure Theory.* Number 135 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2012.

[19] V. Milman and G. Schechtman. *Asymptotic Theory of Finite-dimensional Normed Spaces. With an appendix by M. Gromov*, volume 1200 of *Lecture Notes in Mathematics*. Springer, Berlin, 1986.

[20] R. O'Donnell. *Analysis of Boolean functions.* Cambridge University Press, 2014.

[21] J. Park and H. T. Pham. A proof of the Kahn-Kalai conjecture. *arXiv preprint arXiv:2203.17207*, 2022.

[22] H. Plünnecke. Eine zahlentheoretische Anwendung der Graphentheorie. *J. Reine Angew. Math*, 243:171–183, 1970.

[23] M. Przykucki and A. Roberts. Vertex-isoperimetric stability in the hypercube. *J. of Comb. Theory Ser. A*, 172, 2020.

[24] J. Radcliffe and E. Veomett. Vertex isoperimetric inequalities for a family of graphs on $\mathbb{Z}^k$. *Electron. J. Combin.*, 19(2):P45, 2012.

[25] I. Ruzsa. Sets of sums and commutative graphs. *Studia Sci. Math. Hungar*, 30:127–148, 1995.

[26] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Publ. Math. I.H.E.S.*, 81:73–203, 1995.

[27] D. Wang and P. Wang. Discrete isoperimetric problems. *SIAM J. Appl. Math.*, 32:860–870, 1977.

[28] G. Wulff. Zur Frage der Geschwindigkeit des Wachsrurms und der Auflösung der Kristallflächen. *Z. Kristallogr.*, 34:449–530, 1901.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Nonrepetitive colorings of $\mathbb{R}^d$

(Extended abstract)

Kathleen Barsse* Daniel Gonçalves† Matthieu Rosenfeld‡

Abstract

The results of Thue state that there exists an infinite sequence over 3 symbols without 2 identical adjacent blocks, which we call a 2-nonrepetitive sequence, and also that there exists an infinite sequence over 2 symbols without 3 identical adjacent blocks, which is a 3-nonrepetitive sequence. An *r-repetition* is defined as a sequence of symbols consisting of r identical adjacent blocks, and a sequence is said to be *r-nonrepetitive* if none of its subsequences are r-repetitions. Here, we study colorings of Euclidean spaces related to the work of Thue. A coloring of $\mathbb{R}^d$ is said to be *r-nonrepetitive* of no sequence of colors derived from a set of collinear points at distance 1 is an r-repetition. In this case, the coloring is said to *avoid* r-repetitions. It was proved in [9] that there exists a coloring of the plane that avoids 2-repetitions using 18 colors, and conversely, it was proved in [3] that there exists a coloring of the plane that avoids 43-repetitions using only 2 colors. We specifically study r-nonrepetitive colorings for fixed number of colors : for a fixed number of colors k and dimension d, the aim is to determine the minimum multiplicity of repetition r such that there exists an r-nonrepetitive coloring of $\mathbb{R}^d$ using k colors.

We prove that the plane, $\mathbb{R}^2$, admits a 2- and a 3-coloring avoiding 33- and 18-repetitions, respectively.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-016`

*École Normale Supérieure Paris-Saclay, Gif-sur-Yvette, France. E-mail: `kathleen.barsse@ens-paris-saclay.fr`.
†LIRMM, Univ. Montpellier, CNRS, Montpellier, France. E-mail: `daniel.goncalves@lirmm.fr`.
‡LIRMM, Univ. Montpellier, CNRS, Montpellier, France. E-mail: `matthieu.rosenfeld@lirmm.fr`.

1 Introduction

The Hadwiger-Nelson problem asks for the minimum number of colors required to color the Euclidean plane such that any two points at distance 1 are colored differently. This is called the chromatic number of the plane, and is denoted as $\chi(\mathbb{R}^2)$. The answer to this problem is unknown, but it was proved that $5 \leq \chi(\mathbb{R}^2) \leq 7$ [1, 2, 7]. We study colorings of Euclidean spaces that are connected to the Hadwiger-Nelson problem and where the goal is to avoid specific patterns on straight lines.

An *r-repetition* is a finite sequence of symbols consisting of r identical blocks, where a *block* is a subsequence of consecutive terms. A sequence is r-nonrepetitive if none of its subsequences of consecutive terms are r-repetitions. For instance, the word *hotshots* is a 2-repetition and the word *minimize* is 2-nonrepetitive. The results of Thue state that there exists an infinite 2-nonrepetitive sequence over 3 symbols and an infinite 3-nonrepetitive sequence over 2 symbols. We study the Euclidean variant of Thue sequences introduced by Grytczuk *et al.* [5]. A *straight path* is defined as a sequence of collinear points of $\mathbb{R}^d$, where consecutive points are at distance 1. A coloring of $\mathbb{R}^d$ is *r-nonrepetitive* if for each straight path in $\mathbb{R}^d$, the sequence of the colors of its points is r-nonrepetitive. For fixed integers d and r, the aim is to find the minimum number of colors for which there exists an r-nonrepetitive coloring of $\mathbb{R}^d$. Let $\pi_r(\mathbb{R}^d)$ denote that number.

One easily deduces from Thue's result that $\pi_2(\mathbb{R}) = 3$ and $\pi_3(\mathbb{R}) = 2$. The problem is more difficult for higher dimensions. Colorings of Euclidean spaces that avoid 2-repetitions are called *square-free* colorings. It was proven in [9] that there exists a square-free coloring of the plane that uses 18 colors, which means that $\pi_2(\mathbb{R}^2) \leq 18$. The problem of determining $\pi_2(\mathbb{R}^2)$ is connected to the Hadwiger-Nelson problem in the following way. If a coloring of the plane is 2-nonrepetitive, then 2 points at distance 1 must be colored differently, so at least $\chi(\mathbb{R}^2)$ colors are required. Therefore $5 \leq \chi(\mathbb{R}^2) \leq \pi_2(\mathbb{R}^2) \leq 18$.

Dębski *et al.* studied r-nonrepetitive colorings for larger values of r [3]. More specifically, they gave a proof that for any $d \in \mathbb{N}$, there exists $r = r(d)$ such that $\pi_r(\mathbb{R}^d) = 2$. In other words, for large enough values of r, the problem can be solved with the least possible number of colors. In particular, for $d = 2$, the minimum value of r for which $\pi_r(\mathbb{R}^2) = 2$ is unknown, but the paper provides a proof that $\pi_{43}(\mathbb{R}^2) = 2$ and $\pi_{24}(\mathbb{R}^2) \leq 3$. For smaller values of r, it is known that $\pi_6(\mathbb{R}^2) \leq 4$ and $\pi_3(\mathbb{R}^2) \leq 9$ [4, 9].

We prove that there exists a 33-nonrepetitive coloring of $\mathbb{R}^2$ with 2 colors, that is, $\pi_{33}(\mathbb{R}^2) = 2$. We also prove that $\pi_{18}(\mathbb{R}^2) \leq 3$. Our improvements rely on two main ingredients. First, we provide a better bound on the number of *pathable sequences of hypercubes*. This quantity already played a crucial role in the proof from [3]. Secondly, the proof from [3] uses the Lovász Local Lemma, which we replace with a counting method that yields slightly better bounds in this setting. This argument was first used for nonrepetitive colorings of graphs [6] and was later presented in the more general context of hypergraph coloring [8].

2 Pathable sequences

A standard technique in problems related to colorings of Euclidean spaces is to define a regular tiling of that space and assign the same color to all the points of each tile. The proof of the result from [3] uses a partition of $\mathbb{R}^d$ into hypercubes of diameter 1. We will also use this partition. More precisely, each hypercube is a set of the form $\{(x_1, ..., x_d) \in \mathbb{R}^d : \forall j \in [\![1, d]\!], i_j \leq x_j\sqrt{d} < i_j + 1\}$, with $(i_1, ..., i_d) \in \mathbb{Z}^d$. This way, any two points at distance 1 are always in different hypercubes. Let $\mathcal{H}$ denote the set of hypercubes from this partition.

We call a sequence $(\alpha_0, ..., \alpha_{\ell-1})$ of hypercubes ℓ-pathable if there exists a straight path $(q_0, ..., q_{\ell-1})$ in $\mathbb{R}^d$ with $q_i \in \alpha_i$ for each i (See Figure 1). For a fixed cube H, $D_d(\ell)$ is defined as the number of ℓ-pathable sequences in $\mathbb{R}^d$ containing H (each pair $(\alpha_0, ..., \alpha_{\ell-1})$ and $(\alpha_{\ell-1}, ..., \alpha_0)$ is counted as a single sequence).

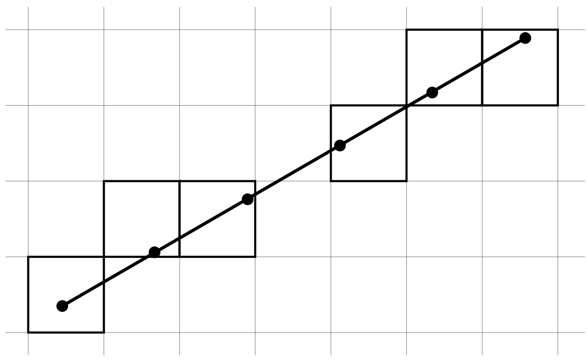


Figure 1: A 6-pathable sequence in $\mathbb{R}^2$.

It is know that $D_d(l) = O(l^{3d})$ [3]. We improve this upper bound for $d = 2$.

Lemma 1. *The number of ℓ-pathable sequences in $\mathbb{R}^2$ is bounded as follows,*

$$D_2(\ell) \leq \frac{2\sqrt{2}}{3}\ell^5 + (2 - \frac{2\sqrt{2}}{3})\ell^3 - 2\ell^2 \, .$$

3 Calculations with the counting argument

In this section, we provide a condition similar to [3, Lemma 2.3]. It provides a condition on r and the number of colors k that ensures that there exists an r-nonrepetitive coloring of $\mathbb{R}^d$ using k colors. However, the condition of Lemma 3, can be proven to be weaker, that is, whenever the condition of [3, Lemma 2.3] holds then our lemma automatically holds with $\beta = k2^{-1/(r-1)}$. In practice, this leads to a slightly better bound for our results.

In the proof of Lemma 2, we will consider an arbitrary subset S of $\mathbb{R}^d$ consisting of finitely many hypercubes from the partition. This method directly shows that there exist exponentially many valid hypercube colorings, with respect to the number of hypercubes in S.

Lemma 2. *Let r, k and d be integers. For every set S of hypercubes, let $\mathcal{C}(S)$ be the set of r-nonrepetitive hypercube colorings of S with k colors.*

If there exists $\beta > 1$ such that

$$k \geq \beta + \sum_{s=1}^{\infty} D_d(rs) \times \beta^{1-(r-1)s},$$

then for every set S of n hypercubes of the partition of $\mathbb{R}^d$ and for every hypercube $H \in S$,

$$|\mathcal{C}(S)| \geq \beta |\mathcal{C}(S-H)|.$$

Remark that $\beta > 1$ and that according to Corollary 2.5 from [3], $D_d(rs) = O((rs)^{3d})$, so the sum in this Lemma is always well-defined.

Proof. We proceed by induction on $n = |S|$. This is true for $n = 1$ because $S - H = \emptyset$. Fix $n \geq 2$ and assume that the result holds for every $i < n$. Let S be a set of n hypercubes and H a hypercube of S. Our induction hypothesis implies that for all $R \subseteq S - H$,

$$\mathcal{C}(S-H-R) \leq \frac{\mathcal{C}(S-H)}{\beta^{|R|}}. \tag{1}$$

Let F be the set of colorings of S that are r-nonrepetitive on $S-H$ but for which there is an r-repetition on S. Then

$$|\mathcal{C}(S)| = k|\mathcal{C}(S-H)| - |F|. \tag{2}$$

Let $s \in \mathbb{N}^*$ and $\alpha = (\alpha_1, ..., \alpha_{rs})$ be a pathable sequence such that $H = \alpha_i$, for some $i \in \{1, \ldots, rs\}$. We define F_α as the subset of F for which there is an r-repetition of length rs on that sequence. Without loss of generality, we assume that $i \geq s+1$. We consider a coloring $\phi \in F_\alpha$. By definition of F, the sequence of colors on α is an r-repetition, and the restriction of ϕ to $S - (\alpha_{s+1}, ..., \alpha_{rs})$ is r-nonrepetitive because $H \in \{\alpha_{s+1}, ..., \alpha_{rs}\}$. Therefore, ϕ is uniquely determined by its restriction to $S - \{\alpha_{s+1}, ..., \alpha_{rs}\}$ and $|F_\alpha| \leq |\mathcal{C}(S - \{\alpha_{s+1}, ..., \alpha_{rs}\}|$. By equation (1), this implies,

$$|F_\alpha| \leq \frac{1}{\beta^{(r-1)s-1}} |\mathcal{C}(S-H)|.$$

Let F_{rs} be the subset of F for which there is an r-repetition of length rs. Recall that $D_d(rs)$ is the number of pathable sequences of length rs containing H. Then,

$$|F_{rs}| \leq D_d(rs) \frac{1}{\beta^{(r-1)s-1}} |\mathcal{C}(S-H)|.$$

Now, by summing over all s, and by using our main hypothesis

$$|F| = \left| \bigcup_{s=1}^{\infty} F_{rs} \right| \leq \sum_{s=1}^{\infty} |F_{rs}| \leq \sum_{s=1}^{\infty} D_d(rs) \frac{1}{\beta^{(r-1)s-1}} |\mathcal{C}(S-H)| \leq |\mathcal{C}(S-H)|(k-\beta).$$

Using this bound inside equation (2),

$$|\mathcal{C}(S)| = k|\mathcal{C}(S-H)| - |F| \geq \beta|\mathcal{C}(S-H)|$$

which concludes our induction. □

For each subset S of $\mathbb{R}^d$ consisting of n hypercubes, $|\mathcal{C}(S)| \geq \beta^{n-1}k$. This means that any finite arbitrary subset of hypercubes of the partition of $\mathbb{R}^d$ can be r-nonrepetitively colored. By compacity (e.g., see the proof of Lemma 2.3 from [3]) there exists an r-nonrepetitive coloring of $\mathbb{R}^d$.

Lemma 3. *For every integers r, k and d, if there exists $\beta > 1$ such that*

$$k \geq \beta + \sum_{s=1}^{\infty} D_d(rs) \times \beta^{1-(r-1)s}$$

then $\pi_r(\mathbb{R}^d) \leq k$.

4 Proof of the main results and conclusion

We can now use the bound from Lemma 1 to verify the conditions of Lemma 3 for well-chosen values of r, β and k. In particular, one can verify that the condition of Lemma 3 holds for $r = 33$, $\beta = 19/10$ and $k = 2$ which implies the following result.

Theorem 4. *There exists a 2-coloring of the plane avoiding 33-repetitions.*

Let $r(d)$ denote the least positive integer such that $\pi_{r(d)}(\mathbb{R}^d) = 2$. We proved that $r(2) \leq 33$, which improves the bound $r(2) \leq 43$ proved in [3]. However, this result probably isn't optimal, since the best known lower bound is $r(2) \geq 3$ which is a consequence of the results of Thue. This means that $r(2)$ lies between 3 and 33. In fact, it is conjectured in [3] that $r(2) = 4$.

Similarly, one can verify that the condition of Lemma 3 holds for $r = 18$, $\beta = 8/3$ and $k = 3$ which implies the following result.

Theorem 5. *There exists a 3-coloring of the plane avoiding 18-repetitions.*

Again the value 18 is an improvement from 24 but is probably still not optimal.

References

[1] P. Brass, W. Moser, and J. Pach. *Research Problems in Discrete Geometry.* Springer, New York, 2005.

[2] A. de Grey. The chromatic number of the plane is at least 5. *Geombinatorics*, 28:18–31, 2018.

[3] M. Dębski, J. Grytczuk, B. Nayar, U. Pastwa, J. Sokół, M. Tuczyński, P. Wenus, and K. Węsek. Avoiding multiple repetitions in euclidean spaces. *SIAM Journal on Discrete Mathematics*, 34(1):40–52, 2020.

[4] M. Dębski, U. Pastwa, and K. Węsek. Grasshopper avoidance of patterns. *Electron. J. Combin.*, 23:1–16, 2016.

[5] J. Grytczuk, K. Kosiński, and M. Zmarz. Nonrepetitive colorings of line arrangements. *European Journal of Combinatorics*, 51:275–279, 2016.

[6] M. Rosenfeld. Another approach to non-repetitive colorings of graphs of bounded degree. *Electronic Journal of Combinatorics*, 27(3), 2020.

[7] A. Soifer. *The Mathematical Coloring Book: Mathematics of Coloring and the Colorful Life of Its Creators.* Springer, New York, 2008.

[8] I. M. Wanless and D. R. Wood. A general framework for hypergraph coloring. *SIAM Journal on Discrete Mathematics*, 36(3):1663–1677, 2022.

[9] P. Wenus and K. Węsek. Nonrepetitive and pattern-free colorings of the plane. *European Journal of Combinatorics*, 54:21–34, 2016.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Exact antichain saturation numbers via a generalisation of a result of Lehman-Ron

(Extended abstract)

Paul Bastide * Carla Groenland † Hugo Jacob ‡ Tom Johnston §

Abstract

For given positive integers k and n, a family $\mathcal{F}$ of subsets of $\{1,\ldots,n\}$ is k-antichain saturated if it does not contain an antichain of size k, but adding any set to $\mathcal{F}$ creates an antichain of size k. We use $\mathrm{sat}^*(n,k)$ to denote the smallest size of such a family. For all k and sufficiently large n, we determine the exact value of $\mathrm{sat}^*(n,k)$. Our result implies that $\mathrm{sat}^*(n,k) = n(k-1) - \Theta(k \log k)$, which confirms several conjectures on antichain saturation. Previously, exact values for $\mathrm{sat}^*(n,k)$ were only known for k up to 6.

We also prove a strengthening of a result of Lehman-Ron which may be of independent interest. We show that given m disjoint chains in the Boolean lattice, we can create m disjoint skipless chains that cover the same elements (where we call a chain skipless if any two consecutive elements differ in size by exactly one).

The complete version of the paper can be found here [4].

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-017`

*LaBRI - University of Bordeaux, `paul.bastide@ens-rennes.fr`

†Utrecht University, `c.e.groenland@uu.nl`, This project has received funding from the European Union's Horizon 2020 research and innovation programme under the ERC grant CRACKNP (number 853234) and the Marie Skłodowska-Curie grant GRAPHCOSY (number 101063180). Views and opinions expressed are however those of the author(s) only.

‡ENS Paris-Saclay, `hjacob@ens-paris-saclay.fr`

§University of Bristol and Heilbronn Institute for Mathematical Research, `tom.johnston@bristol.ac.uk`

Many powerful results have been proved over the years concerning the structure of chains and antichains in the Boolean lattice, e.g. [10, 14, 20, 21, 22]. For example, it is well-known that the Boolean lattice admits a symmetric chain decomposition [1, 9], and in fact these chains may be taken to be *skipless* (or *saturated*): every chain $C_1 \subsetneq \cdots \subsetneq C_r \subseteq [n] = \{1, \dots, n\}$ has the property that $|C_{i+1}| = |C_i| + 1$ for all $i \in [r-1]$. Skipless chains have also been studied in other contexts such as in [3, 6, 16].

Given sets $X_1, \dots, X_m$ from layer r and sets $Y_1, \dots, Y_m$ from layer s such that $X_i \subseteq Y_i$, it need not be possible to find disjoint skipless chains $C^1, \dots, C^m$ linking X_1 to Y_1, X_2 to Y_2 etc. However, it was shown by Lehman and Ron [15] in 2001 that there always exist m disjoint skipless chains that *cover* the sets $X_1, \dots, X_m$ and $Y_1, \dots, Y_m$.

Theorem 1 (Lehman-Ron [15]). *Let integers $1 \le s < r \le n$ and subsets $X_1, \dots, X_m, Y_1, \dots, Y_m \subseteq [n]$ be given with $|X_i| = s$, $|Y_i| = r$ and $X_i \subseteq Y_i$ for all $i \in [m]$. Then there exist m disjoint skipless chains that cover $\{X_1, \dots, X_m, Y_1, \dots, Y_m\}$.*

It is natural to ask if a stronger statement holds. For example, what happens if we allow the sets to come from different layers, or ask that the chains go via some elements from layers between layer r and layer s? Is it possible to cover any m disjoint chains with m disjoint skipless chains, or can we force the use of an additional chain? We show that m chains always suffice. For a family $\mathcal{F}$, we say that $\mathcal{F}$ admits a chain decomposition into m chains if there exits m disjoint chains $C_1, \dots, C_m$ that covers $\mathcal{F}$.

Theorem 2. *Suppose that $\mathcal{F} \subseteq 2^{[n]}$ admits a chain decomposition into m chains. Then there exist disjoint skipless chains $C^1, \dots, C^m$ such that $\mathcal{F} \subseteq \bigcup_{i=1}^m C^i$.*

Proof overview. The core of the proof, despite being slightly more technical, follows a method similar to the one used in [15]. It uses multiple inductive arguments to reduce the problem to a well-structured instance. From there it is possible to uses Menger's theorem [18] to deduce connectivity properties.

The building blocks for our induction can are as follows (see Fig. 1 for an example):

Claim 3. *Let $s \le r \le n$ be integers. Let $C^1, \dots, C^m$ be disjoint chains, such that for all $i \in [m-1]$, the chain C^i starts in layer s and ends in layer r. Suppose that C^m starts in $A \in \binom{[n]}{\le s}$ and ends in $B \in \binom{[n]}{r}$. Then there exist m disjoint chains $D^1, \dots, D^m$ with the following three properties.*

1. *For $i \in [m-1]$, the chain D^i starts in the sth layer, ends in the rth layer and is skipless.*

2. *The chain D^m starts at A and intersects the ith layer for all $i \in [s+1, r]$.*

3. *The chains $D^1, \dots, D^m$ cover the elements in $C^1, \dots, C^m$.*

□

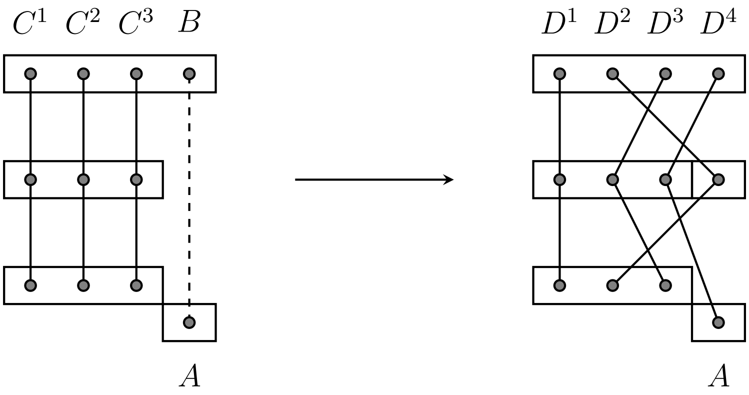


Figure 1: Representation of Claim 3 (case $r = s + 2$ and $m = 3$).

Theorem 2 was already known in the special case that the union $\mathcal{F}$ of the chains we wish to cover is a convex set system (i.e. if $X, Y \in \mathcal{F}$ and $X \subseteq Z \subseteq Y$, then $Z \in \mathcal{F}$) [6]. In this case, the chains can be taken to partition $\mathcal{F}$ as any additional sets must be at the ends of the chains.

Although we believe Theorem 2 to be of interest in its own right, our initial motivation came from the area of induced poset saturation where we use Theorem 2 to easily settle various conjectures concerning the asymptotics of antichain saturation numbers. With more work, we are in fact able to go well beyond the conjectures and pinpoint the exact values.

For given positive integers k and n, a family $\mathcal{F}$ of subsets of $[n]$ is *k-antichain saturated* if it does not contain an antichain of size k, but for all $X \subseteq [n]$ with $X \notin \mathcal{F}$, the family $\mathcal{F} \cup \{X\}$ does contain an antichain of size k. We denote the size of the smallest such family by $\mathrm{sat}^*(n, k)$.

In the literature, this is also sometimes denoted $\mathrm{sat}^*(n, \mathcal{A}_k)$, where $\mathcal{A}_k$ is the poset consisting of k incomparable elements. This is called an *induced* saturation number: it is the size of the smallest set system which is saturated in terms of not containing $\mathcal{A}_k$ as an *induced* subposet. Such saturation numbers for the Boolean lattice were introduced by Gerbner, Keszegh, Lemons, Palmer, Pálvölgyi and Patkós [8] and have been investigated for a variety of posets, for example for the butterfly [11], the diamond [12] and the chain [19]. We refer to [13] for a nice overview.

Ferrara, Kay, Kramer, Martin, Reiniger, Smith and Sullivan [7] were the first to study the particular case of the antichain and made the following conjecture.

Conjecture 1 ([7]). *For $k \geq 3$, $\mathrm{sat}^*(n, k) \sim (k-1)n$ as $n \to \infty$.*

The upper bound is easy to see: for all $i \in [n]$, a k-antichain saturated family can contain at most $k - 1$ subsets of size i since two subsets of the same size are incomparable. Moreover, a k-antichain saturated family must always exist since we can start with the empty family and greedily add subsets until it is no longer possible to do so without creating an antichain of size k.

Martin, Smith and Walker [17] proved the lower bound

$$\mathrm{sat}^*(n,k) \geq \left(1 - \frac{1}{\log_2(k-1)}\right) \frac{(k-1)n}{\log_2(k-1)}$$

for $k \geq 4$ and n sufficiently large. The exact values for $k = 2, 3$ and 4 were shown to be $n+1$, $2n$ and $3n-1$ respectively in [7], the exact values for $k = 5$ and $k = 6$ were recently determined to be $4n-2$ and $5n-5$ respectively by Đanković and Ivan [2]. They also strengthened Conjecture 1 as follows, and proposed two weaker conjectures implied by this conjecture.

Conjecture 2 ([2]). $\mathrm{sat}^*(n,k) = n(k-1) - O_k(1)$.

We show later how all the conjectures mentioned above are easily derived from Theorem 2. In particular, we will show the following corollary.

Corollary 4. *There exist constants $c_1, c_2 > 0$ such that for all $k \geq 4$ and n sufficiently large,*

$$n(k-1) - c_1 k \log k \leq \mathrm{sat}^*(n,k) \leq n(k-1) - c_2 k \log k.$$

In general, obtaining exact saturation numbers is a notoriously difficult problem, and for the antichain exact numbers were only known for k up to 6. Our main result determines the exact value of $\mathrm{sat}^*(n,k)$ for all values of k and n where n is large enough relative to k . We note that n need not be excessively large compared to k and it certainly suffices to assume $n \geq 6 \log k + 1$ for example. Determining the exact values is considerably more involved than just determining the asymptotics, and we require some more definitions just to state the value of the numbers.

Given a natural number k, let ℓ be the smallest integer j such that $\binom{j}{\lfloor j/2 \rfloor} \geq k-1$. Note that when $n < \ell$, there are no antichains of size k in $2^{[n]}$ and $\mathcal{F}$ must contain every set (i.e. $\mathrm{sat}^*(n,k) = 2^n$).

Let $\mathcal{C}(m,t)$ denote the initial segment of layer t of size m when the sets are in colexicographic order. For a family of sets $\mathcal{A}$ from the same layer, let $\nu(\mathcal{A})$ be the size of the maximum matching from $\mathcal{A}$ to its shadow $\partial\mathcal{A}$, and recursively define $c_0, c_1, \ldots, c_{\lfloor \ell/2 \rfloor}$ as follows. Let $c_{\lfloor \ell/2 \rfloor} = k-1$. For $0 \leq t < \lfloor \ell/2 \rfloor$, let $c_t = \nu\left(\mathcal{C}(c_{t+1}, t+1)\right)$.

Theorem 5. *Let $n, k \geq 4$ be integers and let ℓ and $c_0, \ldots, c_{\lfloor \ell/2 \rfloor}$ be as defined above. If $n < \ell$, then $\mathrm{sat}^*(n,k) = 2^n$. If $n \geq \ell$, then*

$$\mathrm{sat}^*(n,k) \geq 2 \sum_{t=0}^{\lfloor \ell/2 \rfloor} c_t + (k-1)(n-1-2\lfloor \ell/2 \rfloor).$$

Moreover, equality holds when $n \geq 2\ell + 1$.

Given the form of the bound in Theorem 5, one might be tempted to suggest that the best approach is to take each layer $t \leq \lfloor \ell/2 \rfloor$ to be an initial segment of colex of the appropriate size, but this is not the case in general. While such an example would have the optimal size, it may already contain an antichain of size k. For example, one can check there is an antichain of size 262 in $\mathcal{C}(261,5)\cup\mathcal{C}(219,4)$, and this approach would not work for $k = 262$.

For infinitely many values of k, a matching upper bound to Theorem 5 was already known [7] which works for all $n \geq \ell+1$. It gives the following corollary.

Corollary 6. *Let ℓ, k, n be integers such that $\binom{\ell}{\lfloor \ell/2 \rfloor} = k-1$. If $n \leq \ell$ then $\mathrm{sat}^*(n,k) = 2^n$. If $n \geq \ell+1$, then*

$$\mathrm{sat}^*(n,k) = 2\sum_{j=0}^{\lfloor \ell/2 \rfloor}\binom{\ell}{j} + (k-1)(n-1-2\lfloor \ell/2 \rfloor).$$

In particular, whenever $k-1$ is a central binomial coefficient (i.e. $k = 3, 4, 7, 11, 21, 36, \dots$) the value of $\mathrm{sat}^*(n,k)$ is determined for all n.

We now explain how Corollary 4 follows from Theorem 2. The upper bound was already known, and we prove a lower bound of $\mathrm{sat}^*(n,k) \geq (n+1-2\ell)(k-1)$ for n sufficiently large. (Recall that ℓ is the smallest j such that $\binom{j}{\lfloor j/2 \rfloor} \geq k-1$, so $\ell = \Theta(\log k)$.)

By Dilworth's theorem [5], having a chain decomposition of size at most $k-1$ is equivalent to not containing any antichain of size k. Suppose that $\mathcal{F} \subseteq 2^{[n]}$ is k-antichain saturated and so admits a decomposition into $k-1$ chains. By Theorem 2, there are $k-1$ disjoint skipless chains $C^1, \dots, C^{k-1}$ that cover the elements of $\mathcal{F}$; since $\mathcal{F}$ is saturated, this must form a chain decomposition of $\mathcal{F}$. It suffices to show that every chain must contain a set of size at most ℓ and a set of size at least $n-\ell$. Suppose the smallest element X of some chain C^i has size $|X| > \ell$, then all subsets Y of X must be present in $\mathcal{F}$ since otherwise we may extend C^i to include Y (and that would mean that $\mathcal{F}\cup\{Y\}$ can also be covered by $k-1$ chains, contradicting the fact that $\mathcal{F}$ is k-antichain saturated). There are at least $k-1$ subsets of X of size $\lfloor \ell/2 \rfloor$, and these cannot all be covered by the other $k-2$ chains. Since each chain contains an element of size at most ℓ and one of size at least $n-\ell$, the bound follows immediately from the fact that the chains are skipless.

In order to prove the exact lower bound of Theorem 5, it is needed to examine what happens on layers $1, \dots, \ell$. This is considerably more delicate and for this we use an auxiliary result concerning the matching number of the colex order, explained in details in the complete version of the paper [4]. There, we also give an explicit construction of a k-antichain saturated system $\mathcal{F}$ which matches our lower bound on each layer provided n is sufficiently large. This construction was already known for the special case $k-1 = \binom{\ell}{\lfloor \ell/2 \rfloor}$, and we apply it recursively for other values of k. The recursion requires special care and depends on a particular way of writing $k-1$ as a sum of binomial coefficients. This notation can be used to write exact values for the matching numbers c_t from Theorem 5.

References

[1] Martin Aigner. Lexicographic matching in Boolean algebras. *Journal of Combinatorial Theory, Series B*, 14:187–194, 1973.

[2] Irina Đanković and Maria-Romina Ivan. Saturation for small antichains. *arXiv preprint arXiv:2205.07392*, 2022.

[3] Béla Bajnok and Shahriar Shahriari. Long symmetric chains in the boolean lattice. *Journal of Combinatorial Theory, Series A*, 75(1):44–54, 1996.

[4] Paul Bastide, Carla Groenland, Hugo Jacob, and Tom Johnston. Exact antichain saturation numbers via a generalisation of a result of Lehman-Ron. *https://arxiv.org/abs/2207.07391*.

[5] Robert P. Dilworth. A decomposition theorem for partially ordered sets. *Annals of Mathematics*, pages 161–166, 1950.

[6] Dwight Duffus, David Howard, and Imre Leader. The width of downsets. *European Journal of Combinatorics*, 79:46–59, 2019.

[7] Michael Ferrara, Bill Kay, Lucas Kramer, Ryan R Martin, Benjamin Reiniger, Heather C Smith, and Eric Sullivan. The saturation number of induced subposets of the Boolean lattice. *Discrete Mathematics*, 340(10):2479–2487, 2017.

[8] Dániel Gerbner, Balázs Keszegh, Nathan Lemons, Cory Palmer, Dömötör Pálvölgyi, and Balázs Patkós. Saturating Sperner families. *Graphs and Combinatorics*, 29(5):1355–1364, 2013.

[9] Curtis Greene and Daniel J. Kleitman. Strong versions of Sperner's theorem. *Journal of Combinatorial Theory, Series A*, 20(1):80–88, 1976.

[10] Jerrold R. Griggs. Collections of subsets with the Sperner property. *Transactions of the American Mathematical Society*, 269(2):575–591, 1982.

[11] Maria-Romina Ivan. Saturation for the butterfly poset. *Mathematika*, 66:806–817, 2020.

[12] Maria-Romina Ivan. Minimal diamond-saturated families. *Contemporary Mathematics*, 3(2), 2022.

[13] Balázs Keszegh, Nathan Lemons, Ryan R. Martin, Dömötör Pálvölgyi, and Balázs Patkós. Induced and non-induced poset saturation problems. *Journal of Combinatorial Theory, Series A*, 184:105497, 2021.

[14] D. J. Kleitman. On an extremal property of antichains in partial orders. the lym property and some of its implications and applications. In M. Hall and J. H. van Lint, editors, *Combinatorics*, pages 277–290. Springer Netherlands, 1975.

[15] Eric Lehman and Dana Ron. On disjoint chains of subsets. *Journal of Combinatorial Theory, Series A*, 94(2):399–404, 2001.

[16] Mark J. Logan. Sperner theory in a difference of boolean lattices. *Discrete Mathematics*, 257:501–512, 2002.

[17] Ryan R Martin, Heather C Smith, and Shanise Walker. Improved bounds for induced poset saturation. *The Electronic Journal of Combinatorics*, 27(2):P2.31, 2020.

[18] Karl Menger. Zur allgemeinen kurventheorie. *Fundamenta Mathematicae*, 10(1):96–115, 1927.

[19] Natasha Morrison, Jonathan A. Noel, and Alex Scott. On saturated k-Sperner systems. *The Electronic Journal of Combinatorics*, 21(3):P3.22, 2014.

[20] Michael Saks. A short proof of the existence of k-saturated partitions of partially ordered sets. *Advances in Mathematics*, 33:207–2011, 1979.

[21] Emanuel Sperner. Ein Satz über Untermengen einer endlichen Menge. *Mathematische Zeitschrift*, 27(1):544–548, 1928.

[22] Benny Sudakov, István Tomon, and Adam Zsolt Wagner. Uniform chain decompositions and applications. *Random Structures & Algorithms*, 60(2):261–286, 2022.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Chromatic number of intersection graphs of segments with two slopes*

(Extended abstract)

Sudatta Bhattacharya Zdeněk Dvořák Fariba Noorizadeh†

Abstract

A *d-dir* graph is an intersection graph of segments, where the segments have at most d different slopes. It is easy to see that a d-dir graph with clique number ω has chromatic number at most $d\omega$. We study the chromatic number of 2-dir graphs in more detail, proving that this upper bound is tight even in the fractional coloring setting.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-018`

The clique number (the size of the largest complete subgraph) is clearly a lower bound on the chromatic number of a graph. It is natural to ask whether the chromatic number can be also bounded from above as a function of the clique number. In particular, a graph G is called *perfect* if G as well as all its induced subgraphs have chromatic number equal to clique number. Graphs from many interesting graph classes, such as bipartite graphs, their linegraphs, interval and chordal graphs, and their complements, are known to be perfect. A celebrated strong perfect graph theorem [5] states that a graph is perfect if and only if it avoids odd holes and antiholes.

Relaxing the notion of perfectness, Gyárfás [11] introduced the notion of χ-boundedness: A graph class $\mathcal{G}$ is χ-*bounded* if there exists a function f such that $\chi(G) \le f(\omega(G))$ for every $G \in \mathcal{G}$. Not all graphs are χ-bounded—Erdős [8] famously proved that there exist graphs with arbitrarily large chromatic number and arbitrarily large girth. However, many interesting graph classes have this property. For example, in the geometric setting,

*Supported by project 22-17398S (Flows and cycles in graphs on surfaces) of Czech Science Foundation.

†Computer Science Institute, Charles University, Prague, Czech Republic. E-mail: `{sudatta,rakdver, f.noorizadeh}@iuuk.mff.cuni.cz`

intersection graphs of scaled and translated copies of any compact convex shape [13], circle graphs (intersection graphs of the chords of a circle) [10, 6], intersection graphs of unit-length segments in the plane [18] and intersection graphs of axis-aligned rectangles in the plane [1] are χ-bounded (on the other hand, triangle-free intersection graphs of boxes in $\mathbb{R}^3$ have unbounded chromatic number [2]).

In this paper, we are interested in a variant of intersection graphs of straight line segments in the plane. A graph G is a *segment intersection graph* if we can assign to each vertex of G a segment in the plane so that distinct vertices $u, v \in V(G)$ are adjacent iff the corresponding segments intersect in at least one point. Answering in negative an old question of Erdős, Pawlik et al. [16] proved that there exist triangle-free segment intersection graphs with arbitrarily large chromatic number. However, it is easy to see that such graphs must contain segments of many different slopes. Indeed, suppose each segment of a segment intersection graph G has one of at most d slopes (such graphs are called *d-dir*). The segments in each direction induce an interval graph, and since interval graphs are perfect, we conclude that $\chi(G) \leq d \cdot \omega(G)$. Hence, for every d, the class of d-dir graphs is χ-bounded.

For context, all planar graphs are segment intersection graphs [4], bipartite planar graphs are 2-dir [7, 12], and 3-colorable planar graphs are 3-dir [9]. Moreover, West [19] conjectured all planar graphs are 4-dir. Let us remark that recognizing whether a graph is a d-dir graph is NP-complete for any $d \geq 2$, while recognizing the segment intersection graphs in general is $\exists\mathbb{R}$-complete, and thus NP-hard and known to be in PSPACE [14].

There are very few χ-bounded graph classes where tight bounds on the chromatic number in terms of the clique number are known. This motivates us to ask whether the trivial bound on the chromatic number of d-dir graphs can be improved. In this paper, we focus on the basic case $d = 2$. Without loss of generality, each 2-dir graph G is represented by horizontal and vertical segments. A *row* of G is a horizontal line that contains at least one of the horizontal segments of G. To get a more detailed understanding, we consider 2-dir graphs with bounded number of rows: Let $\mathcal{D}_{k,t}$ denote the class of all 2-dir graphs of clique number at most t and with at most k rows, and let $\mathcal{D}_t = \bigcup_{k=1}^{\infty} \mathcal{D}_{k,t}$ be the class of all 2-dir graphs of clique number at most t.

It turns out to be convenient to consider *fractional chromatic number* in addition to standard chromatic number. A *fractional coloring* of a graph G is a function φ that to each vertex assigns a set of measure one (in any measure space), such that $\varphi(u) \cap \varphi(v) = \emptyset$ for every $uv \in E(G)$. The *span* of φ is the measure of $\bigcup_{v \in V(G)} \varphi(v)$. The *fractional chromatic number* $\chi_f(G)$ of G is the infimum of the real numbers c such that G has a fractional coloring of span at most c. It is known that the infimum is actually a minimum and that it is always rational [17]. Clearly $\chi_f(G) \leq \chi(G)$ for every graph G. Moreover, it is known that there exist graphs with fractional chromatic number arbitrarily close to 2 and with arbitrarily large chromatic number [15].

For a class of graphs $\mathcal{G}$, let $\chi(\mathcal{G})$ and $\chi_f(\mathcal{G})$ denote the supremum of chromatic and fractional chromatic numbers, respectively, of graphs from $\mathcal{G}$. Our results are based on a detailed investigation of the (fractional) chromatic number of triangle-free 2-dir graphs with a given number of rows, summarized in the following theorem.

Theorem 1. *The following bounds hold:*

- $\chi(\mathcal{D}_{1,2}) = \chi_f(\mathcal{D}_{1,2}) = 2$.
- $\chi(\mathcal{D}_{2,2}) = 3$ *and* $\chi_f(\mathcal{D}_{2,2}) = 5/2$.
- $\chi(\mathcal{D}_{3,2}) = 3$ *and* $\chi_f(\mathcal{D}_{3,2}) = 11/4$.
- $\chi(\mathcal{D}_{k,2}) = 4$ *for every* $k \geq 4$ *and* $\chi_f(\mathcal{D}_{4,2}) = 3$.
- $\chi_f(\mathcal{D}_{k,2}) \leq 4 - \frac{1}{2^{k-1}}$ *for every* $k \geq 1$.
- *For all integers* $r, m \geq 1$, $\chi_f(\mathcal{D}_{rm,2}) \geq 4 - \frac{1}{m} - \frac{1}{2^{r-1}}$.

Consequently, $\chi(\mathcal{D}_2) = \chi_f(\mathcal{D}_2) = 4$*, but there are no triangle-free 2-dir graphs with fractional chromatic number exactly* 4*.*

Let us remark that the fact that $\chi_f(\mathcal{D}_2) = 4$ also follows from the construction of [3], who gave construction of 2-dir triangle-free graphs G with $|V(G)|/\alpha(G)$ arbitrarily close to 4.

For a positive integer b, the *b-blowup* of a graph G is the graph H obtained from G by replacing each vertex by a clique of size b and each edge uv by a complete bipartite graph between the cliques replacing u and v. Note that clique number, fractional chromatic number, and being a d-dir graph all behave predictably after a blowup (this is not the case for the ordinary chromatic number, motivating our focus on its fractional version): We have $\omega(H) = b \cdot \omega(G)$ and $\chi_f(H) = b \cdot \chi_f(G)$, and if G is a d-dir graph, then H is a d-dir graph as well.

Since $\chi_f(\mathcal{D}_2) = 4$, for every even integer t and $\varepsilon > 0$, there exists a triangle-free 2-dir graph with fractional chromatic number at least $4 - \frac{2\varepsilon}{t}$, and applying the $(t/2)$-blowup operation results in a graph of clique number t and fractional chromatic number at least $2t - \varepsilon$. This gives our main result: For 2-dir graphs, the trivial upper bound $\chi(\mathcal{D}_t) \leq 2t$ cannot be improved when the clique number t is even.

Corollary 2. *For every even* t*, we have* $\chi(\mathcal{D}_t) = \chi_f(\mathcal{D}_t) = 2t$*.*

For odd t, this only gives a bound $2t - 2 \leq \chi_f(\mathcal{D}_t) \leq \chi(\mathcal{D}_t) \leq 2t$. We suspect the upper bound is tight in this case as well. Finally, we conjecture this is the case for d-dir graphs in general.

Conjecture 3. *For all positive integers* d *and* t *and real* $\varepsilon > 0$*, there exists a* d*-dir graph* G *of clique number* t *whose fractional chromatic number is at least* $dt - \varepsilon$*.*

Because of the blowup operation, to prove this for even t one only needs to consider the case $t = 2$, i.e., triangle-free d-dir graphs. In the rest of this extended abstract, let us outline the construction showing the lower bounds on the fractional chromatic number stated in Theorem 1.

Lower bounds on the fractional chromatic number

Throughout the rest of the paper, let μ be Lebesgue measure on subsets of real numbers. We will work with a slightly more general notion of fractional coloring. Let G be a graph and let $f : V(G) \to \mathbb{R}_0^+$ be an arbitrary function. An *f-fractional coloring* of G is a function φ that to each vertex $v \in V(G)$ assigns a set of real numbers such that $\mu(\varphi(v)) = f(v)$ and $\varphi(u) \cap \varphi(v) = \emptyset$ for every $uv \in E(G)$. The *f-fractional chromatic number* of G is the infimum of the spans of its f-fractional colorings.

We say that a 2-dir graph is *horizontally trivial* if each of its rows contains exactly one horizontal segment v and this segment v intersects all vertical segments that intersect the row. In other words, the segment v can be extended arbitrarily far in each direction along the row without changing the intersection graph. The motivation for requiring different amounts of colors at each vertex comes from the following observation.

Lemma 4. *Let $r, m \geq 1$ be integers and let G be a horizontally trivial triangle-free 2-dir graph with at most m rows. Let $f : V(G) \to \mathbb{R}^+$ be defined by setting $f(v) = 1$ for each vertical segment v and $f(v) = 2 - \frac{1}{2^{r-1}}$ for each horizontal segment v. Then G has f-fractional chromatic number at most $\chi_f(\mathcal{D}_{rm,2})$.*

Proof. Without loss of generality, we can assume that for $i \in \{0, \ldots, m-1\}$, G has a horizontal segment h_i with endpoints $(0, i)$ and $(1, i)$, and that each vertical segment has the x-coordinate strictly between 0 and 1. Moreover, we can assume that for each endpoint $p = (x, y)$ of a vertical segment, there exists an integer j such that $j - 1/r < y < j$: For any integer j, we can shift all endpoints with $j - 1 < y < j$ to this interval without changing the intersection graph; and if an endpoint p has y-coordinate exactly j, then since G is triangle-free, it is an endpoint of exactly one vertical segment v, and thus we can shift p so that its y-coordinate becomes less than j (if p is the bottom endpoint of v) or more than $j + 1 - 1/r$ (if p is the top endpoint of v).

Let G' be the triangle-free 2-dir graph with rm rows obtained as follows: We copy each vertical segment of G 2^{rm-1} times and shift the copies by 0, 1, …, $2^{rm-1} - 1$ to the right. For $i = 0, \ldots, rm - 1$, G' has a row ℓ_i with y-coordinate $y_i = i/r$, and this row contains 2^i horizontal segments $v_{i,j}$ for $j \in \{0, \ldots, 2^i - 1\}$, where $v_{i,j}$ has endpoints $(j \cdot 2^{rm-1-i}, y_i)$ and $((j + 1) \cdot 2^{rm-1-i}, y_i)$. We view the horizontal segments as arranged in a tree-like fashion, and if $i < rm - 1$, we say that $v_{i+1,2j}$ and $v_{i+1,2j+1}$ are the *children* of $v_{i,j}$; note that the projection of $v_{i,j}$ on the x-axis is the union of the projections of its children. Note also that by the assumptions on the y-coordinates of the endpoints of the vertical segments of G, for $k \in \{0, \ldots, m-1\}$, the rows ℓ_{kr}, ℓ_{kr+1}, …, ℓ_{kr+r-1} intersect exactly the same vertical segments of G'.

Since $G' \in \mathcal{D}_{rm,2}$, G' has a fractional coloring φ of span at most $\chi_f(\mathcal{D}_{rm,2})$. For $k = 1, \ldots, m$, let us choose a horizontal segment $u_k = v_{kr-1,j_k}$ in the row ℓ_{kr-1} and a set C_k of colors of measure $2 - \frac{1}{2^{r-1}}$ as follows:

- If $k = 1$, then let $u_{k,0} = v_{0,0}$, otherwise let $u_{k,0}$ be a child of u_{k-1} chosen arbitrarily, and let $C_{k,0} = \varphi(u_{k,0})$.

- For $i = 1, \dots, r-1$, choose $u_{k,i} = u'$ among the children u' and u'' of $u_{k,i-1}$ so that $\mu(C_{k,i-1} \cap \varphi(u')) \le \mu(C_{k,i-1} \cap \varphi(u''))$. Since $u'u'' \in E(G')$, $\varphi(u')$ and $\varphi(u'')$ are disjoint, and thus $\mu(C_{k,i-1} \cap \varphi(u')) \le \frac{1}{2}\mu(C_{k,i-1})$. Hence, $\mu(C_{k,i-1} \cup \varphi(u')) = \mu(C_{k,i-1}) + 1 - \mu(C_{k,i-1} \cap \varphi(u')) \ge \frac{1}{2}\mu(C_{k,i-1}) + 1$. We let $C_{k,i}$ be a subset of $C_{k,i-1} \cup \varphi(u')$ of measure exactly $2 - \frac{1}{2^i}$. Such a set exists, since by induction, we have $\frac{1}{2}\mu(C_{k,i-1}) + 1 = \frac{1}{2}\big(2 - \frac{1}{2^{i-1}}\big) + 1 = 2 - \frac{1}{2^i}$.
- Let $u_k = u_{k,r-1}$ and $C_k = C_{k,r-1}$.

Consider now the assignment φ' of sets of colors to vertices of G defined as follows: For $k = 1, \dots, m$, the horizontal segment h_{k-1} with endpoints $(0, k-1)$ and $(1, k-1)$ gets $\varphi'(h_{k-1}) = C_k$. Hence, $\mu(\varphi'(h_{k-1})) = 2 - \frac{1}{2^{r-1}} = f(h_{k-1})$. For each vertical segment v, let v' be its copy in G' whose projection on the x-axis is contained inside the projection of u_m (and thus also inside the projection of $u_1, \dots, u_{m-1}$), and let $\varphi'(v) = \varphi(v')$. Note that if v' intersects u_k for some $k \in \{1, \dots, m\}$, then it also intersects all horizontal segments whose color sets contribute to C_k, and thus $\varphi'(v)$ is disjoint from $C_k = \varphi'(h_{k-1})$. Consequently, φ' is an f-fractional coloring of G whose span is at most the span of φ, and thus at most $\chi_f(\mathcal{D}_{rm,2})$ □

We use duality to prove lower bounds on the fractional chromatic number. A function $\gamma : V(G) \to \mathbb{R}_0^+$ is a *fractional clique* in a graph G if $\sum_{v \in I} \gamma(v) \le 1$ for every independent set I in G. The *f-weight* of γ is defined as $\sum_{v \in V(G)} \gamma(v) f(v)$. It is a well-known consequence of linear programming duality that the f-fractional chromatic number is equal to the maximum f-weight of a fractional clique [17].

The *2-universal 2-dir graph with m rows* is the horizontally trivial triangle-free 2-dir graph with rows $\ell_1, \dots, \ell_m$ in order according to the y-coordinate and with m vertical lines, where the first vertical line contains a segment intersecting all rows and for $i \in \{2, \dots, m\}$, the i-th vertical line contains two intersecting vertical segments, one of them intersecting $\ell_1, \dots, \ell_{i-1}$ and the other one intersecting $\ell_i, \dots, \ell_m$.

Lemma 5. *Let $m \ge 1$ be an integer, let $\varepsilon > 0$ be a real number, let G be the 2-universal 2-dir graph with m rows, and let $f : V(G) \to \mathbb{R}^+$ be defined by setting $f(v) = 1$ for each vertical segment v and $f(v) = 2 - \varepsilon$ for each horizontal segment v. Then G has f-fractional chromatic number at least $4 - \frac{1}{m} - \varepsilon$.*

Proof. Let us define $\gamma(v) = \frac{1}{m}$ for every $v \in V(G)$. We claim that γ is a fractional clique. Indeed, consider any independent set I. If I does not contain any horizontal segment, then I contains at most one vertical segment from each vertical line, and thus $|I| \le m$. Otherwise, let j_1 be the minimum index such that I contains the horizontal segment from row ℓ_{j_1} and let j_2 be the maximum such such index; then I contains at most $j_2 - j_1 + 1$ horizontal segments. Moreover, I cannot contain the vertical segment from the first vertical line or from the i-th vertical line for any $i \in \{j_1 + 1, \dots, j_2\}$, and thus $|I| \le (j_2 - j_1 + 1) + m - (j_2 - j_1 + 1) = m$. Therefore, $\sum_{v \in I} \gamma(v) \le \frac{1}{m}|I| \le 1$.

The f-weight of γ is $\frac{1}{m}(m \cdot (2 - \varepsilon) + 2m - 1) = 4 - \frac{1}{m} - \varepsilon$, establishing the desired lower bound on the f-fractional chromatic number of G. □

Since the 2-universal 2-dir graph is horizontally trivial, we can combine Lemmas 4 and 5 to obtain the lower bound from Theorem 1.

Corollary 6. *For all integers $r, m \geq 1$, $\chi_f(\mathcal{D}_{rm,2}) \geq 4 - \frac{1}{m} - \frac{1}{2^{r-1}}$.*

References

[1] E. Asplund and B. Grünbaum. On a colouring problem. *Math. Scand.*, 8:181–188, 1960.

[2] James Perkins Burling. On coloring problems of families of prototypes (phd thesis). *University of Colorado, Boulder*, 1965.

[3] Marco Caoduro, Jana Cslovjecsek, Michał Pilipczuk, and Karol Węgrzycki. Independence number of intersection graphs of axis-parallel segments. *arXiv*, 2205.15189, 2022.

[4] Jérémie Chalopin and Daniel Gonçalves. Every planar graph is the intersection graph of segments in the plane. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 631–638, 2009.

[5] M. Chudnovsky, N. Robertson, P. Seymour, and R. Thomas. The strong perfect graph theorem. *Annals of Mathematics*, 164:51–229, 2006.

[6] James Davies and Rose McCarty. Circle graphs are quadratically χ-bounded. *Bulletin of the London Mathematical Society*, 53(3), 2021.

[7] Hubert de Fraysseix, Patrice Ossona de Mendez, and Janos Pach. Representation of planar graphs by segments. *Intuitive Geometry*, 63:109–117, 1991.

[8] P. Erdős. Graph theory and probability. *Canad. J. Math*, 11:34–38, 1959.

[9] Daniel Gonçalves. 3-colorable planar graphs have an intersection segment representation using 3 slopes. In *Graph-Theoretic Concepts in Computer Science: 45th International Workshop, WG 2019, Vall de Núria, Spain, June 19–21, 2019, Revised Papers 45*, pages 351–363. Springer, 2019.

[10] András Gyárfás. On the chromatic number of multiple interval graphs and overlap graphs. *Discrete mathematics*, 55(2):161–166, 1985.

[11] András Gyárfás. Problems from the world surrounding perfect graphs. *Applicationes Mathematicae*, 19(3-4):413–441, 1987.

[12] I Ben-Arroyo Hartman, Ilan Newman, and Ran Ziv. On grid intersection graphs. *Discrete Mathematics*, 87(1):41–52, 1991.

[13] Seog-Jin Kim, Alexandr Kostochka, and Kittikorn Nakprasit. On the chromatic number of intersection graphs of convex sets in the plane. *the electronic journal of combinatorics*, 11(1):R52, 2004.

[14] Jan Kratochvíl and Jiří Matoušek. Intersection graphs of segments. *Journal of Combinatorial Theory, Series B*, 62(2):289–315, 1994.

[15] László Lovász. Kneser's conjecture, chromatic number, and homotopy. *Journal of Combinatorial Theory, Series A*, 25:319–324, 1978.

[16] Arkadiusz Pawlik, Jakub Kozik, Tomasz Krawczyk, Michał Lasoń, Piotr Micek, William T Trotter, and Bartosz Walczak. Triangle-free intersection graphs of line segments with large chromatic number. *Journal of Combinatorial Theory, Series B*, 105:6–10, 2014.

[17] Edward R. Scheinerman and Daniel H. Ullman. *Fractional graph theory.* Dover Publications Inc., Mineola, NY, 2011.

[18] Andrew Suk. Coloring intersection graphs of x-monotone curves in the plane. *Combinatorica*, 34(4):487–505, 2014.

[19] D. West. Open problems. *SIAM J. Discrete Math. Newslett.*, 2:10–12, 1991.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Big Ramsey Degrees in the Metric Setting

(Extended abstract)

Tristan Bice* Noé de Rancourt† Matěj Konečný‡ Jan Hubička‡

Abstract

Oscillation stability is an important concept in Banach space theory which happens to be closely connected to discrete Ramsey theory. For example, Gowers proved oscillation stability for the Banach space c_0 using his now famous Ramsey theorem for FIN_k as the key ingredient. We develop the theory behind this connection and introduce the notion of compact big Ramsey degrees, extending the theory of (discrete) big Ramsey degrees. We prove existence of compact big Ramsey degrees for the Banach space ℓ_∞ and the Urysohn sphere, with an explicit characterization in the case of ℓ_∞.

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-019

1 Introduction

A *discrete structure* is a structure in a relational language without unary predicates (e.g. orders, graphs, hypergraphs, etc.). Given two discrete structures X and Y in the same language we will denote by $\binom{X}{Y}$ the set of all embeddings $Y \to X$. Given a discrete structure X and a finite substructure A of X we say that *A has a finite big Ramsey degree* in X if there exists $t \geqslant 1$ such that every finite colouring of $\binom{X}{A}$ attains at most t colors

*Institute of Mathematics, Czech Academy of Sciences, Žitná 25, Prague 1, Czech Republic. E-mail: tristan.bice@gmail.com. Supported by GAČR project 22-07833K and RVO: 67985840.

†Univ. Lille, CNRS, UMR 8524 – Laboratoire Paul Painlevé, F-59000 Lille, France. E-mail: nderancour@univ-lille.fr. Supported by the Labex CEMPI (ANR-11-LABX-0007-01).

‡Department of Applied Mathematics (KAM), Charles University, Prague, Czech Republic. E-mail: {matej|hubicka}@kam.mff.cuni.cz. Supported by the project 21-10775S of the Czech Science Foundation (GAČR) and a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 810115).

on $\binom{f[X]}{A}$ for some well-chosen $f \in \binom{X}{X}$. In this case, the *big Ramsey degree* of A in X is the least such t. We say that X *has finite big Ramsey degrees* if every finite substructure of X has a finite big Ramsey degree in X.

The infinite Ramsey theorem says that all big Ramsey degrees in $(\mathbb{N}, <)$ are equal to 1. However, this is usually not the case. By Hjorth [10], no infinite homogeneous structure has all big Ramsey degrees equal to 1. However, they can still be finite. The task of identifying such structures was initiated by an unpublished result by Laver who proved finiteness of big Ramsey degrees of $(\mathbb{Q}, <)$, followed by their exact computation by Devlin [5]. Since then there has been a lot of progress [1, 2, 3, 4, 6, 8, 11, 13, 15, 16, 21]. See also Dobrinen's survey [7].

Recall that c_0 is the Banach space of all sequences $x\colon \mathbb{N} \to \mathbb{R}$ tending to zero at infinity, with the norm $||x||_\infty := \sup_{n\in\mathbb{N}} |x(n)|$, and that the *unit sphere* of a Banach space X is $S_X := \{x \in X \mid \|x\| = 1\}$. Also recall that the *Urysohn sphere* $\mathbb{S}$ is the unique complete and separable metric space of diameter 1 containing copies of all finite metric spaces of diameter at most 1 which is *homogeneous*, that is, every isometry between finite subsets of $\mathbb{S}$ extends to an onto isometry of $\mathbb{S}$. By a *compactum*, we mean a compact metric space.

A discrete structure X is *indivisible* if the big Ramsey degree of a vertex in X is equal to 1. Mostly motivated by the *distortion problem* from Banach space theory, the two first indivisibility-like results for metric structures (called *oscillation stability results* by Banach space theorists) have been proved by Gowers [9], and Nguyen Van Thé and Sauer [17]. Here, Lipschitz maps can be seen as continuous colourings, and compactness is the right metric analogue of finiteness.

Theorem 1.1 (Gowers [9]). *Let K be a compactum and $\chi\colon S_{c_0} \to K$ be a Lipschitz map. For every $\varepsilon > 0$, there exists a linear isometric copy X of c_0 in c_0 such that* $\operatorname{diam}(\chi(S_X)) \leqslant \varepsilon$.

Theorem 1.2 (Nguyen Van Thé–Sauer [17]). *Let K be a compactum and $\chi\colon \mathbb{S} \to K$ be a Lipschitz map. For every $\varepsilon > 0$, there exists an isometric copy X of $\mathbb{S}$ in $\mathbb{S}$ such that* $\operatorname{diam}(\chi(X)) \leqslant \varepsilon$.

The proof of Theorem 1.1 is based on discrete approximations and was the reason why Gowers proved his now well-known FIN_k theorem, which is the main ingredient in the proof. The proof of Theorem 1.2 is also using discrete approximations (following a combinatorial strategy which was proposed earlier by Nguyen Van Thé and Lopez-Abad [14]) combined with indivisibility results for metric spaces with finitely many distances.

The similarity between those results and indivisibility results makes it is natural to ask if a suitable version of big Ramsey degrees could be defined for metric structures. This was addressed by Kechris, Pestov, and Todorcevic [12, § 11(F)] who suggested a definition which, however, is rather restrictive and fails to capture most interesting structures beyond the discrete ones. Our goal is to provide a more general notion, and to demonstrate its suitability on examples such as the Banach space ℓ_∞ and the Urysohn sphere. Our work on the Urysohn sphere builds on results on big Ramsey degrees of homogeneous structures with forbidden cycles announced at Eurocomb 2021 [3].

Our motivations are twofold. First, Zucker [20] extended the KPT correspondence [12] to big Ramsey degrees, giving a correspondence between them and some dynamical invariants of automorphism groups. Our extension to the metric setting could allow us to study the same dynamical invariants for the automorphism groups of metric structures; no tool is currently available for studying those. Second, our methods could lead to a systematical study of the distortion phenomenon in Banach space theory, closely related to oscillation stability and not yet well understood. For instance, Odell and Schlumprecht's solution to the distortion problem [18] shows that the analogue of Theorem 1.1 fails for the separable Hilbert space. Metric big Ramsey degrees could help to express a quantitative and optimal version of their result.

2 Compact big Ramsey degrees

We first review, in a more general setting, some results on discrete big Ramsey degrees to motivate our definitions in the metric case. Let M be a monoid acting by injections on a set X. The action $M \curvearrowright X$ *has a finite big Ramsey degree* if there exists $t \geqslant 1$ such that every colouring of X with finitely many colours takes at most t values on a set of the form $p \cdot X$, $p \in M$. In this case, the *big Ramsey degree* of the action is the least such t. Observe that if Y is a discrete structure and $A \subseteq Y$ a finite substructure, then taking $M := \binom{Y}{Y}$ and $X := \binom{Y}{A}$ and considering the action by left-composition, we recover the classical notion of the big Ramsey degree of A in Y. For $k \geqslant 1$, denote by $[k]$ the set $\{1, \dots, k\}$.

Definition 2.1. Fix $M \curvearrowright X$ as above, and $k \geqslant 1$. A colouring $\chi \colon X \to [k]$ is:

- *persistent* if for every $p \in M$, $\chi(p \cdot X) = [k]$;
- *universal* if for every $l \geqslant 1$, every colouring $\psi \colon X \to [l]$ and every $p \in M$, there exists $q \in M$ and $f \colon [k] \to [l]$ such that $\psi \restriction_{pq \cdot X} = f \circ \chi \restriction_{pq \cdot X}$;
- a *big Ramsey colouring* (or a *canonical partition*, following [13]) if it is both persistent and universal.

The proof of the following fact is elementary.

Proposition 2.2. *Suppose that the action $M \curvearrowright X$ has a finite big Ramsey degree. Then it admits a big Ramsey colouring. Moreover, the number of colours of such a colouring is always equal to the big Ramsey degree of the action.*

In the metric metric setting, one has a monoid M acting by (non-necessarily onto) isometries on a complete metric space X. Inspired by Theorems 1.1 and 1.2, we define a *colouring* of X as a 1-Lipschitz map $X \to K$, where K is a compactum (the Lipschitz constant 1 is here to ensure some rigidity). We will also allow some ε-approximation in our results. For K, L compacta we put $K \leqslant L$ if there exists a 1-Lispchitz surjection $L \to K$. This quasiordering is meant to "replace" the order on $\mathbb{N}$. It is a classical fact that $K \leqslant L$ and $L \leqslant K$ if and only if K and L are isometric. If $\chi, \psi \colon X \to K$ are maps, we will denote by $d_\infty(f, g)$ the supremum distance between f and g.

Definition 2.3. Say that a compactum K is:

- *universal* with respect to the action $M \curvearrowright X$ if for every compactum L, every colouring $\psi \colon X \to L$, and every $\varepsilon > 0$, there exists $q \in M$, a colouring $\chi \colon X \to K$, and a 1-Lipschitz map $f \colon K \to L$ such that $d_\infty(\psi \restriction_{q \cdot X}, f \circ \chi \restriction_{q \cdot X}) \leqslant \varepsilon$;
- the *big Ramsey degree of the action* $M \curvearrowright X$ if K is a $\leqslant$-least universal compactum.

Say that the action $M \curvearrowright X$ *has a compact big Ramsey degree* if it admits a big Ramsey degree in the above sense.

The big Ramsey degree of an action, if it exists, is unique, up to isometry.

Definition 2.4. Say that a colouring $\chi \colon X \to K$ is:

- *persistent* if for every $p \in M$, $\chi(p \cdot X)$ is dense in K;
- *universal* if for every compactum L, every colouring $\psi \colon X \to L$, every $p \in M$ and every $\varepsilon > 0$, there exists $q \in M$ and a 1-Lipschitz map $f \colon K \to L$ such that $d_\infty(\psi \restriction_{pq \cdot X}, f \circ \chi \restriction_{pq \cdot X}) \leqslant \varepsilon$;
- a *big Ramsey colouring* if it is both persistent and universal.

Proposition 2.5. *Suppose that $\chi \colon X \to K$ is a big Ramsey colouring for the action $M \curvearrowright X$. Then K is the big Ramsey degree of this action.*

Proposition 2.6. *Consider the following statements:*

(1) the action $M \curvearrowright X$ admits a universal compactum;

(2) the action $M \curvearrowright X$ has a compact big Ramsey degree;

(3) the action $M \curvearrowright X$ admits a universal colouring;

(4) the action $M \curvearrowright X$ admits a big Ramsey colouring.

Then the following implications hold: (4) $\implies$ (3) $\implies$ (2) $\implies$ (1).

While the analogues of the above implications are equivalent in the discrete setting, we do not know whether any of the reverse implications holds in the metric setting. The most relevant notion seems to be the existence of a big Ramsey colouring as, in the discrete setting, it is the closest to Zucker's condition for getting interesting dynamical consequences [20]. Also, in all metric examples for which we have been able to prove the existence of a universal compactum, we could also prove the existence of a big Ramsey colouring.

We end this section mentioning that, by endowing discrete structures with the metric where any two distinct points are at distance 1, we can "embed" the classical discrete setting for big Ramsey degrees in our metric setting, thus making the discrete setting a particular case of the metric setting.

3 Banach spaces

In this section we study big Ramsey degrees of the spaces ℓ_p and c_0. Instead of colouring (linear isometric) embeddings of finite-dimensional subspaces into the whole space, we will equivalently colour finite tuples of elements of its unit sphere, which makes the presentation easier. Given a Banach space X and $d \geqslant 1$, the set $(S_X)^d$ will be endowed with the supremum distance. We denote by $\mathrm{Emb}(X)$ the monoid of all linear isometric embeddings of X into itself.

For $1 \leqslant p < \infty$ and a sequence $x\colon \mathbb{N} \to \mathbb{R}$, we let $\|x\|_p := \left(\sum_{n\in\mathbb{N}} |x(n)|^p\right)^{\frac{1}{p}}$, and $\|x\|_\infty := \sup_{n\in\mathbb{N}} |x(n)|$; and for $1 \leqslant p \leqslant \infty$, we denote by ℓ_p the Banach space of all sequences $x\colon \mathbb{N} \to \mathbb{R}$ such that $\|x\|_p < \infty$, endowed with the norm $\|\cdot\|_p$. The space c_0 is a particular subspace of ℓ_∞.

Gowers' theorem 1.1 is equivalent to saying that the action $\mathrm{Emb}(c_0) \curvearrowright S_{c_0}$ admits a big Ramsey degree which is a singleton. However, the situation is different in higher arities. To see this, colour a pair $(x, y) \in S_{c_0}^2$ of disjointly supported vectors by the number of times their supports intertwine. Then every block-subspace of c_0 meets an infinite number of colours. While this "colouring" is neither Lipschitz nor compactum-valued, the idea can be developed to prove the following result.

Theorem 3.1. *Let $d \geqslant 2$. Then the action $\mathrm{Emb}(c_0) \curvearrowright (S_{c_0})^d$ does not admit a universal compactum.*

As mentioned in the introduction, Odell and Schlumprecht [18] proved that the separable Hilbert space ℓ_2 does not satisfy an analogue of Theorem 1.1. In fact, their paper immediately implies a stronger conclusion.

Theorem 3.2. *Let $d \geqslant 1$ and $1 \leqslant p < \infty$. Then the action $\mathrm{Emb}(\ell_p) \curvearrowright (S_{\ell_p})^d$ does not admit a universal compactum.*

This means that for an optimal version of the Odell–Schlumprecht theorem (if it exists), a theory of noncompact big Ramsey degrees would be needed; we believe that such a theory can be developed, and keep it in mind for a future project.

We now turn to ℓ_∞. In the rest of this section, we fix $d \geqslant 1$ and consider the action $\mathrm{Emb}(\ell_\infty) \curvearrowright (S_{\ell_\infty})^d$. Classical arguments from Banach space theory show that, even when $d = 1$, this action does not admit a universal compactum. However, the proof involves a diagonal argument based on the Axiom of Choice. In such cases, imposing a definability restriction on colourings often allows one to get positive results (see e.g. [19]). The right topology is the *weak-* topology* (here, we refer to the one we get when seeing ℓ_∞ as the dual of ℓ_1). We can then define the notions of a *definable big Ramsey degree* and a *definable big Ramsey colouring* for the above action by considering, in Definitions 2.3 and 2.4, only colourings that are Borel, or even Suslin–measurable, when $(S_{\ell_\infty})^d$ is endowed with the d-th power of the weak-* topology. All results stated in Section 2 remain valid for these definable notions, and it turns out that we can prove the existence of a definable big Ramsey colouring for the action $\mathrm{Emb}(\ell_\infty) \curvearrowright (S_{\ell_\infty})^d$. In order to state our result, preliminary definitions are needed.

Put $\mathrm{B}_d \coloneqq [-1,1]^d$, endowed with the supremum metric. The entries of a tuple $x \in (S_{\ell_\infty})^d$ will be denoted by $x_1, \dots, x_d$. We use a functional notation for elements of ℓ_∞, so that for $n \in \mathbb{N}$ and $i \in [d]$, the n-th entry of the vector x_i will be denoted by $x_i(n)$. We can finally let, for each $n \in \mathbb{N}$, $x(n)$ be the d-tuple $(x_1(n), \dots, x_d(n))$; it is an element of B_d. In this way, elements of $(S_{\ell_\infty})^d$ can be seen as maps $\mathbb{N} \to \mathrm{B}_d$. If X is a metric space, denote by $\mathcal{K}(X)$ the set of all nonempty compact subsets of X, and endow it with the *Hausdorff metric* d_H defined by $d_H(K,L) \coloneqq \max(\sup_{x\in K} d(x,L), \sup_{y \in L} d(y,K))$. Denote by $\mathcal{SCK}(\mathrm{B}_d)$ the set of all nonempty symmetric, convex and compact subsets of B_d, and see it as a metric subspace of $\mathcal{K}(\mathrm{B}_d)$. If $A \subseteq \mathrm{B}_d$, denote by $\mathrm{sc}(A)$ the symmetric convex hull of the set A.

Definition 3.3. A *d-pumpkin* is a compact subset $\mathcal{P} \subseteq \mathcal{SCK}(\mathrm{B}_d)$ such that $\{0\} \in \mathcal{P}$, there exists $C \in \mathcal{P}$ such that for all $i \in [d]$, $\mathrm{proj}_i(C) = [-1,1]$, and the inclusion induces a dense linear order on $\mathcal{P}$. We denote by Pum_d the set of all d-pumpkins, seen as a subset of $\mathcal{K}(\mathcal{SCK}(\mathrm{B}_d))$.

A d-pumpkin can be seen as a continuously growing symmetric compact convex subset of B_d, starting at $\{0\}$ and such that the final step of the evolution touches all faces of the cube B_d. It can be shown that the metric space Pum_d is compact.

Definition 3.4. For $x \in (S_{\ell_\infty})^d$, let:

$$\mathrm{PP}_d(x) \coloneqq \Big\{ \mathrm{sc}\{x(0), \dots, x(n-1), tx(n)\} \,\Big|\, n \in \mathbb{N}, t \in [0,1] \Big\} \cup \Big\{ \overline{\mathrm{sc}\{x(n) \mid n \in \mathbb{N}\}} \Big\}.$$

This defines a definable colouring $\mathrm{PP}_d \coloneqq (S_{\ell_\infty})^d \to \mathrm{Pum}_d$.

Intuitively, sets $\mathrm{sc}\{x(0), \dots, x(n-1)\}$, $n \in \mathbb{N}$, must be steps of the evolution of the pumpkin $\mathrm{PP}_d(x)$, and the set $\overline{\mathrm{sc}\{x(n) \mid n \in \mathbb{N}\}}$ must be its final step. Between those steps, we "fill in the holes" in an affine way.

Theorem 3.5. PP_d *is a definable big Ramsey colouring of the action* $\mathrm{Emb}(\ell_\infty) \curvearrowright (S_{\ell_\infty})^d$. *In particular,* Pum_d *is the definable big Ramsey degree of this action.*

It is easy to see that Pum_1 is a singleton. Thus, as a corollary of Theorem 3.5, we get the following oscillation stability result for ℓ_∞, analogous to Theorem 1.1.

Corollary 3.6. *Let K be a compactum and $\chi\colon S_{\ell_\infty} \to K$ be a Lipschitz map that is also Borel (or Suslin-measurable) for the weak-* topology. Then for every $\varepsilon > 0$, there exists a linear isometric copy X of ℓ_∞ in itself such that* $\mathrm{diam}(\chi(S_X)) \leqslant \varepsilon$.

The proof of Theorem 3.5 is based on the Carlson–Simpson theorem. The natural presentation of S_{ℓ_∞} as a set of infinite words over the alphabet $[-1,1]$ makes its use particularly simple. Another ingredient in the proof is an analysis of the form of linear isometric copies of ℓ_∞ in itself, based on elementary Banach space theoretic tools.

4 The Urysohn sphere

Recall that $\mathbb{S}$ is the Urysohn sphere. As for Banach spaces, we will consider colourings of tuples from $\mathbb{S}$ rather than embeddings of finite substructures. For each $d \geqslant 1$, endow the d-th power $\mathbb{S}^d$ with the supremum metric. Denote by $\mathrm{Emb}(\mathbb{S})$ the monoid of all (non-necessarily surjective) isometries of $\mathbb{S}$ into itself.

Theorem 4.1. *For every $d \geqslant 1$, the action $\mathrm{Emb}(\mathbb{S}) \curvearrowright \mathbb{S}^d$ admits a big Ramsey colouring.*

Our proof method is based on ideas developed in [3] for proving finiteness of the big Ramsey degrees of discrete versions of the Urysohn sphere. We don't work directly on $\mathbb{S}$ itself but on a metric space $\mathbb{T}$ that is bi-embeddable with it. This metric space is a well-enough behaved space of sequences, allowing us the use of the Carlson–Simpson theorem. Our proof allows us to recover the fact that the big Ramsey degree of the action $\mathrm{Emb}(\mathbb{S}) \curvearrowright \mathbb{S}$ is a singleton, thus giving a new and short proof of Theorem 1.2, based on very different tools than the original proof. However, as soon as $d > 1$, part of our proof relies on a non-constructive argument, and we are currently not able to characterize the big Ramsey degree completely. We are only able to give an upper bound of the big Ramsey degree in the sense of the quasiordering $\leqslant$, as a quotient of $\mathbb{T}^d$ by an action of the monoid of rigid surjections $\mathbb{N} \to \mathbb{N}$.

Acknowledgements. The authors greatly benefited from useful discussions with A. Bartoš and S. Todorcevic in writing the present paper.

References

[1] Martin Balko, David Chodounský, Natasha Dobrinen, Jan Hubička, Matěj Konečný, Lluis Vena, and Andy Zucker. Exact big Ramsey degrees via coding trees. arXiv:2110.08409, 2021.

[2] Martin Balko, David Chodounský, Natasha Dobrinen, Jan Hubička, Matěj Konečný, Lluis Vena, and Andy Zucker. Characterisation of the big Ramsey degrees of the generic partial order. arXiv:2303.10088, 2023.

[3] Martin Balko, David Chodounský, Jan Hubička, Matěj Konečný, Jaroslav Nešetřil, and Lluís Vena. Big Ramsey degrees and forbidden cycles. In Jaroslav Nešetřil, Guillem Perarnau, Juanjo Rué, and Oriol Serra, editors, *Extended Abstracts EuroComb 2021*, pages 436–441, Cham, 2021. Springer International Publishing.

[4] Martin Balko, David Chodounský, Jan Hubička, Matěj Konečný, and Lluis Vena. Big Ramsey degrees of 3-uniform hypergraphs are finite. *Combinatorica*, 42(2):659–672, 2022.

[5] Denis Campau Devlin. *Some partition theorems and ultrafilters on ω*. ProQuest LLC, Ann Arbor, MI, 1980. Thesis (Ph.D.)–Dartmouth College.

[6] Natasha Dobrinen. The Ramsey theory of the universal homogeneous triangle-free graph. *J. Math. Log.*, 20(2):2050012, 75, 2020.

[7] Natasha Dobrinen. Ramsey theory of homogeneous structures: current trends and open problems, 2021. To appear in the Proceedings of the International Congress of Mathematicians—2022.

[8] Natasha Dobrinen. The Ramsey theory of Henson graphs. *J. Math. Log.*, 23(1):2250018, 2023.

[9] W. T. Gowers. Lipschitz functions on classical spaces. *European J. Combin.*, 13(3):141–151, 1992.

[10] Greg Hjorth. An oscillation theorem for groups of isometries. *Geom. Funct. Anal.*, 18(2):489–521, 2008.

[11] Jan Hubička. Big Ramsey degrees using parameter spaces. arXiv:2009.00967, 2023.

[12] A. S. Kechris, V. G. Pestov, and S. Todorcevic. Fraïssé limits, Ramsey theory, and topological dynamics of automorphism groups. *Geom. Funct. Anal.*, 15(1):106–189, 2005.

[13] C. Laflamme, N. W. Sauer, and V. Vuksanovic. Canonical partitions of universal structures. *Combinatorica*, 26(2):183–205, 2006.

[14] Jordi Lopez-Abad and Lionel Nguyen Van Thé. The oscillation stability problem for the Urysohn sphere: A combinatorial approach. *Topology and its Applications*, 155(14):1516–1530, 2008.

[15] Dragan Mašulović. Finite big Ramsey degrees in universal structures. *Journal of Combinatorial Theory, Series A*, 170:105137, 2020.

[16] Lionel Nguyen Van Thé. Big Ramsey degrees and divisibility in classes of ultrametric spaces. *Canadian Mathematical Bulletin*, 51(3):413–423, 2008.

[17] Lionel Nguyen Van Thé and Norbert W. Sauer. The Urysohn sphere is oscillation stable. *Geom. Funct. Anal.*, 19(2):536–557, 2009.

[18] Edward Odell and Thomas Schlumprecht. The distortion problem. *Acta Math.*, 173(2):259–281, 1994.

[19] Stevo Todorcevic. *Introduction to Ramsey spaces*, volume 174 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2010.

[20] Andy Zucker. Big Ramsey degrees and topological dynamics. *Groups Geom. Dyn.*, 13(1):235–276, 2019.

[21] Andy Zucker. On big Ramsey degrees for binary free amalgamation classes. *Advances in Mathematics*, 408:108585, 2022.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

On a recolouring version of Hadwiger's conjecture

(Extended abstract)

Marthe Bonamy*† Marc Heinrich *‡ Clément Legrand-Duchesne *† Jonathan Narboni *§

Abstract

We prove that for any $\varepsilon > 0$, for any large enough t, there is a graph G that admits no K_t-minor but admits a $(\frac{3}{2} - \varepsilon)t$-colouring that is "frozen" with respect to Kempe changes, i.e. any two colour classes induce a connected component. This disproves three conjectures of Las Vergnas and Meyniel from 1981.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-020`

1 Introduction

In an attempt to prove the Four Colour Theorem in 1879, Kempe [7] introduced an elementary operation on the colourings[1] of a graph that became known as a Kempe change. Given a k-colouring α of a graph G, a *Kempe chain* is a maximal bichromatic component[2]. A *Kempe change* in α corresponds to swapping the two colours of a Kempe chain so as to obtain another k-colouring. Two k-colourings are *Kempe equivalent* if one can be obtained from the other through a series of Kempe changes.

*The authors are supported by ANR project GrR (ANR-18-CE40-0032)
†CNRS, LaBRI, Université de Bordeaux, Bordeaux, France.
‡University of Leeds, United Kingdom.
§Theoretical Computer Science Department, Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland.
[1]Throughout this paper, all colourings are proper, i.e. no two vertices with the same colour are adjacent.
[2]If a vertex of G is coloured 1 and has no neighbour coloured 2 in α, then it forms a Kempe chain of size 1.

The study of Kempe changes has a vast history, see e.g. [13] for a comprehensive overview or [3] for a recent result on general graphs. We refer the curious reader to the relevant chapter of a 2013 survey by Cereceda [19]. Kempe equivalence falls within the wider setting of combinatorial reconfiguration, which [19] is also an excellent introduction to. Perhaps surprisingly, Kempe equivalence has direct applications in approximate counting and applications in statistical physics (see e.g. [15, 14] for nice overviews). Closer to graph theory, Kempe equivalence can be studied with a goal of obtaining a random colouring by applying random walks and rapidly mixing Markov chains, see e.g. [20].

Kempe changes were introduced as a mere tool, and are decisive in the proof of Vizing's edge colouring theorem [21]. However, the equivalence class they define on the set of k-colourings is itself highly interesting. In which cases is there a single equivalence class? In which cases does every equivalence class contain a colouring that uses the minimum number of colours? Vizing conjectured in 1965 [22] that the second scenario should be true in every line graph, no matter the choice of k. Despite partial results [1, 2], this conjecture remains wildly open.

In the setting of planar graphs, Meyniel proved in 1977 [12] that all 5-colourings form a unique Kempe equivalence class. The result was then extended to all K_5-minor-free graphs in 1979 by Las Vergnas and Meyniel [11]. They conjectured the following, which can be seen as a reconfiguration counterpoint to Hadwiger's conjecture, though it neither implies it nor is implied by it.

Conjecture 1.1 (Conjecture A in [11])**.** *For every t, all the t-colourings of a graph with no K_t-minor form a single equivalence class.*

They also proposed a related conjecture that is weaker assuming Hadwiger's conjecture holds.

Conjecture 1.2 (Conjecture A' in [11])**.** *For every t and every graph with no K_t-minor, every equivalence class of t-colourings contains some $(t-1)$-colouring.*

Here, we disprove both Conjectures 1.1 and 1.2, as follows.

Theorem 1.3. *For every $\varepsilon > 0$ and for any large enough t, there is a graph with no K_t-minor, whose $(\frac{3}{2} - \varepsilon)t$-colourings are not all Kempe equivalent.*

In fact, we prove that for every $\varepsilon > 0$ and for any large enough t, there is a graph G that does not admit a K_t-minor but admits a $(\frac{3}{2} - \varepsilon)t$-colouring that is *frozen*; Any pair of colours induce a connected component, so that no Kempe change can modify the colour partition. To obtain Theorem 1.3, we then argue that the graph admits a colouring with a different colour partition. The notion of frozen k-colouring is related to the notion of *quasi-K_p-minor*, introduced in [11]. A graph G admits a K_p-minor if it admits p non-empty, pairwise disjoint and connected *bags* $B_1, \dots, B_p \subset V(G)$ such that for any $i \neq j$, there is an edge between some vertex in B_i and some vertex in B_j. For the notion of quasi-K_p-minor, we drop the restriction that each B_i should induce a connected subgraph of G, and replace it with the condition that for any $i \neq j$, the set $B_i \cup B_j$ induces a connected subgraph of

G. If the graph G admits a frozen p-colouring, then it trivially admits a quasi-K_p-minor[3], while the converse may not be true. If all p-colourings of a graph form a single equivalence class, then either there is no frozen p-colouring or there is a unique p-colouring of the graph up to colour permutation. The latter situation in a graph with no K_p-minor would disprove Hadwiger's conjecture, so Las Vergnas and Meyniel conjectured that there is no frozen p-colouring in that case. Namely, they conjectured the following.

Conjecture 1.4 (Conjecture C in [11]). *For any t, any graph that admits a quasi-K_t-minor admits a K_t-minor.*

Conjecture 1.4 was proved for increasing values of t, and is now known to hold for $t \leq 10$ [5, 16, 10]. As discussed above, we strongly disprove Conjecture 1.4 for large t. It is unclear how large t needs to be for a counter-example.

Theorem 1.5. *For every $\varepsilon > 0$ and for any large enough t, there is a graph G that admits a quasi-K_t-minor but does not admit a $K_{\left(\frac{2}{3}+\varepsilon\right)t}$-minor.*

We later became aware a similar construction already appeared in [4].

Trivially, every graph that admits a quasi-K_{2t}-minor admits a K_t-minor. We leave the following two open questions, noting that $\frac{2}{3} \geq c \geq \frac{1}{2}$ and $c' \geq \frac{3}{2}$.

Question 1.6. What is the infimum c such that for any large enough t, there is a graph G that admits a quasi-K_t-minor but no K_{ct}-minor?

Question 1.7. Is there a constant c' such that for every t, all the $c' \cdot t$-colourings of a graph with no K_t-minor form a single equivalence class?

In the 1980's, [8, 9] and [17] proved independently that a graph with no K_t-minor has degeneracy $O(t\sqrt{\log t})$, since improved only by a constant factor [18, 23, 6]. Since all the k-colorings of d-degenerate graphs are equivalent for $k > d$ [11], this gives the best upper bound known so far for Question 1.7.

2 Construction

Let $n \in \mathbb{N}$ and let $\eta > 0$. We build a random graph G_n on vertex set $\{a_1, \ldots, a_n, b_1, \ldots, b_n\}$: for every $i \neq j$ independently, we select one pair uniformly at random among $\{(a_i, a_j), (a_i, b_j), (b_i, a_j), (b_i, b_j)\}$ and add the three other pairs as edges to the graph G_n.

Note that the sets $\{a_i, b_i\}_{1 \leq i \leq n}$ form a quasi-K_n-minor, as for every $i \neq j$, the set $\{a_i, b_i, a_j, b_j\}$ induces a path on four vertices in G_n, hence is connected.

Our goal is to argue that if n is sufficiently large then with high probability the graph G_n does not admit any $K_{\left(\frac{2}{3}+\eta\right)n}$-minor. This will yield Theorem 1.5. To additionally obtain Theorem 1.3, we need to argue that with high probability, G_n admits an n-colouring with a different colour partition than the natural one, where the colour classes are of the form $\{a_i, b_i\}$. Informally, we can observe that each of $\{a_1, \ldots, a_n\}$ and $\{b_1, \ldots, b_n\}$ induces a

[3] One bag for each colour class.

graph behaving like a graph in $\mathcal{G}_{n,\frac{3}{4}}$ (i.e. each edge exists with probability $\frac{3}{4}$) though the two processes are not independent. This argument indicates that $\chi(G_n) = O(\frac{n}{\log n})$, but we prefer a simpler, more pedestrian approach.

Assume that for some i, j, k, ℓ, none of the edges $a_i b_j$, $a_j b_k$, $a_k b_\ell$ and $a_\ell b_i$ exist. Then the graph G_n admits an n-colouring α where $\alpha(a_p) = \alpha(b_p) = p$ for every $p \notin \{i, j, k, \ell\}$ and $\alpha(a_i) = \alpha(b_j) = i$, $\alpha(a_j) = \alpha(b_k) = j$, $\alpha(a_k) = \alpha(b_\ell) = k$ and $\alpha(a_\ell) = \alpha(b_i) = \ell$ (see Figure 1). Since every quadruple (i, j, k, ℓ) has a positive and constant probability of satisfying this property, G_n contains such a quadruple with overwhelmingly high probability when n is large.

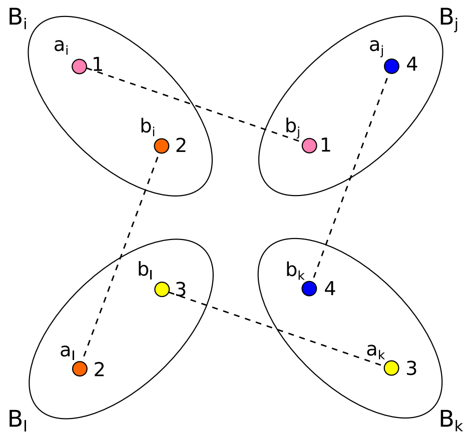


Figure 1: A different n-colouring given an appropriate quadruple.

We are now ready to prove that the probability that G_n admits a $K_{(\frac{2}{3}+\eta)n}$-minor tends to 0 as n grows to infinity. We consider three types of K_p-minors in G, depending on the size of the bags involved. If every bag is of size 1, we say that it is a *simple* K_p-minor – in fact, it is a subgraph. If every bag is of size 2, we say it is a *double* K_p-minor. If every bag is of size at least 3, we say it is a *triple* K_p-minor. We prove three claims, as follows.

Claim 2.1. *For any $\varepsilon > 0$, $\mathbb{P}(G_n$ contains a simple $K_{\varepsilon n}$-minor$) \to 0$ as $n \to \infty$.*

Claim 2.2. *For any $\varepsilon > 0$, $\mathbb{P}(G_n$ contains a double $K_{\varepsilon n}$-minor$) \to 0$ as $n \to \infty$.*

Claim 2.3. *G_n does not contain a triple $K_{\frac{2}{3}n+1}$-minor.*

Claims 2.1, 2.2 and 2.3 are proved in Sections 2.1, 2.2 and 2.3, respectively. If a graph admits a K_p-minor, then in particular it admits a simple K_a-minor, a double K_b-minor and a triple K_c-minor such that $a + b + c \geq p$. Combining Claims 2.1, 2.2 and 2.3, we derive the desired conclusion.

2.1 No large simple minor

Proof of Claim 2.1. Let S be a subset of k vertices of G_n. The probability that S induces a clique in G_n is at most $\left(\frac{3}{4}\right)^{\binom{k}{2}}$. Indeed, if $\{a_i, b_i\} \subseteq S$ for some i, then the probability is 0. Otherwise, $|S \cap \{a_i, b_i\}| \leq 1$ for every i, so we have $G[S] \in \mathcal{G}_{k,\frac{3}{4}}$, i.e. edges exist independently with probability $\frac{3}{4}$. Therefore, the probability that S induces a clique is $\left(\frac{3}{4}\right)^{\binom{k}{2}}$. By union-bound, the probability that some subset on k vertices induces a clique is at most $\binom{2n}{k} \cdot \left(\frac{3}{4}\right)^{\binom{k}{2}}$. For any $\varepsilon > 0$, we note that $\binom{2n}{\varepsilon n} \leq 2^{2n}$. Therefore, the probability that G_n contains a simple $K_{\varepsilon n}$-minor is at most $2^{2n} \cdot \left(\frac{3}{4}\right)^{\binom{\varepsilon n}{2}}$, which tends to 0 as n grows to infinity. ⌟

2.2 No large double minor

Proof of Claim 2.2. Let S' be a subset of k pairwise disjoint pairs of vertices in G_n such that for every i, at most one of $\{a_i, b_i\}$ is involved in S'.

We consider the probability that $G_n/_{S'}$ induces a clique, where $G_n/_{S'}$ is defined as the graph obtained from G_n by considering only vertices involved in some pair of S' and identifying the vertices in each pair.

We consider two distinct pairs $(x_1, y_1), (x_2, y_2)$ of S'. Without loss of generality, $\{x_1, x_2, y_1, y_2\} = \{a_i, a_j, a_k, a_\ell\}$ for some i, j, k, ℓ. The probability that there is an edge between $\{x_1, y_1\}$ and $\{x_2, y_2\}$ is $1 - \left(\frac{1}{4}\right)^4$. In other words, $\mathbb{P}(E((x_1, y_1), (x_2, y_2)) = \emptyset) = \left(\frac{1}{4}\right)^4$ and since at most one of $\{a_i, b_i\}$ is involved in S' for all i, all such events are mutually independent. Therefore, the probability that S' yields a quasi-$K_{|S'|}$-minor is $\left(1 - \left(\frac{1}{4}\right)^4\right)^{\binom{|S'|}{2}}$.

For any $\varepsilon' > 0$, the number of candidates for S' is at most $\binom{2n}{2\varepsilon' n}$ (the number of choices for a ground set of $2\varepsilon' n$ vertices) times $(2\varepsilon' n)!$ (a rough upper bound on the number of ways to pair them). Note that $\binom{2n}{2\varepsilon' n} \cdot (2\varepsilon' n)! \leq (2n)^{2\varepsilon' n}$. We derive that the probability that there is a set S' of size $\varepsilon' n$ such that $G_n/_{S'} = K_{|S'|}$ is at most $(2n)^{2\varepsilon' n} \cdot \left(1 - \left(\frac{1}{4}\right)^4\right)^{\binom{\varepsilon' n}{2}}$, which tends to 0 as n grows large.

Consider a double K_k-minor S of G_n. Note that no pair in S is equal to $\{a_i, b_i\}$ (for any i), as every bag induces a connected subgraph in G_n. We build greedily a maximal subset $S' \subseteq S$ such that S' involves at most one vertex out of every set of type $\{a_i, b_i\}$. Note that $|S'| \geq \frac{|S|}{3}$. By taking $\varepsilon' = \frac{\varepsilon}{3}$ in the above analysis, we obtain that the probability that there is a set S of εn pairs that induces a quasi-$K_{|S|}$-minor tends to 0 as n grows large. ⌟

2.3 No large triple minor

Proof of Claim 2.3. The graph G_n has $2n$ vertices, and a triple K_k-minor involves at least $3k$ vertices. It follows that if G_n contains a triple K_k-minor then $k \leq \frac{2n}{3}$. ⌟

Acknowledgements

The authors thank Vincent Delecroix for helpful discussions, as well as Zixia Song and Raphael Steiner for pointing out additional relevant references, specifically [10, 16] and [4].

References

[1] Armen S. Asratian. A note on transformations of edge colorings of bipartite graphs. *Journal of Combinatorial Theory, Series B*, 99(5):814–818, 2009.

[2] Armen S. Asratian and Carl Johan Casselgren. Solution of Vizing's problem on interchanges for the case of graphs with maximum degree 4 and related results. *Journal of Graph Theory*, 82(4):350–373, 2016.

[3] Marthe Bonamy, Nicolas Bousquet, Carl Feghali, and Matthew Johnson. On a conjecture of Mohar concerning Kempe equivalence of regular graphs. *Journal of Combinatorial Theory, Series B*, 135:179–199, 2019.

[4] Thomas Böhme, Alexandr Kostochka, and Andrew Thomason. Hadwiger numbers and over-dominating colourings. *Discrete Mathematics*, 310(20):2662–2665, 2010. Graph Theory — Dedicated to Carsten Thomassen on his 60th Birthday.

[5] Leif K. Jørgensen. Contractions to K8. *Journal of Graph Theory*, 18(5):431–448, 1994.

[6] Tom Kelly and Luke Postle. A local epsilon version of reed's conjecture. *Journal of Combinatorial Theory, Series B*, 141:181–222, 2020.

[7] Alfred B. Kempe. On the geographical problem of the four colours. *American journal of mathematics*, 2(3):193–200, 1879.

[8] Alexandr V Kostochka. The minimum hadwiger number for graphs with a given mean degree of vertices. *Metody Diskret. Analiz.*, (38):37–58, 1982.

[9] Alexandr V. Kostochka. Lower bound of the hadwiger number of graphs by their average degree. *Combinatorica*, 4(4):307–316, 1984.

[10] Matthias Kriesell. A note on uniquely 10-colorable graphs. *Journal of Graph Theory*, 98(1):24–26, 2021.

[11] Michel Las Vergnas and Henri Meyniel. Kempe classes and the Hadwiger conjecture. *Journal of Combinatorial Theory, Series B*, 31(1):95–104, 1981.

[12] Henry Meyniel. Les 5-colorations d'un graphe planaire forment une classe de commutation unique. *J. Comb. Theory, Ser. B*, 24:251–257, 1978.

[13] Bojan Mohar. Kempe equivalence of colorings. In *Graph Theory in Paris*, pages 287–297. Springer, 2006.

[14] Bojan Mohar and Jesús Salas. A new Kempe invariant and the (non)-ergodicity of the Wang–Swendsen–Kotecký algorithm. *Journal of Physics A: Mathematical and Theoretical*, 42(22):225204, 2009.

[15] Alan D. Sokal. A personal list of unsolved problems concerning lattice gases and antiferromagnetic Potts models. *arXiv preprint cond-mat/0004231*, 2000.

[16] Zi-Xia Song and Robin Thomas. The extremal function for k9 minors. *Journal of Combinatorial Theory, Series B*, 96(2):240–252, 2006.

[17] Andrew Thomason. An extremal function for contractions of graphs. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 95, pages 261–265. Cambridge University Press, 1984.

[18] Andrew Thomason. The extremal function for complete minors. *Journal of Combinatorial Theory, Series B*, 81(2):318–338, 2001.

[19] Jan van den Heuvel. The complexity of change. *Surveys in combinatorics*, 409(2013):127–160, 2013.

[20] Eric Vigoda. Improved bounds for sampling colorings. *Journal of Mathematical Physics*, 41(3):1555–1569, 2000.

[21] Vadim G. Vizing. On an estimate of the chromatic class of a p-graph. *Discret Analiz*, 3:25–30, 1964.

[22] Vadim G. Vizing. Some unsolved problems in graph theory. *Russian Mathematical Surveys*, 23(6):125, 1968.

[23] David R Wood. A note on hadwiger's conjecture. *arXiv preprint arXiv:1304.6510*, 2013.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

The Localization game on locally finite trees

(Extended abstract)

Anthony Bonato* Florian Lehner† Trent G. Marbach‡ JD Nir§

Abstract

We study the Localization game on locally finite graphs and trees, where each vertex has finite degree. As in finite graphs, we prove that any locally finite graph contains a subdivision where one cop can capture the robber. In contrast to the finite case, for n a positive integer, we construct a locally finite tree with localization number n for any choice of n. Such trees contain uncountably many ends, and we show this is necessary by proving that graphs with countably many ends have localization number at most 2. We finish with questions on characterizing the localization number of locally finite trees.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-021`

1 Introduction

Pursuit-evasion games are most commonly studied on finite graphs, but various studies have also considered the infinite case, such as [3, 12, 14, 15]; see also Chapter 7 of [6]. In this extended abstract, we present the first study of the Localization game on locally finite

*Department of Mathematics, Toronto Metropolitan University, Canada. E-mail: abonato@torontomu.ca. Supported by NSERC.
†Department of Mathematics, University of Auckland, New Zealand. E-mail: florian.lehner@auckland.ac.nz.
‡Department of Mathematics, Toronto Metropolitan University, Canada. E-mail: trent.marbach@torontomu.ca.
§Department of Mathematics, Toronto Metropolitan University, Canada. E-mail: jd.nir@torontomu.ca.

graphs, where the degree of each vertex is finite. Unlike in the finite case, the robber may avoid capture in locally finite graphs by moving along an infinite path.

The Localization game was first introduced for one cop by Seager [16, 17], and was subsequently studied in several papers such as [1, 4, 5, 7, 8, 9]. The game consists of two players playing on a graph. One player controls a set of k *cops* and the other controls a single *robber*. The players play over a sequence of discrete time-steps; a *round* of the game is a move by the cops and the subsequent move by the robber. The players move on alternate time-steps, with the robber going first. The robber occupies a vertex of the graph, and when the robber is ready to move during a round, they may move to a neighboring vertex or remain on their current vertex. The cops' move is a placement of cops on a set of vertices. Note that the cops are not limited to moving to neighboring vertices. In each round, the cops occupy a set of vertices $u_1, u_2, \dots, u_k$, and each cop sends out a *cop probe*. Each cop probe returns the distance from u_i to the robber. The cops win if they have a strategy to determine, after a finite number of rounds, the location of the robber, at which time we say that the cops *capture* the robber. We assume the robber is *omniscient*, in the sense that they know the entire strategy for the cops. The robber wins by evading capture indefinitely. For a graph G, the *localization number* of G, written $\zeta(G)$, is the smallest cardinal for which k cops have a winning strategy. As locally finite graphs are countable, $\zeta(G)$ is either a positive integer or the first infinite cardinal, $\aleph_0$.

We present new results on the localization number of locally finite (that is, every vertex has finitely many neighbors) graphs and trees, paying particular attention to whether results persist or change from the finite case. As in finite graphs, we prove that any locally finite graph contains a subdivision where one cop can capture the robber. In contrast to the finite case, we construct a locally finite tree with localization number n for any choice of n, where n is a positive integer or $\aleph_0$. These constructions contain uncountably many "infinite branches" or ends. We show that, as in the finite case, trees with countably many ends have localization number of at most two. We close with open questions about characterizing the localization number of locally finite trees.

All graphs considered are simple, connected, and locally finite. The reader is directed to [2, 11] for additional background on graph theory and infinite graphs.

2 Results

Although determining the localization number for general graphs is **NP**-hard [7], the following theorem of Seager characterizes the Localization game on finite trees. Let T_3 be the tree depicted in Figure 1.

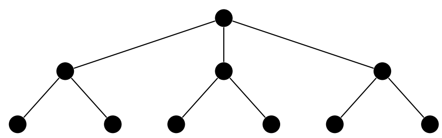


Figure 1: The graph T_3.

Theorem 1 ([17]). *If T is a finite tree, then $\zeta(T) = 1$ if and only if T is T_3-free, and otherwise, $\zeta(T) = 2$.*

The Localization game on locally finite trees has received far less attention. While proving a result for finite graphs, Haselgrave, Johnson, and Koch gave the first theorem extending the Localization game to an infinite tree.

Theorem 2 ([13]). *The infinite Δ-regular tree T_Δ satisfies $\zeta(T_\Delta) \geq \lfloor \frac{\Delta^2}{4} \rfloor$.*

As a consequence of Theorems 1 and 2, we note that locally finite trees offer a richer spectrum of localization numbers than finite trees. In our first contribution, we show that for any choice of n, including $\aleph_0$, there is a locally finite tree with localization number n.

Theorem 3. *If n is a positive integer or $n = \aleph_0$, then there is a locally finite tree T with $\zeta(T) = n$.*

While the full proof of Theorem 3 will be given in the full paper, we sketch it here. Fix $n > 2$ an integer. For locally finite trees T with $\zeta(T) = n$, consider the subdivision of the infinite $n(n-1)$-regular tree, where each edge is subdivided $n-1$ times. A set of n cops can spend $n-1$ rounds determining which subtree contains the robber. This allows them to move n vertices towards the robber, who can only move $n-1$ away, so the cops eventually overtake and capture the robber. The robber can evade $n-1$ cops by playing on an unprobed branch for at least n rounds. This guarantees the robber can choose any fixed distance d to stay from every probe, avoiding capture. For the case of $n = \aleph_0$, note that Theorem 2 allows us to construct a graph on which the robber can evade any finite number of cops.

Theorem 2 tells us the infinite $n(n-1)$-regular tree requires $\Omega(n^4)$ cops, but subdividing reduced the number of required cops to n. This technique was studied in finite graphs, where it is known that every finite graph G has a subdivision G' such that $\zeta(G') = 1$; see [10]. An analogous result holds for locally finite graphs.

Theorem 4. *For every locally finite graph G, there is a subdivision G' of G such that $\zeta(G') = 1$.*

Unlike in the approach given in [10] in the finite case, we subdivide different edges a different number of times. We defer the complete proof to the full paper.

Locally finite trees have infinite paths where the robber may evade capture. Bearing this in mind, we use the setting of ends to formalize our approach to the localization number of locally finite trees. A *ray* is an infinite one-way path. An *end* is an equivalence class of rays with the property that for any finite set of vertices S, each equivalent ray is in the same component of $G - S$. Different ends can be separated by removing finitely many vertices, matching our concept of separate infinite branches. The theory of ends in general locally finite graphs is complex (see [11]); in the context of locally finite trees, however, we may view an end as an infinite branch of the tree with finite subtrees attached to each vertex.

The locally finite trees considered so far all contain the infinite binary tree as a minor, and such graphs have uncountably many ends. Containing the infinite binary tree as a minor is equivalent to having uncountably-many ends [11].

In the full paper, we will prove the following.

Theorem 5. *If T is a locally finite tree with finitely many ends, then $\zeta(T) \leq 2$.*

Perhaps surprisingly, two cops still have a winning strategy in case there are countably many ends.

Theorem 6. *If T is a locally finite tree with countably many ends, then $\zeta(T) \leq 2$.*

For the proof of Theorem 6, we use transfinite induction on a certain ordinal labeling of ends. The base case for the induction uses Theorem 5.

Proof. Given a locally finite rooted tree T and the corresponding tree order where for $u, v \in V(T)$, $u < v$ if and only if u is on the unique path from v to the root, a *recursive pruning* is a labeling of the vertices of T by ordinals where the collection of vertices that receive ordinal α are those which, after removing all vertices with label $\beta < \alpha$, have up-closures that form chains. In other words, after pruning all vertices labeled so far, assign label α to all vertices after the point where any path or ray starting at the root stops branching. We let T_α be the tree resulting from pruning all vertices with label $\beta < \alpha$, and note that the process of recursive pruning ensures T_α is connected for all α.

For more background on recursive prunings, we direct the reader to [11, Chapter 8]. Trees have a recursive pruning if and only if they do not contain a subdivision of the infinite binary tree [11, Proposition 8.5.1]; such trees are the only examples of "infinite branching" where some vertices cannot be labeled. Thus, every rooted tree with countably many ends has a recursive pruning.

Consider an ordinal labeling of the vertices of T by a recursive pruning. Each end of T contains a ray such that the labels of vertices along that ray are weakly decreasing; otherwise, there would be some α for which T_α is not connected. Since decreasing sets of ordinals are finite, each such ray only contains vertices with finitely many labels, and thus, among those labels that occur infinitely often, one must be largest. If ϵ is an end of T, then we call the largest label that occurs infinitely often on the corresponding ray the *end label* of ϵ.

We claim that if an ordinal α is the supremum of the end labels among the (possibly countably many) ends of T, then there are finitely many ends with end label α. After pruning vertices that received label α, the resulting tree $T_{\alpha+1}$ is connected. If $T_{\alpha+1}$ is empty, then T_α contained no vertices of degree greater than two and thus contained at most two ends, each of which had end label α. If $T_{\alpha+1}$ contains finitely many vertices, then as T is locally finite, each vertex was adjacent to finitely many rays where each vertex received label α, and so there are finitely many such ends. Finally, if $T_{\alpha+1}$ contains infinitely many vertices, it must contain an end [11, Proposition 8.2.1], and the vertices on that end must receive a label larger than α, so α was not the supremum of the end labels.

We next show that the supremum α of the end labels is in fact a maximum. Let $\alpha_1, \alpha_2, \ldots$ infinite be a strictly increasing sequence of ordinals such that there are ends $\epsilon_1, \epsilon_2, \ldots$ where ϵ_i has end label α_i. We will show there is at least one end with end label β such that for all i, $\beta > \alpha_i$; therefore, every infinite chain of increasing end labels has a maximal element. By Zorn's lemma, T has an end with maximum end label.

To find an end with end label β, first note that for every end ϵ_i, we can find an infinite ray r_i beginning at some arbitrarily chosen root of T, say v_0, which belongs to the end ϵ_i. As T is locally finite, v_0 has finitely many neighbors. One of these neighbors say v_1, must be contained in r_i for infinitely many i. We repeat this compactness-type argument to find a ray $v_0 v_1 \cdots$ such that each vertex is contained in infinitely many of the r_i. Each v_i must, therefore, receive a label larger than each α_i. Hence, the end containing this ray must have end label $\beta > \alpha_i$ for each i.

We prove that two cops have a winning strategy by transfinite induction on the largest end label in the recursive pruning of T. For the base case, if T contains no ends, or if the largest end label is 1, then given that there are finitely many ends with end label 1, the result follows from Theorem 5.

Assume now that the theorem holds if the largest end label is strictly less than α and let T be a tree with the largest end label α. By implementing a strategy similar to that used to prove Theorem 5, two cops repeatedly restrict the robber's access to ends with label α until the robber is trapped on a subgraph with ends which have label less than α. At this point, the cops have a winning strategy by the induction hypothesis, and the theorem follows. □

3 Further Directions

Given Theorem 6, it is natural to ask if there is a version of Theorem 1 for locally finite trees with countably many ends. Unlike in the finite case, T_3 is not the only obstruction to a locally finite tree having localization number one. One example is the *doubly infinite comb* graph T_1^∞ consisting of a double ray with a leaf attached to each vertex; see Figure 2.

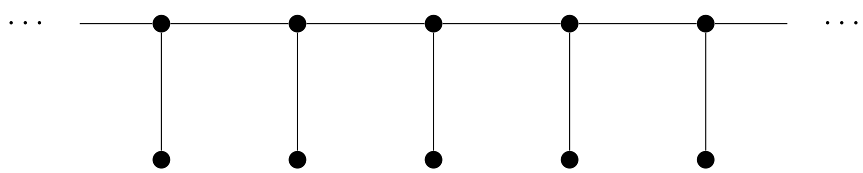


Figure 2: The graph T_1^∞.

The tree T_1^∞ is locally finite, T_3-free tree satisfying $\zeta(T_1^\infty) = 2$. We may show that the tree T_1^∞ is *minimal* in the sense that deleting any edge results in a tree T with $\zeta(T) = 1$. An interesting problem is determining the minimal locally finite trees with countable (or even finite) ends and localization number 2.

References

[1] N.C. Behague, A. Bonato, M.A. Huggan, T.G. Marbach, and B. Pittman, The localization capture time of a graph, *Theoretical Computer Science*, 911:80-91, 2022.

[2] A. Bonato, *An Invitation to Pursuit-Evasion Games and Graph Theory*, American Mathematical Society, Providence, Rhode Island, 2022.

[3] A. Bonato, G. Hahn, and C. Tardif, Large classes of infinite k-cop-win graphs, *Journal of Graph Theory*, 65:334–342, 2010.

[4] A. Bonato, M.A. Huggan, and T. Marbach, The localization number of designs, *Journal of Combinatorial Designs*, 29:175–192, 2021.

[5] A. Bonato and W. Kinnersley, Bounds on the localization number, *Journal of Graph Theory*, 94:1–18, 2020.

[6] A. Bonato and R.J. Nowakowski, *The Game of Cops and Robbers on Graphs*, American Mathematical Society, Providence, Rhode Island, 2011.

[7] B. Bosek, P. Gordinowicz, J. Grytczuk, N. Nisse, J. Sokół, and M. Śleszyńska-Nowak, Localization game on geometric and planar graphs, *Discrete Applied Mathematics*, 251:30–39, 2018.

[8] B. Bosek, P. Gordinowicz, J. Grytczuk, N. Nisse, J. Sokół, and M. Śleszyńska-Nowak, Centroidal localization game, *Electronic Journal of Combinatorics*, 25(4):P4.62, 2018.

[9] A. Brandt, J. Diemunsch, C. Erbes, J. LeGrand, and C. Moffatt, A robber locating strategy for trees, *Discrete Applied Mathematics*, 232:99–106, 2017.

[10] J. Carraher, I. Choi, M. Delcourt, L.H. Erickson, and D.B. West, Locating a robber on a graph via distance queries, *Theoretical Computer Science*, 463:54–61, 2012.

[11] R. Diestel, *Graph Theory*, American Springer Berlin, Heidelberg, Germany, 2017.

[12] G. Hahn, F. Laviolette, N. Sauer, and R.E. Woodrow, On cop-win graphs, *Discrete Mathematics*, 258:27–41, 2002.

[13] J. Haslegrave, R.A. Johnson, and S. Koch, Locating a robber with multiple probes, *Discrete Mathematics*, 341:184–193, 2018.

[14] M.R. Ivan, I. Leader, and M. Walters, Constructible graphs and pursuit, *Theoretical Computer Science*, 930:196–208, 2022.

[15] F. Lehner, Pursuit evasion on infinite graphs, *Theoretical Computer Science*, 655(A):30–40, 2016.

[16] S. Seager, Locating a robber on a graph, *Discrete Mathematics*, 312:3265–3269, 2012.

[17] S. Seager, Locating a backtracking robber on a tree, *Theoretical Computer Science*, **539**:28–37, 2014.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Twin-width and permutations

(Extended abstract)

Édouard Bonnet* Jaroslav Nešetřil† Patrice Ossona de Mendez‡
Sebastian Siebertz§ Stéphan Thomassé*

Abstract

Inspired by a width invariant on permutations defined by Guillemot and Marx, Bonnet, Kim, Thomassé, and Watrigant introduced the twin-width of graphs, which is a parameter describing its structural complexity. This invariant has been further extended to binary structures, in several (basically equivalent) ways. We prove that a class of binary relational structures (that is: edge-colored partially directed graphs) has bounded twin-width if and only if it is a first-order transduction of a proper permutation class. As a by-product, we show that every class with bounded twin-width contains at most $2^{O(n)}$ pairwise non-isomorphic n-vertex graphs.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-022`

1 Introduction

In this paper we consider the graph parameter *twin-width*, defined by Bonnet, Kim, Thomassé and Watrigant [6] as a generalization of an invariant for classes of permutations defined by Guillemot and Marx [9]. Twin-width was recently studied intensively in the context of many structural and algorithmic questions, such as FPT model checking [6], graph enumeration [3], graph coloring [4], and structural properties of matrices and ordered graphs [5].

*Université de Lyon, CNRS, ENS de Lyon, France. E-mails: `edouard.bonnet@ens-lyon.fr`, `stephan.thomasse@ens-lyon.fr`.

†Computer Science Institute of Charles University (IUUK), Praha, Czech Republic. E-mail: `nesetril@iuuk.mff.cuni.cz`.

‡CAMS, EHESS, CNRS UMR 8557, Paris, France. E-mail: `pom@ehess.fr`.

§University of Bremen, Bremen, Germany. E-mail: `siebertz@uni-bremen.de`.

Many well-studied classes of graphs have bounded twin-width: planar graphs, and more generally, any class of graphs excluding a fixed minor, cographs, and more generally, any class of bounded clique-width, etc.

The twin-width of graphs was originally defined using a sequence of 'near-twin' vertex contractions or identifications. Roughly speaking, twin-width measures the accumulated error (recorded via the so-called 'red edges') made by the identifications. To help the reader start forming intuitions, we give a concise definition of the twin-width of a graph.

A *trigraph* is a graph with some edges colored red (while the rest of them are black). A *contraction* (or identification) consists of merging two (non-necessarily adjacent) vertices, say, u, v into a vertex w that is adjacent to a vertex z via a black edge if uz and vz were black edges, or otherwise, via a red edge if at least one of u and v were adjacent to z. The rest of the trigraph does not change. A *contraction sequence* of an n-vertex graph G is a sequence of trigraphs $G = G_n, \dots, G_1$ such that G_i is obtained from G_{i+1} by performing one contraction (observe that G_1 is the 1-vertex graph). A *d-sequence* is a contraction sequence where all the trigraphs have red degree at most d. The *twin-width* of G is then the minimum integer d such that G admits a d-sequence. See Figure 1 for an example of a graph admitting a 2-sequence.

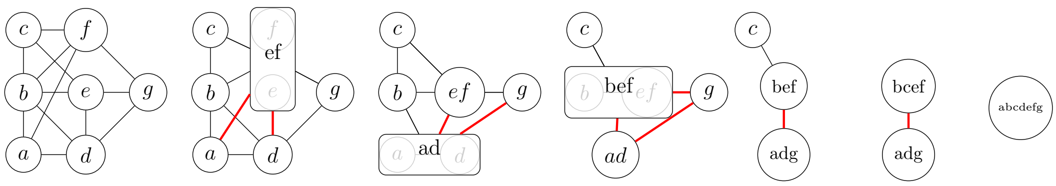


Figure 1: A 2-sequence witnessing that the initial graph has twin-width at most 2.

Our main result deals with relational structures in a way which is consistent (for boundedness) with other definitions used in the literature (see [6], for example).

We show that twin-width can be concisely expressed by special structures, which we call *twin-models*. Twin-models are rooted trees augmented by a set of transversal edges that satisfies two simple properties: minimality and consistency. These properties imply that every twin-model admits a *ranking*, from which we can compute a *width*. The twin-width of a structure then coincides with the optimal width of a ranked twin-model of the structure. While this connection is technical, twin-models provide a simple way to handle classes of binary structures with bounded twin-width. Note that an informal precursor of ranked twin-models appears in [4] in the form of the so-called *ordered union trees* and the realization that the edge set of graphs of twin-width at most d can be partitioned into $O_d(n)$ bicliques where both sides of each biclique are a discrete interval along a unique fixed vertex ordering. The main novelty in the (ranked) twin-models lies in the axiomatization of *legal* sets of transversal edges, which is indispensable to their logical treatment.

This paper is a combination of model-theoretic tools (relational structures, interpretations, transductions), structural graph theory and theory of permutations. Here, by a permutation, we mean a relational structure consisting of two linear orders on the same set (see [1] for a discussion on representations of permutations). Note that this type of representation is particularly adapted to the study of patterns in permutations. The following is the main result of this paper:

Theorem. *A class of binary relational structures has bounded twin-width if and only if it is a first-order transduction of a proper permutation class.*

We recall that a proper permutation class is a set of permutations closed under subpermutations that excludes at least one permutation. Transductions provide a model theoretical tool to encode relational structures (or classes of relational structures) inside other (classes of) relational structures.

The fact that any class of graphs with bounded twin-width is just a transduction of a very simple class (a proper permutation class) is surprising at first glance, and it nicely complements another model theoretic characterization of classes of bounded twin-width: a class of graphs has bounded twin-width if and only if it is the reduct of a dependent class of ordered graphs [5]. It can also be thought of as scaling up the fact that classes of bounded rank-width coincide with transductions of tree orders, and classes of bounded linear rank-width, with transductions of linear orders [8]. On the other hand, twin-models are interesting objects per se and in a way present one of the most permissive forms of width parameters related to trees. Note that for other classes of sparse structures (such as nowhere dense classes or classes with bounded expansion) we do not have such concrete models.

The main result implies that every relational structure on n elements from a class with bounded twin-width can be encoded in a permutation on at most kn elements for some number k. It is then a consequence of [10] that every class of relational structures with bounded twin-width contains at most c^n non-isomorphic structures with n vertices, hence is small (i.e., contains at most $c^n\,n!$ labeled structures with n elements). This extends the main result of [3] while not using the "versatile twin-width" machinery (but only the preservation of bounded twin-width by transductions proved in [6]). This also extends a similar property for proper minor-closed classes of graphs, which can be derived from the boundedness of book thickness, as noticed by McDiarmid (see the concluding remarks of [2]).

The proof of our main result is surprisingly complex and proceeds in several steps, which perhaps add new aspects to the rich spectrum of structures related to twin-width. The basic steps can be outlined as follows (the relevant terminology is formally introduced in the full version of the paper [7]).

We start with a class $\mathscr{C}_0$ of binary relational structures with bounded twin-width. We derive a class $\mathscr{T}$ of twin-models (tree-like representations of the structures using rooted binary trees and transversal binary relations). Replacing the rooted binary trees of the twin-models by binary tree orders, we get a class $\mathscr{F}$ of so-called full twin-models, which we prove has bounded twin-width. This class can be used to retrieve $\mathscr{C}_0$ as a transduction, that is by means of a logical encoding. Using a transduction pairing (generalizing the notion of a bijective encoding) between binary tree orders $\mathscr{O}$ and rooted binary trees ordered by a preorder $\mathscr{Y}^<$ we derive a transduction pairing of the class of full twin-models $\mathscr{F}$ with a class $\mathscr{T}^<$ of ordered twin-models. From the property that the class $\mathscr{G}$ of the Gaifman graphs of the twin-models in $\mathscr{T}$ is degenerate (and has bounded twin-width), we prove a transduction pairing of $\mathscr{T}$ and $\mathscr{G}$, from which we derive a transduction pairing of $\mathscr{T}^<$ and the class $\mathscr{G}^<$ of ordered Gaifman graphs of the ordered twin-models. As a composition of a transduction pairing of $\mathscr{G}^<$ with a class $\mathscr{E}^<$ of ordered binary structures, in which each binary relation induces a pseudoforest and a transduction pairing of $\mathscr{E}^<$ with a class $\mathscr{P}$ of permutations we define a transduction pairing of $\mathscr{G}^<$ and $\mathscr{P}$. As $\mathscr{G}^<$ has bounded twin-width (as it is a transduction of a class with bounded twin-width) we infer that $\mathscr{P}$ avoids a least one pattern. Following the backward transductions, we eventually deduce that $\mathscr{C}_0$ is a transduction of the hereditary closure $\overline{\mathscr{P}}$ of $\mathscr{P}$, which is a proper permutation class.

This proof may be schematically outlined by Figure 2.

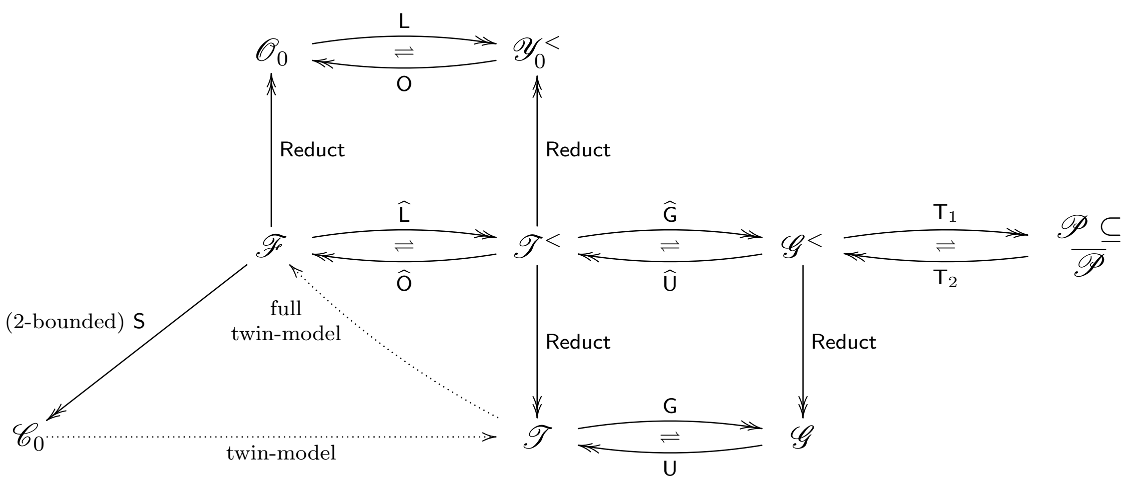


Figure 2: Relations between the classes of structures involved in the proof of the main result.

The full transformation of a graph G into a permutation σ and the inverse transformation (obtained as a transduction) are displayed on Figure 3 on an example.

The full version of this paper is available on arXiv [7].

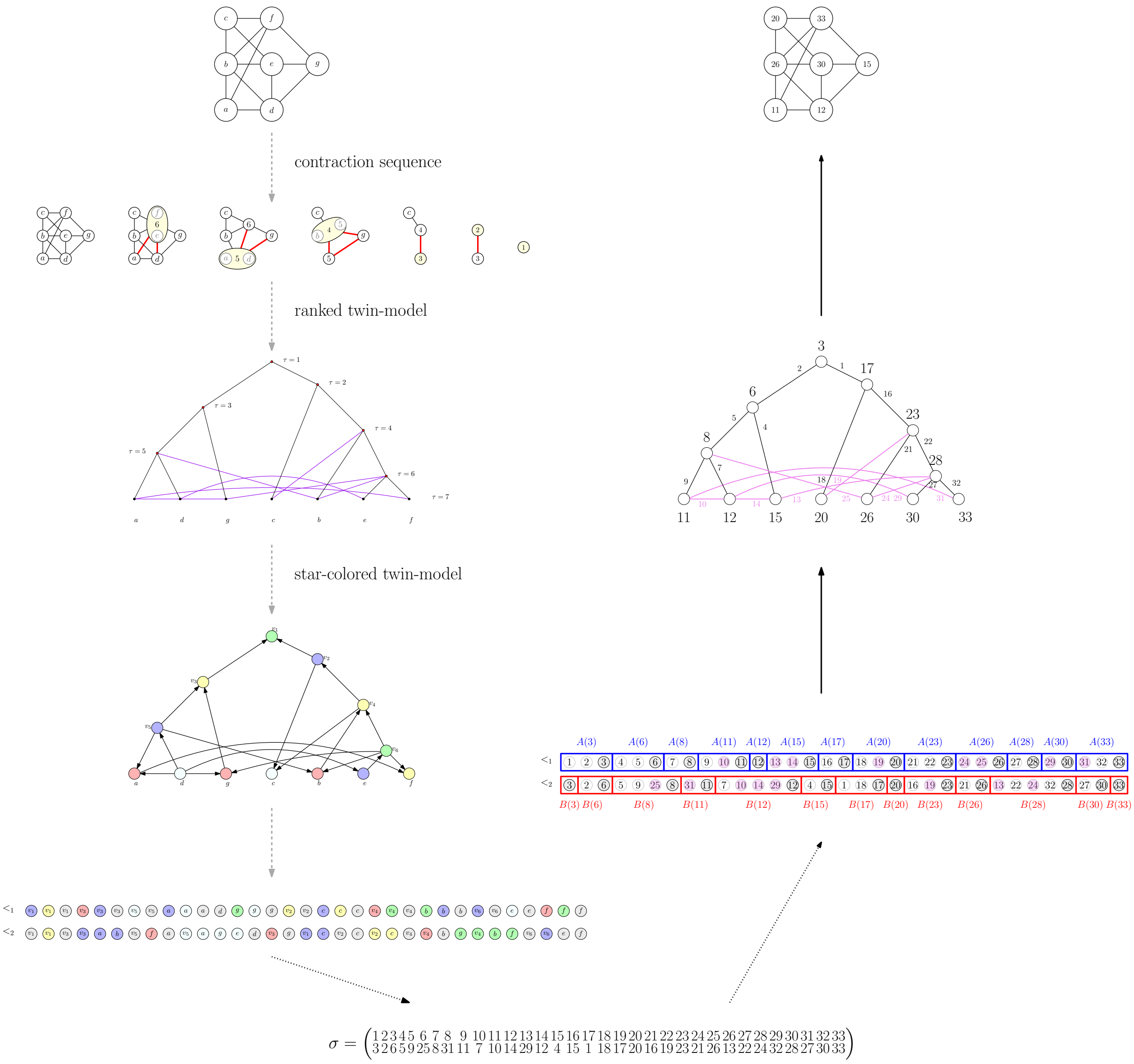


Figure 3: From a graph G to a permutation σ, and back.

References

[1] Michael Albert, Mathilde Bouvel, and Valentin Féray. Two first-order logics of permutations. *Journal of Combinatorial Theory, Series A*, 171:105158, 2020.

[2] Olivier Bernardi, Marc Noy, and Dominic Welsh. Growth constants of minor-closed classes of graphs. *Journal of Combinatorial Theory, Series B*, 100(5):468–484, 2010. `doi:10.1016/j.jctb.2010.03.001`.

[3] Édouard Bonnet, Colin Geniet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width II: small classes. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1977–1996, 2021. `doi:10.1137/1.9781611976465.118`.

[4] Édouard Bonnet, Colin Geniet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width III: Max Independent Set, Min Dominating Set, and Coloring. In Nikhil Bansal, Emanuela Merelli, and James Worrell, editors, *48th International Colloquium on Automata, Languages, and Programming, ICALP 2021, July 12-16, 2021, Glasgow, Scotland (Virtual Conference)*, volume 198 of *LIPIcs*, pages 35:1–35:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. URL: `https://doi.org/10.4230/LIPIcs.ICALP.2021.35`, `doi:10.4230/LIPIcs.ICALP.2021.35`.

[5] Édouard Bonnet, Ugo Giocanti, Patrice Ossona de Mendez, Pierre Simon, Stéphan Thomassé, and Szymon Toruńczyk. Twin-width IV: ordered graphs and matrices, 2021. `arXiv:2102.03117`.

[6] Édouard Bonnet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width I: tractable FO model checking. In *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020*, pages 601–612. IEEE, 2020. URL: `https://doi.org/10.1109/FOCS46700.2020.00062`, `doi:10.1109/FOCS46700.2020.00062`.

[7] Édouard Bonnet, Jaroslav Nešetřil, Patrice Ossona de Mendez, Sebastian Siebertz, and Stéphan Thomassé. Twin-width and permutations. arXiv:2102.06880 [cs.LO], 2021.

[8] Thomas Colcombet. A combinatorial theorem for trees. In Lars Arge, Christian Cachin, Tomasz Jurdzinski, and Andrzej Tarlecki, editors, *Automata, Languages and Programming, 34th International Colloquium, ICALP 2007, Wroclaw, Poland, July 9-13, 2007, Proceedings*, volume 4596 of *Lecture Notes in Computer Science*, pages 901–912. Springer, 2007. URL: `https://doi.org/10.1007/978-3-540-73420-8_77`, `doi:10.1007/978-3-540-73420-8\_77`.

[9] Sylvain Guillemot and Dániel Marx. Finding small patterns in permutations in linear time. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 82–101, 2014. URL: `https://doi.org/10.1137/1.9781611973402.7`, `doi:10.1137/1.9781611973402.7`.

[10] Adam Marcus and Gábor Tardos. Excluded permutation matrices and the Stanley-Wilf conjecture. *J. Comb. Theory, Ser. A*, 107(1):153–160, 2004. URL: `https://doi.org/10.1016/j.jcta.2004.04.002`, `doi:10.1016/j.jcta.2004.04.002`.

This paper is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 810115 – Dynasnet) and from the German Research Foundation (DFG) with grant agreement No 444419611.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Independent dominating sets in planar triangulations

(Extended abstract)

Fábio Botler* Cristina G. Fernandes† Juan Gutiérrez‡§

Abstract

In 1996, Matheson and Tarjan proved that every planar triangulation on n vertices contains a dominating set of size at most $n/3$, and conjectured that this upper bound can be reduced to $n/4$ when n is sufficiently large. In this paper, we consider the analogous problem for independent dominating sets: What is the minimum ε for which every planar triangulation on n vertices contains an independent dominating set of size at most εn? We prove that $2/7 \leq \varepsilon \leq 3/8$.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-023`

1 Introduction

Let S be a set of vertices of a graph G. We say that S is *dominating* if each vertex of G is in S or is a neighbor of some vertex in S; and that S is *independent* if no two vertices in S

*Programa de Engenharia de Sistemas e Computação. Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa em Engenharia. Universidade Federal do Rio de Janeiro, Brazil. E-mail: `fbotler@cos.ufrj.br`. Supported by CNPq (Proc. 423395/2018-1 and 304315/2022-2), FAPERJ (Proc. 211.305/2019 and 201.334/2022) and CAPES-PRINT (Proc. 88887.695773/2022-00).

†Departamento de Ciência da Computação. Universidade de São Paulo, Brazil. E-mail: `cris@ime.usp.br`. Supported by CNPq (Proc. 308116/2016-0 and 423833/2018-9) and FAPESP (Proc. 2019/13364-7).

‡Departamento de Ciencia de la Computación. Universidad de Ingeniería y Tecnología (UTEC), Perú. E-mail: `jgutierreza@utec.edu.pe`. Supported by Movilizaciones para Investigación AmSud, PLANarity and distance IN Graph theory (E070-2021-01-Nro.6997) and Fondo Semilla UTEC 871075-2022.

§This research has been partially supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil – CAPES – Finance Code 001. FAPERJ and FAPESP are, respectively, Research Foundations of Rio de Janeiro and São Paulo, Brazil. CNPq is the National Council for Scientific and Technological Development of Brazil.

are adjacent. In particular, any maximal independent set in G is dominating. We denote by $\gamma(G)$ (resp. $\iota(G)$) the cardinality of a minimum dominating (resp. independent dominating) set of G. Note that $\gamma(G) \leq \iota(G)$. Such parameters are known as the *domination number* and the *independent domination number* of G, respectively, and their calculations are known to be NP-complete problems even on planar bipartite graphs with maximum degree 3 [13, Corollary 3]. Therefore, it is natural to explore such parameters in special classes of graphs, or to look for upper and lower bounds on them.

In this paper, we focus on *planar* graphs, i.e., graphs that can be drawn in the plane so that intersections of edges happen only at their ends. By a *plane* graph we mean a planar graph together with a fixed planar drawing of it. For general terminology on planar graphs we refer to the book of Diestel [3]. In particular, a *planar triangulation* is a plane graph in which each face is bounded by a triangle; and a *triangulated disk* or a *near triangulation* is a plane graph in which each face, except possibly its outer face, is bounded by a triangle. In 1996, Matheson and Tarjan [8] proved that every triangulated disk G on n vertices satisfies $\gamma(G) \leq n/3$, and posed the following conjecture.

Conjecture 1 (Matheson–Tarjan, 1996)**.** *For every sufficiently large planar triangulation G on n vertices, we have $\gamma(G) \leq n/4$.*

Several partial results has been proved for Conjecture 1 [6, 7, 9, 10]. The best known general result is due to Špacapan [12], who proved that $\gamma(G) \leq 17n/53$ for every planar triangulation G on $n \geq 6$ vertices. Related results for maximal outerplanar graphs has been given in [2, 11].

We are interested in the analogous problem for the independent domination number: What is the minimum ε such that $\iota(G) \leq \varepsilon n$ for every planar triangulation G on n vertices? In contrast to the domination number, this parameter has not received so much attention on planar triangulations. It is known that, for any planar graph G on $n \geq 10$ vertices, $\iota(G) < 3n/4$ [5, Theorem 6] and that $\iota(G) \leq n/2$ if G is planar and $\delta(G) \geq 2$ [5, Theorem 8]. For an excellent survey on independent dominating sets, see [4].

Now, note that since every Eulerian planar triangulation has chromatic number 3, they contain three disjoint independent dominating sets. Goddard and Henning [5, Question 1] asked whether such three sets exist in any planar triangulation. In particular, this would imply that $\iota(G) \leq n/3$ for every n-vertex planar triangulation G. We state the later statement as a conjecture.

Conjecture 2. *For every planar triangulation G on n vertices, we have $\iota(G) \leq n/3$.*

Our main contribution is the following theorem.

Theorem 3. *For every planar triangulation G on n vertices, we have $\iota(G) < 3n/8$. Moreover, if $\delta(G) \geq 5$, then $\iota(G) \leq n/3$.*

We also show that the bound $3n/8$ cannot be reduced below $2n/7$, by presenting an infinite family of planar triangulations G for which $\iota(G) \geq 2n/7$ (see Theorem 5). Note that this improves an observation of Goddard and Henning [5, Figure 6], who presented an infinite family of planar triangulations G for which $\iota(G) \geq 5n/19$.

As a starting example, we prove that $\iota(G) \leq 2n/5$ for every planar triangulation G on n vertices due to a relation between r-dynamic and acyclic colorings as follows. A *k-coloring* of a graph G is a partition of $V(G)$ into k independent sets. Each part in such a partition is called a *color class*. A coloring of G is *r-dynamic* if each vertex v has neighbors in at least $\min\{r, d(v)\}$ color classes, where $d(v)$ denotes the degree of v in G; and a coloring of G is *acyclic* if the union of any two of its color classes induces a forest. We use the following result of Goddard and Henning [5, Lemma 4].

Lemma 4 (Goddard–Henning, 2020). *For every graph G on n vertices with $\delta(G) \geq r$ for which there is an r-dynamic k-coloring, we have $\iota(G) \leq (k-r)\,n/k$.*

Borodin [1] showed that every planar graph admits an acyclic 5-coloring. Let G be a planar triangulation and χ_a be an acyclic 5-coloring of G. Note that $\delta(G) \geq 3$ and the neighborhood of each vertex contains a cycle; hence, because χ_a is an acyclic coloring, every vertex has neighbors in at least three color classes of χ_a. Therefore, χ_a is 3-dynamic and, by Lemma 4, we have $\iota(G) \leq 2n/5$.

2 An improved upper bound

In this section, we prove the main theorem of this paper, Theorem 3. For that, we introduce a concept and settle some notation. For a vertex v in G, denote by $N(v)$ the set of neighbors of v in G. For a vertex set S of G, denote by $N(S)$ the set of all neighbors of vertices in S (that may also include vertices of S), and by $N[S]$ the closed neighborhood of S, that is, $N[S] = N(S) \cup S$.

Proof of Theorem 3. Let G be a planar triangulation on n vertices. The celebrated Four-Color Theorem assures that there exists a 4-coloring for G [3, Theorem 5.1.1]. Let C_1, C_2, C_3, C_4 be the color classes in such a coloring. Each C_i is an independent set. If some C_i is empty, then each of the other color classes is non-empty and dominating. Therefore, the smallest of the three non-empty color classes is an independent dominating set of size at most $n/3 < 3n/8$.

So suppose each C_i is non-empty. For each i, let U_i be the set of vertices that are not dominated by C_i. Note that U_1, U_2, U_3, and U_4 are pairwise disjoint, as the neighborhood of any vertex is colored with at least two colors, distinct from the color used in the vertex. We start by proving a stronger statement on U_1, U_2, U_3, and U_4, namely that

$$N[U_i] \cap U_j \;=\; \emptyset \qquad \text{if } i \neq j. \tag{1}$$

Indeed, by contradiction, say $v \in N(U_1) \cap U_2$. Then $v \in C_3 \cup C_4$. Let u be a neighbor of v in U_1. As $v \in U_2$, we conclude that $u \in N(U_2) \cap U_1$, and hence $u \in C_3 \cup C_4$ also. Because G is a planar triangulation, and u and v are adjacent, u and v have a common neighbor, say w. Since either $v \in C_3$ and $u \in C_4$, or $v \in C_4$ and $u \in C_3$, vertex w has either color 1 or 2, contradicting the fact that $v \in U_2$ and $u \in U_1$.

Let S_i be an independent dominating set of $G[U_i]$ and let $S = S_1 \cup S_2 \cup S_3 \cup S_4$. Because each S_i is a subset of U_i, by (1), the set S is independent in G. Let $H = G - S$, and consider the plane embedding of H induced by G. Because G is a planar triangulation, S is independent, and all neighbors of S are in H, each vertex of S lies in a different face of H. Moreover, the boundary of each face of H is a cycle, hence H is connected. Note that, for each $v \in U_i$, the set $N(v)$ is colored with two colors. Thus, the boundary of each face of H in which a vertex from S lies has an even number of vertices.

Let G' be a connected plane graph on n' vertices and m' edges. For $i \geq 3$, let f_i be the number of faces in G' with i edges on their boundary. Then, $2m' = \sum_{i\geq 3} i\, f_i$ and the number of faces in G' is $f' = \sum_{i\geq 3} f_i$. By Euler's formula [3, Theorem 4.2.9], we have that $n' - m' + f' = 2$, which implies that $n' - (\sum_{i\geq 3} i\, f_i)/2 + \sum_{i\geq 3} f_i = n' - \sum_{i\geq 3} \frac{(i-2)}{2} f_i = 2$. Hence, $f_4 + 2\sum_{i\geq 6} f_i \leq \sum_{i\geq 4} \frac{(i-2)}{2} f_i \leq \sum_{i\geq 3} \frac{(i-2)}{2} f_i = n' - 2$. So, if $G' = H$, then there are at most $f_4 + \sum_{i\geq 6} f_i$ vertices in S, because vertices of S lie in faces with an even number of vertices on their boundary. Thus, $2|S| \leq 2(f_4 + \sum_{i\geq 6} f_i) \leq n' - 2 + f_4$. Moreover, $n' + |S| = n$ because $H = G - S$. Joining the last two inequalities, we conclude that $3|S| \leq n - 2 + f_4$. Hence, each set $C_i \cup S_i$ is an independent dominating set and

$$\sum_{i=1}^{4}(|C_i| + |S_i|) = n + |S| \leq \frac{4n - 2 + f_4}{3}.$$

So the smallest of these four independent dominating sets has size less than $n/3 + f_4/12$. The number of faces in G is $2n - 4$, and there are $4f_4$ faces of G incident to the vertices of degree 4 in S. Therefore $f_4 \leq (2n-4)/4 < n/2$, and $\iota(G) < n/3 + n/24 = 3n/8$. Moreover, if $\delta(G) \geq 5$, then $f_4 = 0$ and $\iota(G) \leq n/3$. □

3 A lower bound

As far as we know, Theorem 3 might not be tight: we do not know a family of planar triangulations G on n vertices with $\iota(G)$ approaching $3n/8$. We improve the previous lower bound on ε given by Goddard and Henning [5, Figure 6] in the next result.

Theorem 5. *There is an infinite family $\mathcal{F}$ of planar triangulations such that $\iota(G) = 2n/7$ for every $G \in \mathcal{F}$, where n is the number of vertices in G.*

Proof. Consider the diamond graph depicted in Figure 1(a). Let us describe a family $\mathcal{F}$ of planar triangulations using this graph. Each planar triangulation in $\mathcal{F}$ consists of a circular chain of such diamond graphs, as depicted in Figure 1(b), with edges added to result in a planar triangulation. The planar triangulation G_k obtained in this way with k diamond graphs has $n = 7k$ vertices. The squared vertices in Figure 1(b) show an independent dominating set with $2k = 2n/7$ vertices. Note that any independent dominating set in such a planar triangulation G_k must contain at least two vertices in each diamond graph, therefore $\iota(G_k) = 2k = 2n/7$. □

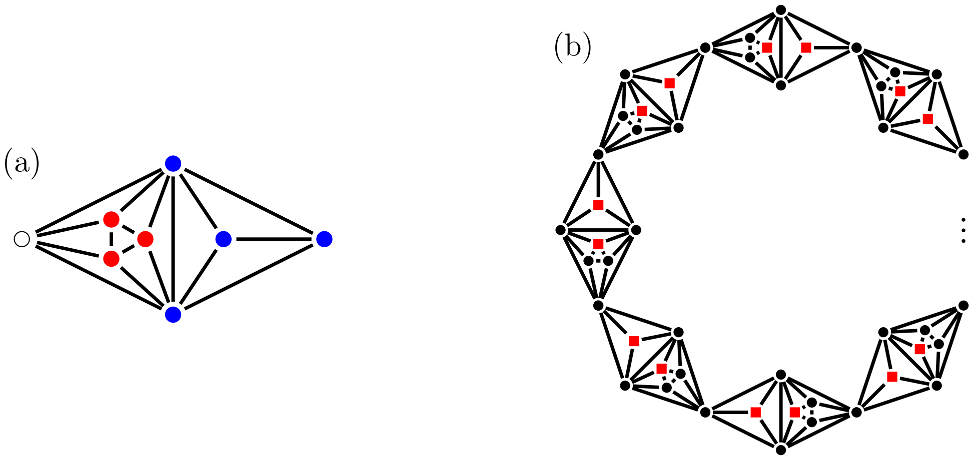


Figure 1: (a) A gadget consisting of seven vertices: the white vertex is not part of the gadget. Any independent dominating set has one of the red vertices, otherwise, being independent, it cannot dominate the three red vertices. Analogously, any independent dominating set has one of the blue vertices, otherwise it does not dominate the middle blue vertex. (b) A graph consisting of a circular chain of gadgets. In each copy of the gadget, two of its vertices are needed in any independent dominating set. The red squared vertices form an independent set with exactly two vertices in each gadget.

4 Further results and concluding remarks

In this section, we explore a few families of planar triangulations for which we can obtain better bounds on their independent domination number. A planar 3-tree is a planar triangulation that can be obtained from a triangle by repeatedly choosing one of its faces and adding a new vertex inside of it while joining this new vertex to the three vertices of the face. It is not hard to prove that any planar 3-tree admits a 4-coloring in which each of its color classes is dominating. Thus $\iota(G) \leq n/4$ for every planar 3-tree on n vertices.

As we observed before, Conjecture 2 is valid for any Eulerian planar triangulation. We can prove a better bound for a particular class of Eulerian triangulations, which we call *recursive Eulerian triangulations*, and define as follows. A recursive Eulerian triangulation is either a triangle, or a triangulation obtained from a recursive Eulerian triangulation by selecting a face and drawing a triangle inside of it while joining each vertex of the selected triangle to the ends of a different edge of the new triangle.

Theorem 6. *For every recursive Eulerian triangulation G on $n \geq 9$ vertices, $\iota(G) < \frac{13n}{42}$.*

Finally, if every vertex of a planar triangulation G on n vertices has odd degree, then every color class of a 4-coloring of G is dominating, hence $\iota(G) \leq n/4$. We can extend this result and show that if G has at least αn odd-degree vertices, then $\iota(G) \leq (2-\alpha)n/4$, which improves the bound in Theorem 3 when $\alpha \geq 2/7$. Also, Conjecture 2 holds for any n-vertex planar triangulation with at least $2n/3$ odd-degree vertices.

References

[1] Oleg V. Borodin. On acyclic colorings of planar graphs. *Discrete Math.*, 25(3):211–236, 1979.

[2] Christiane N. Campos and Yoshiko Wakabayashi. On dominating sets of maximal outerplanar graphs. *Discrete Appl. Math.*, 161(3):330–335, 2013.

[3] Reinhard Diestel. *Graph Theory*, volume 173 of *Graduate Texts in Mathematics*. Springer, Heidelberg, fourth edition, 2010.

[4] Wayne Goddard and Michael A. Henning. Independent domination in graphs: a survey and recent results. *Discrete Math.*, 313(7):839–854, 2013.

[5] Wayne Goddard and Michael A. Henning. Independent domination, colorings and the fractional idomatic number of a graph. *Appl. Math. Comput.*, 382:125340, 8, 2020.

[6] Erika L. C. King and Michael J. Pelsmajer. Dominating sets in plane triangulations. *Discrete Math.*, 310(17-18):2221–2230, 2010.

[7] Zepeng Li, Enqiang Zhu, Zehui Shao, and Jin Xu. On dominating sets of maximal outerplanar and planar graphs. *Discrete Appl. Math.*, 198:164–169, 2016.

[8] Lesley R. Matheson and Robert E. Tarjan. Dominating sets in planar graphs. *European J. Combin.*, 17(6):565–568, 1996.

[9] Michael D. Plummer, Dong Ye, and Xiaoya Zha. Dominating plane triangulations. *Discrete Appl. Math.*, 211:175–182, 2016.

[10] Michael D. Plummer, Dong Ye, and Xiaoya Zha. Dominating maximal outerplane graphs and Hamiltonian plane triangulations. *Discrete Appl. Math.*, 282:162–167, 2020.

[11] Shin-ichi Tokunaga. Dominating sets of maximal outerplanar graphs. *Discrete Appl. Math.*, 161(18):3097–3099, 2013.

[12] Simon Špacapan. The domination number of plane triangulations. *J. Combin. Theory Ser. B*, 143:42–64, 2020.

[13] Igor E. Zverovich and Vadim E. Zverovich. An induced subgraph characterization of domination perfect graphs. *J. Graph Theory*, 20(3):375–395, 1995.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Biclique immersions in graphs with independence number 2 *

(Extended abstract)

F. Botler[1] A. Jiménez[2] C. N. Lintzmayer[3] A. Pastine[4]
D. A. Quiroz[2] M. Sambinelli[3]

Abstract

The analogue of Hadwiger's Conjecture for the immersion relation states that every graph G contains an immersion of $K_{\chi(G)}$. For graphs with independence number 2, this is equivalent to stating that every such n-vertex graph contains an immersion of $K_{\lceil n/2 \rceil}$. We show that every n-vertex graph with independence number 2 contains every complete bipartite graph on $\lceil n/2 \rceil$ vertices as an immersion.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-024`

*This research has been partially supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil – CAPES – Finance Code 001, by MATHAMSUD MATH190013, and by grant 2019/13364-7 of the São Paulo Research Foundation (FAPESP). F. Botler is supported by the National Council for Scientific and Technological Development of Brazil, CNPq (423395/2018-1), the Research Foundation of Rio de Janeiro, FAPERJ (211.305/2019 and 201.334/2022), and CAPES-PRINT (88887.695773/2022-00). A. Jiménez is partially supported by ANID/Fondecyt Regular 1220071, MATHAMSUD MATH210008, and ANID/PCI/REDES 190071. C. N. Lintzmayer is partially supported by CNPq (Proc. 312026/2021-8). A. Pastine is partially supported by UNSL, grants PROICO 03-0918 and PROIPRO 03-1720, and by ANPCyT grants PICT-2020-SERIEA-04064 and PICT-2020-SERIEA-00549. D.A. Quiroz is partially supported by ANID/Fondecyt Iniciación en Investigación 11201251 and MATHAMSUD MATH210008.

[1]Universidade Federal do Rio de Janeiro – Brasil. [2]Universidad de Valparaíso – Chile. [3]Universidade Federal do ABC – Brazil. [4]CONICET & Universidad Nacional de San Luis – Argentina. E-mails: `fbotler@cos.ufrj.br`, `andrea.jimenez@uv.cl`, `carla.negri@ufabc.edu.br`, `agpastine@unsl.edu.ar`, `daniel.quiroz@uv.cl`, `m.sambinelli@ufabc.edu.br`

1 Introduction

A central problem in graph theory is guaranteeing dense substructures in graphs with a given chromatic number. Hadwiger's Conjecture [13] is one of the most important examples of this pursuit, stating that every loopless graph G contains the complete graph $K_{\chi(G)}$ as a minor (where $\chi(G)$ is the chromatic number of G), thus aiming to generalize the Four Color Theorem. This difficult conjecture is known to hold whenever $\chi(G) \leq 6$ [27], and it is open for the remaining values. Thus, a natural approach is to study whether it holds whenever G is restricted to a particular class of graphs. A class of graphs that has received particular attention (and yet remains open) is that of graphs with independence number 2. A recent survey of Seymour [28] emphasizes the importance of this case, which was first remarked by Mader (see [24]). Plummer, Stiebitz, and Toft [24] gave an equivalent formulation of Hadwiger's Conjecture for such graphs: every n-vertex graph with independence number 2 contains a minor of $K_{\lceil n/2 \rceil}$. Before that, in 1982, Duchet and Meyniel [8] had shown a result that implies that every such graph contains a minor of $K_{\lceil n/3 \rceil}$. Despite much work, see e.g. [11, 14, 15, 32], it is still open whether there is a constant $c > 1/3$ such that every graph with independence number 2 contains a minor of $K_{\lceil cn \rceil}$. Given the difficulty to obtain a clique minor on $\lceil n/2 \rceil$ vertices, Norin and Seymour [23] recently turned into finding dense minors on this amount of vertices. They proved that every n-vertex graph with independence number 2 contains a (simple) minor of a graph H on $\lceil n/2 \rceil$ vertices and $0.98688 \cdot \binom{|V(H)|}{2} - o(n^2)$ edges.

The focus of this paper is a conjecture related to Hadwiger's, concerned with finding graph immersions in graphs with a given chromatic number; this type of substructure is defined as follows. To *split off* a pair of adjacent edges uv,vw amounts to deleting those two edges and adding the edge uw. A graph G is said to contain an *immersion* of another graph H if H can be obtained from a subgraph of G by splitting off pairs of edges and deleting isolated vertices. Notice then that if G contains H as a subdivision, it contains H as an immersion (and as a minor). Immersions have received increased attention in recent years, see e.g. [6, 9, 20, 21, 22, 31], particularly since Robertson and Seymour [26] proved that graphs are well-quasi-ordered by the immersion relation. Much of this attention has been centered around the following conjecture of Abu-Khzam and Langston [1], which is the immersion-analog of Hadwiger's Conjecture.

Conjecture 1 (Abu-Khzam and Langston [1])**.** *Every loopless graph G contains an immersion of $K_{\chi(G)}$.*

The above conjecture holds whenever $\chi(G) \leq 4$ because Hajós' subdivision conjecture holds in this case, actually giving a subdivision of $K_{\chi(G)}$ [7]. The cases where $\chi(G) \in \{5, 6, 7\}$ were verified independently by Lescure and Meyniel [19] and by DeVos, Kawarabayashi, Mohar, and Okamura [5]. In general, a result of Gauthier, Le, and Wollan [12] guarantees that every graph G contains an immersion of a clique on $\lceil \frac{\chi(G)-4}{3.54} \rceil$ vertices. This result improves on theorems due to Dvořák and Yepremyan [10] and DeVos, Dvořák, Fox, McDonald, Mohar, and Scheide [4].

The case of graphs with independence number 2 has also received attention in regard to Conjecture 1. In particular, Vergara [30] showed that, for such graphs, Conjecture 1 is equivalent to the following conjecture.

Conjecture 2 (Vergara [30])**.** *Every n-vertex graph with independence number 2 contains an immersion of $K_{\lceil n/2 \rceil}$.*

As evidence for her conjecture, Vergara proved that every n-vertex graph with independence number 2 contains an immersion of $K_{\lceil n/3 \rceil}$. This was later improved by Gauthier *et al.* [12], who showed that every such graph contains an immersion of $K_{2\lfloor n/5 \rfloor}$. This last result was extended to graphs with arbitrary independence number [3]. Additionally, Vergara's Conjecture has been verified for graphs with small forbidden subgraphs [25]. The main contribution of this paper is the following result, which states that graphs with independence number 2 contain an immersion of every complete bipartite graph on $\lceil n/2 \rceil$ vertices.

Theorem 3. *Let G be an n-vertex graph with independence number 2, and $\ell \leq \lceil n/2 \rceil - 1$ be a positive integer. Then G contains an immersion of $K_{\ell, \lceil n/2 \rceil - \ell}$.*

Using an argument due to Plummer *et al.* [24] we can show that this implies the following.

Corollary 4. *Let G be a graph with independence number 2, and $1 \leq \ell \leq \chi(G) - 1$. Then G contains an immersion of $K_{\ell, \chi(G) - \ell}$.*

This result leads us to make the following conjecture, which holds trivially when $\ell = 1$.

Conjecture 5. *If $1 \leq \ell \leq \chi(G) - 1$, then G contains an immersion of $K_{\ell, \chi(G) - \ell}$.*

We denote by $K_{a,b,c}$ the graph that admits a partition into parts of sizes a, b, and c such that any pair of these parts induces a complete bipartite graph. In addition to Corollary 4, as evidence for Conjecture 5, we can prove the following strengthening of the case $\ell = 2$.

Proposition 6. *If $\chi(G) \geq 3$, then G contains $K_{1,1,\chi(G)-2}$ as an immersion.*

We note that Conjecture 5 has its parallel in the minor order. Woodall [33] and, independently, Seymour (see [18]), proposed the following conjecture: every graph G with $\ell \leq \chi(G) - 1$ contains a minor of $K_{\ell, \chi(G) - \ell}$. In [33], Woodall showed that (the list-coloring strengthening of) his conjecture holds whenever $\ell \leq 2$. Kostochka and Prince [18] showed that the case $\ell = 3$ holds as long as $\chi(G) \geq 6503$. Kostochka [16] proved it for every ℓ as long as $\chi(G)$ is very large in comparison to ℓ, and later [17] improved this so that $\chi(G)$ could be polynomial in ℓ, namely, whenever $\chi(G) > 5(200\ell \log_2(200\ell))^3 + \ell$. In fact, the results in [16, 17, 18] obtain the full join $K^*_{\ell, \chi(G) - \ell}$, which is the graph obtained from the disjoint union of a K_ℓ and an independent set on $\chi(G) - \ell$ vertices by adding all of the possible edges between them. This and the above-cited result of Norin and Seymour leads us to make the following conjecture.

Conjecture 7. *Let G be an n-vertex graph with independence number 2, and $1 \leq \ell \leq \lceil n/2 \rceil - 1$. Then G contains a minor of $K_{\ell, \lceil n/2 \rceil - \ell}$.*

Note that the result of Kostochka leaves open the balanced case, thus not implying Conjecture 7. Moreover, it is not hard to build a graph that is denser than the minor obtained by the result of Norin and Seymour, and yet does not contain $K_{\lfloor n/4 \rfloor, \lceil n/4 \rceil}$: take a complete graph on $\lceil n/2 \rceil$ vertices and delete $\lfloor n/4 \rfloor + 1$ edges incident to the same vertex. Thus Conjecture 7 is not implied by this result either.

The rest of the paper is organized as follows. In Section 1.1 we give a few definitions and present an interesting lemma that is used to prove Theorem 3 in Section 2. Due to space limitations, we only present a sketch of the proof. We refer the interested reader to [2] for its details.

1.1 Preliminaries and notation

Let G be a graph. For $v \in V(G)$ and $S \subseteq V(G)$, we define $E(v, S) = \{vu \in E(G) \colon u \in S\}$. If A and B are disjoint sets, we let $K_{A,B}$ be the complete bipartite graph with bipartition (A, B). A substructure (subgraph, immersion, minor, etc.) is a *clique* if it is a complete graph, and is a *biclique* if it is a complete bipartite graph. In a manner that is equivalent to the definition given in the introduction, we say that a graph G *contains an immersion of H* if there exists an injection $f \colon V(H) \to V(G)$ and a collection of edge-disjoint paths in G, one for each edge of H, such that the path P_{uv} corresponding to the edge uv has endpoints $f(u)$ and $f(v)$.

Finally, in our proof, we make use of the following lemma, which we believe to be interesting by itself, and that we could not find in the literature.

Lemma 8. *Let $j \leq k$ be positive integers, and let $C_1, \ldots, C_j \subseteq [n]$ be sets of size k. Let A be a set of size k disjoint from $[n]$. Then there are disjoint matchings $M_1, \ldots, M_j$ in $K_{A,[n]}$ such that M_i matches A with C_i for every $i \in [j]$.*

2 Outline of the proof of Theorem 3

Indeed, we can consider graphs with independence number at most 2. The proof follows by induction on $n + \ell$. Let G be an n-vertex graph with $\alpha(G) \leq 2$ and let $\ell \leq \lceil n/2 \rceil - 1$ be a positive integer. Note that the result is easy when $n \leq 4$, so we can assume $n \geq 5$. Note also that it suffices to prove the statement in the case G is *edge-critical*, i.e., that the removal of any edge of G increases its independence number. Now, if $n \leq 4\ell - 2$, then $\lceil n/2 \rceil - \ell \leq 2\ell - 1 - \ell < \ell$. Thus, by induction there is an immersion of $K_{\ell', \lceil n/2 \rceil - \ell'}$ in G, where $\ell' = \lceil n/2 \rceil - \ell$. But this is the desired immersion because $K_{\lceil n/2 \rceil - \ell, \lceil n/2 \rceil - \lceil n/2 \rceil + \ell}$ is isomorphic to $K_{\lceil n/2 \rceil - \ell, \ell}$. Thus, from now on, we assume that $n \geq 4\ell - 1$.

Now, suppose that G contains two non-adjacent vertices, say x and y, with at least $\ell - 1$ common neighbors, and let $G' = G - x - y$. If $\ell \leq \lceil n/2 \rceil - 2 = \lceil (n-2)/2 \rceil - 1$, the induction hypothesis guarantees that G' contains an immersion of $K_{\ell, \lceil (n-2)/2 \rceil - \ell}$, which we call H'. Otherwise, if we have $\ell = \lceil n/2 \rceil - 1$ we let H' be an arbitrary set of ℓ vertices. Let L and B be the parts of H' having size ℓ and $\lceil n/2 \rceil - 1 - \ell$, respectively, and let $R = V(G') \setminus (L \cup B)$. As $\alpha(G) = 2$, every vertex in G' is either adjacent to x or to y in G. This is true in particular for the vertices in L. In what follows, we add either x or y to B, in order to obtain the desired immersion of $K_{\ell,\lceil n/2 \rceil - \ell}$. This is immediate if x or y is adjacent to every vertex in L. Thus we may assume that $|E(y, L)|, |E(x, L)| < \ell$. Now, let L_x (resp. L_y) be the set of vertices in L adjacent to x but not to y (resp. to y but not to x); L_c be the set of vertices in L adjacent to both x and y; and O_c be the set of vertices adjacent to both x and y that are not in L. As x is adjacent to every vertex in $L_x \cup L_c$, it is enough to find (edge-disjoint) paths from x to L_y without using edges of H'. Notice that $|L_y| + |L_c| = |E(y, L)| \leq \ell - 1$ and that, by hypothesis, we have $|L_c| + |O_c| \geq \ell - 1$. Thus $|O_c| \geq |L_y|$. Let $O_c = \{o_1, o_2, \ldots, o_{|O_c|}\}$ and $L_y = \{\ell_1, \ell_2, \ldots, \ell_{|L_y|}\}$. For $1 \leq i \leq |L_y|$, we take the path $x o_i y \ell_i$. These paths are as desired. Therefore, we may assume that $|N(u) \cap N(v)| \leq \ell - 2$ for every pair of non-adjacent vertices u, v.

2.1 Consequences of edge-criticality

Recall that G is edge-critical, meaning that the removal of any edge $uv \in E(G)$ creates an independent set of size 3. Hence, for such an edge there is a vertex w that is not adjacent to both u and v. We formalize this argument in the following claim.

Claim 9. *For any $u, v \in V(G)$, we have $uv \in E(G)$ if and only if $N[u] \cup N[v] \neq V(G)$.*

For the rest of the proof, we fix two non-adjacent vertices x and y, and partition $V(G)$ as follows:

- ▷ $C = N(x) \cap N(y)$, the set of common neighbors of x and y;
- ▷ $X = \overline{N[y]}$, the set of non-neighbors of y excluding y, which contains x; and
- ▷ $Y = \overline{N[x]}$, the set of non-neighbors of x excluding x, which contains y.

We observe that $|C| \leq \ell - 2$, and that each of X and Y induces a complete subgraph of G, otherwise we could find an independent set of size 3. Moreover, the edge-criticality of G yields the following claim.

Claim 10. *For every vertex $a \in C$, we have $X, Y \not\subseteq N(a)$.*

2.2 Key vertex sets

Let $X_C \subseteq X$ (resp. $Y_C \subseteq Y$) be the set containing vertices $v \in X$ (resp. $v \in Y$) for which $C \subset N(v)$, and put $\overline{X}_C = X \setminus X_C$ (resp. $\overline{Y}_C = Y \setminus Y_C$). Now, given a vertex a in C, we denote by X_a (resp. Y_a) the set of vertices in X (resp. in Y) that are adjacent to a, and put $\overline{X}_a = X \setminus X_a$ and $\overline{Y}_a = Y \setminus Y_a$. Notice that if $v \in \overline{X}_a$ and $w \in \overline{Y}_a$, then v and w must be adjacent, as the independence number of G is 2. Thus we get $K_{\overline{X}_a, \overline{Y}_a}$ as a subgraph of G.

Note that $X_C \subseteq X_a$ and $\overline{X}_C \supseteq \overline{X}_a$ (resp. $Y_C \subseteq Y_a$ and $\overline{Y}_C \supseteq \overline{Y}_a$) for every $a \in C$. Indeed, we have $X_C = \bigcap_{a \in C} X_a$ (resp. $Y_C = \bigcap_{a \in C} Y_a$) and $\overline{X}_C = \bigcup_{a \in C} \overline{X}_a$ (resp. $\overline{Y}_C = \bigcup_{a \in C} \overline{Y}_a$). The following claim gives bounds and relations on the sizes of some of these sets. This control is the key to build the desired immersion.

Claim 11.

1. *Given $a \in C$, we have that $|X_C| \leq |X_a| \leq \ell - 2$ and $|Y_C| \leq |Y_a| \leq \ell - 2$. Furthermore, we have $|\overline{X}_a| \geq \lceil n/2 \rceil - |Y| + 3$ and $|\overline{Y}_a| \geq \lceil n/2 \rceil - |X| + 3$.*
2. *For every $v \in \overline{X}_C$ (resp. $w \in \overline{Y}_C$), we have $|N(v) \cap \overline{Y}_C| > \lceil n/2 \rceil - |X|$ (resp. $|N(w) \cap \overline{X}_C| > \lceil n/2 \rceil - |Y|$).*

2.3 Constructing the immersion

The rest of the proof is divided into two cases which depend on the sizes of $\overline{X}_C$ and $\overline{Y}_C$. The sets that form the bipartition of the immersion depend on which case we are dealing with. The construction requires more care in the case one of $\overline{X}_C, \overline{Y}_C$ is large, which we sketch here. Say, without loss of generality, that $|\overline{X}_C| \geq \ell$. For the rest of the proof, we fix $a \in C$. By Claim 11(1), we can choose $Y^* \subset \overline{Y}_a$ with $|Y^*| = \lceil n/2 \rceil - |X|$, and since $|\overline{X}_C| \geq \ell$, we can choose $X^* \subset \overline{X}_C \setminus \overline{X}_a$ with $|X^*| = \ell - |\overline{X}_a|$. Using Claim 11(1) again, we can show that

$$|X^*| \leq |Y^*|. \tag{1}$$

Since X, Y, and $\overline{X}_a \cup \overline{Y}_a$ induce cliques, G contains all edges joining (i) vertices in $\overline{X}_a$ to vertices in Y^*; (ii) vertices in $\overline{X}_a$ to vertices in $X_a \setminus X^*$; and (iii) vertices in X^* to vertices in $X_a \setminus X^*$. It remains to find edge-disjoint paths joining vertices in X^* to vertices in Y^*. For these paths, we only use edges that are incident to vertices in Y and not to vertices in $\overline{X}_a$; this assures that they are disjoint from the edges already used. Let $X^* = \{v_1, v_2, \ldots, v_{|X^*|}\}$ and $Y^* = \{y_1, y_2, \ldots, y_{|Y^*|}\}$. The first step is to use Lemma 8 to find paths joining each vertex v_i to all vertices in Y^* allowing edges between vertices of Y^* to be used at most twice. Nevertheless, the intersections are relatively few and with a combination of different techniques, we are able to fix them and obtain the desired immersion.

Claim 12. *For each $i \in \{1, 2, \ldots, |X^*|\}$, there is a subgraph $K(v_i)$ which contains an immersion of K_{v_i, Y^*} and satisfies that:*

i) each path of $K(v_i)$ with an endpoint in v_i has length at most 2;

ii) for each path $v_i z y_j$ in $K(v_i)$ we have $z \in \overline{Y}_C$; and

iii) if $i \neq j$ and $uw \in E(K(v_i)) \cap E(K(v_j))$, then there is no $r \notin \{i, j\}$ such that $uw \in E(K(v_r))$, and one path containing uw ends at u while the other ends at w.

Proof. Note that, since $X^* \subseteq \overline{X}_C$, Claim 11(2) assures that for each $i \in \{1, 2, \ldots, |X^*|\}$, we have $|N(v_i) \cap \overline{Y}_C| > \lceil n/2 \rceil - |X| = |Y^*|$. In order to use Lemma 8, we define, for each such i, a set $N_i \subset N(v_i) \cap \overline{Y}_C$ with $|N_i| = |Y^*|$, and a set of auxiliary vertices $A = \{a_1, a_2, \ldots, a_{|Y^*|}\}$ with $N(a_j) = \overline{Y}_C$. By (1), we can apply Lemma 8 to $N_1, \ldots, N_{|X^*|}$ together with A to obtain disjoint matchings $M_1, M_2, \ldots, M_{|X^*|}$ such that M_i matches A to N_i, for $i \in \{1, \ldots, |X^*|\}$. Let $M_i = \{z_{i,1}a_1, \ldots, z_{i,|Y^*|}a_{|Y^*|}\}$ where $z_{i,j} \in N_i$ for every i, j.

For each $v_i \in X^*$, we obtain $K(v_i)$ by using y_j whenever a_j is used in a matching. In other words, for every $1 \leq j \leq |Y^*|$, if $z_{i,j}a_j \in M_i$, then we use the path $v_i z_{i,j} y_j$. Notice that $z_{i,j}$ could be y_j itself. When that is the case, we use the path $v_i y_j$. Formally, we define

$$P(i,j) = \begin{cases} v_i z_{i,j} y_j & \text{if } y_j \neq z_{i,j} \\ v_i y_j & \text{if } y_j = z_{i,j}\,. \end{cases}$$

Notice that $P(i,j)$ may not be edge-disjoint from $P(i,k)$ if $k \neq j$, but this can only happen if $P(i,j) = v_i y_k y_j$ and $P(i,k) = v_i y_j y_k$. If that is the case, we redefine $P(i,j)$ as $v_i y_j$ and $P(i,k)$ as $v_i y_k$. Thus, after doing all the necessary changes, we can assume that $P(i,j)$ is disjoint of $P(i,k)$ whenever $j \neq k$. Finally we define $K(v_i) = \bigcup_{j=1}^{|Y^*|} P(i,j)$ and since the $P(i,j)$'s are edge disjoint each $K(v_i)$ contains an immersion of K_{v_i,Y^*} and clearly satisfies items *i*) and *ii*).

Furthermore, as $M_1, \ldots, M_{|X^*|}$ are disjoint matchings, if $uw \in E(K(v_i)) \cap E(K(v_j))$ for some pair $i \neq j$, then it must be that $u, w \in Y^*$. Let $u = y_h$ and $w = y_k$. Then either $z_{i,h} = y_k$ or $z_{i,k} = y_h$. Assume, w.l.o.g., that $z_{i,h} = y_k$. This means that $z_{i,h}a_h = y_k a_h \in M_i$. Thus $y_k a_h \notin M_r$ for $r \neq i$. This, in turn, implies that $z_{j,k} = y_h$, which means that $y_h a_k \in M_j$ and $y_h a_k \notin M_r$ for $r \neq j$. This proves *iii*). □

Let $K(v_1), \ldots, K(v_{|X^*|})$ be the subgraphs given by Claim 12. We would like the v_i, Y^*-paths on these subgraphs to be the X^*, Y^*-paths in our immersion. Yet, if $i \neq j$, $K(v_i)$ might not be edge disjoint from $K(v_j)$. Fortunately, by Claim 12 *iii*) the intersections are restricted, which we can show to be sufficient for fixing them.

References

[1] Faisal N. Abu-Khzam and Michael A. Langston. Graph coloring and the immersion order. In *Computing and combinatorics*, volume 2697 of *Lecture Notes in Comput. Sci.*, pages 394–403. Springer, Berlin, 2003.

[2] F. Botler, A. Jiménez, C. N. Lintzmayer, A. Pastine, D. A. Quiroz, and M. Sambinelli. Biclique immersions in graphs with independence number 2. *arXiv:2303.06483 [math.CO]*, 2023.

[3] Sebastián Bustamante, Daniel A. Quiroz, Maya Stein, and José Zamora. Clique immersions and independence number. *European J. Combin.*, 106:Paper No. 103550, 9, 2022.

[4] Matt DeVos, Zdeněk Dvořák, Jacob Fox, Jessica McDonald, Bojan Mohar, and Diego Scheide. A minimum degree condition forcing complete graph immersion. *Combinatorica*, 34:279–298, 2014.

[5] Matt DeVos, Ken-ichi Kawarabayashi, Bojan Mohar, and Haruko Okamura. Immersing small complete graphs. *Ars Math. Contemp.*, 3(2):139–146, 2010.

[6] Matt DeVos, Jessica McDonald, Bojan Mohar, and Diego Scheide. Immersing complete digraphs. *European J. Combin.*, 33(6):1294–1302, 2012.

[7] G. A. Dirac. A property of 4-chromatic graphs and some remarks on critical graphs. *J. London Math. Soc.*, 27:85–92, 1952.

[8] P. Duchet and H. Meyniel. On Hadwiger's number and the stability number. *Ann. Discrete Math.*, 13:71–73, 1982.

[9] Zdeněk Dvořák and Paul Wollan. A structure theorem for strong immersions. *J. Graph Theory*, 83(2):152–163, 2016.

[10] Zdeněk Dvořák and Liana Yepremyan. Complete graph immersions and minimum degree. *J. Graph Theory*, 88(1):211–221, 2018.

[11] Jacob Fox. Complete minors and independence number. *SIAM J. Discrete Math.*, 24(4):1313–1321, 2010.

[12] Gregory Gauthier, Tien-Nam Le, and Paul Wollan. Forcing clique immersions through chromatic number. *European J. Combin.*, 81:98–118, 2019.

[13] H. Hadwiger. Über eine Klassifikation der Streckenkomplexe. *Vierteljschr. Naturforsch. Ges. Zürich*, 88:133–142, 1943.

[14] Ken-ichi Kawarabayashi, Michael D. Plummer, and Bjarne Toft. Improvements of the theorem of Duchet and Meyniel on Hadwiger's conjecture. *J. Combin. Theory Ser. B*, 95(1):152–167, 2005.

[15] Ken-ichi Kawarabayashi and Zi-Xia Song. Independence number and clique minors. *J. Graph Theory*, 56(3):219–226, 2007.

[16] A. V. Kostochka. On $K_{s,t}$ minors in $(s+t)$-chromatic graphs. *J. Graph Theory*, 65(4):343–350, 2010.

[17] A. V. Kostochka. $K_{s,t}$ minors in $(s+t)$-chromatic graphs, II. *J. Graph Theory*, 75(4):377–386, 2014.

[18] A. V. Kostochka and N. Prince. Dense graphs have $K_{3,t}$ minors. *Discrete Math.*, 310(20):2637–2654, 2010.

[19] F. Lescure and H. Meyniel. On a problem upon configurations contained in graphs with given chromatic number. *Ann. Discrete Math.*, 41:325–331, 1989.

[20] Chun-Hung Liu. Immersion and clustered coloring. *J. Combin. Theory, Ser. B*, 158(1):252–282, 2023.

[21] Chun-Hung Liu and Irene Muzi. Well-quasi-ordering digraphs with no long alternating paths by the strong immersion relation. *J. Combin. Theory, Ser. B*, 158(1):210–251, 2023.

[22] Dániel Marx and Paul Wollan. Immersions in highly edge connected graphs. *SIAM J. Discrete Math.*, 28(1):503–520, 2014.

[23] Sergey Norin and Paul Seymour. Dense minors of graphs with independence number two, 2022. arXiv:2206.00186 [math.CO].

[24] Michael D. Plummer, Michael Stiebitz, and Bjarne Toft. On a special case of Hadwiger's conjecture. *Discuss. Math. Graph Theory*, 23(2):333–363, 2003.

[25] Daniel A. Quiroz. Clique immersions in graphs of independence number two with certain forbidden subgraphs. *Discrete Math.*, 344(6):Paper No. 112365, 9, 2021.

[26] Neil Robertson and Paul Seymour. Graph minors XXIII. Nash-Williams' immersion conjecture. *J. Combin. Theory, Ser. B*, 100(2):181–205, 2010.

[27] Neil Robertson, Paul Seymour, and Robin Thomas. Hadwiger's conjecture for K_6-free graphs. *Combinatorica*, 13(3):279–361, 1993.

[28] Paul Seymour. Hadwiger's conjecture. In *Open problems in mathematics*, pages 417–437. Springer, [Cham], 2016.

[29] Matěj Stehlík. Critical graphs with connected complements. *J. Combin. Theory Ser. B*, 89(2):189–194, 2003.

[30] Sylvia Vergara. Complete graph immersions in dense graphs. *Discrete Math.*, 340(5):1019–1027, 2017.

[31] Paul Wollan. The structure of graphs not admitting a fixed immersion. *J. Combin. Theory, Ser. B*, 110(1):47–66, 2016.

[32] David R. Wood. Independent sets in graphs with an excluded clique minor. *Discrete Math. Theor. Comput. Sci.*, 9(1):171–175, 2007.

[33] Douglas R. Woodall. List colourings of graphs. In *Surveys in combinatorics, 2001 (Sussex)*, volume 288 of *London Math. Soc. Lecture Note Ser.*, pages 269–301. Cambridge Univ. Press, Cambridge, 2001.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

A resolution of the Kohayakawa–Kreuter conjecture for the majority of cases

(Extended abstract)

Candida Bowtell* Robert Hancock† Joseph Hyde‡

Abstract

For graphs $G, H_1, \ldots, H_r$, write $G \to (H_1, \ldots, H_r)$ to denote the property that whenever we r-colour the edges of G, there is a monochromatic copy of H_i in colour i for some $i \in \{1, \ldots, r\}$. Mousset, Nenadov and Samotij proved an upper bound on the threshold function for the property that $G(n,p) \to (H_1, \ldots, H_r)$, thereby resolving the 1-statement of the Kohayakawa–Kreuter conjecture. We extend upon the many partial results for the 0-statement, by resolving it for a large number of cases, which in particular includes (but is not limited to) when $r \geq 3$, when H_2 is strictly 2-balanced and not bipartite, or when H_1 and H_2 have the same 2-densities.

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-025

1 Introduction

Let $r \in \mathbb{N}$ and $G, H_1, \ldots, H_r$ be graphs. We write $G \to (H_1, \ldots, H_r)$ to denote the property that whenever we colour the edges of G with colours from the set $[r] := \{1, \ldots, r\}$ there exists $i \in [r]$ and a copy of H_i in G monochromatic in colour i. Ramsey's theorem states that for any $H_1, \ldots, H_r$, there exists n_0 such that for all $n \geq n_0$, $K_n \to (H_1, \ldots, H_r)$. Since

*Mathematics Institute, University of Warwick, United Kingdom, Candy.Bowtell@warwick.ac.uk, research supported by ERC Starting Grant 947978 and Philip Leverhulme Prize PLP–2020–183.

†Institut für Informatik, Heidelberg University, Germany, hancock@informatik.uni-heidelberg.de, research supported by a Humboldt Research Fellowship.

‡Mathematics and Statistics, University of Victoria, Victoria, B.C., Canada, josephhyde@uvic.ca, research supported by the UK Research and Innovation Future Leaders Fellowship MR/S016325/1 and ERC Advanced Grant 101020255.

the property $G \to (H_1, \dots, H_r)$ is monotone, a result of Bollobás and Thomason [1] implies that there must exist a threshold function p_0 for the property that the binomial random graph $G(n,p)$ (which has n vertices and contains each possible edge independently with probability p) satisfies $G(n,p) \to (H_1, \dots, H_r)$. Rödl and Ruciński [11, 12, 13] famously located the threshold for the symmetric case, while Kohayakawa and Kreuter [4] gave a conjecture for the threshold for the asymmetric case.

1.1 Notation

Before we can state these thresholds we require some notation. Let $G = (V, E)$ be a graph. We denote the number of vertices in G by $v_G := |V(G)|$ and the number of edges in G by $e_G := |E(G)|$. Moreover, for graphs H_1 and H_2 we let $v_1 := |V(H_1)|$, $e_1 := |E(H_1)|$, $v_2 := |V(H_2)|$ and $e_2 := |E(H_2)|$.

Let H be a graph. We define

$$d(H) := \begin{cases} e_H/v_H & \text{if } v_H \geq 1, \\ 0 & \text{otherwise;} \end{cases}$$
$$m(H) := \max\{d(J) : J \subseteq H\}.$$

We define the *arboricity* (also known as the 1-density measure) by

$$d_1(H) := \begin{cases} e_H/(v_H - 1) & \text{if } v_H \geq 2, \\ 0 & \text{otherwise;} \end{cases}$$
$$ar(H) = m_1(H) := \max\{d_1(J) : J \subseteq H\}.$$

In [11], Rödl and Ruciński introduced the following so-called *2-density measure*:

$$d_2(H) := \begin{cases} (e_H - 1)/(v_H - 2) & \text{if } H \text{ is non-empty with } v_H \geq 3, \\ 1/2 & \text{if } H \cong K_2, \\ 0 & \text{otherwise;} \end{cases}$$
$$m_2(H) := \max\{d_2(J) : J \subseteq H\}.$$

We say that H is *strictly 2-balanced* if for all proper subgraphs $J \subset H$, we have $d_2(J) < m_2(H)$.

Regarding asymmetric Ramsey properties, in [4], Kohayakawa and Kreuter introduced the following asymmetric versions of d_2 and m_2. Let H_1 and H_2 be any graphs, and define

$$d_2(H_1, H_2) := \begin{cases} \frac{e_1}{v_1 - 2 + \frac{1}{m_2(H_2)}} & \text{if } H_2 \text{ is non-empty and } v_1 \geq 2, \\ 0 & \text{otherwise;} \end{cases}$$
$$m_2(H_1, H_2) := \max\{d_2(J, H_2) : J \subseteq H_1\}.$$

We say that H_1 is *strictly balanced w.r.t.* $d_2(\cdot, H_2)$ if for all proper subgraphs $J \subset H_1$ we have $d_2(J, H_2) < m_2(H_1, H_2)$.

The relevance of strictly balanced graphs is as follows. Let H_1, H_2 be graphs with $m_2(H_1) \geq m_2(H_2)$. We call (H_1, H_2) a *heart* if

- H_2 is strictly 2-balanced,
- when $m_2(H_1) = m_2(H_2)$, H_1 is strictly 2-balanced,
- when $m_2(H_1) > m_2(H_2)$, H_1 is strictly balanced w.r.t. $d_2(\cdot, H_2)$.

It is easy to show that for any pair of graphs (H_1, H_2) with $m_2(H_1) \geq m_2(H_2)$, there exists a heart (H_1', H_2') with

- $H_i' \subseteq H_i$ for each $i \in [2]$,
- $m_2(H_2') = m_2(H_2)$,
- $m_2(H_1', H_2') = m_2(H_1, H_2)$ if $m_2(H_1) > m_2(H_2)$, and
- $m_2(H_1') = m_2(H_1)$ if $m_2(H_1) = m_2(H_2)$.

We call this pair a *heart of* (H_1, H_2). Now observe that in order to prove that $G \not\to (H_1, H_2)$, it suffices to prove that $G \not\to (H_1', H_2')$ for some heart (H_1', H_2') of (H_1, H_2), since any colouring avoiding a monochromatic copy of a subgraph of some H clearly avoids a monochromatic copy of H itself.

1.2 Previous and new results

We can now state the aforementioned symmetric random Ramsey theorem and asymmetric random Ramsey conjecture.

Theorem 1.1 (Rödl and Ruciński [11, 12, 13]). *Let $r \geq 2$ and let H be a non-empty graph such that at least one component of H is not a star or, when $r = 2$, a path on 3 edges. Then there exist positive constants $b, B > 0$ such that*

$$\lim_{n\to\infty} \mathbb{P}[G_{n,p} \to (\underbrace{H, \ldots, H}_{r \text{ times}})] = \begin{cases} 0 & \text{if } p \leq bn^{-1/m_2(H)}, \\ 1 & \text{if } p \geq Bn^{-1/m_2(H)}. \end{cases}$$

Note that the assumption on the structure of H is necessary, see e.g. [8] for details.

Conjecture 1.2 (Kohayakawa and Kreuter [4]). *Let $r \geq 2$ and suppose that $H_1, \ldots, H_r$ are non-empty graphs such that $m_2(H_1) \geq m_2(H_2) \geq \cdots \geq m_2(H_r)$ and $m_2(H_2) > 1$. Then there exist constants $b, B > 0$ such that*

$$\lim_{n\to\infty} \mathbb{P}[G_{n,p} \to (H_1, \ldots, H_r)] = \begin{cases} 0 & \text{if } p \leq bn^{-1/m_2(H_1,H_2)}, \\ 1 & \text{if } p \geq Bn^{-1/m_2(H_1,H_2)}. \end{cases}$$

This statement of the conjecture involves a slight rephrasing of the original statement as per [8], generalising from the case $r = 2$ and including the assumption of Kohayakawa, Schacht and Spöhel [5] that $m_2(H_2) > 1$. This is in order to avoid possible complications arising from H_2 (and/or H_1) being certain forests, such as those excluded in the statement of Theorem 1.1.

The progress on Conjecture 1.2 so far is as follows.

Theorem 1.3. *The* 1*-statement of Conjecture 1.2 holds* ([8])*. Further, the* 0*-statement of Conjecture 1.2 holds for* $(H_1, \dots, H_r)$ *in each of the following cases. For some heart* (H_1', H_2') *of* (H_1, H_2)*, we have that:*

(i) H_1' *and* H_2' *are both cycles* ([4])*;*

(ii) H_1' *and* H_2' *are both cliques* ([7])*;*

(iii) H_1' *is a clique and* H_2' *is a cycle* ([6])*;*

(iv) H_1' *and* H_2' *are a pair of regular graphs, excluding the cases when (a)* H_1' *is a clique and* H_2' *is a cycle; (b)* H_2' *is a cycle and* $v_1' \geq v_2'$*; (c)* $(H_1', H_2') = (K_3, K_{3,3})$ ([3])*.*

Note that (i)–(iii) above were only stated for (H_1, H_2) of the precise form of (H_1', H_2') stated (i.e. for (i), with H_1 and H_2 themselves both cycles). However, note that each such pair is a heart itself, so the theorem, via the remark at the end of Section 1.1, extends to the cases indicated.

Our main result is that we can vastly extend the number of cases for which the 0-statement holds:

Theorem 1.4. *The* 0*-statement of Conjecture 1.2 holds for the following cases:*

(i) When $r \geq 3$*, i.e. we have at least* 3 *graphs* H_1, H_2, H_3*;*

(ii) When $m_2(H_1) = m_2(H_2)$*;*

(iii) When there exists a heart (H_1', H_2') *of* (H_1, H_2) *such that* $\chi(H_2') \geq 3$ *or* $m(H_2') > 2$ *or* $ar(H_2') > 2$*.*

We remark that Kuperwasser and Samotij announced independently a proof of case (ii) above at Random Structures and Algorithms 2021/2022.

In the rest of this extended abstract, we shall outline the proof strategy of Theorem 1.4.

2 Proof strategy

Suppose that G is a graph with constant size and $m(G) \leq m_2(H_1, H_2)$. It is easy to show that for $p = cn^{-1/m_2(H_1,H_2)}$, with at least constant probability, G will appear as a subgraph of $G_{n,p}$. Therefore, it better be the case that $G \not\to (H_1, H_2)$. It is natural to ask whether this is in fact the only obstruction in proving a 0-statement.

Question 2.1. *Does it suffice to prove that for all G, H_1, H_2 with $m_2(H_1) \geq m_2(H_2) > 1$ and $m(G) \leq m_2(H_1, H_2)$ we have $G \not\to (H_1, H_2)$, in order to prove the 0-statement of Conjecture 1.2?*

Further, recalling the definition of hearts earlier, observe that we only need to prove such a statement for (H_1, H_2) which are hearts.

In the symmetric setting, the answer to this question is yes. Additionally, in [9], Nenadov et al. showed that this same phenomenon occurs for a number of symmetric Ramsey-style properties. Therefore, naturally, there have been attempts to answer this question in the asymmetric setting. The first result on this question was given by Gugelmann et al. [2], who additionally proved their result extends to the setting of k-uniform hypergraphs.

Theorem 2.2 ([2])**.** *Let (H_1, H_2) be a heart. If*

(i) a certain family of graphs $\mathcal{F}(H_1, H_2)$ is so-called 'asymmetric-balanced',

(ii) for all G such that $m(G) \leq m_2(H_1, H_2)$ then $G \not\to (H_1, H_2)$,

then the 0-statement holds for any pair of graphs with heart (H_1, H_2).

See [2] for the precise description of property (i). The next major step was made by the third author, who was inspired by the proof techniques used in [7] for proving Theorem 1.3(ii).

Theorem 2.3 ([3])**.** *Let (H_1, H_2) be a heart. If there exists $\varepsilon > 0$ such that*

(i) a certain family of graphs $\hat{\mathcal{A}}(H_1, H_2, \varepsilon)$ is finite,

(ii) for all $G \in \hat{\mathcal{A}}(H_1, H_2, \varepsilon)$ (which in particular satisfy $m(G) \leq m_2(H_1, H_2) + \varepsilon$) we have $G \not\to (H_1, H_2)$,

then the 0-statement holds for any pair of graphs with heart (H_1, H_2).

By streamlining the approach of the third author, we are able to prove the desired results.

Theorem 2.4. *Let (H_1, H_2) be a heart. There exists a family $\hat{\mathcal{B}}(H_1, H_2) \subseteq \hat{\mathcal{A}}(H_1, H_2, 0)$ such that if*

(i) $\hat{\mathcal{B}}(H_1, H_2)$ is finite,

(ii) for all $G \in \hat{\mathcal{B}}(H_1, H_2)$ we have $G \not\to (H_1, H_2)$,

then the 0-statement holds for any pair of graphs with heart (H_1, H_2).

Theorem 2.5. *Let (H_1, H_2) be a heart. Then $\hat{\mathcal{B}}(H_1, H_2)$ is finite.*

Combining Theorems 2.4 and 2.5 shows that the answer to Question 2.1 is yes. In the next section we will give a description of the families $\hat{\mathcal{A}}(H_1, H_2, \varepsilon)$ and $\hat{\mathcal{B}}(H_1, H_2)$ in the case of $m_2(H_1) = m_2(H_2)$.

Now it suffices to prove the colouring result contained in Question 2.1. Note that in the symmetric setting, this colouring result holds in all cases and has a short proof (see e.g. Theorem 3.2 in [10]). By appropriately generalising this result, we can prove the following cases of the asymmetric statement.

Lemma 2.6. *For all G, H_1, H_2 with $m_2(H_1) \geq m_2(H_2) > 1$ and $m(G) \leq m_2(H_1, H_2)$, we have $G \not\to (H_1, \ldots, H_r)$ if any of the following conditions are satisfied:*

(i) We have $r \geq 3$, i.e. at least 3 graphs H_1, H_2, H_3;

(ii) We have $m_1(H_2) = m_2(H_2)$;

(iii) We have $\chi(H_2) \geq 3$ or $m(H_2) > 2$ or $ar(H_2) > 2$.

Theorem 1.4 immediately follows from Theorems 2.4 and 2.5 combined with Lemma 2.6. Note that the assumption in Theorem 2.5 that (H_1, H_2) is a heart is actually necessary. This leads to the slightly technical nature of the set of graphs given in Theorem 1.4.

2.1 More details on Theorems 2.4 and 2.5

In the case where $m_2(H_1) = m_2(H_2)$, we have $G \in \hat{\mathcal{A}}(H_1, H_2, \varepsilon)$ if G satisfies

- every edge $e = E(R) \cap E(L)$ for some $R \cong H_1$ and $L \cong H_2$, where $L, R \subseteq G$;
- $m(G) \leq m_2(H_1, H_2) + \varepsilon$;
- G is 2-connected.

For $\ell \geq 4$, define $C_\ell^{K_4}$ to be the graph on 3ℓ vertices and 6ℓ edges obtained by taking a cycle C_ℓ and extending each of its edges to a copy of K_4. This graph satisfies that every edge is the intersection of two triangles, $m(C_\ell^{K_4}) = m_2(K_3, K_3) = 2$ and is 2-connected, and therefore the family $\hat{\mathcal{A}}(K_3, K_3, \varepsilon)$ is not finite for any $\varepsilon > 0$.

The key idea is to refine the family $\hat{\mathcal{A}}(H_1, H_2, \varepsilon)$ so that graphs such as $C_\ell^{K_4}$ are excluded. We now describe $\hat{\mathcal{B}}(H_1, H_2)$ in the case where $m_2(H_1) = m_2(H_2)$. Call an edge e *open in G* if $e \neq E(R) \cap E(L)$ for any $R \cong H_1$ and $L \cong H_2$, where $R, L \subseteq G$. Define $\lambda(G) := v_G - e_G / m_2(H_1, H_2)$. We have that $\hat{\mathcal{B}}(H_1, H_2)$ is the collection of all outputs G that can be returned in the running of algorithm Grow-$\hat{\mathcal{B}}$-Alt (see the figure below) which additionally satisfy $m(G) \leq m_2(H_1, H_2)$.

It is not too hard to see that $\hat{\mathcal{B}}(H_1, H_2) \subseteq \hat{\mathcal{A}}(H_1, H_2, 0)$, and further, $C_\ell^{K_4} \notin \hat{\mathcal{B}}(K_3, K_3)$.

The proof of Theorem 2.4 follows from a careful analysis which is very similar to the proof of Theorem 2.3 in [3]. So finally, we summarise how we prove Theorem 2.5.

procedure Grow-$\hat{\mathcal{B}}$-Alt(H_1, H_2)
  $G_0 \leftarrow H_1$
  $i \leftarrow 0$
  while $\lambda(G_i) \geq 0$ **do**
    if $\exists e \in E(G_i)$ s.t. e is open **then**
      $\{L, R\} \leftarrow$ any pair $\{L, R\}$ s.t. $L \cong H_2$, $R \cong H_1$ and $E(L) \cap E(R) = \{e\}$
      $G_{i+1} \leftarrow G_i \cup L \cup R$
      $i \leftarrow i + 1$
    else
      return G_i
      $e \leftarrow$ any edge of G_i
      $\{L, R\} \leftarrow$ any pair $\{L, R\}$ s.t. $L \cong H_2$, $R \cong H_1$, $E(L) \cap E(R) = \{e\}$ and $E(L) \cup E(R) \not\subseteq E(G_i)$
      $G_{i+1} \leftarrow G_i \cup L \cup R$
      $i \leftarrow i + 1$
    end if
  end while
end procedure

Figure 1: The implementation of algorithm Grow-$\hat{\mathcal{B}}$-Alt.

Let $\eta(G)$ be the number of open edges in G. First note that if G is an output of the algorithm Grow-$\hat{\mathcal{B}}$-Alt, then it satisfies $\lambda(G) \geq 0$ and $\eta(G) = 0$. Suppose the following is true:

There exist constants $\kappa, x, y > 0$ depending only on H_1 and H_2 such that in each iteration of the algorithm Grow-$\hat{\mathcal{B}}$-Alt, we either have:

(I) $\lambda(G_i) \leq \lambda(G_{i-1}) - \kappa$ and $\eta(G_i) \geq \eta(G_{i-1}) - x$;

(II) $\lambda(G_i) = \lambda(G_{i-1})$ and $\eta(G_i) \geq \eta(G_{i-1}) + y$.

Then, letting T_1^i and T_2^i count the number of iterations of type I and type II, respectively, to construct G_i, we obtain $\eta(G_i) \geq T_2^i \cdot y - T_1^i \cdot x$. Overall this implies that the number of outputs of the algorithm Grow-$\hat{\mathcal{B}}$-Alt is finite, as required. This is the essence of how we prove finiteness of $\hat{\mathcal{B}}$, however our actual approach involves more technical definitions we wish to avoid here.

For the case where $m_2(H_1) > m_2(H_2)$, the algorithm describing the family $\hat{\mathcal{B}}(H_1, H_2)$ is more complicated, but the overall idea is the same.

References

[1] B. Bollobás and A.G. Thomason, Threshold functions, *Combinatorica* **7** (1987), 35–38.

[2] L. Gugelmann, R. Nenadov, Y. Person, N. Škorić, A. Steger and H. Thomas, Symmetric and asymmetric Ramsey properties in random hypergraphs, *Forum Math. Sigma* **5** (2017).

[3] J. Hyde, Towards the 0-statement of the Kohayakawa-Kreuter conjecture, *Comb. Probab. Comput.* **32** (2021), 225–268.

[4] Y. Kohayakawa and B. Kreuter, Threshold functions for asymmetric Ramsey properties involving cycles, *Random Structures Algorithms* **11** (1997), 245–276.

[5] Y. Kohayakawa, M. Schacht and R. Spöhel, Upper bounds on probability thresholds for asymmetric Ramsey properties, *Random Structures Algorithms* **44(1)** (2014), 1–28.

[6] A. Liebenau, L. Mattos, W. Mendonça and J. Skokan, Asymmetric Ramsey properties of random graphs for cliques and cycles, *Random Structures Algorithms*, to appear.

[7] M. Marciniszyn, J. Skokan, R. Spöhel and A. Steger, Asymmetric Ramsey properties of random graphs involving cliques, *Random Structures Algorithms* **34(4)** (2009), 419–453.

[8] F. Mousset, R. Nenadov and W. Samotij, Towards the Kohayakawa–Kreuter conjecture on asymmetric Ramsey properties, *Combin. Probab. Comput.* **29(6)** (2020), 943–955.

[9] R. Nenadov, Y. Person, N. Škorić and A. Steger, An algorithmic framework for obtaining lower bounds for random Ramsey problems, *J. Combin. Theory Ser. B* **124** (2017), 1–38.

[10] R. Nenadov and A. Steger, A short proof of the random Ramsey theorem, *Combin. Probab. Comput.* **25(1)** (2016), 130-144.

[11] V. Rödl and A. Ruciński, Lower bounds on probability thresholds for Ramsey properties, *Combinatorics, Paul Erdos is eighty, Vol. 1, Bolyai Soc. Math. Stud., János Bolyai Math. Soc., Budapest* (1993), 317–346.

[12] V. Rödl and A. Ruciński, Random graphs with monochromatic triangles in every edge coloring, *Random Structures Algorithms* **5** (1994), 253–270.

[13] V. Rödl and A. Ruciński, Threshold functions for Ramsey properties, *J. Amer. Math. Soc.* **8** (1995), 917–942.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Tight path, what is it (Ramsey-)good for? Absolutely (almost) nothing!

(Extended abstract)

Simona Boyadzhiyska* Allan Lo*

Abstract

Given a pair of k-uniform hypergraphs (G, H), the *Ramsey number* of (G, H), denoted by $R(G, H)$, is the smallest integer n such that in every red/blue-colouring of the edges of $K_n^{(k)}$ there exists a red copy of G or a blue copy of H. Burr showed that, for any pair of graphs (G, H), where G is large and connected, the Ramsey number $R(G, H)$ is bounded below by $(v(G) - 1)(\chi(H) - 1) + \sigma(H)$, where $\sigma(H)$ stands for the minimum size of a colour class over all proper $\chi(H)$-colourings of H. Together with Erdős, he then asked when this lower bound is attained, introducing the notion of Ramsey goodness and its systematic study. We say that G is H*-good* if the Ramsey number of (G, H) is equal to the general lower bound. Among other results, it was shown by Burr that, for any graph H, every sufficiently long path is H-good.

Our goal is to explore the notion of Ramsey goodness in the setting of 3-uniform hypergraphs. Motivated by Burr's result concerning paths and a recent result of Balogh, Clemen, Skokan, and Wagner, we ask: what 3-graphs H is a (long) tight path good for? We demonstrate that, in stark contrast to the graph case, long tight paths are generally not H-good for various types of 3-graphs H. Even more, we show that the ratio $R(P_n, H)/n$ for a pair (P_n, H) consisting of a tight path on n vertices and a 3-graph H cannot in general be bounded above by *any* function depending only on $\chi(H)$. We complement these negative results with a positive one, determining the Ramsey number asymptotically for pairs (P_n, H) when H belongs to a certain family of hypergraphs.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-026`

*School of Mathematics, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK. E-mail: `s.s.boyadzhiyska@bham.ac.uk, s.a.lo@bham.ac.uk`. The research leading to these results was supported by EPSRC, grant no. EP/V002279/1 (A. Lo) and EP/V048287/1 (A. Lo and S. Boyadzhiyska). There are no additional data beyond that contained within the main manuscript.

1 Introduction

A k-uniform hypergraph H, or a k-graph for short, consists of a (finite) set $V(H)$ of vertices and a set $E(H)$ of k-element subsets of $V(H)$, called (hyper)edges. Given k-graphs G and H, the *Ramsey number* of the pair (G, H), denoted by $R(G, H)$, is the smallest integer n such that, in every red/blue-colouring of the edges of the complete k-graph $K_n^{(k)}$, we can find a red copy of G or a blue copy of H. Ramsey's seminal result [14] implies that $R(G, H)$ is finite for any pair of k-graphs G and H. Since then, the study of Ramsey numbers has become a prominent area of research in combinatorics and has inspired the development of many powerful tools in the field (see for example [9, 13] and the references therein).

Even in the simplest setting, when the uniformity is two, Ramsey numbers are often notoriously difficult to understand. The most well-studied case is when $G = H = K_t$. It is known from the early work of Erdős [10] and Erdős and Szekeres [11] that, up to lower order terms, $2^{t/2} \leq R(K_t, K_t) \leq 2^{2t}$ as $t \to \infty$; these bounds remained essentially best possible for several decades, until very recently Campos, Griffiths, Morris, and Sahasrabudhe [6] announced the first exponential improvement in the upper bound.

Apart from demonstrating the difficulty of understanding Ramsey numbers, this example shows that Ramsey numbers can grow very quickly compared to $v(G)$ and $v(H)$. It is then natural to ask: how *small* can Ramsey numbers be? Here we will always assume that G is connected. As shown by Burr [4], following a slightly weaker observation by Chvátal and Harary [8], for any G and H with $v(G) \geq \sigma(H)$[1], we have

$$R(G, H) \geq (v(G) - 1)(\chi(H) - 1) + \sigma(H). \tag{1}$$

Indeed, colour the complete graph of order $(v(G) - 1)(\chi(H) - 1) + \sigma(H) - 1$ so that the red edges form $\chi(H)$ cliques, one of order $\sigma(H) - 1$ and the rest of order $v(G) - 1$; it is not difficult to check that there is neither a red copy of G nor a blue copy of H in this colouring. A classic result of Chvátal [7] shows that the bound in (1) is attained with equality when the pair consists of a tree and a complete graph. Motivated by this result, Burr [4] and Burr and Erdős [5] investigated what other pairs have this property, introducing the notion of *Ramsey goodness*. More precisely, a graph G is said to be *H-good* if the lower bound in (1) is attained for the pair (G, H). Since its introduction this notion has received considerable attention (see [9, Section 2.5] and the references therein for some history and results). Typically in this line of research H is thought of as a fixed graph and the task is to identify what properties make a (sufficiently large) graph H-good. Several conjectures were made (for example, by Burr [4] and Burr and Erdős [5]), suggesting that, for a fixed graph H, every sufficiently sparse large graph G should be H-good. These conjectures turned out to be false in general, as shown by Brandt [3]. On the other hand, it is known that there are some families of graphs such that every sufficiently large member is H-good for every H. In particular, Burr [4] showed that, for any graph H, any sufficiently long path is H-good.

[1] Here $\sigma(H)$ is the smallest possible size of a colour class in a proper colouring of H using $\chi(H)$ colours.

More generally, Allen, Brightwell, and Skokan [1] showed that, for every fixed H, every large graph with bounded bandwidth is H-good.

We are interested in exploring the notion of Ramsey goodness for hypergraphs. Again, if G is connected[2], the lower bound in (1)[3] holds for all k-graphs H with $v(G) \geq \sigma(H)$. We say that G is *H-good* if equality holds in (1).

The study of Ramsey goodness in hypergraphs was first undertaken by Balogh, Clemen, Skokan, and Wagner [2] and was motivated by a question of Conlon. The n-vertex 3-uniform *tight path* P_n consists of n vertices $v_1, \ldots, v_n$ and hyperedges given by $v_i v_{i+1} v_{i+2}$ for all $i \in [n-2]$. Letting $\mathbb{F}$ denote the *Fano plane*, that is, the unique 3-graph on seven vertices in which every pair of vertices is contained in a unique edge, Conlon asked what 3-graphs are $\mathbb{F}$-good. Balogh, Clemen, Skokan, and Wagner [2] made progress towards answering this question by showing that any sufficiently long tight path is $\mathbb{F}$-good. In light of their work and Burr's result for paths in the graph case [4], we seek to identify what hypergraphs a tight path is good for. We focus specifically on 3-uniform hypergraphs.

2 Results

For a 3-graph H, we say that H is *Ramsey-good for tight paths* if every sufficiently long tight path is H-good. Perhaps surprisingly, it turns out that there are plenty of classes of 3-graphs which are not Ramsey-good for tight paths. It is not difficult to check that the Fano plane $\mathbb{F}$ can be properly 3-coloured so that each hyperedge intersects precisely two different colour classes and that $\chi(\mathbb{F}) = 3$ and $\sigma(\mathbb{F}) = 1$. The first property is crucial, as demonstrated by the following proposition.

Proposition 2.1. *Let H be a 3-graph with $\chi(H) = 3$ and $n \geq 3\sigma(H) + 3$. Assume that in every proper 3-colouring of H, there exists an edge intersecting all three colour classes. Then $R(P_n, H) \geq 2(n-1) + \lfloor \frac{1}{3}n \rfloor > 2(n-1) + \sigma(H)$.*

Proof. Let $N = 2(n-1) + \lfloor \frac{1}{3}n \rfloor - 1$. We first partition the vertex set of $K = K_N^{(3)}$ into sets V_1, V_2, V_3 satisfying $|V_1| = n - 1 = |V_2|$ and $|V_3| = \lfloor \frac{1}{3}n \rfloor - 1$. We then colour every hyperedge intersecting exactly two different sets V_i blue and every other hyperedge red.

Suppose there exists a red tight path P on n vertices. Then P contains a matching of size $\lfloor \frac{1}{3}n \rfloor > |V_3|$, so one of the matching edges does not intersect V_3. This edge must then be fully contained in some V_i for $i \in [2]$, which in turn implies that P is fully contained in this V_i. Hence $v(P) \leq |V_i| < n$, a contradiction.

To see why there is no blue copy of H, note that, since $\chi(H) = 3$, any blue copy of H in K must intersect all three sets V_i. But in every proper 3-colouring of H some edge intersects all three colour classes. Since all edges intersecting all three sets V_i are red, there cannot exist a blue copy of H in K. $\square$

[2]We say that G is connected if G is not a disjoint union of two smaller hypergraphs.

[3]As usual, a proper colouring of a hypergraph H is a colouring of the vertices of H such that no edge of H is monochromatic; $\chi(H)$ is the minimum number of colours in a proper colouring of H, and $\sigma(H)$ is the smallest possible size of a colour class in a proper colouring of H using $\chi(H)$ colours.

It is possible to obtain a result similar to Proposition 2.1 also when $\chi(H) > 3$. Thus, from now on, we concentrate on hypergraphs H that have at least one $\chi(H)$-colouring in which every edge intersects precisely two different colour classes. In fact, we restrict our attention to a special subclass of hypergraphs of this kind, which we define below.

Definition 1. Let $\chi \geq 1$ be an integer and T_χ be a tournament on $[\chi]$. We say that a 3-graph H is a *tournament hypergraph associated to* T_χ if $V(H)$ can be partitioned into sets $A_1 \cup \cdots \cup A_\chi$ so that $E(H) = \{xyz : x, y \in A_i, z \in A_j, (i,j) \in E(T_\chi)\}$, that is, the edge set of H consists of precisely those triples containing two vertices from some set A_i and a third vertex from some set A_j, where (i,j) is an arc of T_χ. For an integer $m \geq 1$, we write $H(T_\chi, m)$ for a tournament hypergraph associated to T_χ in which each vertex class A_i has size m.

Let $\chi \geq 1$ be an integer, T_χ be a non-transitive tournament on $[\chi]$, and $H = H(T_\chi, m)$. It turns out that, in this case, not only is H not Ramsey-good for tight paths, but in fact the ratio $R(P_n, H)/n$ cannot be bounded above by any function depending only on χ. This is the content of the next proposition.

Proposition 2.2. *Let $\chi \geq 3$ and $m \geq 2$ be integers and T_χ be a non-transitive tournament on $[\chi]$. Let $n, t \geq 1$ be integers such that $\lfloor \frac{3t}{2} \rfloor + 1 < n$. Then $R(P_n, H(T_\chi, m)) \geq (m-1)t+1$.*

Proof. Let $N = (m-1)t$. We partition the vertex set of $K = K_N^{(3)}$ into sets $V_1, \ldots, V_{m-1}$ satisfying $|V_i| = t$ for all $i \in [m-1]$. We then colour every hyperedge xyz with $x, y \in V_i$ and $z \in V_j$ for $1 \leq i \leq j \leq m-1$ red and every other hyperedge blue.

It is not difficult to see that a red tight path in this colouring has at most $n-1$ vertices. Indeed, any red tight path must contain either vertices from a single V_i, in which case it has at most $t < n$ vertices, or b vertices from a set V_i and at most $\lfloor \frac{b}{2} \rfloor + 1$ vertices from $V_{i+1} \cup \cdots \cup V_{m-1}$, in which case its number of vertices cannot exceed $t + \lfloor \frac{t}{2} \rfloor + 1 < n$.

Now suppose there is a blue copy H' of H in K with vertex classes $W_1, \ldots, W_\chi$. For each $j \in [\chi]$, we have $|W_j| = m$, and thus there exists an index $k_j \in [m-1]$ such that $|W_j \cap V_{k_j}| \geq 2$. Note that, since the edges fully contained in a single set V_i are red, for every arc (j, ℓ) of T_χ, no set V_i can contain three vertices x, y, z with $x, y \in W_j$ and $z \in W_\ell$. Therefore, all k_j are distinct. But then by the definition of our colouring $H'\left[\bigcup_{j \in [\chi]} V_{k_j}\right]$ is a tournament hypergraph associated to a transitive tournament, which contradicts the fact that H is associated to a non-transitive tournament. □

Observe that the proof of Proposition 2.2 shows that the same result holds if H is associated (in a similar way as in Definition 1) to any digraph containing a cycle.

The situation is fairly different when H is a tournament hypergraph associated to a *transitive* tournament. We write TT_ℓ for the transitive tournament on $[\ell]$. Once again, H is generally not Ramsey-good for tight paths, but as we will soon see, in this case $R(P_n, H)/n$ *can* be bounded above by a function depending only on $\chi(H)$. Given an integer $\ell \geq 1$, let $\vec{R}(\ell)$ denote the smallest integer N such that any tournament on at least N vertices contains a copy of TT_ℓ. It is well known that $\vec{R}(\ell)$ is finite for any $\ell \geq 1$.

Proposition 2.3. *Let $\chi \geq 3$ be an integer, $R = \vec{R}(\chi)$, and $m \geq R$. Then $H = H(TT_\chi, m)$ satisfies $R(P_n, H) \geq \left(\frac{2}{3}n - 6\right)(R-1) + 1 = (1+o(1))\frac{2}{3}(R-1)n$ as $n \to \infty$.*

Proof. Let T_{R-1} be a tournament on vertex set $[R-1]$ that does not contain a copy of TT_χ, which exists by the definition of R. Let $N = \left(\left\lfloor \frac{2}{3}n \right\rfloor - 5\right)(R-1) \geq \left(\frac{2}{3}n - 6\right)(R-1)$ and $K = K_N^{(3)}$. Partition the vertex set of K into sets $V_1, \ldots, V_{R-1}$ with $|V_i| = \left\lfloor \frac{2}{3}n \right\rfloor - 5$ for all $i \in [R-1]$. We now assign the colour red to all edges that are fully contained in a single set V_i and all edges of the form xyz for $x \in V_i$ and $y, z \in V_j$, where (i, j) is an arc of T_{R-1}. All remaining edges are coloured blue. Note in particular that the blue edges intersecting precisely two vertex classes form a copy of $H(T_{R-1}, \left\lfloor \frac{2}{3}n \right\rfloor - 5)$.

Using a similar argument as in the proof of Proposition 2.2, we conclude that there is no red tight path on n vertices. Suppose there exists a blue copy H' of H with vertex classes $W_1, \ldots, W_\chi$. Since $|W_j| \geq R$ for each $j \in [\chi]$, there exists an integer $k_j \in [R-1]$ such that $|W_j \cap V_{k_j}| \geq 2$. As before, all of these k_j are distinct. But then the hypergraph $H'\left[\bigcup_{j \in [\chi]} V_{k_j}\right]$ is a tournament hypergraph associated to TT_χ. But T_{R-1} does not contain a copy of TT_χ, a contradiction. □

It turns out that the lower bound in Proposition 2.3 is asymptotically tight as $n \to \infty$. More precisely, we are able to prove the following theorem.

Theorem 2.4. *Given integers $\chi \geq 2$ and $m \geq 2$ and a real number $\varepsilon > 0$, there exists an integer $n_0 = n_0(\chi, m, \varepsilon)$ such that, for all $n \geq n_0$,*

$$R(P_n, H(TT_\chi, m)) \leq \begin{cases} (1+\varepsilon)n & \text{if } \chi = 2, \\ \left(\frac{2}{3} + \varepsilon\right)(\vec{R}(\chi) - 1)n & \text{if } \chi \geq 3. \end{cases}$$

Since $\vec{R}(3) = 4$, Proposition 2.3 and Theorem 2.4 imply that $R(P_n, H(TT_3, m)) = (2 + o(1))n$ as $n \to \infty$. This means that P_n is *asymptotically* $H(TT_3, m)$-good as $n \to \infty$. In particular, since the Fano plane is a subhypergraph of $H(TT_3, 4)$, Theorem 2.4 extends the result of Balogh, Clemen, Skokan, and Wagner [2] asymptotically to a large family of 3-graphs.

We provide a brief sketch of the proof of Theorem 2.4. Some of the ideas resemble those used in [2]. The proof uses induction on the chromatic number χ. We outline the induction step. Suppose $\chi \geq 3$ and that the theorem holds for $\chi - 1$. Let $\varepsilon > 0$ and $m \geq 2$ be given and $H = H(TT_\chi, m)$. Set $N = \left(\frac{2}{3} + \varepsilon\right)(\vec{R}(\chi) - 1)n$ and suppose there is a colouring of $K_N^{(3)}$ with no red copy of P_n and no blue copy of H.

We first find a red tight path P of length approximately $\frac{2}{3}n$ with a special property: there exist disjoint intervals $I_1, \ldots, I_c$ covering most vertices of P such that the vertices of each interval induce a red clique. Our task is then to absorb more vertices from the rest of $K_N^{(3)}$ in between the vertices of each interval I_j. A key idea here is that, since the vertices of each I_j form a clique, we can change the order in which they appear on the path. We then go through the intervals I_j in turn and repeatedly apply the induction hypothesis to $K_N^{(3)} \setminus V(P)$ to find copies of $H(TT_{\chi-1}, m')$ for some appropriately chosen large constant m'. For each such copy H' of $H(TT_{\chi-1}, m')$, either there will be a lot of blue edges with two vertices in I_j and a third vertex in $V(H')$, in which case we can embed a copy of H, or we will find enough red edges of this kind to allow us to absorb a number of vertices from $V(H')$ into the interval I_j (after possibly rearranging the vertices of I_j). Eventually, unless we find a blue copy of H, we will be able to absorb approximately $\frac{1}{2}|I_j|$ vertices into each interval I_j, resulting in a tight path of total length at least n.

3 Conclusion and open problems

A number of natural questions arise from our work.

First of all, it would be interesting to determine the Ramsey numbers of more pairs of the form (P_n, H), for instance, those discussed in Propositions 2.1 and 2.2, at least asymptotically. Similarly, a natural way to improve Theorem 2.4 and Proposition 2.3 is to remove the error term and prove a precise result.

In a slightly different direction, in the examples given in Propositions 2.2 and 2.3, our tournament hypergraphs are fairly dense. It would be interesting to consider subhypergraphs of tournament hypergraphs and investigate how sparse such a subgraph H can be made before (P_n, H) meets the lower bound. We are able to find reasonably sparse hypergraphs H, albeit not subgraphs of tournament hypergraphs, such that (P_n, H) exceeds the general lower bound.

A third possible direction for further research is to consider higher uniformities. Do long k-uniform tight paths behave similarly to 2-uniform paths or 3-uniform tight paths as k increases? We note here that we tried to generalise the result of Balogh, Clemen, Skokan, and Wagner in a different direction, by replacing the Fano plane by a higher-order projective plane $\mathbb{F}^q$ for some prime power q. Surprisingly, long tight paths are generally not $\mathbb{F}^q$-good. It is a simple exercise to show that, when $q \geq 3$, we have $\chi(\mathbb{F}^q) = 2$. Then a result of Keevash and Zhao [12] allows us to build colourings showing that $\mathbb{F}^q$ is not Ramsey-good for tight paths for an infinite sequence of values of q.

References

[1] P. Allen, G. Brightwell, and J. Skokan. Ramsey-goodness – and otherwise. *Combinatorica*, 33:125–160, 2013.

[2] J. Balogh, F. C. Clemen, J. Skokan, and A. Z. Wagner. The Ramsey number of the Fano plane versus the tight path. *Electron. J. Combin.*, 27(1):P1.60, 2020.

[3] S. Brandt. *Expanding graphs and Ramsey numbers.* Freie Univ., Fachbereich Mathematik und Informatik, 1996.

[4] S. A. Burr. Ramsey numbers involving graphs with long suspended paths. *J. London Math. Soc.*, 2(3):405–413, 1981.

[5] S. A. Burr and P. Erdős. Generalizations of a Ramsey-theoretic result of Chvátal. *J. Graph Theory*, 7(1):39–51, 1983.

[6] M. Campos, S. Griffiths, R. Morris, and J. Sahasrabudhe. An exponential improvement for diagonal Ramsey. *arXiv:2303.09521*, 2023.

[7] V. Chvátal. Tree-complete graph Ramsey numbers. *J. Graph Theory*, 1(1):93, 1977.

[8] V. Chvátal and F. Harary. Generalized Ramsey theory for graphs. III. Small off-diagonal numbers. *Pacific J. Math.*, 41(2):335–345, 1972.

[9] D. Conlon, J. Fox, and B. Sudakov. Recent developments in graph Ramsey theory. In *Surveys in combinatorics 2015*, volume 424 of *London Math. Soc. Lecture Note Ser.*, pages 49–118. Cambridge Univ. Press, Cambridge, 2015.

[10] P. Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.*, 53:292–294, 1947.

[11] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.

[12] P. Keevash and Y. Zhao. Codegree problems for projective geometries. *J. Combin. Theory Ser. B*, 97(6):919–928, 2007.

[13] D. Mubayi and A. Suk. A survey of hypergraph Ramsey problems. *Discrete Math. Appl.*, 165:405–428, 2020.

[14] F. P. Ramsey. On a problem in formal logic. *Proc. Lond. Math. Soc.*, 30:264–286, 1930.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Effective bounds for induced size-Ramsey numbers of cycles

(Extended abstract)

Domagoj Bradač* Nemanja Draganić* Benny Sudakov*

Abstract

The induced size-Ramsey number $\hat{r}^k_{\text{ind}}(H)$ of a graph H is the smallest number of edges a (host) graph G can have such that for any k-coloring of its edges, there exists a monochromatic copy of H which is an induced subgraph of G. In 1995, in their seminal paper, Haxell, Kohayakawa and Łuczak showed that for cycles, these numbers are linear for any constant number of colours, i.e., $\hat{r}^k_{\text{ind}}(C_n) \leq Cn$ for some $C = C(k)$. The constant C comes from the use of the regularity lemma, and has a tower type dependence on k. In this paper we significantly improve these bounds, showing that $\hat{r}^k_{\text{ind}}(C_n) \leq O(k^{102})n$ when n is even, thus obtaining only a polynomial dependence of C on k. We also prove $\hat{r}^k_{\text{ind}}(C_n) \leq e^{O(k \log k)}n$ for odd n, which almost matches the lower bound of $e^{\Omega(k)}n$. Finally, we show that the ordinary (non-induced) size-Ramsey number satisfies $\hat{r}^k(C_n) = e^{O(k)}n$ for odd n. This substantially improves the best previous result of $e^{O(k^2)}n$, and is best possible, up to the implied constant in the exponent. To achieve our results, we present a new host graph construction which, roughly speaking, reduces our task to finding a cycle of approximate given length in a graph with local sparsity.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-027`

1 Introduction

The Ramsey number $r^k(H)$ of a graph is the smallest integer n such that every k-coloring of the edges of K_n contains a monochromatic copy of H. The notion of Ramsey numbers

*Department of Mathematics, ETH, Zürich, Switzerland. Research supported in part by SNSF grant 200021_196965. E-mail: `{domagoj.bradac, nemanja.draganic, benjamin.sudakov}@math.ethz.ch`.

is one of the most central notions in combinatorics and it has been studied extensively since Ramsey [23] showed their existence for every graph H. Motivated by this definition, we say that a graph G is k-Ramsey for a graph H if any k-coloring of the edges of (the *host graph*) G, contains a monochromatic copy of H, and we write $G \xrightarrow{k} H$. Using this notation, we have that $r^k(H) = \min\{|V(G)| : G \xrightarrow{k} H\}$.

The notion of Ramsey numbers is measuring the minimality of the host graph in terms of the number of vertices. Are there graphs G with significantly fewer edges than the clique on $r^k(H)$ vertices that are k-Ramsey for H? This general question is captured by the notion of size-Ramsey numbers introduced in 1978 by Erdős, Faudree, Rousseau and Schelp [11]. The size-Ramsey number of a graph H is defined as $\hat{r}^k = \min\{E(G)|G \xrightarrow{k} H\}$. In the last few decades, there has been extensive research on this notion, see, e.g., [3].

One of the main goals is to understand which classes of graphs have size-Ramsey numbers which are linear in their number of edges. Beck [2] showed that this is true for paths, which was later extended to all bounded-degree trees by Friedman and Pippenger [14]. It is also known that logarithmic subdivisions of bounded degree graphs have linear size-Ramsey numbers [6], as well as bounded degree graphs with bounded treewidth [18]. Given all of the mentioned results, it might be tempting to assume that all graphs of bounded degree have linear size-Ramsey numbers. In an elegant paper of Rödl and Szemerédi [25], it was shown that this is not the case. Indeed, they showed that there exist n-vertex cubic graphs which have size-Ramsey numbers at least $n \log^c n$, for a small constant $c > 0$. This bound has only very recently been improved to $cne^{c\sqrt{\log n}}$ for some $c > 0$ by Tikhomirov [26]. For more results see [7] and references therein.

A related studied notion is that of induced size-Ramsey numbers. Given a graph H, the induced size-Ramsey number $\hat{r}^k_{ind}(H)$ is the smallest number of edges a graph G can have such that any k-coloring of G contains a monochromatic copy of H which is an induced subgraph of G. The existence of these numbers is an important generalisation of Ramsey's theorem, proved independently by Deuber [4], Erdős, Hajnal, and Pósa [12], and Rödl [24]. Naturally, this concept is much harder to understand for most classes of target graphs H and much less precise bounds are known than for the (non-induced) size-Ramsey number.

Indeed, already for bounded degree trees we know that the size-Ramsey number is linear in their number of vertices, whereas for its induced counterpart we have no good bounds while we have every reason to believe that the answer should also be linear. Further, the best general upper bound on $\hat{r}^2_{ind}(H)$ for n-vertex graphs H is obtained by Conlon, Fox and Sudakov [19], and is of the order $2^{O(n \log n)}$, while Erdős [10] conjectured that $\hat{r}^2_{ind}(H) \leq 2^{cn}$. In comparison, the bound for Ramsey numbers (and hence also for size-Ramsey numbers) is known to be exponential in the number of vertices of the target graph. Further, it is known that the size-Ramsey number of n-vertex graphs with degree bounded by a constant d, is between $ne^{\Omega(\sqrt{\log n})}$ and $O(n^{2-\frac{1}{d}+\varepsilon})$, proven by Tikhomirov [26], and by Kohayakawa, Rödl, Schacht, and Szemerédi [20], respectively. On the other hand, the best upper bound on the induced size-Ramsey number of these graphs, proved by Fox and Sudakov [13] is of the order $n^{O(d \log d)}$, while the best lower bound is still the bound for the (non-induced) Ramsey number of those graphs, which is often the state of the art for such questions.

For paths it is known that $\Omega(k^2)n \leq \hat{r}^k(P_n) \leq O(k^2 \log k)n$ (see [9, 21] for the lower bound and [22, 8] for the upper bound). In the induced case, by a recent result of Draganić, Krivelevich and Glock [5], we have that $\hat{r}^k_{ind}(P_n) \leq O(k^3 \log^4 k)n$. For cycles, the discrepancy between the size-Ramsey and the induced size-Ramsey number is significantly larger. Indeed, by a recent result of Javadi and Miralaei [17], which improved another recent result by Javadi, Khoeini, Omidi and Pokrovskiy [16], we have $\hat{r}^k(C_n) = O(k^{120} \log^2 k)n$ for even n, and $\hat{r}^k(C_n) = O(2^{16k^2+2\log k})n$ for odd n. On the other hand, the only known upper bound on the induced size-Ramsey numbers of cycles was obtained in the seminal paper of Haxell, Kohayakawa and Łuczak [15]. Their proof uses a technically very involved argument relying on the use of the Sparse Regularity lemma and therefore shows that $\hat{r}^k_{ind}(C_n) \leq Cn$ where $C = C(k)$ has a tower type dependence on k.

In this paper, we prove the following theorem which quite significantly improves the tower-type bounds of Haxell, Kohayakawa and Łuczak.

Theorem 1.1. *For any integer $k \geq 1$, there exists $n_0(k)$ such that for all $n \geq n_0(k)$, the following holds.*

a) If n is even, then $\hat{r}^k_{ind}(C_n) = O(k^{102})n$.

b) If n is odd, then $\hat{r}^k_{ind}(C_n) = e^{O(k \log k)}n$.

While the focus of this paper is on induced size-Ramsey numbers of cycles, our method can be also used to substantially improve the upper bound for the non-induced case as well. Our next result gives an essentially tight estimate for the size-Ramsey numbers of odd cycles.

Theorem 1.2. *For any integer $k \geq 1$, there exists $n_0(k)$ such that for all $n \geq n_0(k)$, we have $\hat{r}^k(C_n) = e^{O(k)}n$.*

The best known lower bound for size-Ramsey numbers of even cycles comes from the bound for paths, which is of the order $\Omega(k^2)n$ [9, 22]. In the odd case, there is a simple construction of a coloring which gives a lower bound of $2^{k-1}n$ (see [17]), showing that the second result in Theorem 1.1 is tight up to an $O(\log k)$ factor in the exponent, while the bound in Theorem 1.2 is tight up to a constant factor in the exponent.

We remark that, as in [15], our proofs can easily be adapted to provide monochromatic induced cycles of all (even) lengths between $C \log n$ and n for some constant C depending only on k. We also note that our bound on the size-Ramsey number of even cycles $\hat{r}^k(C_n) \leq \hat{r}^k_{ind}(C_n) = O(k^{102})n$ can be further improved significantly, using the same methods, but we chose not to present that here.

2 Proof outline

The main idea behind our proof is the following: consider a binomial random graph $G \sim \mathcal{G}(N, C/N)$, where $N = C'n$ and C, C' are appropriately chosen large constants. Let G be adversarially k-edge-colored. Then, it is easier to find an induced monochromatic cycle of length in $[0.9n, 1.1n]$, say, then of length precisely n. Our host graph is constructed to take advantage of this.

In the rest of the outline we focus on the proof of the induced odd case (Theorem 1.1 b)) and at the end we outline the changes needed for the other two statements.

Given k, we find a fixed "gadget" graph $F = F(k)$ which is k-induced-Ramsey for a 5-cycle. We denote $s = v(F)$. We construct an s-uniform N-vertex hypergraph H by taking CN random hyperedges. We clean H so it does not have any short Berge cycles so, in particular, it is linear. Then we construct our host graph Γ by placing an isomorphic copy of F inside every hyperedge of H. By definition, inside every copy of F, there is a monochromatic induced copy of C_5. The main object we work with will be an auxiliary k-edge-coloured graph G on the same vertex set as Γ. For each placed copy of F in Γ, in G we put an edge between a single pair of vertices which are at distance 2 in one of the induced monochromatic copies of C_5 in the copy of F, and colour this edge with the colour of that cycle.

Now, suppose we find a monochromatic, say red, cycle Q of length $\ell \in [n/3, n/2]$ in G. By definition, each edge of Q corresponds to an induced 5-cycle in Γ, where the endpoints of the edge are at distance 2 in the cycle. For each of these 5-cycles, we can choose either a path of length 2 or a path of length 3 in G to obtain a red cycle Q' of length exactly n in Γ (see Figure 1). The main technical difficulty is in obtaining certain properties of Q such that the resulting cycle Q' is induced in Γ.

More precisely, the following will be sufficient. Recall that every edge $e \in E(G)$ comes from a hyperedge in H which we denote by $h(e)$. Suppose Q is a cycle in G with edges $e_1, \dots, e_\ell$ such that no hyperedge apart from $h(e_1), \dots, h(e_\ell)$ in H intersects $\bigcup_{i\in[\ell]} h(e_i)$ in more than one vertex. Further, suppose that each $h(e_i)$ only intersects $h(e_{i-1})$ and $h(e_{i+1})$ among the mentioned hyperedges. Then, it is not difficult to see that the cycle Q' obtained as above is induced in Γ. We will call such a cycle Q *good*.

Let us now explain how to find an induced monochromatic cycle of length between $n/2$ and $n/3$ in a k-edge-colored graph $G \sim \mathcal{G}(N, C/N)$ with $N = C'n$ for some large constants C, C'. Our real task is more involved as we require a stronger condition on the found cycle as discussed above, since we are not working with a binomial random graph. However, most of the ideas can be described through the lens of this simpler problem.

We now sketch how to find a monochromatic induced cycle of length between $n/2$ and $n/3$ in $G \sim \mathcal{G}(N, C/N)$. The proof strategy is illustrated in Figure 2. By standard results, it is not difficult to clean G without losing many edges, so that it has no cycles of length $O(1)$. Further, we also know that it is locally sparse, that is, all sets U of size $|U| \le \varepsilon N$ span at most $\frac{3}{2}|U|$ edges, where $\varepsilon > 0$ is some constant depending on C. We consider the subgraph corresponding to the densest colour class, say red and using a result of Krivelevich [21], we find inside it a large expanding subgraph G'. Draganić, Glock and Krivelevich [5] showed using a modified DFS algorithm that under the given assumptions, G' has a red induced path P of length $2n/5$ and we adapt their argument to our setting. Given such a red induced path of length $2n/5$, from the endpoints we construct two trees

T_1, T_2 each of depth $O(\log N)$ and with $\Omega(\varepsilon N)$ leaves. Moreover, we do it in such a way that any path containing the initial endpoints is good, i.e. if there is a red edge connecting two vertices in different trees, it closes a good cycle in G'. Let $W = V(P) \cup V(T_1) \cup V(T_2)$ and remove from it a large constant number of the last layers in T_1 and T_2, so that the resulting W is small enough compared to the leaf sets of T_1 and T_2. Denote by R_1 and R_2 the vertices in the deleted layers in T_1 and T_2, respectively. Finally, using the expanding properties of G', we may expand from the sets R_1 and R_2, while avoiding vertices which are incident to W until the two balls around R_1 and R_2 of large enough constant diameter intersect, and thus we close a cycle of desired length. Using the girth assumption on our graph it is not difficult to show that this cycle is induced.

Let us now comment on the differences in the proofs for the three different statements. In the odd induced case, we can take F to be Alon's [1] celebrated construction of a dense pseudorandom triangle-free graph on $e^{\Theta(k \log k)}$ vertices. We will prove that, every k-edge-colouring of that graph will contain an induced monochromatic C_5. However, when n is even, we can instead take F to be k-induced-Ramsey for a 6-cycle with only $O(k^6)$ vertices by taking a sufficiently dense bipartite C_4-free graph. Again, in each copy of F, we find a monochromatic induced 6-cycle and connect two vertices at distance 2 on the cycle to form our auxiliary graph. The same argument as above shows that given a monochromatic cycle of length ℓ in the auxiliary graph G, we can find a monochromatic cycle of any even length between 2ℓ and 4ℓ in Γ. Finally, for the odd non-induced case, we can take F to be the complete graph on $2^k + 1$ vertices. It is easy to see that any k-edge coloring of that graph has a monochromatic odd cycle. For simplicity, we take the most common length L among those cycles, and for each of these L-cycles, we connect two vertices at distance $(L-1)/2$ on the cycle to form the auxiliary graph. Then, a monochromatic good cycle of length between $2n/(L-1)$ and $2n/(L+1)$ in the auxiliary graph yields a monochromatic cycle of length n in our host graph. This required extra precision in the length of the good cycle in the auxiliary graph will only cost us a factor of $2^{O(k)}$ in the number of copies of F we use in our construction.

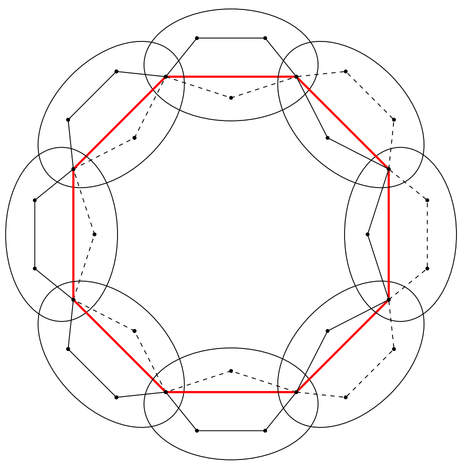


Figure 1: Transforming an 8-cycle in the auxiliary graph (thick red edges) into a 21-cycle in the original graph by using 5 paths of length 3 and 3 paths of length 2.

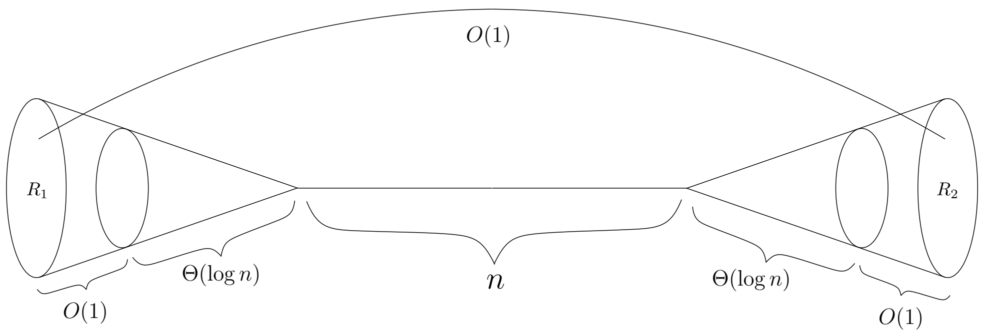


Figure 2: Building an induced cycle

References

[1] Noga Alon. Explicit Ramsey graphs and orthonormal labelings. *The Electronic Journal of Combinatorics*, pages R12–R12, 1994.

[2] József Beck. On size Ramsey number of paths, trees, and circuits. I. *Journal of Graph Theory*, 7(1):115–129, 1983.

[3] D. Conlon, J. Fox and B. Sudakov, Recent developments in graph Ramsey theory. In: *Surveys in Combinatorics* 2015, Cambridge University Press, 2015, 49–118.

[4] W. Deuber, A generalization of Ramsey's theorem. In: *Infinite and Finite Sets*, Vol. 1, Colloquia Mathematica Societatis János Bolyai, Vol. 10, North-Holland, Amsterdam/London, 1975, 323–332.

[5] Nemanja Draganić, Stefan Glock, and Michael Krivelevich. Short proofs for long induced paths. *Combinatorics, Probability and Computing*, 31(5):870–878, 2022.

[6] Nemanja Draganić, Michael Krivelevich, and Rajko Nenadov. Rolling backwards can move you forward: on embedding problems in sparse expanders. *Transactions of the American Mathematical Society*, 2022.

[7] Nemanja Draganić and Kalina Petrova. Size-Ramsey numbers of graphs with maximum degree three. *arXiv preprint arXiv:2207.05048*, 2022.

[8] Andrzej Dudek and Paweł Prałat. An alternative proof of the linearity of the size-Ramsey number of paths. *Combinatorics, Probability and Computing*, 24(3):551–555, 2015.

[9] Andrzej Dudek and Paweł Prałat. On some multicolor Ramsey properties of random graphs. *SIAM Journal on Discrete Mathematics*, 31(3):2079–2092, 2017.

[10] Paul Erdős. Problems and results on finite and infinite graphs. In *Recent Advances in Graph Theory. Proceedings of the Second Czechoslovak Symposium*, pages 183–192. Academia, Prague, 1975.

[11] Paul Erdős, Ralph J. Faudree, Cecil C. Rousseau, and Richard H. Schelp. The size Ramsey number. *Periodica Mathematica Hungarica*, 9(1-2):145–161, 1978.

[12] P. Erdős, A. Hajnal, and L. Pósa. Strong embeddings of graphs into colored graphs. In: *Infinite and Finite Sets*, Vol. 1, Colloquia Mathematica Societatis János Bolyai, Vol. 10, North-Holland, Amsterdam/London, 1975, 585–595.

[13] Jacob Fox and Benny Sudakov. Induced Ramsey-type theorems. *Advances in Mathematics*, 219(6):1771–1800, 2008.

[14] Joel Friedman and Nicholas Pippenger. Expanding graphs contain all small trees. *Combinatorica*, 7(1):71–76, 1987.

[15] Penny E. Haxell, Yoshiharu Kohayakawa, and Tomasz Łuczak. The induced size-Ramsey number of cycles. *Combinatorics, Probability and Computing*, 4(3):217–239, 1995.

[16] Ramin Javadi, Farideh Khoeini, Gholam Reza Omidi, and Alexey Pokrovskiy. On the size-Ramsey number of cycles. *Combinatorics, Probability and Computing*, 28(6):871–880, 2019.

[17] Ramin Javadi and Meysam Miralaei. The multicolor size-Ramsey numbers of cycles. *Journal of Combinatorial Theory, Series B*, 158:264–285, 2023.

[18] Nina Kamčev, Anita Liebenau, David R. Wood, and Liana Yepremyan. The size Ramsey number of graphs with bounded treewidth. *SIAM Journal on Discrete Mathematics*, 35(1):281–293, 2021.

[19] Yoshiharu Kohayakawa, Hans Jürgen Promel, and Vojtěch Rödl. Induced Ramsey numbers. *Combinatorica*, 18(3):373–404, 1998.

[20] Yoshiharu Kohayakawa, Vojtěch Rödl, Mathias Schacht, and Endre Szemerédi. Sparse partition universal graphs for graphs of bounded degree. *Advances in Mathematics*, 226(6):5041–5065, 2011.

[21] Michael Krivelevich. Expanders—how to find them, and what to find in them. In: *Surveys in combinatorics 2019*, volume 456 of *London Mathematical Society Lecture Note Series*, pages 115–142. Cambridge University Press, Cambridge, 2019.

[22] Michael Krivelevich. Long cycles in locally expanding graphs, with applications. *Combinatorica*, 39(1):135–151, 2019.

[23] Frank P. Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, 30(4):264–286, 1929.

[24] Vojtěch Rödl. The dimension of a graph and generalized Ramsey theorems, Master's thesis, Charles University, 1973.

[25] Vojtěch Rödl and Endre Szemerédi. On size Ramsey numbers of graphs with bounded degree. *Combinatorica*, 20(2):257–262, 2000.

[26] Konstantin Tikhomirov. On bounded degree graphs with large size-Ramsey numbers. *arXiv preprint arXiv:2210.05818*, 2022.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Single-conflict colorings of degenerate graphs

(Extended abstract)

Peter Bradshaw* Tomáš Masařík†‡

Abstract

We consider the *single-conflict coloring* problem, in which each edge of a graph receives a forbidden ordered color pair. The task is to find a vertex coloring such that no two adjacent vertices receive a pair of colors forbidden at an edge joining them. We show that for any assignment of forbidden color pairs to the edges of a d-degenerate graph G on n vertices of edge-multiplicity at most $\log \log n$, $O(\sqrt{d} \log n)$ colors are always enough to color the vertices of G in a way that avoids every forbidden color pair. This answers a question of Dvořák, Esperet, Kang, and Ozeki for simple graphs (Journal of Graph Theory 2021).

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-028`

1 Introduction

We consider graphs without loops and possibly with parallel edges. A *coloring* of a graph G is a function $\phi : V(G) \to C$ that assigns a color from some color set C to each vertex of

*Department of Mathematics, Simon Fraser University, Burnaby, BC, Canada & Department of Mathematics, University of Illinois Urbana-Champaign, Urbana, Illinois, USA. E-mail: pb38@illinois.edu.

†Department of Mathematics, Simon Fraser University, Burnaby, BC, Canada & Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Poland. E-mail: masarik@mimuw.edu.pl.
T.M. completed a part of this work while being supported by a postdoctoral fellowship at the Simon Fraser University through NSERC grants R611450 and R611368. He did a part of this work supported by project BOBR that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 948057).

‡The full preprint version of this paper is available on arXiv [3].

G. In this paper, we consider a special version of graph colorings known as *single-conflict colorings*, defined as follows. Let G be a graph, and let C be a color set. Let f be a function such that for each edge $e \in E(G)$ with endpoints u and v, f maps the triple (u, e, v) to a forbidden color pair (c_1, c_2), and f maps the triple (v, e, u) to the reverse forbidden color pair (c_2, c_1). Then, we say that a (not necessarily proper) coloring $\phi : V(G) \to C$ is a single-conflict coloring with respect to f and C if $f(u, e, v) \neq (\phi(u), \phi(v))$ for each edge $e = uv$ of G. We call the image of a triple (u, e, v) under f a *conflict*, and we call f a *conflict function*. If k is the minimum integer for which a graph G always has a single-conflict coloring for any color set C of size k and any conflict function f, then we say that k is the *single-conflict chromatic number* of G, and we write $\chi_{\not\leftrightarrow}(G) = k$.

1.1 Background

Dvořák and Postle [5] first introduced a concept similar to single-conflict coloring, called DP-coloring. Independently, single-conflict coloring was considered by Fraigniaud, Heinrich, and Kosowski [6][1], and the notion of single-conflict chromatic number was later introduced by Dvořák, Esperet, Kang, and Ozeki [4]. In [4], the authors proved the following:

Theorem 1.1 ([4]). *If G is a graph of maximum degree Δ, then $\chi_{\not\leftrightarrow}(G) \leq \left\lceil \sqrt{e(2\Delta - 1)} \right\rceil$.*

In fact, a stronger bound than in Theorem 1.1 with the leading multiplicative constant of 2 (instead of $\sqrt{2e}$) can be shown. Surprisingly, the factor of 2 is asymptotically sharp, as shown very recently by Groenland, Kaiser, Treffers, Wales [8]. For simple graphs, an even better coefficient of $1 + o(1)$ holds, which can be derived from a result of Kang and Kelly [11], or from an independent result Glock and Sudakov [7]. The last three cited results were stated for independent transversals. Another noteworthy result concerning the single-conflict chromatic number of graphs on surfaces was shown in [4].

Theorem 1.2 ([4]). *If G is a simple graph of Euler genus g, then $\chi_{\not\leftrightarrow}(G) = O((g + 1)^{1/4} \log(g + 2))$.*

Furthermore, the authors of [4] show that a graph of average degree $\overline{d}$ has a single-conflict chromatic number of at least $\left\lfloor \sqrt{\frac{\overline{d}}{\log \overline{d}}} \right\rfloor$.

The notion of a single-conflict coloring is a generalization of several graph coloring variants. Most immediately, single-conflict colorings generalize the notion of proper colorings as follows. Given a graph G, let $G^{(k)}$ denote the graph obtained from G by replacing each edge of G with k parallel edges. Then, $\chi(G) \leq k$ if and only if $G^{(k)}$ has a single-conflict coloring with a set C of k colors when each set of k parallel edges in $G^{(k)}$ is assigned k distinct monochromatic conflicts.

A single-conflict coloring is also a generalization of a *DP-coloring*, first introduced by Dvořák and Postle [5] under the name of *correspondence coloring*. One may define a DP-coloring of a graph G as a single-conflict coloring of a graph G' on $V(G)$ which is obtained

[1] Although [6] was published before [5], the arXiv version of [5] appears approximately two months before that of [6].

as follows. First, for each edge $uv \in E(G)$, select a matching M_{uv} in the complete bipartite graph $C \times C$. Then, for each edge $(c_1, c_2) \in M_{uv}$, give G' an edge e with endpoints u, v and a conflict $f(u, e, v) = (c_1, c_2)$. In this way, the single-conflict coloring problem can represent every instance of a DP-coloring problem.

Furthermore, a single-conflict coloring is a generalization of an earlier concept known as an *adapted coloring*, introduced by Hell and Zhu [10], which is defined as follows. Given a graph G with a (not necessarily proper) edge coloring ψ, an *adapted coloring* on G is a (not necessarily proper) vertex coloring ϕ of G in which no edge e is colored the same color as both of its endpoints u and v—that is, $\neg(\psi(e) = \phi(u) = \phi(v))$. In other words, if $e \in E(G)$ is a `red` edge, then both endpoints of e may not be colored `red`, but both endpoints of e may be colored, say, `blue`, and the endpoints of e may also be colored with two different colors. The *adaptable chromatic number* of G, written $\chi_{ad}(G)$, is the minimum integer m such that for any edge coloring of G using a set C of m colors, there exists an adapted vertex coloring of G using colors of C. It is easy to see that an adapted coloring on G is a single-conflict coloring on G when each edge $e \in E(G)$ is assigned the monochromatic conflict $(\psi(e), \psi(e))$. Therefore, for every graph G, $\chi_{ad}(G) \leq \chi_{\not\leftrightarrow}(G)$.

For graphs G of maximum degree Δ, Molloy and Thron [15] show that $\chi_{ad}(G) \leq (1 + o(1))\sqrt{\Delta}$. Molloy [13] shows furthermore that graphs G with chromatic number $\chi(G)$ satisfy $\chi_{ad}(G) \geq (1 + o(1))\sqrt{\chi(G)}$, implying that $\sqrt{\chi(G)}$, $\chi_{ad}(G)$, $\chi_{\not\leftrightarrow}(G)$, and $\sqrt{\Delta}$ all only differ by a constant factor for graphs G satisfying $\chi(G) = \Theta(\Delta)$. The parameters $\sqrt{\chi(G)}$, $\chi_{ad}(G)$, and $\chi_{\not\leftrightarrow}(G)$ can also differ by a constant factor even when $\chi(G)$ is not of the form $\Theta(\Delta)$. For instance, for graphs G of maximum degree Δ without cycles of length 3 or 4, the parameters $\sqrt{\chi(G)}$, $\chi_{\not\leftrightarrow}(G)$, and $\chi_{ad}(G)$ are all of the form $O\left(\sqrt{\frac{\Delta}{\log \Delta}}\right)$ [2, 14], and since randomly constructed Δ-regular graphs of girth 5 often have chromatic number as high as $\frac{1}{2}\frac{\Delta}{\log \Delta}$ [14], this upper bound is often tight.

We note that adapted colorings are equivalent to the notion of *cooperative colorings*, which are defined as follows. Given a family $\mathcal{G} = \{G_1, \ldots, G_k\}$ of graphs on a common vertex set V, a cooperative coloring on $\mathcal{G}$ is defined as a family of sets $R_1, \ldots, R_k \subseteq V$ such that for each $1 \leq i \leq k$, R_i is an independent set of G_i, and $V = \bigcup_{i=1}^k R_i$. A cooperative coloring problem may be translated into an adapted coloring problem by coloring the edges of each graph $G_i \in \mathcal{G}$ with the color i and then considering the union of all graphs in $\mathcal{G}$. Overall, this gives us the following observation.

Observation 1.3. *Given a family $\mathcal{G} = \{G_1, \ldots, G_k\}$ of graphs on a common vertex set, the cooperative coloring problem on $\mathcal{G}$ is equivalent to the adapted coloring problem on the edge-colored graph $G = \bigcup_{i=1}^k G_i$ in which each edge originally from G_i is colored with the color i.*

It is straightforward to show that Theorem 1.1 implies that a graph family $\mathcal{G}$ containing k graphs of maximum degree Δ on a common vertex set V has a cooperative coloring whenever $k \geq 2e\Delta$. In fact, Haxell [9] showed earlier that it is sufficient to let $k \geq 2\Delta$.

1.2 Our results

We have seen that for graphs G of maximum degree Δ, $\chi_{\not\leftrightarrow}(G) = O(\sqrt{\Delta})$. However, it is natural to ask whether we can obtain a better upper bound when G has bounded degeneracy. For example, in the related problem of cooperative coloring, Aharoni, Berger, Chudnovsky, Havet, and Jiang [1] obtained the following improved result by considering families of 1-degenerate graphs, i.e. forests:

Theorem 1.4 ([1])**.** *If $\mathcal{T}$ is a family of forests of maximum degree Δ on a common vertex set V, then there exists a value $k = (1 + o(1)) \log_{4/3} \Delta$ such that if $|\mathcal{T}| \geq k$, then $\mathcal{T}$ has a cooperative coloring.*

One key tool used to prove Theorem 1.4 is an application of the Lovász Local Lemma in which each vertex $v \in V$ receives a random *inventory* S_v of colors from a color set C indexing the forests in $\mathcal{T}$, and then a color c is deleted from S_v if c also belongs to the inventory S_w of the parent w of v in the forest of $\mathcal{T}$ indexed by c. Dvořák, Esperet, Kang, and Ozeki [4] also posed the following question, asking whether this upper bound can be improved for certain degenerate graphs.

Question 1.5. *Suppose G is a d-degenerate graph on n vertices. Is it true that $\chi_{\not\leftrightarrow}(G) = O(\sqrt{d} \log n)$?*

The authors remarked that a positive answer to Question 1.5 would give an alternative proof of Theorem 1.2. In this paper, we will prove the following theorem, which shows the upper bound of $\chi_{\not\leftrightarrow}(G) = O(\sqrt{\Delta})$ can often be improved for graphs of bounded degeneracy.

Theorem 1.6. *If G is a d-degenerate graph with maximum degree Δ and edge-multiplicity at most μ, then*

$$\chi_{\not\leftrightarrow}(G) \leq \left\lceil \sqrt{d} \cdot 2^{\mu/2+2} \sqrt{\mu} \sqrt{1 + \log((d+1)\Delta)} \right\rceil .$$

Theorem 1.6 gives a large class of d-degenerate graphs G satisfying $\chi_{\not\leftrightarrow}(G) = O(d^{\frac{1}{2}+o(1)})$, containing in particular those d-degenerate simple graphs G with maximum degree $\Delta = \exp(d^{o(1)})$. This upper bound is close to best possible, since Molloy [13] shows that d-degenerate graphs G of chromatic number $d+1$ satisfy $\chi_{\not\leftrightarrow}(G) \geq \chi_{ad}(G) \geq (1+o(1))\sqrt{d+1}$. By applying the argument used for Theorem 1.6 to simple graphs, we also obtain the following theorem.

Theorem 1.7. *If G is a d-degenerate simple graph of maximum degree Δ, then*

$$\chi_{\not\leftrightarrow}(G) \leq \left\lceil 2\sqrt{d\,[1 + \log((d+1)\Delta)]} \right\rceil .$$

Theorem 1.7 immediately answers Question 1.5 for simple graphs and thus also implies Theorem 1.2. In fact, the result that we will prove is slightly stronger than Theorem 1.7, and we will obtain the following corollary, which generalizes Theorem 1.4 at the expense of a constant factor.

Corollary 1.8. *Let $\mathcal{G}$ be a family of k graphs on a common vertex set V. Suppose each graph $G \in \mathcal{G}$ is at most d-degenerate and of maximum degree Δ. Then, whenever $k \geq 13(1 + d\log(d\Delta))$, $\mathcal{G}$ has a cooperative coloring.*

One natural question is whether the logarithmic factors are necessary in these upper bounds. While we are unable to answer these questions exactly, we note that an upper bound of less than $d+1$ is unachievable, as Kostochka and Zhu [12] give examples of d-degenerate graphs G that satisfy $\chi_{ad}(G) = d+1$. Additionally, Question 1.5 remains open for graphs of large edge-multiplicity.

2 Uniquely restrictive conflicts

It is well known that an oriented graph with maximum out-degree d is d-degenerate. Therefore, rather than working directly with d-degenerate graphs, we will consider the larger class of oriented graphs of maximum out-degree d. Given an oriented graph G, we write $A(G)$ for the set of arcs of G. For a vertex $v \in V(G)$, we write $A^+(v)$ for the set of arcs outgoing from v, and we write $A^-(v)$ for the set of arcs incoming to v. Given an arc $e = uv$ in an oriented graph G, and given a conflict function f on G, we will often write $f(e) = f(u, e, v)$.

Consider a color set C and an oriented graph G with a conflict function f. First, given a vertex $v \in V(G)$ and an arc $e \in A(G)$ containing v and second endpoint u, we say that the (v, e) *conflict color* is the first color appearing in the ordered pair $f(v, e, u)$. We write $\mathrm{cc}(v, e)$ for the (v, e) conflict color. Then, we have the following definition.

Definition 2.1. Let $w \in V(G)$. Suppose that for each parallel arc pair $e_1, e_2 \in A^-(w)$ satisfying $\mathrm{cc}(w, e_1) = \mathrm{cc}(w, e_2)$, it holds that $\mathrm{cc}(v, e_1) = \mathrm{cc}(v, e_2)$, where v is the second endpoint of e_1 and e_2. Then, we say f is *uniquely restrictive* at w. Furthermore, if f is uniquely restrictive at each $w \in V(G)$, then we simply say that f is *uniquely restrictive*.

An informal way of describing unique restrictiveness would be to say that if we color a vertex $w \in V(G)$ with some color, say `red`, then we only want this choice of `red` at w to contribute to the exclusion of at most one color possibility at each in-neighbor of w. We note that unique restrictiveness is a rather natural idea, as the conflict functions that represent adapted coloring and proper coloring problems are uniquely restrictive; indeed, in both of these settings, choosing the color `red` at a vertex v can only contribute to the exclusion of the color `red` at neighbors of v. Furthermore, DP-coloring problems always give uniquely restrictive conflict functions when represented as single-conflict coloring problems since the conflicts between any two vertices form a matching in $C \times C$.

We will also use the following form of the Lovász Local Lemma.

Theorem 2.2 ([16])**.** *Let $\mathcal{B}$ be a set of bad events. Suppose that each event $B \in \mathcal{B}$ occurs with probability at most p, and suppose further that each event $B \in \mathcal{B}$ is mutually independent with all but at most d other events $B' \in \mathcal{B}$. If $ep(d+1) \leq 1$, then with positive probability, no bad event in $\mathcal{B}$ occurs.*

With these preliminaries in place, we have the following theorem, which gives an upper bound on the number of colors needed for a single-conflict coloring of a d-degenerate graph whose conflict function is uniquely restrictive. Since any conflict function on a simple graph is uniquely restrictive, the following theorem implies Theorem 1.7 and hence gives an affirmative answer to Question 1.5. Our main tool for this theorem will be the application of the Lovász Local Lemma used by Aharoni, Berger, Chudnovsky, Havet, and Jiang [1] in which each vertex receives a random inventory of colors.

Theorem 2.3. *Let G be an oriented graph of maximum degree Δ with a maximum out-degree of at most d. Let C be a set of k colors, and let each arc $e \in A(G)$ have an associated conflict $f(e) \in C^2$. If f is uniquely restrictive, and if $k \geq 2\sqrt{d\,[1 + \log((d+1)\Delta)]}$, then G has a single-conflict coloring with respect to f and C.*

Proof. First, we note that since every subgraph of G has an average degree of at most $2d$, G is $(2d)$-degenerate and hence has a single-conflict coloring whenever $k \geq 2d+1$. Therefore, we may assume in our proof that $k \leq 2d$.

First, for each vertex $v \in V(G)$, we define a color inventory S_v, and for each color $c \in C$, we add c to S_v independently with probability $p = \frac{k}{2d} \leq 1$. Next, we let S'_v be a copy of S_v. (We will need these copies for technical reasons related to the Lovász Local Lemma.) Then, for each vertex $v \in V(G)$, we consider each outgoing arc e of v, and we write $e = (v, w)$. If, for some color $c \in S_v$, we have $f(e) \in \{(c, c') : c' \in S_w\}$, then we delete c from S'_v. In other words, if the color c at v contributes to the forbidden pair $f(v, w) = (c, c')$ of an outgoing arc $(v, w) \in A^+(v)$, and if $c' \in S_w$, then we delete c from S'_v. Then, for each vertex $v \in V(G)$, we let B_v denote the bad event that after this process, S'_v is empty. We observe that if no bad event occurs, then we may arbitrarily color each vertex v with a color from S'_v to obtain a single-conflict coloring of G. Indeed, if some arc (v, w) is colored with a forbidden pair (c, c') where $c \in S'_v$ and $c' \in S'_v$, then it must follow that c was actually deleted from S'_v, a contradiction.

Now, given a vertex $v \in V(G)$, we calculate the probability that the bad event B_v occurs. For a given color $c \in C$, we write b_c for the number of arcs $e \in A^+(v)$ for which $c = \mathrm{cc}(v, e)$. If c does not belong to S'_v, then either c was never added to S_v, or c was added to S_v and then deleted from S'_v. The probability that c was never added to S_v is equal to $1-p$, and the probability that c was added to S_v and then deleted from S'_v is at most $b_c p^2$. Therefore, the total probability that $c \notin S'_v$ is at most $1 - p + b_c p^2$. Furthermore, since f is uniquely restrictive, the probabilities of any two given colors being absent from S'_v are independent. Therefore, the probability of the bad event B_v is at most

$$\prod_{c \in C} \left(1 - \left(p - b_c p^2\right)\right) < \exp\left(-\sum_{c \in C} \left(p - b_c p^2\right)\right) = \exp\left(-pk + p^2 \sum_{c \in C} b_c\right) = \exp\left(-pk + p^2 d\right).$$

Substituting $p = \frac{k}{2d}$, we see that $\Pr(B_v) < \exp\left(-\frac{k^2}{4d}\right)$. Furthermore, as the bad event B_v involves $d+1$ vertices (namely v and at most d out-neighbors of v), each of maximum degree Δ, B_v is dependent with fewer than $(d+1)\Delta$ other bad events. Note that since we use unmodified inventories S_w to determine whether the copy S'_v is empty, we prevent the dependencies of B_v from spreading past the out-neighbors of v. Therefore, using the Lovász Local Lemma (Theorem 2.2), we see that G receives a single-conflict coloring with positive probability as long as $e(d+1)\Delta \exp\left(-\frac{k^2}{4d}\right) \le 1$. This inequality holds whenever $k \ge 2\sqrt{d[1+\log((d+1)\Delta)]}$, which completes the proof. □

Using Theorem 2.3, we can prove Corollary 1.8, which gives an upper bound on the number of colors needed for a cooperative coloring of a family of degenerate graphs. The proof is available in the full version on arXiv [3].

If G does not have parallel edges, then any conflict function $f : E(G) \to C^2$ must be uniquely restrictive. Then, Theorem 2.3 tells us that $\chi_{\leftrightarrow}(G) \le 2\left\lceil \sqrt{d\,(1+\log((d+1)\Delta))} \right\rceil$, which gives an affirmative answer to Question 1.5 for simple graphs.

3 General conflicts

Given an oriented graph G with a conflict function f, we define the *restrictiveness* of f at v as the maximum value r_v for which there exists an r_v-tuple of parallel arcs in $A^+(v)$ whose conflicts form a set $\{(c_1, c^*), (c_2, c^*), \ldots, (c_{r_v}, c^*)\}$, where the first entry in each conflict corresponds to v, where $c^* \in C$ is any single color, and where $c_1, \ldots, c_{r_v}$ are all distinct colors. Then, we say that the *restrictiveness* of f is the maximum restrictiveness r_v of f at v, taken over all vertices $v \in V(G)$. The restrictiveness r of a uniquely restrictive conflict function satisfies $r = 1$. If f is a conflict function on a graph G of edge-multiplicity at most μ, then the restrictiveness r of f satisfies $r \le \mu$.

Theorem 2.3 gives an upper bound on number of colors needed for a single-conflict coloring given a conflict function with restrictiveness $r = 1$. In this section, we will show in the following theorem that we can also find an upper bound on the number of colors needed for a single-conflict coloring given a conflict function whose restrictiveness r is known but may be greater than 1. Since $r \le \mu$ for any graph G with edge multiplicity at most μ, the following theorem (proven in the full version [3]) also proves Theorem 1.6, giving an upper bound for $\chi_{\leftrightarrow}(G)$ of d-degenerate graphs G with small edge-multiplicity.

Theorem 3.1. *Let G be an oriented graph of maximum degree Δ with a maximum out-degree of at most d. Let C be a set of k colors, and let each arc $e \in A(G)$ have an associated conflict $f(e)$. If the restrictiveness of f is at most r, and if*

$$k \ge \sqrt{d} \cdot 2^{r/2+2}\sqrt{r}\sqrt{1+\log((d+1)\Delta)},$$

then G has a single-conflict coloring with respect to f and C.

Acknowledgements. The authors would like to thank Ladislav Stacho and Jana Novotná for helpful discussions on the topic.

References

[1] Ron Aharoni, Eli Berger, Maria Chudnovsky, Frédéric Havet, and Zilin Jiang. Co-operative colorings of trees and bipartite graphs. *Electron. J. Combin.*, 27(1):Paper 1.41, 2020. doi:10.37236/8111.

[2] Jurgen Aliaj and Michael Molloy. Adaptable and conflict colouring multigraphs with no cycles of length three or four, 2021. arXiv:2107.04253.

[3] Peter Bradshaw and Tomáš Masařík. Single-conflict colorings of degenerate graphs, 2021. arXiv:2112.06333.

[4] Zdeněk Dvořák, Louis Esperet, Ross J. Kang, and Kenta Ozeki. Single-conflict colouring. *J. Graph Theory*, 97(1):148–160, 2021. doi:10.1002/jgt.22646.

[5] Zdeněk Dvořák and Luke Postle. Correspondence coloring and its application to list-coloring planar graphs without cycles of lengths 4 to 8. *J. Combin. Theory Ser. B*, 129:38–54, 2018. doi:10.1016/j.jctb.2017.09.001.

[6] Pierre Fraigniaud, Marc Heinrich, and Adrian Kosowski. Local conflict coloring. In *57th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2016*, pages 625–634. IEEE Computer Soc., Los Alamitos, CA, 2016. doi:10.1109/FOCS.2016.73.

[7] Stefan Glock and Benny Sudakov. An average degree condition for independent transversals. *J. Combin. Theory Ser. B*, 154:370–391, may 2022. doi:10.1016/j.jctb.2022.01.004.

[8] Carla Groenland, Tomáš Kaiser, Oscar Treffers, and Matthew Wales. Graphs of low average degree without independent transversals. *J. Graph Theory*, 102(2):374–387, February 2023. doi:10.1002/jgt.22876.

[9] Penny E. Haxell. A note on vertex list colouring. *Combin. Probab. Comput.*, 10(4):345–347, 2001. doi:10.1017/S0963548301004758.

[10] Pavol Hell and Xuding Zhu. On the adaptable chromatic number of graphs. *European J. Combin.*, 29(4):912–921, 2008. doi:10.1016/j.ejc.2007.11.015.

[11] Ross J Kang and Tom Kelly. Colorings, transversals, and local sparsity. *Random Struct. Algorithms*, 61(1):173–192, August 2022. doi:10.1002/rsa.21051.

[12] Alexandr V. Kostochka and Xuding Zhu. Adapted list coloring of graphs and hypergraphs. *SIAM J. Discrete Math.*, 22(1):398–408, 2008. doi:10.1137/070698385.

[13] Michael Molloy. The adaptable chromatic number and the chromatic number. *J. Graph Theory*, 84(1):53–56, 2017. doi:10.1002/jgt.22011.

[14] Michael Molloy. The list chromatic number of graphs with small clique number. *J. Combin. Theory Ser. B*, 134:264–284, 2019. doi:10.1016/j.jctb.2018.06.007.

[15] Michael Molloy and Giovanna Thron. An asymptotically tight bound on the adaptable chromatic number. *J. Graph Theory*, 71(3):331–351, 2012. doi:10.1002/jgt.20649.

[16] Joel Spencer. Asymptotic lower bounds for Ramsey functions. *Discrete Math.*, 20(1):69–76, 1977/78. doi:10.1016/0012-365X(77)90044-9.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Monadic NIP in monotone classes of relational structures

(Extended abstract)

Samuel Braunfeld* Anuj Dawar† Ioannis Eleftheriadis‡
Aris Papadopoulos§

Abstract

We prove that for any monotone class of finite relational structures, the first-order theory of the class is NIP in the sense of stability theory if, and only if, the collection of Gaifman graphs of structures in this class is nowhere dense. This generalises results previously known for graphs to relational structures and answers an open question posed by Adler and Adler (2014). The result is established by the application of Ramsey-theoretic techniques and shows that the property of being NIP is highly robust for monotone classes. We also show that the model-checking problem for first-order logic is intractable on any monotone class of structures that is not (monadically) NIP. This is a contribution towards the conjecture that the hereditary classes of structures admitting fixed-parameter tractable model-checking are precisely those that are monadically NIP.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-029`

*Computer Science Institute of Charles University (IUUK), Praha, Czech Republic. E-mail: `sbraunfeld@iuuk.mff.cuni.cz`. Supported by Project 21-10775S of the Czech Science Foundation (GAČR), European Union's Horizon 2020 research and innovation programme (grant agreement No 810115 – Dynasnet).

†Department of Computer Science and Technology, University of Cambridge, UK. E-mail: `anuj.dawar@cl.cam.ac.uk`. Supported by EPSRC grant EP/T007257/1.

‡Department of Computer Science and Technology, University of Cambridge, UK. E-mail: `ie257@cam.ac.uk`. Supported by a George and Marie Vergottis Scholarship awarded through Cambridge Trust, an Onassis Foundation Scholarship, and a Robert Sansom Studentship.

§School of Mathematics, Univesity of Leeds, UK. E-mail: mmadp@leeds.ac.uk. Supported by a Leeds Doctoral Scholarship.

1 Introduction

The development of stability theory in classical model theory, originating with Shelah's classification programme fifty years ago [12, 2], has sought to distinguish *tame* first-order theories from *wild* ones. A key discovery is that combinatorial configurations serve as dividing lines in this classification.

Separately, in the development of finite model theory, there has been in interest in investigating *tame* classes of finite structures. Here tameness can refer to *algorithmic tameness*, meaning that algorithmic problems that are intractable in general may be tractable on a tame class; or it can refer to *model-theoretic tameness*, meaning that the class enjoys some desirable model-theoretic properties that are absent in the class of all finite structures. See [4] for an exposition of these notions of tameness. The tame classes that arise in this context are often based on notions taken from the study of sparse graphs [9] and usually extended to classes of relational structures beyond graphs by applying them to the *Gaifman graphs* of such structures.

In the context of algorithmic tameness of sparse classes, this line of work culminated in the major result of Grohe et al. [7] showing that the problem of model checking first-order sentences is fixed-parameter tractable (FPT) on any class of graphs that is *nowhere dense*. This generalized a sequence of earlier results showing the tractability of the model checking problem on classes of graphs satisfying other notions of sparsity. Moreover, it is also known [8] that this is the limit of tractability for *monotone* classes of graphs. That is to say that (under reasonable assumptions) any monotone class of graphs in which first-order model checking is FPT is necessarily nowhere dense. These results underline the centrality of the notion of nowhere denseness in the study of sparse graph classes.

A significant line of recent research has sought to generalize the methods and results on tame sparse classes of graphs to more general classes that are not necessarily sparse. Interestingly, this has tied together notions of tameness arising in finite model theory and those in classical model theory. Notions arising from stability theory play an increasingly important role in these considerations (see [10, 6], for example). Central to this connection is the realisation that for well-studied notions of sparseness in graphs, the first-order theory of a sparse class $\mathcal{C}$ is stable. Thus, stability-theoretic notions of tameness, applied to the theory of a class of finite structures, generalize the notions of tameness emerging from the theory of sparsity.

A key result connecting the two directions is that a monotone class of finite graphs is stable if, and only if, it is nowhere dense. This connection between stability and combinatorial sparsity was established in the context of infinite graphs by Podewski and Ziegler [11] and extended to classes of finite graphs by Adler and Adler [1]. Indeed, for monotone classes of graphs, stability is a rather robust concept as the theory of such a class is stable if, and only if, it is NIP, and these conditions on monotone classes are in turn equivalent to it being monadically stable and monadically NIP.

A question posed by Adler and Adler is whether their result can be extended from graphs to structures in any finite relational language. We settle this question in the present paper by establishing Theorem 1 below. In the following $\mathsf{Gaif}(\mathcal{C})$ denotes the collection of Gaifman graphs of structures in the class $\mathcal{C}$. Note that the extension from graphs to relational structures requires considerable combinatorial machinery in the form of Ramsey-theoretic results, which we detail in later sections. We also relate the characterization to the tractability of the classes. In summary, our key results are stated in the following theorem.

Theorem 1. *Let $\mathcal{C}$ be a monotone class of finite structures in a finite relational language. Then, the following are equivalent:*

1. *$\mathcal{C}$ is NIP;*
2. *$\mathcal{C}$ is monadically NIP;*
3. *$\mathcal{C}$ is stable;*
4. *$\mathcal{C}$ is monadically stable;*
5. $\mathsf{Gaif}(\mathcal{C})$ *is nowhere dense; and*
6. *(assuming* $\mathsf{AW}[*] \neq \mathsf{FPT}$*) $\mathcal{C}$ admits fixed-parameter tractable model checking.*

Moreover, the equivalence of the first six notions also holds for classes containing infinite structures.

The equivalence of the first four notions for any monotone class $\mathcal{C}$ is due to Braunfeld and Laskowski [3]. The equivalence of the fifth and sixth notions follows by results in sparsity theory (see [9]). We, therefore, establish the equivalence of the first with the fifth and the sixth. More precisely, we show that if $\mathsf{Gaif}(\mathcal{C})$ is not nowhere dense, then $\mathcal{C}$ admits a formula with the independence property. That nowhere density of $\mathsf{Gaif}(\mathcal{C})$ implies tractability is implicit in [7]. We establish the converse of this statement here.

2 Preliminaries

We assume familiarity with first-order logic and the basic concepts of model theory and graph theory. Throughout this paper, $\mathcal{L}$ denotes a finite, first-order, relational language. Given an $\mathcal{L}$-structure M, we write $\mathsf{Gaif}(M)$ for the *Gaifman graph* of M, i.e. the graph on domain M with the property that two elements are adjacent if and only if they appear together in a relation of M. For a class $\mathcal{C}$ of $\mathcal{L}$-structures, we write $\mathsf{Gaif}(\mathcal{C})$ for the class of graphs $\{\mathsf{Gaif}(M) : M \in \mathcal{C}\}$. We say that a class $\mathcal{C}$ is monotone if it is closed under weak substructures, i.e. if $(M, R^M)_{R\in\mathcal{L}} \in \mathcal{C}$ then $(M', R^{M'})_{R\in\mathcal{L}} \in \mathcal{C}$ for any $M' \subseteq M$ and $R^{M'} \subseteq R^M$.

We say that a class $\mathcal{C}$ of $\mathcal{L}$-structures is *NIP* (Not the Independence Property) if there is no $\mathcal{L}$-formula $\phi(\bar{x}, \bar{y})$ satisfying that for all bipartite graphs $G = (U, V; E) \in \mathfrak{B}$ there is some $M_G \in \mathcal{C}$ and sequences of tuples $(\bar{a}_i)_{i\in U}$ and $(\bar{b}_j)_{j\in V}$ such that:

$$M_G \models \phi(\bar{a}_i, \bar{b}_j) \text{ if, and only if, } (i, j) \in E.$$

Moreover, a class $\mathcal{C}$ of graphs is said to be *nowhere dense* if for every $r \in \mathbb{N}$ there is some $n \in \mathbb{N}$ such that for all $G \in \mathcal{C}$ we have that $K_n^{(r)}$ is not a subgraph of G, and otherwise, $\mathcal{C}$ is *somewhere dense*.

By the model checking problem on a class $\mathcal{C}$, we refer to the parametrised decision problem whereby, given a structure $M \in \mathcal{C}$ and an FO-sentence ϕ whose depth acts as parameter, we want to decide if M satisfies ϕ. We say that the model checking problem on a class $\mathcal{C}$ is *fixed-parameter tractable*, if there is an algorithm that decides on input (M, ϕ) whether $M \models \phi$, in time $f(|\phi|) \cdot |M|^{\mathcal{O}(1)}$ for some computable function f. Model checking on the class of all graphs is complete with respect to the complexity class $\mathsf{AW}[*]$, which is conjectured to strictly contain the class FPT. We shall assume throughout that $\mathsf{AW}[*] \neq \mathsf{FPT}$.

3 Main results

Here, we sketch the proofs of implications (1) $\implies$ (5) and (6) $\implies$ (5) from Theorem 1. We first prove that for any monotone class $\mathcal{C}$ of relational structures whose Gaifman class is somewhere dense, there is a formula which codes the edge relation of all bipartite graphs uniformly over $\mathcal{C}$. We work towards this theorem via a preparatory lemma, which has the benefit of applying to classes that are not necessarily monotone. Intuitively, this tells us that in any class of relational structures $\mathcal{C}$ whose Gaifman class is somewhere dense, there is a primitive positive formula which codes the edge relation of any complete bipartite graph with "sufficiently disjoint" witnesses.

Lemma 1. *Let $\mathcal{C}$ be a class of $\mathcal{L}$-structures such that $\mathsf{Gaif}(\mathcal{C})$ is somewhere dense. Then there is a primitive positive formula $\phi(\bar{x}, \bar{y}, \bar{z}) = \exists \bar{w} \psi(\bar{x}, \bar{y}, \bar{z}, \bar{w})$ with parameters $\bar{p}$, and for every $n \in \mathbb{N}$ there is some $M_n \in \mathcal{C}$ and tuples $(\bar{a}_i)_{i \in [n]}$, $(\bar{b}_j)_{j \in [n]}$, $(\bar{c}_{i,j})_{(i,j) \in [n]^2}$, $(\bar{d}_{i,j})_{i,j \in [n]^2}$ from M_n such that the following hold for all $i, i', j, j' \in [n]$:*

1. *$M_n \models \psi(\bar{a}_i, \bar{b}_j, \bar{c}_{i,j}, \bar{d}_{i,j})$;*
2. *$\bar{a}_i(k) \neq \bar{a}_{i'}(k)$, for $i \neq i'$ and all $k \in [|\bar{x}|]$;*
3. *$\bar{b}_j(k) \neq \bar{b}_{j'}(k)$, for $j \neq j'$ and all $k \in [|\bar{y}|]$;*
4. *$\bar{c}_{i,j}(k) \neq \bar{c}_{i',j'}(k)$ and $\bar{c}_{i,j}(k) \neq \bar{c}_{i,j}(l)$, for $(i,j) \neq (i',j')$ and all $k \neq l$ from $[|\bar{z}|]$;*
5. *$\bar{d}_{i,j}(k) \neq \bar{d}_{i',j'}(k)$, for $(i,j) \neq (i',j')$ and all $k \in [|\bar{w}|]$.*

The proof of this lemma is a combinatorial argument, resting on few applications of different Ramsey theorems. First, we ensure that the subdivided edges coming from the assumption that $\mathsf{Gaif}(\mathcal{C})$ is somewhere dense are witnessed in every M_n by the same sequence of relations $R_1, \ldots, R_k$; this requires an application of the finite Ramsey theorem (for local uniformity within each structure), and of the pigeonhole principle (for global uniformity for the whole class). Next, by consecutive applications of the canonical Erdős-Rado theorem (see [5]), we may obtain a finite set of elements that are common in all such relations, and ensure that the remaining elements on these are essentially distinct. The level of disjointedness achieved is precisely what allows us, under the additional assumption that $\mathcal{C}$ is monotone, to remove relations so as to turn the encoded complete bipartite graphs into arbitrary bipartite graphs and violate NIP. Consequently, we establish the following.

Theorem 2. *Let $\mathcal{C}$ be a monotone class of $\mathcal{L}$-structures such that* $\mathsf{Gaif}(\mathcal{C})$ *is somewhere dense. Then there is a primitive positive formula $\phi(\bar{x}, \bar{y}) = \exists \bar{w} \psi(\bar{x}, \bar{y}, \bar{w})$ with parameters $\bar{p}$ and for each bipartite graph $G = (U, V; E)$ there is some $M_G \in \mathcal{C}$ and sequences of tuples $(\bar{a}_u)_{u \in U}$ $(\bar{b}_v)_{v \in V}$, $(\bar{h}_{u,v})_{(u,v) \in E}$ from M_G such that:*

1. *$M_G \models \phi(\bar{a}_u, \bar{b}_v)$ if, and only if, $(u, v) \in E$ (so, in particular $\mathcal{C}$ is not NIP);*
2. *If $(u, v) \in E$ then $M_G \models \psi(\bar{a}_u, \bar{b}_v, \bar{h}_{u,v})$;*
3. *The equality type of $\bar{p}_{u,v} = \bar{a}_u{}^\frown \bar{b}_v{}^\frown \bar{h}_{u,v}$ is constant for all $(u, v) \in E(G)$;*
4. *Any two tuples in $\{\bar{a}_u, \bar{b}_v, \bar{h}_{u,v} : u \in U, v \in V\}$ are disjoint and do not intersect the parameters $\bar{p}$.*

Next, we prove that any monotone class of relational structures whose Gaifman class is somewhere dense polynomially interprets the class of all bipartite graphs, and is therefore intractable. Towards this, we first strengthen Theorem 2 to obtain a "simple path formula" that performs the encoding; this is essentially a primitive positive formula $\phi(\bar{x}, \bar{y})$ that describes a sequence of relation $R_1, \ldots, R_k$, with the property that $\bar{x} \subseteq R_1$ and $\bar{y} \subseteq R_k$. Moreover, having full control over the equality type of the elements in M_G allows to obtain a polynomial-time construction of M_G from G.

Lemma 2. *Let $\mathcal{C}$ be a monotone class of $\mathcal{L}$-structures such that* $\mathsf{Gaif}(\mathcal{C})$ *is somewhere dense. Then there is a simple path formula $\phi(\bar{x}, \bar{y})$ with parameters $\bar{p}$ and a polynomial time computable function $\Phi : \mathfrak{B} \to \mathcal{C}$, such that for each bipartite graph $G = (U, V; E) \in \mathfrak{B}$ there are tuples $(\bar{a}_u)_{u \in U}$ $(\bar{b}_v)_{v \in V}$, $(\bar{h}_{u,v})_{(u,v) \in E}$ from $\Phi(G)$ satisfying:*

$$\Phi(G) \models \phi(\bar{a}_u, \bar{b}_v) \text{ if, and only if, } (u, v) \in E.$$

With this, we proceed to show intractability for monotone classes with somewhere Gaifman class. Our proof is essentially based on the proof of [8, Theorem 6.1], which covers the case of graphs. There, the aim is to definably distinguish the native points of an r-subdivided graph G from the subdivision points. The idea is to distinguish points by their degrees; however, while all subdivision points have degree two, other points in G may as well have degree two. To address this, we first pre-process G to obtain a graph G' by adding two pendant vertices to each non-isolated vertex. Then, G is definably recovered from G', and moreover, given an r-subdivision of G', we can definably distinguish the subdivision points and the remaining points by their degrees. Our construction is essentially the same, although the degree of a subdivision point is bounded by the length of paths in the subdivision, rather than by two.

Theorem 3. *Let $\mathcal{C}$ be a monotone class of $\mathcal{L}$-structures such that* $\mathsf{Gaif}(\mathcal{C})$ *is somewhere dense, and assume that* $\mathsf{AW}[*] \neq \mathsf{FPT}$. *Then* FO *model-checking on $\mathcal{C}$ is not fixed-parameter tractable.*

References

[1] Hans Adler and Isolde Adler. Interpreting nowhere dense graph classes as a classical notion of model theory. *European Journal of Combinatorics*, 36:322–330, 2014.

[2] John T. Baldwin. *Fundamentals of Stability Theory.* Perspectives in Logic. Cambridge University Press, 2017.

[3] Samuel Braunfeld and Michael C. Laskowski. Existential characterizations of monadic NIP, 2022.

[4] Anuj Dawar. Finite model theory on tame classes of structures. In *MFCS*, volume 4708 of *Lecture Notes in Computer Science*, pages 2–12. Springer, 2007.

[5] Paul Erdős and Richard Rado. A combinatorial theorem. *Journal of the London Mathematical Society*, 1(4):249–255, 1950.

[6] Jakub Gajarský, Michal Pilipczuk, and Szymon Torunczyk. Stable graphs of bounded twin-width. In Christel Baier and Dana Fisman, editors, *LICS '22: 37th Annual ACM/IEEE Symposium on Logic in Computer Science, Haifa, Israel, August 2 - 5, 2022*, pages 39:1–39:12. ACM, 2022.

[7] Martin Grohe, Stephan Kreutzer, and Sebastian Siebertz. Deciding first-order properties of nowhere dense graphs. *J. ACM*, 64(3), jun 2017.

[8] Stephan Kreutzer and Anuj Dawar. Parameterized complexity of first-order logic. *Electron. Colloquium Comput. Complex.*, TR09-131, 2009.

[9] Jaroslav Nešetřil and Patrice Ossona De Mendez. *Sparsity: graphs, structures, and algorithms*, volume 28. Springer Science & Business Media, 2012.

[10] Michal Pilipczuk, Sebastian Siebertz, and Szymon Torunczyk. Parameterized circuit complexity of model-checking on sparse structures. In Anuj Dawar and Erich Grädel, editors, *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2018, Oxford, UK, July 09-12, 2018*, pages 789–798. ACM, 2018.

[11] Klaus-Peter Podewski and Martin Ziegler. Stable graphs. *Fund. Math*, 100(2):101–107, 1978.

[12] Saharon Shelah. *Classification theory and the number of nonisomorphic models*, volume 92 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam-New York, 1978.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Decomposition horizons: from graph sparsity to model-theoretic dividing lines

(Extended abstract)

Samuel Braunfeld* Jaroslav Nešetřil* Patrice Ossona de Mendez† Sebastian Siebertz‡

Abstract

Low treedepth decompositions are central to the structural characterizations of bounded expansion classes and nowhere dense classes, and the core of main algorithmic properties of these classes, including fixed-parameter (quasi) linear-time algorithms checking whether a fixed graph F is an induced subgraph of the input graph G. These decompositions have been extended to structurally bounded expansion classes and structurally nowhere dense classes, where low treedepth decompositions are replaced by low shrubdepth decompositions. In the emerging framework of a structural graph theory for hereditary classes of structures based on tools from model theory, it is natural to ask how these decompositions behave with the fundamental model theoretical notions of dependence (alias NIP) and stability.

In this work, we prove that the model theoretical notions of NIP and stable classes are transported by decompositions. Precisely: Let $\mathscr{C}$ be a hereditary class of graphs. Assume that for every p there is a hereditary NIP class $\mathscr{D}_p$ with the property that the vertex set of every graph $G \in \mathscr{C}$ can be partitioned into $N_p = N_p(G)$ parts in such a way that the union of any p parts induce a subgraph in $\mathscr{D}_p$ and $\log N_p(G) \in o(\log |G|)$. We prove that then $\mathscr{C}$ is (monadically) NIP. Similarly, if every $\mathscr{D}_p$ is stable, then $\mathscr{C}$ is (monadically) stable. Results of this type lead to the definition of decomposition horizons as closure operators. We establish some of their basic properties and provide several further examples of decomposition horizons.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-030`

*Computer Science Institute of Charles University (IUUK), Praha, Czech Republic.
E-mails: `sbraunfeld@iuuk.mff.cuni.cz`, `nesetril@iuuk.mff.cuni.cz`.
†CAMS (CNRS, UMR 8557), Paris, France. E-mail:`pom@ehess.fr`.
‡University of Bremen, Bremen, Germany. E-mail: `siebertz@uni-bremen.de`.

1 Introduction and Previous Work

In the late 90's, Baker [2] introduced the shifting strategy, allowing a linear time approximation scheme for independent sets on planar graphs. The idea is to start a breadth-first search at a vertex v of a planar graph, which partitions the vertex set of the graph into layers $L_1, \ldots, L_h$ and to fix an integer D. Then, for given $s \in [D]$, by deleting all the layers L_i with $i \equiv s \bmod D$, one gets a graph with treewidth bounded by $3D$, on which a maximum independent set can be found in linear time. Considering all the possible values of s, we obtain a $(1 + 1/D)$-approximate solution of the problem. Note that grouping the layers L_i with i in a same class modulo D yields a partition of the vertex set into D parts $V_0, \ldots, V_{D-1}$ such that the union of any $p < D$ of them induces a subgraph with treewidth at most $3p + 4$.

This approach was further developed by DeVos et al. [7], who proved in particular that for every proper minor closed class of graphs $\mathscr{C}$ and every integer p, there exists an integer N_p such that the vertex set of every graph $G \in \mathscr{C}$ can be partitioned into N_p parts, each p of them inducing a subgraph with treewidth at most $p - 1$.

This result has been further extended by two of the authors of the present paper in a characterization of both bounded expansion classes and nowhere dense classes. Before stating these results, recall that the *treedepth* of a graph G is the minimum depth of a rooted forest F, such that G is a subgraph of the closure of F (the graph obtained from F by adding edges between each vertex and its ancestors). With this definition, the characterization theorems read as follows.

Theorem 1.1 ([15]). *A class $\mathscr{C}$ has bounded expansion if and only if, for every parameter p, there is an integer N_p such that the vertex set of each graph $G \in \mathscr{C}$ can be partitioned into at most N_p parts, each p of them inducing a subgraph with treedepth at most p.*

Theorem 1.2 (see [16,17]). *A class $\mathscr{C}$ is nowhere dense if and only if, for every parameter p and for every graph $G \in \mathscr{C}$ there is an integer $N_p(G) \in |G|^{o(1)}$, such that the vertex set of G can be partitioned into at most $N_p(G)$ parts, each p of them inducing a subgraph with treedepth at most p.*

The notions of classes with bounded expansion and of nowhere dense classes are central to the study of classes of sparse graphs [16]. Note that the treewidth of a graph is bounded from above by its treedepth and hence by the result of DeVos et al. [7] and Theorem 1.1 every proper minor closed class has bounded expansion. Surprisingly, it appeared that for monotone classes of graphs, the notion of nowhere dense class of graphs coincides with fundamental dividing lines introduced in modern model theory [21]:

Theorem 1.3 ([1]). *For a monotone class of graphs $\mathscr{C}$, the following are equivalent:*

(1) $\mathscr{C}$ is nowhere dense;
(2) $\mathscr{C}$ is stable;
(3) $\mathscr{C}$ is monadically stable;
(4) $\mathscr{C}$ is NIP*;*
(5) $\mathscr{C}$ is monadically NIP.

For general hereditary classes of graphs, we do not have the collapse of the notions of stability, monadic stability, NIP, and monadic NIP stated in Theorem 1.3 for monotone classes. However, we still have the following collapses:

Theorem 1.4 ([5])**.** *A hereditary class of graphs is monadically* NIP *if and only if it is* NIP*. A hereditary class of graphs is monadically stable if and only if it is stable.*

The study of monadic stability and monadic NIP and their relations with first-order transductions [3] opened the way to the study of *structurally sparse* classes of graphs, that is of classes of graphs that are first-order transductions of classes of sparse graphs [6,9,10,18–20]. Intuitively, a (first-order) transduction is a way to construct a set of target graphs from the vertex-colorings of a source graph by fixed first-order formulas, and, by extension, a new class of graphs from a given class of graphs.

Extending Theorem 1.1, first-order transductions of bounded expansion classes have been characterized in terms of low shrubdepth colorings. Recall the following high level characterization of classes with bounded shrubdepth [11, 12]: A class $\mathscr{D}$ has *bounded shrubdepth* if it is a transduction of a class of bounded depth rooted forests.

Theorem 1.5 ([10])**.** *A class $\mathscr{C}$ is a first-order transduction of a class with bounded expansion if and only if, for every parameter p, there is an integer N_p and a class $\mathscr{D}_p$ with bounded shrubdepth, such that the vertex set of each graph $G \in \mathscr{C}$ can be partitioned into at most N_p parts, each p of them inducing a subgraph in $\mathscr{D}_p$.*

Theorem 1.5 can be seen as a generalization of Theorem 1.1 as shrubdepth is a dense analogue of treedepth. On the other hand, only one direction of Theorem 1.2 has been extended to transductions of nowhere dense classes.

Theorem 1.6 ([8])**.** *Let $\mathscr{C}$ be a first-order transduction of a nowhere dense class. Then, for every parameter p there is a class $\mathscr{D}_p$ with bounded shrubdepth, such that for every graph $G \in \mathscr{C}$ there is an integer $N_p(G) \in |G|^{o(1)}$, with the property that the vertex set of G can be partitioned into at most $N_p(G)$ parts, each p of them inducing a subgraph in $\mathscr{D}_p$.*

Similar decompositions, where p parts induce a subgraph with bounded rankwidth were introduced in [13], while classes having such decompositions where p parts induce a subgraph with bounded linear rankwidth were discussed in [20]. However, it was not known whether such classes are monadically NIP. This question, which appears for instance in [20, Figure 3] and again in [19], will get a positive answer as a direct consequence of Theorem 2.1, which is our main result.

The theoretical significance of first-order transductions of nowhere dense classes is witnessed by the following conjecture.

Conjecture 1.7 ([9])**.** *A class of graphs is monadically stable if and only if it is a first-order transduction of a nowhere dense class of graphs.*

We show that Conjecture 1.7 can be refined as follows.

Conjecture 1.8. *For a hereditary class of graphs $\mathscr{C}$, the following properties are equivalent:*
(1) $\mathscr{C}$ is a first-order transduction of a nowhere dense class;
(2) $\mathscr{C}$ admits low shrubdepth decompositions with $n^{o(1)}$ parts;
(3) $\mathscr{C}$ is monadically stable;
(4) $\mathscr{C}$ is stable.

By Theorem 1.6, property (1) implies property (2). That property (2) implies property (3) will follow from our main result (Theorem 2.1). By Theorem 1.4, properties (3) and (4) are equivalent. Closing the chain of implications corresponds to Conjecture 1.7, which we now can decompose into two weaker statements: that property (3) implies property (2), and that property (2) implies property (1).

2 Statement of the results

We show that NIP and stability are fixed under taking decompositions as in Theorems 1.1, 1.2, 1.5 and 1.6.

Theorem 2.1. *Let $\mathscr{C}$ be a hereditary graph class. Suppose that for every parameter p there is an* NIP *(resp. stable) class $\mathscr{D}_p$ such that for every graph $G \in \mathscr{C}$ there is an integer $N_p(G) \in |G|^{o(1)}$, with the property that the vertex set of G can be partitioned into at most $N_p(G)$ parts, each p of them inducing a subgraph in $\mathscr{D}_p$. Then $\mathscr{C}$ is* NIP*(resp. stable).*

In particular, this proves that property (2) implies property (4) in Conjecture 1.8, and so it follows that Conjectures 1.7 and 1.8 are equivalent. As mentioned after Theorem 1.6, this also proves that classes admitting low (linear) rankwidth decompositions are monadically NIP.

To place this theorem in a broader context, we introduce the notion of decomposition horizons. These seem to be of significant independent interest, and we prove some general properties. Theorem 2.1 can then be stated as "NIP and stability are decomposition horizons".

We define a *hereditary class property* to be a downset Π of hereditary graph classes, that is, a set of hereditary classes such that if $\mathscr{C} \in \Pi$ and $\mathscr{D}$ is a hereditary class with $\mathscr{D} \subseteq \mathscr{C}$, then $\mathscr{D} \in \Pi$.

Definition 1. *Let Π be a hereditary class property, let $f : \mathbb{N} \to \mathbb{N}$ be a non-decreasing function and let p be a positive integer. We say that a class $\mathscr{C}$ has an f-bounded* Π-decomposition *with parameter p if there exists $\mathscr{D}_p \in \Pi$ such that, for every graph $G \in \mathscr{C}$, there exists an integer $N \leq f(|G|)$ and a partition $V_1, \dots, V_N$ of the vertex set of G with $G[V_{i_1} \cup \cdots \cup V_{i_p}] \in \mathscr{D}_p$ for all $i_1, \dots, i_p \in [N]$.*

When f is a constant function, we say that $\mathscr{C}$ has a *bounded-size* Π-*decomposition* with parameter p; when f is a function with $f(n) = n^{o(1)}$, we say that $\mathscr{C}$ has a *quasi-bounded-size* Π-*decomposition* with parameter p. If a class $\mathscr{C}$ has a bounded-size (resp. a quasi-bounded-size) Π-decomposition with parameter p for each positive integer p, we say that $\mathscr{C}$ has *bounded-size* Π-*decompositions* (resp. *quasi-bounded-size* Π-*decompositions*).

For instance, by Theorem 1.1 and Theorem 1.2, considering the hereditary class property "bounded treedepth", we have that a class $\mathscr{C}$ has bounded-size bounded treedepth decompositions if and only if it has bounded expansion, and it has quasi-bounded-size bounded treedepth decompositions if and only if it is nowhere dense. With these definition in hand, it is natural to consider the following constructions of graph class properties:

Definition 2. *For a hereditary class property* Π *we define the properties* Π^+ *(resp.* Π^**) as follows:*

- $\mathscr{C} \in \Pi^+$ *if* $\mathscr{C}$ *has bounded-size* Π*-decompositions;*
- $\mathscr{C} \in \Pi^*$ *if* $\mathscr{C}$ *has quasi-bounded-size* Π*-decompositions.*

For every hereditary class property Π, we show that $(\Pi^+)^+ = \Pi^+$ and $(\Pi^*)^+ = \Pi^*$ (but we are not aware of any hereditary (NIP) class property Π, such that $\Pi^* \neq (\Pi^*)^*$). Also, for every two hereditary class properties Π_1 and Π_2, we show in the full paper that $(\Pi_1 \cap \Pi_2)^+ = \Pi_1^+ \cap \Pi_2^+$ and $(\Pi_1 \cap \Pi_2)^* = \Pi_1^* \cap \Pi_2^*$, which suggests that, for every hereditary class property Π, there might exist an inclusion-minimum class Λ with $\Lambda^+ = \Pi^+$. On the other hand, if $(\Pi_i)_{i\in I}$ is a family of hereditary class properties indexed by a set I, then $(\bigcup_{i\in I} \Pi_i)^+ = \bigcup_{i\in I} \Pi_i^+$ and $(\bigcup_{i\in I} \Pi_i)^* = \bigcup_{i\in I} \Pi_i^*$. In particular, the inclusion order of decomposition horizons is a distributive lattice.

Definition 3. *We say that a hereditary class property* Π *is a* decomposition horizon *if* $\Pi^* = \Pi$. *If* Λ *is a hereditary class property, the* decomposition horizon *of* Λ *is the smallest decomposition horizon including* Λ.

For example, the hereditary class property of all hereditary classes excluding a fixed graph H is a decomposition horizon. In the full paper, we also prove that several hereditary class properties are decomposition horizons, including

- the class properties "bounded maximum degree after deletion of at most k vertices",
- the class property "transduction of a class with bounded maximum degree" (this property is equivalent to the model-theoretic property "mutually algebraic" [6], hence to the model-theoretic property "monadic NFCP" [14]),
- the class property "weakly sparse" (i.e. "biclique-free") of classes excluding a fixed biclique as a subgraph,
- the class property "nowhere dense".

Our examples include an infinite countable chain of decomposition horizons (the class properties"bounded maximum degree after deletion of at most k vertices"), witnessing some richness of the inclusion order on decomposition horizons.

While it is natural to conjecture that stable hereditary classes of graphs are exactly those hereditary classes with quasi-bounded-size bounded shrub-depth decompositions, NIP hereditary classes seem to be more elusive. It was proved in [4] that for hereditary classes of ordered graphs, being NIP is equivalent to having bounded twin-width. On the other hand, classes with quasi-bounded-size bounded twin-width decompositions are NIP (as classes with bounded twin-width are NIP) and include transductions of nowhere dense classes (thus, conjecturally, all stable hereditary classes). Hence, it is a natural question whether every NIP hereditary class has quasi-bounded-size bounded twin-width decompositions.

References

[1] H. Adler and I. Adler. Interpreting nowhere dense graph classes as a classical notion of model theory. *European Journal of Combinatorics*, 36:322–330, 2014.

[2] B. S. Baker. Approximation algorithms for NP-complete problems on planar graphs. *Journal of the ACM*, 41(1):153–180, 1994.

[3] J. T. Baldwin and S. Shelah. Second-order quantifiers and the complexity of theories. *Notre Dame Journal of Formal Logic*, 26(3):229–303, 1985.

[4] E. Bonnet, U. Giocanti, P. Ossona de Mendez, P. Simon, S. Thomassé, and S. Toruńczyk. Twin-width IV: ordered graphs and matrices. In *STOC 2022: Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, 2022. `doi:10.1145/3519935.3520037`.

[5] S. Braunfeld and M. C. Laskowski. Existential characterizations of monadic NIP. *arXiv preprint arXiv:2209.05120*, 2022.

[6] S. Braunfeld, J. Nešetřil, P. Ossona de Mendez, and S. Siebertz. On the first-order transduction quasiorder of hereditary classes of graphs. *arXiv preprint arXiv:2208.14412*, 2022.

[7] M. DeVos, G. Ding, B. Oporowski, D. P. Sanders, B. Reed, P. D. Seymour, and D. Vertigan. Excluding any graph as a minor allows a low tree-width 2-coloring. *Journal of Combinatorial Theory, Series B*, 91(1):25–41, 2004.

[8] J. Dreier, J. Gajarský, S. Kiefer, M. Pilipczuk, and S. Toruńczyk. Treelike decompositions for transductions of sparse graphs. *arXiv preprint arXiv:2201.11082*, 2022.

[9] J. Gajarský, P. Hliněný, J. Obdržálek, D. Lokshtanov, and M.S. Ramanujan. A new perspective on FO model checking of dense graph classes. In *Proceedings of the Thirty-First Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 176–184. ACM, 2016. `doi:10.1145/3383206`.

[10] J. Gajarský, S. Kreutzer, J. Nešetřil, P. Ossona de Mendez, M. Pilipczuk, S. Siebertz, and S. Toruńczyk. First-order interpretations of bounded expansion classes. In *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*, volume 107 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 126:1–126:14, 2018.

[11] R. Ganian, P. Hliněný, J. Nešetřil, J. Obdržálek, and P. Ossona de Mendez. Shrubdepth: Capturing height of dense graphs. *Logical Methods in Computer Science*, 15(1), 2019.

[12] Robert Ganian, Petr Hliněný, Jaroslav Nešetřil, Jan Obdržálek, Patrice Ossona de Mendez, and Reshma Ramadurai. When trees grow low: shrubs and fast MSO_1. In *International Symposium on Mathematical Foundations of Computer Science*, pages 419–430. Springer, 2012.

[13] O.-j. Kwon, M. Pilipczuk, and S. Siebertz. On low rank-width colorings. In *Graph-theoretic concepts in computer science*, volume 10520 of *Lecture Notes in Comput. Sci.*, pages 372–385. Springer, Cham, 2017.

[14] M. C. Laskowski. Mutually algebraic structures and expansions by predicates. *The Journal of Symbolic Logic*, 78(1):185–194, 2013.

[15] J. Nešetřil and P. Ossona de Mendez. Grad and classes with bounded expansion I. Decompositions. *European Journal of Combinatorics*, 29(3):760–776, 2008. `doi:10.1016/j.ejc.2006.07.013`.

[16] J. Nešetřil and P. Ossona de Mendez. *Sparsity (Graphs, Structures, and Algorithms)*, volume 28 of *Algorithms and Combinatorics*. Springer, 2012. 465 pages.

[17] J. Nešetřil and P. Ossona de Mendez. On low tree-depth decompositions. *Graphs and Combinatorics*, 31(6):1941–1963, 2015. `doi:10.1007/s00373-015-1569-7`.

[18] J. Nešetřil and P. Ossona de Mendez. Structural sparsity. *Uspekhi Matematicheskikh Nauk*, 71(1):85–116, 2016. (Russian Math. Surveys 71:1 79-107). `doi:10.1070/RM9688`.

[19] J. Nešetřil, P. Ossona de Mendez, M. Pilipczuk, R. Rabinovich, and S. Siebertz. Rankwidth meets stability. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2014–2033, 2021. `doi:10.1137/1.9781611976465.120`.

[20] J. Nešetřil, P. Ossona de Mendez, R. Rabinovich, and S. Siebertz. Linear rankwidth meets stability. In Shuchi Chawla, editor, *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms*, pages 1180–1199, 2020.

[21] S. Shelah. *Classification theory: and the number of non-isomorphic models*, volume 92. Elsevier, 1990.

This paper is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 810115 – DYNASNET) and from the German Research Foundation (DFG) with grant agreement No 444419611.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Countable ultrahomogeneous 2-colored graphs consisting of disjoint unions of cliques

(Extended abstract)

Sofia Brenner* Irene Heinrich†

Abstract

We classify the countable ultrahomogeneous 2-vertex-colored graphs in which the color classes form disjoint unions of cliques. This generalizes work by Jenkinson et. al. [8], Lockett and Truss [11] as well as Rose [13] on ultrahomogeneous n-graphs. As the key aspect in such a classification, we identify a concept called piecewise ultrahomogeneity. We prove that there are two specific graphs whose occurrence essentially dictates whether a graph is piecewise ultrahomogeneous, and we exploit this fact to prove the classification.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-031`

1 Introduction

Ultrahomogeneous structures are relational structures in which every isomorphism between finite substructures can be extended to an automorphism of the entire structure.[1] The extensive study of ultrahomogeneous objects relates various areas of research, such as model theory, permutation group theory and Ramsey theory (see [12] for a survey). A vast collection of ultrahomogeneous classes of relational structures has been classified.

*Department of Mathematics, TU Darmstadt, 64289 Darmstadt, Germany. E-mail: `brenner@mathematik.tu-darmstadt.de`. Supported by ERC, EngageS, grant agreement No. 820148.

†Department of Mathematics, TU Darmstadt, 64289 Darmstadt, Germany. E-mail: `heinrich@mathematik.tu-darmstadt.de`. Supported by ERC, EngageS, grant agreement No. 820148.

[1]Some authors use the term "homogeneous" for this property.

For instance, apart from different classes of graphs, which we discuss below, there exist classification results for partially ordered sets [14], tournaments [9, 2] as well as countably infinite permutations [1].

In this article, we focus on a special class of countable ultrahomogeneous graphs. By work of Sheehan [15] and Gardiner [3] as well as Golfand and Klin [4], the finite ultrahomogeneous graphs are known. Lachlan and Woodrow [10] gave a characterization of the ultrahomogeneous graphs with countably infinitely many vertices. Cherlin [2] asked for a classification of ultrahomogeneous n-graphs, that is, ultrahomogeneous graphs for which the vertex set is partitioned into n subsets which are respected by the partial isomorphisms considered.

Nowadays, one usually thinks of n-graphs as graphs with a vertex-coloring in n colors, and considers isomorphisms preserving colors. Finite ultrahomogeneous vertex-colored graphs were classified in [5]. Every color class in an ultrahomogeneous graph induces a monochromatic ultrahomogeneous graph. In particular, every infinite color class forms an independent set or a disjoint union of cliques, or it induces a Rado graph or a Henson graph (see [10]). Jenkinson et. al. [8] considered vertex-colored graphs in which the color classes form independent sets. Their work was extended by Lockett and Truss [11] who allowed an additional coloring of the edges (while still requiring that every color class forms an independent set). In his dissertation, Rose [13] investigates countable 2-colored graphs. The main part of his work covers the case that one color class forms a disjoint union of cliques and the other one induces a Rado graph or a Henson graph. For the case that both color classes form disjoint unions of cliques, a partial list of possible cases is stated, but not proven.

In this paper, we classify the countable 2-colored ultrahomogeneous graphs for which both color classes form disjoint unions of cliques. We identify a new concept, which we call piecewise ultrahomogeneity, as key aspect in such classifications. An ultrahomogeneous graph whose color classes form disjoint unions of cliques is called piecewise ultrahomogeneous if each subgraph induced by a pair of maximal cliques of distinct color is ultrahomogeneous. As explained by Rose (see [13, Theorem 5.2]), this concept also appears in the dissertation of Jenkinson [7]. We obtain the following characterization of piecewise ultrahomogeneity (see Theorems 4.1 and 5.2):

Theorem A. *Let G be a non-bipartite, countable, 2-colored ultrahomogeneous graph in which the color classes form disjoint unions of cliques and that is not a blow-up. Apart from one degenerate case $F_{2,2}$, the graph G is piecewise ultrahomogeneous if and only if it contains induced subgraphs isomorphic to the graphs Q and $\widetilde{Q}$ depicted in Figure 1.*

We leverage the theorem to completely classify countable 2-colored ultrahomogeneous graphs in which the color classes form disjoint unions of cliques:

Theorem B. *Let G be a countable 2-colored ultrahomogeneous graph in which the color classes form disjoint unions of cliques and that is not a blow-up. Then (after possibly interchanging the colors) exactly one of the following holds:*

(i) (Piecewise ultrahomogeneous, Theorem 6.1) Either both color classes in G form an independent set or a single clique, G belongs to a single biparametric family $\{G_{r,b}\colon r, b \in \mathbb{N} \cup \{\infty\}\}$*, or G is isomorphic to the specific graph* $F_{2,2}$.

(ii) (Not piecewise ultrahomogeneous, Theorem 5.2) The graph G belongs to one of two monoparametric families $\{F^k_{\infty,1}\colon k \in \mathbb{N}_{\geq 2}\}$ *or* $\{F^k_{\infty,2}\colon k \in \mathbb{N}_{\geq 2}\}$*, or it is isomorphic to one of four specific graphs* $F_{2,1}$, $F_{\infty,1}$, $F_{\infty,2}$ *or* $F_{\infty,\infty}$.

This paper is organized as follows: Section 2 contains preliminary results. In Section 3, we recall Fraïssé's theory and study the structure of minimally omitted subgraphs. In Section 4, we introduce the concept of piecewise ultrahomogeneity and prove one implication of Theorem A. In Sections 5 and 6, we classify graphs that are not piecewise ultrahomogeneous and piecewise ultrahomogeneous, respectively, thereby proving Theorems A and B. We conclude with some final remarks in Section 7.

2 Preliminaries

Let G be a (simple) graph. We denote by $V(G)$ the vertex set of G. A *2-colored graph* is a graph whose vertices are colored in two distinct colors, which we usually call "blue" and "red". The graph G is *ultrahomogeneous* if every partial isomorphism between two finite induced subgraphs of G extends to an automorphism of G. In order to shorten our notation, we call a graph G *clique-ultrahomogeneous* (CUH) if G is a countably infinite ultrahomogeneous graph on blue and red vertices such that both color classes are disjoint unions of cliques. Two distinct vertices $v, v' \in V(G)$ are called *twins* if they have the same color and the same neighbors in $G - \{v, v'\}$. The edges in G with endpoints of different color are called *cross edges*. We write $\widetilde{G}$ for the graph obtained from G by complementing the cross edges while maintaining the edges within each color class. Let $\mathcal{R}$ and $\mathcal{B}$ denote the sets of maximal red and blue cliques of G, respectively. By [10], the elements of $\mathcal{R}$ all have the same size (similarly for $\mathcal{B}$). Note that the automorphism group of G permutes the set $\mathcal{R}$. Similarly, it permutes $\mathcal{B}$. From the definition of ultrahomogeneity, we obtain the following statement (also see [5, Lemma 6.1]):

Lemma 2.1. *Let G be a 2-colored ultrahomogeneous graph, and let H be obtained from G by any combination of complementations of the edges within a color class or the cross edges. Then H is ultrahomogeneous.*

Let H be a 2-colored graph in which one color class is an independent set. We call G a *blow-up* of H if G is obtained from H by, for some $i \in \mathbb{N}_{\geq 2} \cup \{\infty\}$, replacing all vertices in this color class by i-cliques and joining their vertices to the neighbors of the original vertex in H. The following property is easily verified (also see [5, Lemma 6.2]):

Lemma 2.2. *A blow-up of a graph H is ultrahomogeneous if and only if H is ultrahomogeneous.*

We call a CUH graph *basic* if $|\mathcal{R}|, |\mathcal{B}| \geq 2$ holds, and G is not a blow-up. By complementation inside the color classes and reduction of blow-ups, which preserves ultrahomogeneity (see Lemmas 2.1 and 2.2), we can always pass to a basic CUH graph. It therefore suffices to consider basic graphs. A 2-colored graph G is called *trivial* if all or none of the possible cross edges in G are present. Concerning the sizes of the color classes, we note the following:

Lemma 2.3. *Let G be a basic CUH graph. If a color class of G is finite, then G is trivial.*

By [8], there exists a unique countably infinite 2-colored ultrahomogeneous graph whose color classes form independent sets and which is generic in the following sense: For every $c \in \{\text{red}, \text{blue}\}$ and all finite disjoint subsets S, T of color c, there exists a vertex of color $c' \neq c$ adjacent to all vertices in S and to none of the vertices in T. This graph is called the *generic bipartite graph* (GB-graph). We frequently use the following classification:

Theorem 2.4 ([8, Theorem 2.2])**.** *Let G be a countable 2-colored ultrahomogeneous graph whose color classes form independent sets. Either G is trivial, the cross edges in G form a perfect matching or its complement, or G is isomorphic to the GB-graph.*

Note that the graphs given in Theorem 2.4 are bipartite.

3 Fraïssé limits and omitted subgraphs

Let L be a countable relational language. An L-structure D is called *ultrahomogeneous* if every isomorphism between finite substructures of D extends to an automorphism of D. The *age* of an L-structure D is the class of all finite L-structures that are isomorphic to induced substructures of D. An *amalgamation class* is a class of finite L-structures which is closed under isomorphism and taking induced substructures, and has the *amalgamation property*: For $J, A_1, A_2 \in \mathcal{A}$ and embeddings $\iota_i\colon J \to A_i$ $(i = 1, 2)$, there exists $A \in \mathcal{A}$ and embeddings $\kappa_i\colon A_i \to A$ $(i = 1, 2)$ such that $\kappa_1 \circ \iota_1 = \kappa_2 \circ \iota_2$ holds. Then A is called an *amalgam* of A_1 and A_2.

Theorem 3.1 (Fraïssé)**.** *Let D be a countable ultrahomogeneous L-structure. Then the age of D is an amalgamation class. Conversely, for every amalgamation class $\mathcal{C}$ of finite L-structures, there exists a countable ultrahomogeneous L-structure D with age $\mathcal{C}$, and D is unique up to isomorphism.*

In the setting of Theorem 3.1, we call D the *Fraïssé limit* of $\mathcal{C}$. Further information on this topic can be found, for example, in [6]. We now return to the special case of countable 2-colored graphs. If H is a finite graph that is isomorphic to an induced subgraph of a graph G, we say that H is *realized* in G. Otherwise, H is *omitted* in G. The graph H is called *minimally omitted* if H is omitted in G and every proper induced subgraph of H is realized in G. The set of minimally omitted subgraphs of G is denoted by $O(G)$. Concerning the structure of the color classes of a graph in $O(G)$, we obtain the following result, which forms the basis for all further arguments:

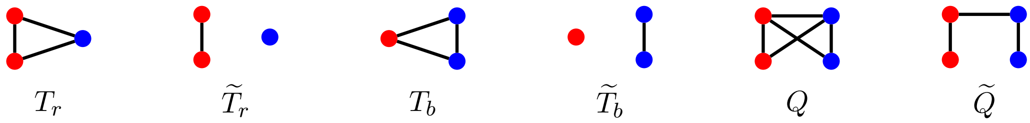


Figure 1: The graphs in $\mathcal{T} \cup \{Q, \widetilde{Q}\}$

Theorem 3.2. *Let G be a CUH graph. If $H \in O(G)$ is not monochromatic, then for every color $c \in \{red,\ blue\}$, one of the following holds:*

(i) The vertices of color c in H form a clique of size at least 3 and they are all twins in H,

(ii) the vertices of color c in H form a 2-clique, or,

(iii) the vertices of color c in H form an independent set.

4 Piecewise ultrahomogeneity

We call a CUH graph G *piecewise ultrahomogeneous* if for every $R \in \mathcal{R}$ and $B \in \mathcal{B}$, the graph $G[R \cup B]$ is ultrahomogeneous. Let T_r and T_b be the triangles containing a single blue vertex and a single red vertex, respectively. We set $\mathcal{T} = \{T_r, \widetilde{T}_r, T_b, \widetilde{T}_b\}$. Moreover, let Q be the graph arising from a complete graph on two red and two blue vertices by omitting one cross edge. The graphs in $\mathcal{T}$ as well as Q and $\widetilde{Q}$ are depicted in Figure 1.

Theorem 4.1. *Let G be a basic CUH graph. If Q and $\widetilde{Q}$ are realized in G, then G is piecewise ultrahomogeneous.*

Proof (Sketch). Let $R \in \mathcal{R}$ and $B \in \mathcal{B}$. We show that $G[R \cup B]$ is the complement of the GB-graph. To this end, consider finite disjoint subsets S and T of R. We need to show that there exists a vertex $v \in B$ adjacent to all vertices in S and to none of the vertices in T. The vertices in $S \cup T$ form a red clique of size $|S| + |T|$. In a series of lemmas, we show that there exists a vertex $v' \in B$ whose neighborhood in $S \cup T$ is a set S' of size $|S|$. Here, the main technical difficulty is to ensure that the blue vertex with the required neighborhood can be found in the clique B. Moreover, we prove that there exists a joint neighbor $b \in B$ of the vertices in $S \cup T$. Now consider the partial isomorphism φ of G obtained by bijectively mapping S' to S and $(S \cup T) \setminus S'$ to T while fixing b. Let $\hat{\varphi}$ be an automorphism of G extending φ. Then $\hat{\varphi}(v) \in B$ is a vertex adjacent to all vertices in S and to none of the vertices in T. For the other color class, one can argue similarly. Hence $G[R \cup B]$ is ultrahomogeneous. □

5 CUH graphs omitting Q or $\widetilde{Q}$

In this section, we classify the basic CUH graphs which omit Q or $\widetilde{Q}$. We first determine the possible graphs in $O(G)$.

Theorem 5.1. *Let G be a basic CUH graph in which Q or $\widetilde{Q}$ is omitted, and which is not isomorphic to one of the graphs in Theorem 2.4. Then every non-monochromatic graph in $O(G)$ is contained in $\mathcal{T} \cup \{Q, \widetilde{Q}\}$. Moreover, we have $T_r \in O(G)$ if and only if $\widetilde{T}_r \in O(G)$ holds, and the same is true for T_b and $\widetilde{T}_b$ as well as Q and $\widetilde{Q}$.*

Proof (Sketch). One first shows that $T_r \in O(G)$ holds if and only if $\widetilde{T}_r \in O(G)$ holds, and that this is the case precisely if the maximal red cliques in G have size 2 (similarly for the blue color class). Moreover, we show that if neither the blue nor the red vertices in G form an independent set, then for every $R \in \mathcal{R}$ and $B \in \mathcal{B}$, there exist partitions $R = R_1 \dot\cup R_2$ and $B = B_1 \dot\cup B_2$ such that $(R_1 \times B_1) \cup (R_2 \times B_2)$ is precisely the set of cross edges between R and B. Using this structure, we show that the maximal monochromatic cliques in G either have size 1 or 2, or they are infinite.

Let $H \in O(G)$ and assume that H is not monochromatic. Using Theorem 3.2, we show that either $H \in \mathcal{T} \cup \{Q, \widetilde{Q}\}$ holds, or that the color classes in H form independent sets of size at least 2 (all other possibilities can be eliminated by using the structure of $G[R \cup B]$ described above). We then show that the second case cannot occur. The main technical difficulty is to show that for a given monochromatic independent set $J \subseteq V(G)$ of color $c \in \{\text{red, blue}\}$ and a vertex v of color $c' \neq c$, there exist sufficiently many maximal cliques of color c' containing vertices with the same neighbors in J as v. Using this property, we can then successively show that G contains a subgraph isomorphic to H. $\square$

Let $\mathcal{C}$ be the class of finite graphs on red and blue vertices whose color classes form disjoint unions of cliques.

Theorem 5.2. *Let G be a basic CUH graph that omits Q or $\widetilde{Q}$, and that is not isomorphic to one of the graphs in Theorem 2.4. Up to exchanging the color classes, one of the following cases arises:*

(i) If $k := |\mathcal{R}|$ is finite, then G is isomorphic to one of the following graphs:

- *(a) The Fraïssé limit $F^k_{\infty,1}$ of the class of graphs in $\mathcal{C}$ that omit a red independent set of size $k+1$ and the blue 2-clique.*
- *(b) The Fraïssé limit $F^k_{\infty,2}$ of the class of graphs in $\mathcal{C}$ that omit a red independent set of size $k+1$, the blue triangle as well as T_b and $\widetilde{T}_b$.*

(ii) Otherwise, G is isomorphic to one of the following graphs:

- *(a) The Fraïssé limit $F_{2,1}$ of the class of graphs in $\mathcal{C}$ that omit the red triangle, the blue 2-clique, T_r and $\widetilde{T}_r$.*
- *(b) The Fraïssé limit $F_{\infty,1}$ of the class of graphs in $\mathcal{C}$ that omit the blue 2-clique.*
- *(c) The Fraïssé limit $F_{2,2}$ of the class of graphs in $\mathcal{C}$ that omit monochromatic triangles and the graphs in $\mathcal{T}$.*

(d) *The Fraïssé limit* $F_{\infty,2}$ *of the class of graphs in* $\mathcal{C}$ *that omit the blue triangle as well as* T_b *and* $\widetilde{T}_b$.

(e) *The Fraïssé limit* $F_{\infty,\infty}$ *of the class of graphs in* $\mathcal{C}$ *that omit* Q *and* $\widetilde{Q}$.

Proof (Sketch). One needs to verify that the given graph classes fulfill the amalgamation property. Proving that the above list is exhaustive is done by using Theorem 5.1. □

We remark that the graph $F_{2,1}$ given in Theorem 5.2 appears to be excluded by the (unproven) enumeration of possible CUH graphs stated in [13]. Combining Theorems 4.1 and 5.2 yields the characterization of piecewise ultrahomogeneity stated in Theorem A.

6 Classification of piecewise ultrahomogeneous graphs

Using the results in Theorem 2.4, we obtain the following classification:

Theorem 6.1. *Let* G *be a basic piecewise ultrahomogeneous CUH graph. Then* G *is isomorphic to one of the graphs in Theorem 2.4, to the specific graph* $F_{2,2}$*, or to the Fraïssé limit* $G_{|\mathcal{R}|,|\mathcal{B}|}$ *of the class* $\mathcal{A}_{|\mathcal{R}|,|\mathcal{B}|}$ *of finite graphs on red and blue vertices in which the red and the blue color class form disjoint unions of at most* $|\mathcal{R}|$ *and* $|\mathcal{B}|$ *cliques, respectively.*

For the proof, one first proves that if G is neither isomorphic to one of the graphs in Theorem 2.4 nor to $F_{2,2}$, then $G[R \cup B]$ is the complement of the GB-graph for every $R \in \mathcal{R}$ and $B \in \mathcal{B}$. Using the structure of the GB-graph, one then shows that every graph in $\mathcal{A}_{|\mathcal{R}|,|\mathcal{B}|}$ is realized in G. By Fraïssé's theorem, this implies $G \cong G_{|\mathcal{R}|,|\mathcal{B}|}$. This completes the classification given in Theorem B.

7 Conclusion

In this paper, we classified the countable 2-colored ultrahomogeneous graphs in which each color class forms a disjoint union of cliques. Our key tool was the concept of piecewise ultrahomogeneity introduced in Section 4. We showed that with one exception, a basic non-bipartite CUH graph is piecewise ultrahomogeneous if and only if two specific graphs appear as induced subgraphs (see Theorem A). Using this result, we obtained the classification of countable 2-colored CUH graphs given in Theorem B.

There are several natural continuations of this paper. For example, it would be interesting to classify edge-colored versions of CUH graphs, extending the work of Lockett and Truss [11]. Moreover, one could investigate n-colored versions of CUH graphs for an arbitrary number $n \in \mathbb{N}$. In both cases, we believe that a suitable generalization of piecewise ultrahomogeneity could play a central role. Just as in the case studied in this paper, one could hope to characterize the piecewise ultrahomogeneous graphs in terms of a small number of induced subgraphs, and then use the classifications of ultrahomogeneous multipartite graphs given in [8] and [11]. Conversely, if a graph fails to be piecewise ultrahomogeneous, its structure might again be very limited.

References

[1] P. J. Cameron. Homogeneous permutations. *Electron. J. Combin.*, 9(2), 2002.

[2] G. Cherlin. *The classification of countable homogeneous directed graphs and countable n-tournaments*, volume 621 of *Mem. Amer. Math. Soc.* Amer. Math. Soc., Providence, 1998.

[3] A. Gardiner. Homogeneous graphs. *J. Comb. Theory, Ser. B*, 20(1):94–102, 1976.

[4] Y. Golfand and M. Klin. On k-regular graphs. *Algorithmic Research in Combinatorics*, 186:76–85, 1978.

[5] I. Heinrich, T. Schneider, and P. Schweitzer. Classification of finite highly regular vertex-coloured graphs. `https://arxiv.org/abs/2012.01058`.

[6] W. Hodges. *Model Theory*, volume 42 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 1993.

[7] T. Jenkinson. The construction and classification of homogeneous structures in model theory. Dissertation, University of Leeds, 2006.

[8] T. Jenkinson, J. K. Truss, and D. Seidel. Countable homogeneous multipartite graphs. *Europ. J. Combin.*, 33(1):82–109, 2012.

[9] A. H. Lachlan. Countable homogeneous tournaments. *Trans. Amer. Math. Soc.*, 284:431–461, 1984.

[10] A. H. Lachlan and R. E. Woodrow. Countable ultrahomogeneous undirected graphs. *Trans. Amer. Math. Soc.*, 262:51–94.

[11] D. C. Lockett and J. K. Truss. Homogeneous coloured multipartite graphs. *Europ. J. Comb.*, 42:217–242, 2014.

[12] D. Macpherson. A survey of homogeneous structures. *Discrete Mathematics*, 311(15):1599–1634, 2011.

[13] S. E. Rose. Classification of countable homogeneous 2-graphs. Dissertation, University of Leeds, 2011.

[14] J. H. Schmerl. Countable homogeneous partially ordered sets. *Algebra Universalis*, 9:317–321, 1979.

[15] J. Sheehan. Smoothly embeddable subgraphs. *J. London Math. Soc.*, s2-9(2):212–218, 1974.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Raising the roof on the threshold for Szemerédi's theorem with random differences

(Extended abstract)

Jop Briët* Davi Castro-Silva†

Abstract

Using recent developments on the theory of locally decodable codes, we prove that the critical size for Szemerédi's theorem with random differences is bounded from above by $N^{1-\frac{2}{k}+o(1)}$ for length-k progressions. This improves the previous best bounds of $N^{1-\frac{1}{\lceil k/2 \rceil}+o(1)}$ for all odd k.

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-032

1 Introduction

Szemerédi [14] proved that dense sets of integers contain arbitrarily long arithmetic progressions, a result which has become a hallmark of additive combinatorics. Multiple proofs of this result were found over the years, using ideas from combinatorics, ergodic theory and Fourier analysis over finite abelian groups.

Furstenberg's ergodic theoretic proof [12] opened the floodgates to a series of powerful generalizations. In particular, it led to versions of Szemerédi's theorem where the common differences for the arithmetic progressions are restricted to very sparse sets. We say that a set $D \subseteq [N]$ is *ℓ-intersective* if any positive-density set $A \subseteq [N]$ contains an $(\ell+1)$-term arithmetic progression with common difference in D. Szemerédi's theorem implies

*CWI & QuSoft, Science Park 123, 1098 XG Amsterdam, The Netherlands. Supported by the Dutch Research Council (NWO) as part of the NETWORKS programme (grant no. 024.002.003).

†CWI & QuSoft, Science Park 123, 1098 XG Amsterdam, The Netherlands. Supported by the Dutch Research Council (NWO) as part of the NETWORKS programme (grant no. 024.002.003).

that for large enough N_0, the set $\{0, 1, \ldots, N_0\}$ is ℓ-intersective for $N \geq N_0$. Non-trivial examples include a special case of a result of Bergelson and Leibman [3] showing that the perfect squares are ℓ-intersective for every ℓ, and a special case of a result of Wooley and Ziegler [17] showing the same for the prime numbers minus one.

The existence of such sparse intersective sets motivated the problem of showing whether, in fact, random sparse sets are typically intersective. The task of making this quantitative falls within the scope of research on threshold phenomena. We say that a property of subsets of $[N]$, given by a family $\mathcal{F} \subseteq 2^{[N]}$, is *monotone* if $A \in \mathcal{F}$ and $A \subseteq B \subseteq [N]$ imply $B \in \mathcal{F}$. The *critical size* $m^* = m^*(N)$ of a property is the least m such that a uniformly random m-element subset of $[N]$ has the property with probability at least 1/2. (This value exists if $\mathcal{F}$ is non-empty and monotone, as this probability then increases monotonically with m). A famous result of Bollobás and Thomason [4] asserts that every monotone property has a threshold function; this is to say that the probability

$$p(m) = \Pr_{A \in \binom{[N]}{m}}[A \in \mathcal{F}]$$

spikes suddenly from $o(1)$ to $1 - o(1)$ when m increases from $o(m^*)$ to $\omega(m^*)$.[1] In general, it is notoriously hard to determine the critical size of a monotone property.

This problem is also wide open for the property of being ℓ-intersective, which is clearly monotone, and for which we denote the critical size by $m^*_\ell(N)$. Bourgain [5] showed that the critical size for 1-intersective sets is given by $m^*_1(N) \asymp \log N$; at present, this is the only case where precise bounds are known. It has been conjectured [11] that $\log N$ is the correct bound for all fixed ℓ, and indeed no better lower bounds are known for $\ell \geq 2$. It was shown by Frantzikinakis, Lesigne and Wierdl [10] and independently by Christ [9] that

$$m^*_2(N) \ll N^{\frac{1}{2}+o(1)}. \tag{1}$$

The same upper bound was later shown to hold for $m^*_3(N)$ by the first author, Dvir and Gopi [6]. More generally, they showed that

$$m^*_\ell(N) \ll N^{1-\frac{1}{\lceil(\ell+1)/2\rceil}+o(1)}, \tag{2}$$

which improved on prior known bounds for all $\ell \geq 3$. The appearance of the peculiar ceiling function in these bounds is due to a reduction for even ℓ to the case $\ell + 1$. The reason for this reduction originates from work on locally decodable error correcting codes [13]. It was shown in [6] that lower bounds on the block length of $(\ell + 1)$-query locally decodable codes (LDCs) imply upper bounds on m^*_ℓ. The bounds (2) then followed directly from the best known LDC bounds; see [7] for a direct proof of (2), however.

For the same reason, a recent breakthrough of Alrabiah et al. [1] on 3-query LDCs immediately implies an improvement of (1) to

$$m^*_2(N) \ll N^{\frac{1}{3}+o(1)}.$$

[1] Our (standard) asymptotic notation is defined as follows. Given a parameter n which grows without bounds and a function $f : \mathbb{R}_+ \to \mathbb{R}_+$, we write: $g(n) = o(f(n))$ to mean $g(n)/f(n) \to 0$; $g(n) = \omega(f(n))$ to mean $g(n)/f(n) \to \infty$; $g(n) \ll f(n)$ to mean that $g(n) \leq Cf(n)$ holds for some constant $C > 0$ and all n; and $g(n) \asymp f(n)$ to mean both $g(n) \ll f(n)$ and $f(n) \ll g(n)$.

For technical reasons, their techniques do not directly generalize to improve the bounds for q-query LDCs with $q \geq 4$, although they could potentially lead to improvements for all odd $q \geq 3$ (but not for even q). Here, we use the ideas of [1] to directly prove upper bounds on m^*_ℓ. Due to the additional arithmetic structure in our problem, it is possible to simplify the exposition and, more importantly, apply the techniques to improve the previous best known bounds for all even $\ell \geq 2$. In particular, we remove the ceiling (raise the roof) in (2).

Theorem 1.1. *For every integer $\ell \geq 2$, we have that*

$$m^*_\ell(N) \ll N^{1-\frac{2}{\ell+1}+o(1)}.$$

2 Outline of the argument

We now give an outline of the proof of Theorem 1.1. Fix an integer $k \geq 3$ and a positive parameter $\varepsilon > 0$, and suppose N is sufficiently large relative to k and ε. Given a sequence of differences $D = (d_1, \ldots, d_m) \in [N]^m$ and some set $A \subseteq [N]$, let $\Lambda_D(A)$ be the normalized count of k-APs with common difference in D which are contained in A:

$$\Lambda_D(A) = \mathbb{E}_{i\in[m]}\mathbb{E}_{x\in[N]} \prod_{\ell=0}^{k-1} A(x+\ell d_i).$$

Let $m \geq 1$ be an integer, and suppose

$$\Pr_{D\in[N]^m}\big(\exists A \subseteq [N] :\ |A| \geq \varepsilon N,\, \Lambda_D(A) = 0\big) \geq 1/2. \tag{3}$$

By a standard averaging argument originally due to Varnavides [16], we can conclude from Szemerédi's theorem that

$$\Lambda_{[N]}(A) \gg_{k,\varepsilon} 1 \quad \text{for all } A \subseteq [N] \text{ with } |A| \geq \varepsilon N \tag{4}$$

(where we identify $[N]$ with the sequence $(1, 2, \ldots, N) \in [N]^N$). Noting that $\mathbb{E}_{D'\in[N]^m}\Lambda_{D'}(A) = \Lambda_{[N]}(A)$, by combining inequalities (3) and (4) we conclude that

$$\mathbb{E}_{D\in[N]^m} \max_{A\subseteq[N]:\,|A|\geq\varepsilon N} \big|\Lambda_D(A) - \mathbb{E}_{D'\in[N]^m}\Lambda_{D'}(A)\big| \gg_{k,\varepsilon} 1.$$

From this last inequality, a simple "symmetrization argument" given in [6] implies

$$\mathbb{E}_{D\in[N]^m}\mathbb{E}_{\sigma\in\{-1,1\}^m} \max_{A\subseteq[N]:\,|A|\geq\varepsilon N} \left|\mathbb{E}_{i\in[m]}\mathbb{E}_{x\in[N]}\, \sigma_i \prod_{\ell=0}^{k-1} A(x+\ell d_i)\right| \gg_{k,\varepsilon} 1;$$

the appearance of the expectation over signs $\sigma \in \{-1,1\}^m$ is crucial to our arguments. By an easy multilinearity argument, we can replace the set $A \subseteq [N]$ (which can be seen as a vector in $\{0,1\}^N$) by a vector $Z \in \{-1,1\}^N$:

$$\mathbb{E}_{D\in[N]^m}\mathbb{E}_{\sigma\in\{-1,1\}^m} \max_{Z\in\{-1,1\}^N} \left|\mathbb{E}_{i\in[m]}\mathbb{E}_{x\in[N]}\, \sigma_i \prod_{\ell=0}^{k-1} Z(x+\ell d_i)\right| \gg_{k,\varepsilon} 1; \tag{5}$$

here and in what follows we use the convention that $Z(y) = 0$ for all $y > N$ when $Z \in \{-1,1\}^N$. The change from $\{0,1\}^N$ to $\{-1,1\}^N$ is a convenient technicality so we can ignore terms which get squared in a product.

This last inequality (5) is what we need to prove the result for even values of k using the arguments we will outline below. For odd values of k, however, this inequality is unsuited due to the odd number of terms inside the product. The main idea from [1] to deal with this case is to apply a "Cauchy-Schwarz trick" to pass from (5) to the inequality

$$\mathbb{E}_{D\in[N]^m}\mathbb{E}_{\sigma\in\{-1,1\}^m} \max_{Z\in\{-1,1\}^N} \sum_{i\in L, j\in R} \sum_{x\in[N]} \sigma_i\sigma_j \prod_{\ell=1}^{k-1} Z(x+\ell d_i)Z(x+\ell d_j) \gg_{k,\varepsilon} m^2 N, \tag{6}$$

where (L, R) is a suitable partition of the index set $[m]$ and we assume (without loss of generality) that m is sufficiently large depending on ε and k.

From now on we assume that k is odd,[2] and write $k = 2r+1$. For $i, j \in [m]$, denote $P_i(x) = \{x+d_i, x+2d_i, \ldots, x+2rd_i\}$ and $P_{ij}(x) = P_i(x) \cup P_j(x)$. From inequality (6) it follows that we can fix a "good" set $D \in [N]^m$ satisfying

$$\mathbb{E}_{\sigma\in\{-1,1\}^m} \max_{Z\in\{-1,1\}^N} \sum_{i\in L, j\in R} \sigma_i\sigma_j \sum_{x\in[N]} \prod_{y\in P_{ij}(x)} Z(y) \gg_{k,\varepsilon} m^2 N \tag{7}$$

and for which we have the technical conditions

$$\big|\big\{i\in L, j\in R:\ |P_{ij}(0)| \neq 4r\big\}\big| \ll_k m^2/N \quad \text{and} \tag{8}$$

$$\max_{x\in[N]} \sum_{i=1}^{m}\sum_{\ell=1}^{2r} \mathbf{1}\{\ell d_i = x\} \ll_k \log N, \tag{9}$$

which are needed to bound the probability of certain bad events later on.

The next key idea is to construct matrices M_{ij} for which the quantity

$$\mathbb{E}_{\sigma\in\{-1,1\}^m} \Big\| \sum_{i\in L, j\in R} \sigma_i\sigma_j M_{ij} \Big\|_{\infty\to 1} \tag{10}$$

is related to the expression on the left-hand side of inequality (7). The reason for doing so is that this allows us to use strong *matrix concentration inequalities*, which can be used to obtain a good upper bound on the expectation (10); this in turn translates to an upper bound on m as a function of N, which is our goal. Such uses of matrix inequalities go back to work of Ben-Aroya, Regev and de Wolf [2], in turn inspired by work of Kerenidis and de Wolf [13] (see also [8]).

The matrices we will construct are indexed by sets of a given size s, where (with hindsight) we choose $s = \lfloor N^{1-2/k} \rfloor$. For $i \in L$, $j \in R$, define the matrix $M_{ij} \in \mathbb{R}^{\binom{[N]}{s}\times\binom{[N]}{s}}$ by

$$M_{ij}(S,T) = \sum_{x\in[N]} \mathbf{1}\big\{|S\cap P_i(x)| = |S\cap P_j(x)| = r,\ S\triangle T = P_{ij}(x)\big\}$$

[2]The even case is similar but simpler. We focus on the odd case here since this is where we obtain new bounds.

if $|P_{ij}(0)| = 4r$, and $M_{ij}(S,T) = 0$ if $|P_{ij}(0)| \neq 4r$. From the definition of this matrix, it is not hard to deduce from inequality (7) a lower bound on the expectation (10): one can show that

$$\mathbb{E}_{\sigma\in\{-1,1\}^m}\Big\|\sum_{i\in L, j\in R}\sigma_i\sigma_j M_{ij}\Big\|_{\infty\to 1} \gg_{k,\varepsilon} \binom{N-4r}{s-2r} m^2 N. \tag{11}$$

Now we need to compute an upper bound for the expectation above. The key ingredient for this is the following non-commutative version of Khintchine's inequality, which can be extracted from a result of Tomczak-Jaegermann [15]:

Theorem 2.1. *Let $n, d \geq 1$ be integers, and let $A_1, \dots, A_n$ be any sequence of $d \times d$ real matrices. Then*

$$\mathbb{E}_{\sigma\in\{-1,1\}^n}\Big\|\sum_{i=1}^n \sigma_i A_i\Big\|_2 \leq 10\sqrt{\log d}\Big(\sum_{i=1}^n \|A_i\|_2^2\Big)^{1/2}.$$

In order to apply this inequality, it is better to collect the matrices M_{ij} into groups and use only one half of the random signs σ_i (another idea from [1]). For $i \in L$, $\sigma_R \in \{-1,1\}^R$, we define the matrix

$$M_i^{\sigma_R} = \sum_{j\in R}\sigma_j M_{ij}.$$

Applying Theorem 2.1 to the expression

$$\mathbb{E}_{\sigma\in\{-1,1\}^L}\Big\|\sum_{i\in L}\sigma_i M_i^{\sigma_R}\Big\|_2$$

(for some fixed $\sigma_R \in \{-1,1\}^R$) and using properties (8) and (9) to bound the sum $\sum_{i\in L}\|M_i^{\sigma_R}\|_2^2$, one can show (with some effort) that

$$\mathbb{E}_{\sigma\in\{-1,1\}^L}\Big\|\sum_{i\in L}\sigma_i M_i^{\sigma_R}\Big\|_2 \ll_{k,\varepsilon} \sqrt{\log\binom{N}{s}}\cdot m^{1/2}(\log N)^k \frac{m}{N^{1-2/k}} \tag{12}$$

holds whenever $m \geq N^{1-2/k}$ (recall that we choose $s = \lfloor N^{1-2/k}\rfloor$).

Finally, we note that

$$\Big\|\sum_{i\in L, j\in R}\sigma_i\sigma_j M_{ij}\Big\|_{\infty\to 1} = \Big\|\sum_{i\in L}\sigma_i M_i^{\sigma_R}\Big\|_{\infty\to 1} \leq \binom{N}{s}\Big\|\sum_{i\in L}\sigma_i M_i^{\sigma_R}\Big\|_2.$$

Averaging over all signs $\sigma \in \{-1,1\}^m$ and combining inequalities (11) and (12), we conclude that $m \ll_{k,\varepsilon} N^{1-2/k}(\log N)^{2k+1}$. As we started with the assumption (3), this shows that $m_{k-1}^*(N) \ll N^{1-2/k}(\log N)^{2k+1}$ as wished.

References

[1] Omar Alrabiah, Venkatesan Guruswami, Pravesh Kothari, and Peter Manohar, *A near-cubic lower bound for 3-query locally decodable codes from semirandom CSP refutation*, 2022, Electronic Colloquium on Computational Complexity (ECCC). Report no. TR22-101.

[2] Avraham Ben-Aroya, Oded Regev, and Ronald de Wolf, *A hypercontractive inequality for matrix-valued functions with applications to quantum computing and ldcs*, 2008 IEEE 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS), IEEE Computer Society, 2008, pp. 477–486.

[3] V. Bergelson and A. Leibman, *Polynomial extensions of van der Waerden's and Szemerédi's theorems*, J. Amer. Math. Soc. **9** (1996), no. 3, 725–753. MR 1325795

[4] B. Bollobás and A. Thomason, *Threshold functions*, Combinatorica **7** (1987), no. 1, 35–38. MR 905149

[5] J. Bourgain, *Ruzsa's problem on sets of recurrence*, Israel J. Math. **59** (1987), no. 2, 150–166. MR 920079

[6] Jop Briët, Zeev Dvir, and Sivakanth Gopi, *Outlaw distributions and locally decodable codes*, Theory Comput. **15** (2019), Paper No. 12, 24. MR 4028880

[7] Jop Briët and Sivakanth Gopi, *Gaussian width bounds with applications to arithmetic progressions in random settings*, Int. Math. Res. Not. IMRN (2020), no. 22, 8673–8696. MR 4216700

[8] Jop Briët, *On embeddings of ℓ_1^k from locally decodable codes*, 2016.

[9] Michael Christ, *On random multilinear operator inequalities*, 2011.

[10] Nikos Frantzikinakis, Emmanuel Lesigne, and Máté Wierdl, *Random sequences and pointwise convergence of multiple ergodic averages*, Indiana Univ. Math. J. **61** (2012), no. 2, 585–617. MR 3043589

[11] ———, *Random differences in Szemerédi's theorem and related results*, J. Anal. Math. **130** (2016), 91–133. MR 3574649

[12] Harry Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256. MR 498471

[13] Iordanis Kerenidis and Ronald de Wolf, *Exponential lower bound for 2-query locally decodable codes via a quantum argument*, J. Comput. System Sci. **69** (2004), no. 3, 395–420, Preliminary version in STOC'03. MR 2087942

[14] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245. MR 369312

[15] Nicole Tomczak-Jaegermann, *The moduli of smoothness and convexity and the Rademacher averages of trace classes* $S_p(1 \leq p < \infty)$, Studia Math. **50** (1974), 163–182. MR 355667

[16] P. Varnavides, *Note on a theorem of Roth*, J. London Math. Soc. **30** (1955), 325–326. MR 76797

[17] Trevor D. Wooley and Tamar D. Ziegler, *Multiple recurrence and convergence along the primes*, Amer. J. Math. **134** (2012), no. 6, 1705–1732. MR 2999293

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Random restrictions of high-rank tensors and polynomial maps

(Extended abstract)

Jop Briët[*] Davi Castro-Silva[†]

Abstract

Motivated by a problem in computational complexity, we consider the behavior of rank functions for tensors and polynomial maps under random coordinate restrictions. We show that, for a broad class of rank functions called *natural rank functions*, random coordinate restriction to a dense set will typically reduce the rank by at most a constant factor.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-033`

1 Introduction

Different but equivalent definitions of matrix rank have been generalized to truly different rank functions for tensors. Although they have proved useful in a variety of applications, the basic theory of these rank functions, describing for instance their interrelations and elementary properties, is still far from complete. Without going into the definitions, we mention below a number of these rank functions to indicate some of the contexts in which they have appeared.

The *slice rank* of a tensor was introduced by Tao [15, 16] to reformulate the breakthrough proof of the cap set conjecture due to Croot, Lev and Pach [3] and Ellenberg and Gijswijt [4]. Slice rank is generalized by the *partition rank*, which was introduced by

[*]CWI & QuSoft, Science Park 123, 1098 XG Amsterdam, The Netherlands. Supported by the Dutch Research Council (NWO) as part of the NETWORKS programme (grant no. 024.002.003).
[†]CWI & QuSoft, Science Park 123, 1098 XG Amsterdam, The Netherlands. Supported by the Dutch Research Council (NWO) as part of the NETWORKS programme (grant no. 024.002.003).

Naslund to prove bounds on the size of subsets of $\mathbb{F}_q^n$ without k-right corners [12], as well as provide exponential improvements on the Erdős–Ginzburg–Ziv constant [11]. The *analytic rank* is based on a measure of equidistribution for multilinear forms associated to tensors over finite fields, and was introduced by Gowers and Wolf to study solutions to linear systems of equations in large subsets of finite vectors spaces [6]. *Geometric rank*, defined and studied by Kopparty, Moshkovitz and Zuiddam in the context of algebraic complexity theory [10], gives a natural analogue of analytic rank for tensors over infinite fields.

Closely related to these rank functions for tensors are notions of rank for multivariate polynomials. A notion of polynomial rank akin to the partition rank of tensors was used already in the '80s by Schmidt in work on algebraic geometry [13], and has since been re-discovered and proven useful on several occasions. Work on the Inverse Theorem for the Gowers uniformity norms led Green and Tao to define the notion of *degree rank* [7], which quantifies how hard it is to express the considered polynomial as a function of lower-degree polynomials; this notion was shown to be closely linked to equidistribution properties of multivariate polynomials over prime fields $\mathbb{F}_p$. Tao and Ziegler [17] later studied the relationship between the degree rank of a polynomial and its *analytic rank*, defined as the (tensor) analytic rank of its associated homogeneous multilinear form, and exploited their close connection in order to prove the general case of the Gowers Inverse Theorem over $\mathbb{F}_p^n$.

Recent work on constant-depth Boolean circuits by Buhrman, Neumann and the present authors gave rise to a problem on equidistribution properties of higher-dimensional polynomial maps under biased input distributions [1]. This motivated a new notion of analytic rank for (high-dimensional) polynomial maps and prompted the study of rank under random coordinate restrictions, which is the topic of this work.

Common to the tensors, polynomials and polynomial maps considered here is that they can be viewed as maps on $\mathbb{F}^X$, where $\mathbb{F}$ is a given field and X is a finite set indexing the variables. The main question we address is whether, if a map ϕ on $\mathbb{F}^X$ has high rank, then most of its coordinate restrictions $\phi_{|I}$ on $\mathbb{F}^I$ also have high rank for dense subsets $I \subseteq X$ (where we also respect the product structure of X in the case of tensors). Our main results show that this is the case for all "natural" rank functions, which include all those mentioned above.

2 The matrix case

It is instructive to first consider the case of matrices, which is simpler and illustrates the spirit of our main results. For a matrix $A \in \mathbb{F}^{n\times n}$ and subsets $I, J \subseteq [n]$, denote by $A_{|I\times J}$ the sub-matrix of A induced by the rows in I and columns in J. Given $\sigma \in (0,1)$, consider a random set $I \subseteq [n]$ containing each element independently with probability σ; we write $I \sim [n]_\sigma$ when I is distributed as such. Note that, if $I \sim [n]_\rho$ and $J \sim [n]_\sigma$ are independent, then $I \cup J \sim [n]_\eta$ with $\eta = 1-(1-\rho)(1-\sigma)$.

Proposition 2.1. *For every $\sigma \in (0,1]$ there exists $\kappa \in (0,1]$ such that for every matrix*

$A \in \mathbb{F}^{n\times n}$ *we have*

$$\Pr_{I\sim[n]_\sigma}\big[\operatorname{rk}(A_{|I\times I}) \geq \kappa\cdot\operatorname{rk}(A)\big] \geq 1 - 2e^{-\kappa\operatorname{rk}(A)}.$$

Proof: Write $\rho = 1-\sqrt{1-\sigma}$ and let $J, J' \sim [n]_\rho$ be independent random sets; note that $J\cup J' \sim [n]_\sigma$. Let $r = \operatorname{rk}(A)$, and fix a set $S \subseteq [n]$ of r linearly independent rows of A. By the Chernoff bound [8], the probability that the set J satisfies $|J\cap S| < \rho r/2$ is at most $e^{-\rho r/8}$.

Now let $B := A_{|(J\cap S)\times[n]}$ be the (random) sub-matrix of A formed by the rows in $J\cap S$. Since its rows are linearly independent, the rank of B is precisely $|J\cap S|$; let $T\subseteq[n]$ be a set of $|J\cap S|$ linearly independent columns of B. Then the probability that $|J'\cap T| < \rho|T|/2$ is at most $e^{-\rho|T|/8}$, and the rank of $B_{|(J\cap S)\times(J'\cap T)} = A_{|(J\cap S)\times(J'\cap T)}$ is equal to $|J'\cap T|$. It follows from the union bound and monotonicity of rank under restrictions that, with probability at least $1-2e^{-\rho^2 r/16}$, the principal sub-matrix of A induced by $J\cup J'$ has rank at least $\rho^2 r/4$. The result now follows since $J\cup J' \sim [n]_\sigma$. □

3 Main results

Here we generalize Proposition 2.1 to tensors and polynomial maps for rank functions that satisfy a few natural properties, namely "sub-additivity", "monotonicity", a "Lipschitz condition" and, in the case of polynomial maps, "symmetry" (see below for the precise definitions). Those functions which satisfy these properties are called *natural rank functions*; we note that all notions of rank mentioned in the Introduction are natural rank functions.

Since our results are independent of the field considered (which can be finite or infinite), we will always denote it by $\mathbb{F}$ and suppress statements of the form "let $\mathbb{F}$ be a field" or "for every field $\mathbb{F}$".

3.1 Tensors

We begin by considering the case of tensors.

Definition 3.1. For finite sets $X_1,\dots,X_d \subset \mathbb{N}$, a d-tensor is a map $T : X_1\times\cdots\times X_d \to \mathbb{F}$. We will associate with any d-tensor T a multilinear map $\mathbb{F}^{X_1}\times\cdots\times\mathbb{F}^{X_d} \to \mathbb{F}$ and an element of $\mathbb{F}^{X_1}\otimes\cdots\otimes\mathbb{F}^{X_d}$ in the obvious way, and also denote these objects by T.

For a tensor T as in Definition 3.1 and subsets $I_1 \subseteq X_1,\dots,I_d\subseteq X_d$, denote $I_{[d]} = I_1\times\cdots\times I_d$ and write $T_{|I_{[d]}}$ for the restriction of T to $I_{[d]}$. If T is viewed as an element of $\mathbb{F}^{X_1}\otimes\cdots\otimes\mathbb{F}^{X_d}$, then $T_{|I_{[d]}}$ is simply a sub-tensor.

We denote the set of d-tensors over $\mathbb{F}$ with finite support by $(\mathbb{F}^\infty)^{\otimes d}$; note that the tensors defined on finite sets naturally embed into this set, and that the rank functions for tensors discussed above are invariant under this embedding. The notions of tensor rank we will consider here are those called *natural rank functions* as defined below:

Definition 3.2. We say that $\mathrm{rk} : (\mathbb{F}^\infty)^{\otimes d} \to \mathbb{R}_+$ is a natural rank function if it satisfies the following properties:

1. Sub-additivity:
   $\mathrm{rk}(T+S) \le \mathrm{rk}(T) + \mathrm{rk}(S)$ for all $T, S \in (\mathbb{F}^\infty)^{\otimes d}$.
2. Monotonicity under restrictions:
   $\mathrm{rk}\left(T_{|I_{[d]}}\right) \le \mathrm{rk}(T)$ for all $T \in (\mathbb{F}^\infty)^{\otimes d}$ and all sets $I_1, \dots, I_d \subset \mathbb{N}$.
3. Restriction Lipschitz property:
   $\mathrm{rk}\left(T_{|J_{[d]}}\right) \le \mathrm{rk}\left(T_{|I_{[d]}}\right) + \sum_{i=1}^d |J_i \setminus I_i|$ for all $T \in (\mathbb{F}^\infty)^{\otimes d}$ and all sets $I_1 \subseteq J_1, \dots, I_d \subseteq J_d$.

Our main result in this setting concerns how natural rank functions behave under random coordinate restrictions. Intuitively, it shows that random restrictions of high-rank tensors will also have high rank with high probability. It can be formally stated as follows:

Theorem 3.3. *For every $d \in \mathbb{N}$ and $\sigma \in (0,1]$, there exist constants $C, \kappa > 0$ such that the following holds. For every natural rank function* $\mathrm{rk} : (\mathbb{F}^\infty)^{\otimes d} \to \mathbb{R}_+$ *and every d-tensor $T \in \bigotimes_{i=1}^d \mathbb{F}^{n_i}$ we have*

$$\Pr_{I_1 \sim [n_1]_\sigma, \dots, I_d \sim [n_d]_\sigma}\left[\mathrm{rk}\left(T_{|I_{[d]}}\right) \ge \kappa \cdot \mathrm{rk}(T)\right] \ge 1 - Ce^{-\kappa\,\mathrm{rk}(T)}.$$

From this theorem one can easily deduce a more symmetric version, which is valid in the standard case of "cubic" tensors where every row is indexed by the same set:

Corollary 3.4. *For every $d \in \mathbb{N}$ and $\sigma \in (0,1]$, there exist constants $C, \kappa > 0$ such that the following holds. For every natural rank function* $\mathrm{rk} : (\mathbb{F}^\infty)^{\otimes d} \to \mathbb{R}_+$ *and every d-tensor $T \in (\mathbb{F}^n)^{\otimes d}$ we have*

$$\Pr_{I \sim [n]_\sigma}\left[\mathrm{rk}\left(T_{|I^d}\right) \ge \kappa \cdot \mathrm{rk}(T)\right] \ge 1 - Ce^{-\kappa\,\mathrm{rk}(T)}.$$

3.2 Polynomial maps

Next we consider the setting of polynomials and higher-dimensional polynomial maps.

Definition 3.5. A polynomial map is an ordered tuple $\phi(x) = \big(f_1(x), \dots, f_k(x)\big)$ of polynomials $f_1, \dots, f_k \in \mathbb{F}[x_1, \dots, x_n]$. We identify with ϕ a map $\mathbb{F}^n \to \mathbb{F}^k$ in the natural way. The degree of ϕ is the maximum degree of the f_i.

For a polynomial map $\phi : \mathbb{F}^n \to \mathbb{F}^k$ and a set $I \subseteq [n]$, define the restriction $\phi_{|I} : \mathbb{F}^I \to \mathbb{F}^k$ to be the map given by $\phi_{|I}(y) = \phi(\bar{y})$, where $\bar{y} \in \mathbb{F}^n$ agrees with y on the coordinates in I and is zero elsewhere.

We denote the space of all polynomial maps $\phi : \mathbb{F}^n \to \mathbb{F}^k$ of degree at most d by $\mathrm{Pol}_{\le d}(\mathbb{F}^n, \mathbb{F}^k)$, and write

$$\mathrm{Pol}_{\le d}(\mathbb{F}^\infty, \mathbb{F}^k) = \bigcup_{n \in \mathbb{N}} \mathrm{Pol}_{\le d}(\mathbb{F}^n, \mathbb{F}^k).$$

The notions of rank we consider are defined below:

Definition 3.6. We say that $\mathrm{rk} : \mathrm{Pol}_{\leq d}(\mathbb{F}^\infty, \mathbb{F}^k) \to \mathbb{R}_+$ is a natural rank function if it satisfies the following properties:

1. Symmetry:
   $\mathrm{rk}(\phi) = \mathrm{rk}(-\phi)$ for all $\phi \in \mathrm{Pol}_{\leq d}(\mathbb{F}^\infty, \mathbb{F}^k)$.

2. Sub-additivity:
   $\mathrm{rk}(\phi + \gamma) \leq \mathrm{rk}(\phi) + \mathrm{rk}(\gamma)$ for all $\phi, \gamma \in \mathrm{Pol}_{\leq d}(\mathbb{F}^\infty, \mathbb{F}^k)$.

3. Monotonicity under restrictions:
   $\mathrm{rk}(\phi_{|I}) \leq \mathrm{rk}(\phi)$ for all $\phi \in \mathrm{Pol}_{\leq d}(\mathbb{F}^\infty, \mathbb{F}^k)$ and all sets $I \subset \mathbb{N}$.

4. Restriction Lipschitz property:
   $\mathrm{rk}(\phi_{|I\cup J}) \leq \mathrm{rk}(\phi_{|I}) + |J|$ for all $\phi \in \mathrm{Pol}_{\leq d}(\mathbb{F}^\infty, \mathbb{F}^k)$ and all sets $I, J \subset \mathbb{N}$.

Our second main result shows that random restrictions of a high-rank polynomial map will also have high rank with high probability. Its formal statement is given as follows:

Theorem 3.7. *For every $d \in \mathbb{N}$ and $\sigma, \varepsilon \in (0, 1]$, there exist constants $\kappa = \kappa(d, \sigma) > 0$ and $R = R(d, \sigma, \varepsilon) \in \mathbb{N}$ such that the following holds. For every natural rank function* rk : $\mathrm{Pol}_{\leq d}(\mathbb{F}^\infty, \mathbb{F}^k) \to \mathbb{R}_+$ *and every map $\phi \in \mathrm{Pol}_{\leq d}(\mathbb{F}^n, \mathbb{F}^k)$ with* $\mathrm{rk}(\phi) \geq R$*, we have*

$$\mathrm{Pr}_{I\sim[n]_\sigma}\big[\,\mathrm{rk}(\phi_{|I}) \geq \kappa \cdot \mathrm{rk}(\phi)\big] \geq 1 - \varepsilon.$$

3.3 The proofs

Whereas the proof of the matrix case (Proposition 2.1) uses in an essential way the fact that a rank-r matrix contains a full-rank $r \times r$ submatrix, an analogous property is not known to be true in general for tensors and polynomial maps. In fact, it was shown by Gowers that such a property is *false* in the case of slice rank for 3-tensors (see [9, Proposition 3.1]). Karam [9] recently studied the extent for which similar but quantitatively weaker properties hold for tensor rank functions, but the quantitative bounds obtained are still insufficient for an argument akin to that of Proposition 2.1 to work.

The proofs of our main theorems must then proceed differently from the simpler case of matrices. Our proof of Theorem 3.3 (the tensor case) uses instead ideas from probability theory, in particular concerning concentration inequalities on product spaces; it relies mainly on an inequality of Talagrand [14, Theorem 3.1.1].

The proof of Theorem 3.7 (for polynomial maps) is again very different from the tensor case, which implicitly makes use of multilinearity; it relies instead on results from the analysis of Boolean functions, in particular Friedgut's Junta Theorem [5], taken together with elementary (but somewhat involved) combinatorial arguments. The full proofs can be found in the full version of our paper [2].

References

[1] Jop Briët, Harry Buhrman, Davi Castro-Silva, and Niels M. P. Neumann, *Noisy decoding by shallow circuits with parities: classical and quantum*, 2023.

[2] Jop Briët and Davi Castro-Silva, *Random restrictions of high-rank tensors and polynomial maps*, 2022.

[3] Ernie Croot, Vsevolod F. Lev, and Péter Pál Pach, *Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small*, Ann. of Math. (2) **185** (2017), no. 1, 331–337. MR 3583357

[4] Jordan S. Ellenberg and Dion Gijswijt, *On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression*, Ann. of Math. (2) **185** (2017), no. 1, 339–343. MR 3583358

[5] Ehud Friedgut, *Boolean functions with low average sensitivity depend on few coordinates*, Combinatorica **18** (1998), no. 1, 27–35. MR 1645642

[6] W. T. Gowers and J. Wolf, *Linear forms and higher-degree uniformity for functions on $\mathbb{F}_p^n$*, Geom. Funct. Anal. **21** (2011), no. 1, 36–69. MR 2773103

[7] Ben Green and Terence Tao, *The distribution of polynomials over finite fields, with applications to the Gowers norms*, Contrib. Discrete Math. **4** (2009), no. 2, 1–36. MR 2592422

[8] Torben Hagerup and Christine Rüb, *A guided tour of Chernoff bounds*, Information Processing Letters **33** (1990), no. 6, 305–308.

[9] Thomas Karam, *High-rank subtensors of high-rank tensors*, 2022, arXiv:2207.08030.

[10] Swastik Kopparty, Guy Moshkovitz, and Jeroen Zuiddam, *Geometric rank of tensors and subrank of matrix multiplication*, 35th Computational Complexity Conference, LIPIcs. Leibniz Int. Proc. Inform., vol. 169, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2020, pp. Art. No. 35, 21.

[11] Eric Naslund, *Exponential bounds for the Erdős-Ginzburg-Ziv constant*, J. Combin. Theory Ser. A **174** (2020), 105185, 19.

[12] ———, *The partition rank of a tensor and k-right corners in $\mathbb{F}_q^n$*, J. Combin. Theory Ser. A **174** (2020), 105190, 25. MR 4078997

[13] Wolfgang M. Schmidt, *The density of integer points on homogeneous varieties*, Acta Math. **154** (1985), no. 3-4, 243–296. MR 781588

[14] Michel Talagrand, *Concentration of measure and isoperimetric inequalities in product spaces*, Inst. Hautes Études Sci. Publ. Math. (1995), no. 81, 73–205. MR 1361756

[15] Terence Tao, https://terrytao.wordpress.com/2016/05/18/a-symmetric-formulation-of-the-croot-lev-pach-ellenberg-gijswijt-capset-bound/, 2016.

[16] Terence Tao and Will Sawin, https://terrytao.wordpress.com/2016/08/24/notes-on-the-slice-rank-of-tensors/, 2016.

[17] Terence Tao and Tamar Ziegler, *The inverse conjecture for the Gowers norm over finite fields in low characteristic*, Ann. Comb. **16** (2012), no. 1, 121–188. MR 2948765

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Strict Erdős-Ko-Rado for simplicial complexes

(Extended abstract)

Denys Bulavka* Russ Woodroofe†

Abstract

We show that the strict Erdős-Ko-Rado property holds for sequentially Cohen-Macaulay near-cones. In particular, this implies that chordal graphs with at least one isolated vertex satisfy the strict Erdős-Ko-Rado property.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-034`

1 Introduction

What is the largest cardinality of a family of pairwise-intersecting sets? A now-classic result of Erdős, Ko, and Rado answers this question if the sets all have the same number of elements, and are otherwise unrestricted.

Theorem 1.1 (Erdős, Ko, and Rado [4]). *Let $r \leq n/2$. If $\mathcal{F}$ is a family of pairwise-intersecting subsets of $[n]$, each with r elements, then $|\mathcal{F}| \leq \binom{n-1}{r-1}$.*

If $|\mathcal{F}|$ achieves the upper bound and $r < n/2$, then $\mathcal{F}$ consists of all the r-element subsets containing some fixed element.

*Department of Applied Mathematics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic. Supported by Grant Schemes at CU, reg. no. CZ.02.2.69/0.0/0.0/19_073/0016935 and by grant no. 21-32817S of the Czech Science Foundation (GAČR). Email: dbulavka@kam.mff.cuni.cz.

†Univerza na Primorskem, Glagoljäka 8, 6000 Koper, Slovenia. Supported in part by the Slovenian Research Agency (research program P1-0285 and research projects J1-9108, N1-0160, J1-2451). Email: russ.woodroofe@famnit.upr.si. `https://osebje.famnit.upr.si/~russ.woodroofe/`.

That is, under the above hypotheses a family of pairwise-intersecting objects of maximal size is given by a family with a common intersection. Moreover, under slightly stronger hypotheses, this is the only such family. Hilton and Milner [8] later gave upper bounds for pairwise-intersecting families that do not all contain a common element.

There are a large number of generalizations of Theorem 1.1. We focus on one in particular. Holroyd and Johnson asked at the 1997 British Combinatorial Conference [13] about whether an analogue of Erdős-Ko-Rado property holds for independent sets in cyclic and similar graphs. Talbot showed the answer to be "yes" in a strong sense.

Theorem 1.2 (Talbot [22])**.** *Let n, k, r be positive integers such that $r \leq n/(k+1)$. Let G be the graph with vertex set $\mathbb{Z}_n$ and edges consisting of those x, y such that $x - y \in \{1, \ldots, k\}$.*

If $\mathcal{F}$ is a family of pairwise-intersecting independent sets of G, each with r elements, then $|\mathcal{F}|$ is smaller than the family $\mathcal{B}$ of all independent sets with r elements containing 0. If $|\mathcal{F}|$ achieves the upper bound and $n \neq 2r + 2$, then $\mathcal{F}$ is $\mathcal{B}$ up to relabeling the vertices.

Holroyd and Talbot asked whether similar results hold for independent sets in other graphs G. There are counterexamples for r around the size of a maximum independent set, but not for somewhat smaller r. Since the collection of independent sets form a simplicial complex, and since it is our main object of study, we introduce it now. A *simplicial complex* K with vertex set V is a set system $K \subseteq 2^V$ that is closed under taking susbset, i.e., if $F \in K$ and $G \subseteq F$ then $G \in K$. In this article we will assume that the vertex set V is equipped with a total order and consequently we can identify it with $[|V|] = \{1, \ldots, |V|\}$ respecting this order. The members of K are called *faces*, the faces of size r are called *r-faces*, the number of r-faces in K is denoted by $f_r(K)$ and the maximal faces with respect to inclusion are called *facets*. The *dimension of a face F* is defined by $\dim(F) = |F| - 1$ and the *dimension of* K by $\dim(K) = \max_{F \in K} \dim(F)$. The *$r$-skeleton of* K, denote by $K^{(r)}$, is the set of faces from K of dimension at most r. The *pure r-skeleton* of K is the simplicial complex given by all the faces from K of dimension r and their subsets. Given a face $F \in K$, the *link of* F in K is defined as the simplicial complex $\mathrm{lk}(F, K) = \{G \in K : G \cap F = \emptyset \text{ and } F \cup G \in K\}$. By $K[S] = \{F \in K : F \subseteq S\}$ we denote the *induced* simplicial complex on the vertices S. For two simplicial complexes K and L, the *join* is defined by $K * L = \{F \sqcup G : F \in K, G \in L\}$, where $\sqcup$ denotes disjoint union.

Conjecture 1.3 (Holroyd and Talbot [10], extended by Borg to arbitrary simplicial complexes [1])**.** *Let K be a simplicial complex whose smallest facet has d vertices, and let $r \leq d/2$. If $\mathcal{F}$ is a family of pairwise-intersecting faces of K, each with r elements, then there is some vertex v of K so that $|\mathcal{F}| \leq f_{r-1}(\mathrm{lk}(v, K))$. If $r < d/2$ and $|\mathcal{F}|$ achieves the upper bound, then $\mathcal{F}$ consists of the faces containing some vertex v.*

If a simplicial complex K satisfies the upper bound of Conjecture 1.3 at a specified value of r, then we say that K is *r-EKR*. If every intersecting family of maximum size has a common intersection, then we say that K is *strictly r-EKR*. We abuse terminology to say that a graph is (strictly) r-EKR if its independence complex has the same property.

There has been considerable work on Conjecture 1.3. Hurlbert and Kamat showed [11] that any chordal graph with an isolated vertex satisfies the upper bound of Conjecture 1.3.

Borg showed [1] that the conjecture holds asymptotically, in the precise sense that if the minimal facet cardinality of a simplicial complex K is at least $(r-1)\binom{3r-3}{2}+r$, then K satisfies r-EKR. Other related works are [10, 9, 19]. Rather than working with arbitrary simplicial complexes we will focus on the so called sequentially Cohen-Macaulay near-cones which we introduce now. A simplicial complex K is a near-cone with apex v if for every $F \in K$ and every $w \in F$ we have that $F \setminus \{w\} \cup \{v\} \in K$. A simplicial complex is called *Cohen-Macaulay over* $\mathbb{F}$ if for every face $F \in K$ we have that $\tilde{H}_i(\mathrm{lk}(F,K),\mathbb{F}) = 0$ for $i < \dim(\mathrm{lk}(F,K))$, that is the reduced homology of every link vanishes on every dimension except possibly the top one. A simplicial complex is said to be *sequentially Cohen-Macaulay over* $\mathbb{F}$ if for every r, the pure r-skeleton of K is Cohen-Macaulay over $\mathbb{F}$. From now on we will assume that the field has characteristic 0 and drop it from the notation. The second author showed more generally [25] that any sequentially Cohen-Macaulay near-cone satisfies the upper bound of Conjecture 1.3. We note here that the independence complex of a graph G is a cone if and only if G has an isolated vertex. Moreover, the class of sequentially Cohen-Macaulay simplicial complexes is a broad class that includes the independence complexes of chordal graphs and many others [2, 17, 24]. Neither Hurlbert and Kamat nor the second author addressed the strict r-EKR property.

The main purpose of the current paper is to fill in this gap. We show:

Theorem 1.4. *Let $2 \le r < d/2$. If the simplicial complex K is a sequentially Cohen-Macaulay near-cone with minimal facet cardinality d, then K is strictly r-EKR, that is the pairwise-intersecting families of maximum size consist of all r-faces containing an apex vertex.*

The novelty of our techniques is to combine algebraic and combinatorial shifting operations. We also make use of some of the ideas behind proofs of the Hilton-Milner theorem [7, 12].

This article is organized as follows, in Section 2 we review the main results needed for the proof. In Section 3 we give the proof of Theorem 1.4.

2 Shifting

A set system $\mathcal{F} \subseteq \binom{[n]}{r}$ is said to be *shifted* if for every $F \in \mathcal{F}$ and $i,j \in [n]$ such that $i < j$, $j \in F$ and $i \notin F$ we have that $F \setminus \{j\} \cup \{i\} \in \mathcal{F}$. In this section we review two operations that assign to a set system another set system that is shifted while preserving several properties of interest.

Given a set system $\mathcal{F} \subseteq \binom{[n]}{r}$, and $F \in \mathcal{F}$. Let $i,j \in [n]$ such that $i < j$, the *combinatorial shift* $\mathrm{Shift}_{i,j}$ is defined by

$$\mathrm{Shift}_{i,j}(F) = \begin{cases} F \setminus \{j\} \cup \{i\} & \text{if } j \in F, i \notin F \text{ and } F \setminus \{j\} \cup \{i\} \notin \mathcal{F}, \\ F & \text{otherwise.} \end{cases}$$

$$\mathrm{Shift}_{i,j}(\mathcal{F}) = \{\mathrm{Shift}_{i,j}(F) \colon F \in \mathcal{F}\}.$$

We will be using the following properties of combinatorial shifting [6].

Theorem 2.1. *Let $\mathcal{F} \subseteq \binom{[n]}{r}$ and $i, j \in [n]$ such that $i < j$.*

1. $|\operatorname{Shift}_{i,j}(\mathcal{F})| = |\mathcal{F}|$.
2. *If $\mathcal{G} \subseteq \mathcal{F}$, then $\operatorname{Shift}_{i,j}(\mathcal{G}) \subseteq \operatorname{Shift}_{i,j}(\mathcal{F})$ where in each case we do the combinatorial shifting according to the respective family.*
3. *If $\mathcal{F}$ is shifted, then $\operatorname{Shift}_{i,j}(\mathcal{F}) = \mathcal{F}$.*
4. *If $\mathcal{F}$ is pairwise-intersecting, then $\operatorname{Shift}_{i,j}(\mathcal{F})$ is pairwise-intersecting.*

By iterating the combinatorial shifting operation we will eventually obtain a set system that is shifted, but the final set system is dependent on the order of the shifts.

Kalai [15] introduced a shifting operation that produces a shifted set system preserving several algebraic properties, so-called *(exterior) algebraic shifting*. This operation assigns to a simplicial complex K a shifted simplicial complex $\Delta(K)$. We would like to point out that in contrast to combinatorial shifting, algebraic shifting works in one step rather than an iterative procedure. Here we merely state the properties we will be using.

Theorem 2.2. *Let K be a simplicial complex.*

1. *[15, Theorem 2.1.2] $\Delta(K)$ is shifted.*
2. *[15, Theorem 2.2.7] $K \subseteq L$, then $\Delta(K) \subseteq \Delta(L)$.*
3. *[15, Theorem 2.1.1] $f_r(K) = f_r(\Delta(K))$.*
4. *[15, Theorem 4.1] K is Cohen-Macaulay then $\Delta(K)$ is Cohen-Macaulay.*
5. *[15, Theorem 6.2] If $\mathcal{F} \subseteq \binom{[n]}{a}$ and $\mathcal{G} \subseteq \binom{[n]}{b}$ are cross-intersecting, then $\Delta(\mathcal{F})$ and $\Delta(\mathcal{G})$ are cross-intersecting.*
6. *[18, Theorem 5.3] If K is a near-cone with apex v, then $\Delta(K) = (1 * \Delta(\operatorname{lk}(v, K))) \cup B$ where $B = \{F \in \Delta(K) \colon 1 \notin F\}$. In particular, $f_r(\operatorname{lk}(1, \Delta(K))) = f_r(\Delta(\operatorname{lk}(v, K))) = f_r(\operatorname{lk}(v, K))$.*

Since the minimal facet cardinality plays a key role in Conjecture 1.3. We need to be able to control its behavior when performing (algebraic) shifting operations. For this purpose we introduce the following definition of *depth of a simplicial complex K*

$$\operatorname{depth} K = \max\{d \colon K^{(d)} \text{ is Cohen-Macaulay}\}.$$

The depth of a simplicial complex is one less than the depth of its Stanley-Reisner ring [21].

Corollary 2.3 ([3])**.** *The minimum facet dimension of $\Delta(K)$ is at least d if and only if $K^{(d)}$ is Cohen-Macaulay over $\mathbb{F}$.*

From the above corollary it follows that $\operatorname{depth} K + 1$ is the minimum facet cardinality of $\Delta(K)$ which is at most the minimal facet cardinality of K. Notice that when K is sequentially Cohen-Macaulay, the minimal facet cardinality of K coincides with $\operatorname{depth} K + 1$ and consequently with the minimal facet cardinality of its algebraic shift $\Delta(K)$.

3 Proof of Theorem 1.4

First, we adapt the approach in [7, 12] for shifted simplicial complexes. For $\{i,j\} \in \binom{[n]}{2}$, let $\mathrm{swap}_{i,j}$ denote the function exchanging vertices i and j.

Lemma 3.1. *Let K be a shifted simplicial complex with minimal facet cardinality d and $\mathcal{F} \subseteq K$ a non-trivial pairwise-intersecting family of r-faces with $r \leq d/2$ of maximal size. Then, there exists a shifted non-trivial pairwise-intersecting family $\mathcal{F}' \subseteq K$ of r-faces such that $|\mathcal{F}| = |\mathcal{F}'|$.*

Proof. Consider $\mathcal{F}$ of maximal size. We apply $\mathrm{Shift}_{i,j}$ repeatedly to $\mathcal{F}$ until it results in a trivial pairwise-intersecting family. Let $\mathrm{Shift}_{s,t}$ be the first shifting operation making the family trivial and let $\mathcal{H}$ be the non-trivial family before applying the last shifting $\mathrm{Shift}_{s,t}$. Because K is shifted, the repeated application of combinatorial shifting to the pairwise-intersecting family keeps the family in the simplicial complex at each step. That is, $\mathrm{Shift}_{i,j}(\mathcal{F}) \subseteq \mathrm{Shift}_{i,j}(K) = K$, where the first inclusion follows from Theorem 2.1.2 and the last step from Theorem 2.1.3.

If $\mathrm{Shift}_{1,s}(\mathcal{H})$ is non-trivial and $\mathrm{Shift}_{2,t} \circ \mathrm{Shift}_{1,s}(\mathcal{H})$ is non-trivial then applying $\mathrm{Shift}_{1,2}$ to this last family gives a trivial one and we are in the same situation as [7, Proposition 1.6], that is $s = 1$ and $t = 2$. If $\mathrm{Shift}_{1,s}(\mathcal{H})$ is trivial while $\mathrm{Shift}_{2,s}(\mathcal{H})$ is non-trivial, then $\mathrm{Shift}_{1,2} \circ \mathrm{Shift}_{2,s}(\mathcal{H})$ is trivial and we are again in the same situation as above. The remaining case is when we have that $\mathrm{Shift}_{1,s}(\mathcal{H})$ and $\mathrm{Shift}_{2,s}(\mathcal{H})$ (or $\mathrm{Shift}_{1,t}(\mathcal{H})$ and $\mathrm{Shift}_{2,t}(\mathcal{H})$) are both trivial. We will use repeatedly the following argument: if $\mathcal{F}$ is non-trivial pairwise-intersecting family of maximal size, and $\mathrm{Shift}_{i,j}(\mathcal{F})$ is trivial then $\{i,j\} \cap F \neq \emptyset$ for all $F \in \mathcal{F}$ and, because of maximality of $|\mathcal{F}|$, we have that

$$\mathcal{T}_{i,j} = \{T \cup \{i,j\} \colon T \in \mathrm{lk}(\{i,j\}, K), |T| = r-2\} \subseteq \mathcal{F}.$$

Let $3 \leq s_0 < s$ such that $\mathcal{H}' = \mathrm{Shift}_{s_0,s}(\mathcal{H})$ is non-trivial, we take the first one if it exists, or we set $s_0 = s$ otherwise. We notice that $s_0 \leq r+1$, since otherwise every member of $\mathcal{H}$ would contain $[r+1]$ or s, since the first option is not possible due to its size then $\mathcal{H}'$ is trivial, which is a contradiction. Since $\mathrm{Shift}_{i,s_0}(\mathcal{H}')$ is trivial for all $i \in [s_0 - 1]$ while $\mathcal{H}'$ is not, a routine computation shows that the following holds: $\mathcal{T}_{i,s_0} \subseteq \mathcal{H}'$ for $i \in [s_0 - 1]$; for each $F \in \mathcal{H}'$ we have that $[s_0 - 1] \subseteq F$ or $s_0 \in F$; for each $i \in [s_0 - 1]$ there exist $F_i \in \mathcal{H}'$ such that $i \notin F_i$ and $s_0 \in F_i$; there exist F_0 such that $[s_0 - 1] \subseteq F_0$ and $s_0 \notin F_0$. Finally, set $\mathcal{G} = \mathrm{swap}_{1,s_0}(\mathcal{H}')$ and $G_i = \mathrm{swap}_{1,s_0}(F_i)$, then $1 \notin G_0$ and $s_0 \notin G_1$.

Claim 1: $\mathcal{G}$ is intersecting. The only non-trivial case to verify is if $F, F' \in \mathcal{H}'$ are such that $F \cap [s_0] = [s_0 - 1]$ and $F' \cap [s_0] = s_0$. Since $\mathcal{H}'$ is intersecting, there exists $x \in F \cap F'$. Then, $x \notin [s_0]$ and consequently it is not affected by swap_{1,s_0}. From this we can conclude that $x \in \mathrm{swap}_{1,s_0}(F) \bigcap \mathrm{swap}_{1,s_0}(F')$.

Claim 2: $\mathcal{G}$ is non-trivial. If $x \in \bigcap \mathcal{G}$, then since $\bigcap \mathcal{H}' = \emptyset$ we must have that $x = 1$ or $x = s_0$. But, $1 \notin G_0$ and $s_0 \notin G_1$.

Claim 3: $\mathcal{G} \subseteq K$. For $F \in \mathcal{H}'$ such that $s_0 \in F$ and $1 \notin F$, $\mathrm{swap}_{1,s_0}(F) \in K$ since K is shifted. For $F \in \mathcal{H}'$ such that $1 \in F$ and $s_0 \notin F$ consider F' given by adding the first r

vertices of $[n] \setminus F$ to F, this is a face since the minimal facet cardinality is $d \geq 2r$ and it contains s_0 since $s_0 \leq r+1$. Consequently $\mathrm{swap}_{1,s_0}(F) \in K$.

Claim 4: $\mathcal{T}_{1,2} \subseteq \mathcal{G}$. Since $s_0 \geq 3$, for $i \in [s_0 - 1]$ we have that $\mathcal{T}_{i,s_0} \subseteq \mathcal{H}'$ and consequently $\mathcal{T}_{1,i} = \mathrm{swap}_{1,s_0}(\mathcal{T}_{i,s_0}) \subseteq \mathrm{swap}_{1,s_0}(\mathcal{H}') = \mathcal{G}$.

Next, we apply repeatedly $\mathrm{Shift}_{i,j}$ with $3 \leq i < j \leq n$ to $\mathcal{G}$ until it is stable under this restricted set of combinatorial shiftings. We notice that this does not change $\mathcal{T}_{1,2}$. Denote the stable set by $\mathcal{G}'$.

Claim 5: $\mathcal{G}'$ is non-trivial. Since $[s_0 - 1] \setminus \{1\} \subseteq G_0 \in \mathcal{G}$, then $[r+1] \setminus \{1\} = [s_0] \setminus 1 \cup \{s_0+1, \ldots, r+1\} \in \mathcal{G}'$. On the other hand, since $\mathcal{T}_{1,2} \subseteq \mathcal{G}'$ we have that $[r+1] \setminus i \in \mathcal{G}'$ for $i \in \{3, \ldots, r+1\}$. Because $1 \in G_2 \in \mathcal{G}$ while $2 \notin G_2$, we also have that $[r+1] \setminus \{2\} \in \mathcal{G}'$. That is, $\binom{[r+1]}{r} \subseteq \mathcal{G}'$.

Consequently applying $\mathrm{Shift}_{i,j}$ for $1 \leq i < j \leq n$ to $\mathcal{G}'$ does not create a trivial family. Finally, we apply $\mathrm{Shift}_{i,j}$ for $1 \leq i < j \leq n$ repeatedly to $\mathcal{G}'$ until it is stable and denote the resulting family by $\mathcal{F}'$. $\square$

We will need the following technical lemma.

Lemma 3.2 ([7, 12])**.** *Let $\mathcal{F}$ be an pairwise-intersecting shifted family. For every $F \in \mathcal{F}$ there exists $l \geq 1$ such that $|F \cap [2l-1]| \geq l$. Moreover, the maximum such $l = l(F)$ satisfies $|F \cap [2l(F)]| = l(F)$.*

The following function was previously defined in [7] in the unrestricted context. We extend it to the setting of simplicial complexes. Let K be a shifted simplicial complex with vertex set $[n]$ and $\mathcal{F} \subseteq K$ a shifted pairwise-intersecting family of r-faces, set

$$\alpha \colon \mathcal{F} \to (\mathrm{st}(1,K) \setminus \mathrm{st}(1, K[[n] \setminus [2, r+1]])) \cup \{2, \ldots, r+1\}$$

given by

$$\alpha(F) = \begin{cases} F & \text{if } 1 \in F \text{ or } [2, r+1] \subseteq F, \\ F \Delta [2l(F)] & \text{otherwise.} \end{cases}$$

The following lemma shows that α is well defined and injective.

Lemma 3.3. *For $F \in \mathcal{F}$ such that $\alpha(F) \neq F$ we have that:* (1) $1 \in \alpha(F)$. (2) $\alpha(F) \notin \mathcal{F}$. (3) $\alpha(F) \cap [2, r+1] \neq \emptyset$. (4) *$\alpha$ is injective.* (5) $\alpha(F) \in K$.

Proof. Properties (1-4) were proved previously [7], we only need to verify property (5). Notice that $d/2 \geq r \geq |F \cap [2l(F)]| = l(F)$. In particular, $|F \cup [2l(F)]| = |F| + |[2l(F)]| - |F \cap [2l(F)]| \leq 2r \leq d$. Because K is shifted with minimal facet cardinality d we have that $F \cup [2l(F)] \in K$ since it is the smallest face, with respect to the partial order, of size $r + l(F)$ containing F. Since $\alpha(F) = F \Delta [2l(F)] \subseteq F \cup [2l(F)]$ we conclude that $\alpha(F) \in K$. $\square$

Proposition 3.4. *Let K be a shifted simplicial complex with vertex set $[n]$ and minimal facet cardinality d and $r \leq d/2$. Let $\mathcal{F} \subseteq K$ be a non-trivial intersecting family of r-faces. We have that*

$$|\mathcal{F}| \leq f_{r-1}(\mathrm{lk}(1,K)) - f_{r-1}(\mathrm{lk}(1, K[[n] \setminus [2, r+1]])) + 1.$$

Moreover, if $2 \leq r < d/2$ *then* $|\mathcal{F}| < f_{r-1}(\mathrm{lk}(1, K))$.

Proof. By Lemma 3.1 we can assume that $\mathcal{F}$ is shifted. The first part follows from the injectivity of α. For the second part, because $2r+1 \leq d$ we have that $\{1, r+2, \ldots, 2r+1\} \setminus \{i\} \subseteq [d] \setminus [2, r+1] \in K$ for $i \in [r+2, 2r+1]$. Consequently $f_{r-1}(\mathrm{lk}(1, K[[n] \setminus [2, r+1]])) \geq r \geq 2$ and the conclusion follows. □

Remark 3.5. It is not hard to show that if K be a near-cone with apex v and minimal facet cardinality d, then for $r \leq d/2$ we have that $K^{(r)} \subseteq v * \mathrm{lk}(v, K)$.

Theorem 3.6. *Let* K *be a near-cone with apex* v, *then* K *is strict* r*-EKR for* $r < \frac{\mathrm{depth}_{\mathbb{F}} K + 1}{2}$.

Proof. Let $\mathcal{F}$ be a non-trivial intersecting family of r-faces. Since $\mathrm{depth}_{\mathbb{F}} K \leq d$ then $r < d/2$. By above remark we can conclude that $\mathcal{F} \subseteq v * \mathrm{lk}(v, K)$. Set $\mathcal{F}(v) = \{F \setminus \{v\} \colon F \in \mathcal{F}, v \in F\}$ and $\mathcal{F}(\bar{v}) = \{F \colon F \in \mathcal{F}, v \notin F\}$. Then $\mathcal{F}(v), \mathcal{F}(\bar{v}) \subseteq \mathrm{lk}(v, K)$ are cross-intersecting and $\mathcal{F}(\bar{v})$ is pairwise-intersecting. By Theorem 2.2.2 and Theorem 2.2.5 we have that $\Delta(\mathcal{F}(v)), \Delta(\mathcal{F}(\bar{v})) \subseteq \Delta(\mathrm{lk}(v, K))$ are cross-intersecting and $\Delta(\mathcal{F}(\bar{v}))$ is pairwise-intersecting. Consequently the family $\mathcal{F}' = \{\{1\} \cup F \colon F \in \Delta\mathcal{F}(v)\} \cup \Delta(\mathcal{F}(\bar{v})) \subseteq 1 * \Delta(\mathrm{lk}(v, K))$ is non-trivial and intersecting. By Theorem 2.2.6, $\Delta(K) = (1 * \Delta(\mathrm{lk}(v, K))) \cup B$ where $B = \{F \in \Delta(K) \colon 1 \notin F\}$, consequently $\mathcal{F}' \subseteq \Delta(K)$ and non-trivial. Moreover, since no member of $\Delta(\mathcal{F}(\bar{v}))$ contains vertex 1 we have that

$$|\mathcal{F}'| = |\Delta(\mathcal{F}(v))| + |\Delta(\mathcal{F}(\bar{v}))| = |\mathcal{F}(v)| + |\mathcal{F}(\bar{v})| = |\mathcal{F}|$$

where we have used Theorem 2.2.3. Since K is shifted, by Theorem 2.2.1, with minimal facet cardinality $\mathrm{depth}_{\mathbb{F}} K + 1$, by Proposition 3.4, we can conclude that $|\mathcal{F}| < f_{r-1}(\mathrm{lk}(1, \Delta K)) = f_{r-1}(\mathrm{lk}(v, K))$ by Theorem 2.2.6. □

As a corrolary we obtain Theorem 1.4.

References

[1] Peter Borg. Extremal t-intersecting sub-families of hereditary families. *J. Lond. Math. Soc. (2)*, 79(1):167–185, 2009.

[2] Anton Dochtermann and Alexander Engström. Algebraic properties of edge ideals via combinatorial topology. *Electron. J. Combin.*, 16(2, Special volume in honor of Anders Björner):Research Paper 2, 24, 2009.

[3] Art M. Duval. Algebraic shifting and sequentially Cohen-Macaulay simplicial complexes. *Electron. J. Combin.*, 3(1):Research Paper 21, approx. 14, 1996.

[4] P. Erdős, Chao Ko, and R. Rado. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser. (2)*, 12:313–320, 1961.

[5] Christopher A. Francisco and Adam Van Tuyl. Sequentially Cohen-Macaulay edge ideals. *Proc. Amer. Math. Soc.*, 135(8):2327–2337, 2007.

[6] Peter Frankl. The shifting technique in extremal set theory. In *Surveys in combinatorics 1987 (New Cross, 1987)*, volume 123 of *London Math. Soc. Lecture Note Ser.*, pages 81–110. Cambridge Univ. Press, Cambridge, 1987.

[7] Peter Frankl. A simple proof of the Hilton-Milner theorem. *Mosc. J. Comb. Number Theory*, 8(2):97–101, 2019.

[8] A. J. W. Hilton and E. C. Milner. Some intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser. (2)*, 18:369–384, 1967.

[9] Fred Holroyd, Claire Spencer, and John Talbot. Compression and Erdős-Ko-Rado graphs. *Discrete Math.*, 293(1-3):155–164, 2005.

[10] Fred Holroyd and John Talbot. Graphs with the Erdős-Ko-Rado property. *Discrete Math.*, 293(1-3):165–176, 2005.

[11] Glenn Hurlbert and Vikram Kamat. Erdős-Ko-Rado theorems for chordal graphs and trees. *J. Combin. Theory Ser. A*, 118(3):829–841, 2011.

[12] Glenn Hurlbert and Vikram Kamat. New injective proofs of the Erdős-Ko-Rado and Hilton-Milner theorems. *Discrete Math.*, 341(6):1749–1754, 2018.

[13] A. Johnson and F. C. Holroyd. *'Problem 25', Problems from the Sixteenth British Combinatorial Conference.* P.J. Cameron, 1997.

[14] Gil Kalai. Hyperconnectivity of graphs. *Graphs Combin.*, 1(1):65–79, 1985.

[15] Gil Kalai. Algebraic shifting. In *Computational commutative algebra and combinatorics (Osaka, 1999)*, volume 33 of *Adv. Stud. Pure Math.*, pages 121–163. Math. Soc. Japan, Tokyo, 2002.

[16] Younjin Kim. Erdős-ko-rado type theorems for simplicial complexes via algebraic shifting. *J. Korean Math. Soc.*, 57(6):1323–1333, 2020.

[17] Susan Morey and Rafael H. Villarreal. Edge ideals: algebraic and combinatorial properties. In *Progress in commutative algebra 1*, pages 85–126. de Gruyter, Berlin, 2012.

[18] Eran Nevo. Algebraic shifting and basic constructions on simplicial complexes. *J. Algebraic Combin.*, 22(4):411–433, 2005.

[19] Jorge Alberto Olarte, Francisco Santos, Jonathan Spreer, and Christian Stump. The EKR property for flag pure simplicial complexes without boundary. *J. Combin. Theory Ser. A*, 172:105205, 29, 2020.

[20] Seyed Amin Seyed Fakhari. Erdős-ko-rado type theorems for simplicial complexes. *Electron. J. Combin.*, 24(2):Paper No. 2.38, 11, 2017.

[21] Dean E. Smith. On the Cohen-Macaulay property in commutative algebra and simplicial topology. *Pacific J. Math.*, 141(1):165–196, 1990.

[22] John Talbot. Intersecting families of separated sets. *J. London Math. Soc. (2)*, 68(1):37–51, 2003.

[23] Larry X. W. Wang. Restricted intersecting families on simplicial complex. *Adv. in Appl. Math.*, 124:Paper No. 102144, 15, 2021.

[24] Russ Woodroofe. Vertex decomposable graphs and obstructions to shellability. *Proc. Amer. Math. Soc.*, 137(10):3235–3246, 2009.

[25] Russ Woodroofe. Erdős-Ko-Rado theorems for simplicial complexes. *J. Combin. Theory Ser. A*, 118(4):1218–1227, 2011.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Exact enumeration of graphs and bipartite graphs with degree constraints

(Extended abstract)

Emma Caizergues* Élie de Panafieu†

Abstract

We provide a new explicit formula enumerating graphs with constraints on their degrees, such as regular graphs, and extend it to bipartite graphs. It relies on generating function manipulations and Hadamard products.
Keywords. regular graphs, exact enumeration, D-finite, differentiably finite

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-035`

Related work. The most famous graphs with degree constraints are k-regular graphs, where all vertices have degree k. There are two natural generalizations: graphs with a given degree sequence, and graphs where all vertices have their degree in a given set. In this article, we consider the later. There is a large literature on the asymptotic enumeration [1, 22, 7] and typical structure of graphs with degree constraints [25, 18, 15, 14, 3, 16, 8]. We focus on exact enumeration. The main result in this field is that the generating function of graphs with their degrees in a given finite set is D-finite, meaning that it is solution of a differential equation with polynomial coefficients. The previous proofs relied on a symmetric function approach [13, 12, 23, 24]. It starts by considering the infinite product $\prod_{1 \leqslant i<j}(1 + x_i x_j)$ representing graphs where the degree of vertex i is the power of x_i. Arguments on the D-finiteness of the scalar product of symmetric functions are then applied. In contrast, we obtain a formula (Theorem 1) for the generating function of those graphs that is *explicit* and uses only a *finite* number of variables (assuming the degrees are bounded). Our approach relies on direct translation of combinatorial properties into

*Nokia Bell Labs, supported by the RandNET project, authors presented in alphabetical order.
†Nokia Bell Labs, supported by the RandNET project

generating function equations (symbolic method [2, 11]), and manipulation of those equations, in particular using Hadamard products. Works of similar spirit include [4, Chapters 3,4,5,7] and [7]. Our expression provides a new proof of D-finiteness, as D-finite series are stable by Hadamard product [21, 27] and evaluation [28]. Although effective algorithms exist [5] to compute the differential equation characterizing the generating function of graphs with degree constraints, they are computationally costly and the differential equation is only known up to $k = 4$ for k-regular graphs. We hope our new formula will allow the computation of differential equations for k-regular graphs with $k \geqslant 5$ and fast enumeration of those graphs [19]. Our results extend to bipartite graphs with different degree sets for the left and right vertices. For bipartite graphs, we used a multidimensional version of the Hadamard product, that has been well studied in the literature [9, 20, 26]. To our knowledge, the asymptotic structure of those graphs has not been investigated [10], and we hope our work will be a step in that direction.

Structure. We enumerate successively several graph-like families with degree constraints: *weighted multigraphs*, *loopless weighted multigraphs*, *weighted graphs* and finally *graphs* in Theorem 1. They are all depicted in Figure 1. The generating function of the first family is expressed directly. Then, to go from each family $\mathcal{A}$ to the next family $\mathcal{B}$, we first express the generating function of $\mathcal{A}$ using the generating function of $\mathcal{B}$, then invert this relation. The extension to bipartite graphs is presented in Theorem 2.

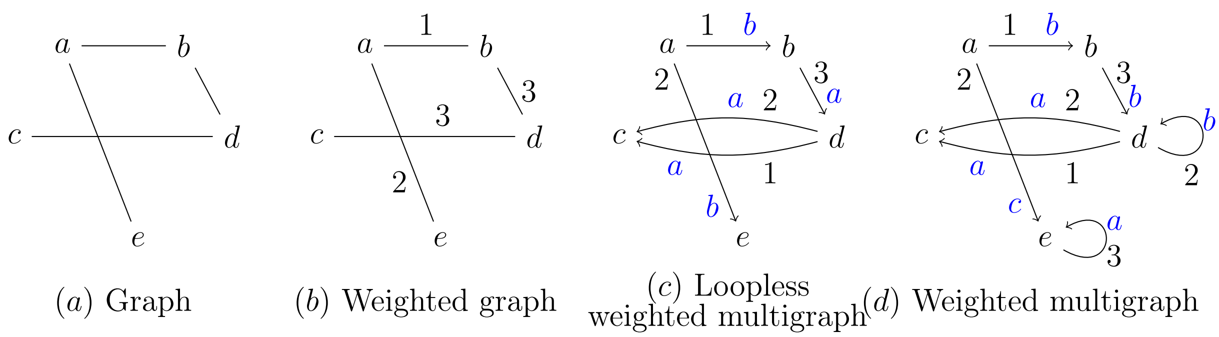


Figure 1: Steps of a possible transformation of a simple graph into a weighted multigraph. Labels are represented by letters, while weights are represented by integers.

Notations. The nth coefficient of a formal power series is denoted by $[z^n]\sum_m a_m z^m = a_n$. The *exponential Hadamard product* ([17, Theorem 3], [2, Section 2.1]) is defined as

$$\left(\sum_n a_n \frac{z^n}{n!}\right) \odot_z \left(\sum_n b_n \frac{z^n}{n!}\right) = \sum_n a_n b_n \frac{z^n}{n!}.$$

We denote by $\odot_{z=1}$ the exponential Hadamard product followed by the evaluation at $z = 1$. Throughout this article, the variable δ_d marks vertices of degree d. We denote by $\boldsymbol{\delta}$ the infinite vector $(\delta_0, \delta_1, \delta_2, \dots)$. The same bold convention extends to other letters. An interesting particular case is, given a set $\mathcal{D}$ of nonnegative integers, to set $\delta_d = 1$ if $d \in \mathcal{D}$, and $\delta_d = 0$ otherwise. That way, only graphs with all their vertices having degree in $\mathcal{D}$ are counted. We associate to those variables the generating function

$$\Delta(z, \boldsymbol{\delta}) = \sum_{d \geqslant 0} \delta_d \frac{z^d}{d!}.$$

In the following, we consider graphs with vertices of degree at most D for some $D \geqslant 0$, so $\delta_j = 0$ whenever $j > D$ and $\Delta(z, \boldsymbol{\delta})$ is a polynomial.

1 Weighted multigraphs

Definition. A weighted multigraph G is a finite sequence

$$G = (V(G), E_1(G), E_2(G), \dots)$$

where $V(G) = \{1, 2, \dots, n(G)\}$ is the set of $n(G)$ vertices, and for all $j \geqslant 1$, $E_j(G)$ is the finite set of $m_j(G)$ edges of weight j

$$E_j(G) = \{(u_{j,1}, v_{j,1}, 1), \dots, (u_{j,m_j(G)}, v_{j,m_j(G)}, m_j(G))\}$$

where each $u_{i,j}$ and $v_{k,\ell}$ belongs to $V(G)$. Thus, vertices are labeled, edges of weight j are labeled (and have their own independent label set), and all edges are oriented. Furthermore, loops and multiple edges are allowed. The degree $\deg(u)$ of a vertex u is defined as the sum of the weights of all edges adjacent to it, counted twice if they are loops. For examples, in Figure 1 (d), vertex e has degree 8.

Generating function. We use the variable z to mark the vertices, and for all $j \geqslant 1$, we use $w_j/2$ to mark each edge of weight j. This factor 1/2 is here for historical reasons only [6, Section 2.3]. Additionally, for each $d \geqslant 0$, the variable δ_d is introduced to mark vertices of degree d. The generating function of weighted multigraphs $\mathrm{WMG}(z, \boldsymbol{w}, \boldsymbol{\delta})$ is then defined as a sum over all weighted multigraphs

$$\mathrm{WMG}(z, \boldsymbol{w}, \boldsymbol{\delta}) = \sum_G \left(\prod_{u=1}^{n(G)} \delta_{\deg(u)} \right) \left(\prod_{j \geqslant 1} \frac{(w_j/2)^{m_j(G)}}{m_j(G)!} \right) \frac{z^{n(G)}}{n(G)!}. \tag{1}$$

Lemma 1. *Let $P_{\mathrm{WMG}}(\boldsymbol{x}, \boldsymbol{\delta})$ denote the polynomial*

$$P_{\mathrm{WMG}}(\boldsymbol{x}, \boldsymbol{\delta}) = \Delta(y, \boldsymbol{\delta}) \odot_{y=1} e^{\sum_{j=1}^{D} x_j y^j},$$

then the generating function of weighted multigraphs is equal to

$$\mathrm{WMG}(z, \boldsymbol{w}, \boldsymbol{\delta}) = e^{\sum_{j=1}^{D} w_j x_j^2/2} \odot_{x_1=1} \cdots \odot_{x_D=1} e^{z P_{\mathrm{WMG}}(\boldsymbol{x}, \boldsymbol{\delta})}.$$

Proof. Any weighted multigraph decomposes uniquely as a set of labeled vertices, each attached to a set of labeled half-edges of weight j, for $j \geqslant 1$. If the vertex has degree d, then the sum of the weights of the half-edges attached to it should be d. Then, using the variable x_j to mark the half-edges of weight j, the generating function of such sets is $P_{\mathrm{WMG}}(\boldsymbol{x}, \boldsymbol{\delta})$. If the multigraph contains m_j edges of weight j, then after cutting them in two, we are left with $2m_j$ half-edges of weight j. The symbolic method [11] implies

$$\mathrm{WMG}(z, \boldsymbol{w}, \boldsymbol{\delta}) = \sum_{\boldsymbol{m}} (2\boldsymbol{m})! [\boldsymbol{x}^{2\boldsymbol{m}}] e^{z P_{\mathrm{WMG}}(\boldsymbol{x}, \boldsymbol{\delta})} \prod_{j \geqslant 1} \frac{(w_j/2)^{m_j}}{m_j!}.$$

This expression is simplified using Hadamard products with the function

$$\sum_{m \geqslant 0} (2m)! \frac{(w/2)^m}{m!} \frac{x^{2m}}{(2m)!} = e^{wx^2/2}.$$

□

A multigraph is *loopless* if it has no edge containing twice the same vertex. The generating function $\mathrm{LWMG}(z, \boldsymbol{w}, \boldsymbol{\delta})$ of those weighted multigraphs is defined by restricting the sum from Equation (1) to them.

Lemma 2. *Let $P_{\mathrm{LWMG}}(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{\delta})$ denote the polynomial*

$$P_{\mathrm{LWMG}}(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{\delta}) = \Delta(y, \boldsymbol{\delta}) \odot_{y=1} e^{\sum_{j=1}^{D} x_j y^j - \sum_{j=1}^{D} w_j y^{2j}/2},$$

then the generating function of loopless weighted multigraphs is equal to

$$\mathrm{LWMG}(z, \boldsymbol{w}, \boldsymbol{\delta}) = e^{\sum_{j=1}^{D} w_j x_j^2/2} \odot_{x_1=1} \cdots \odot_{x_D=1} e^{z P_{\mathrm{LWMG}}(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{\delta})}.$$

Proof. Any weighted multigraph is uniquely obtained by adding a set of loops on each vertex of a loopless weighted multigraph. During this operation, any vertex of degree d becomes a vertex of degree $d + 2k$, for some nonnegative integer k, and a set of weighted loops whose weights sum to k. This corresponds to replacing δ_d with

$$\delta_d(\boldsymbol{\eta}, \boldsymbol{w}) = \sum_{k \geqslant 0} \eta_{d+2k} [y^k] e^{\frac{1}{2} \sum_{d \geqslant 1} w_d y^d} = \sum_{j \geqslant 0} \eta_j [y^j] y^d e^{\frac{1}{2} \sum_{d \geqslant 1} w_d y^{2d}}. \tag{2}$$

Thus, the generating functions of weighted multigraphs and loopless weighted multigraphs are linked by the relation

$$\mathrm{WMG}(z, \boldsymbol{w}, \boldsymbol{\eta}) = \mathrm{LWMG}(z, \boldsymbol{w}, \boldsymbol{\delta}(\boldsymbol{\eta}, \boldsymbol{w}))$$

Inverting Equation (2), we obtain

$$\eta_d(\boldsymbol{\delta}, \boldsymbol{w}) = \sum_{j \geqslant 0} \delta_j [y^j] y^d e^{-\frac{1}{2} \sum_{i \geqslant 1} w_i y^{2i}}$$

so

$$\Delta(z, \boldsymbol{\eta}(\boldsymbol{\delta}, \boldsymbol{w})) = \sum_{d=0}^{D} \sum_{j \geqslant 0} \delta_j [y^j] y^d e^{-\sum_{i=1}^{D} w_i y^{2i}/2} \frac{z^d}{d!} = \Delta(y, \boldsymbol{\delta}) \odot_{y=1} e^{yz - \sum_{i=1}^{D} w_i y^{2i}/2}.$$

Given the expression of $\mathrm{WMG}(z, \boldsymbol{w}, \boldsymbol{\delta})$ from Lemma 1, we deduce

$$\begin{aligned} \mathrm{LWMG}(z, \boldsymbol{w}, \boldsymbol{\delta}) &= \mathrm{WMG}(z, \boldsymbol{w}, \boldsymbol{\eta}(\boldsymbol{\delta}, \boldsymbol{w})) \\ &= e^{\sum_{j=1}^{D} w_j x_j^2/2} \odot_{x_1=1} \cdots \odot_{x_D=1} e^{z \Delta(y, \boldsymbol{\eta}) \odot_{y=1} \exp\left(\sum_{j=1}^{D} x_j y^j\right)}. \end{aligned}$$

Finally, the properties $F(x\,y) \odot_x G(x) = F(x) \odot_x G(x\,y)$ and $F(x) \odot e^x = F(x)$ of the exponential Hadamard product imply

$$\begin{aligned} \Delta(y, \boldsymbol{\eta}) \odot_{y=1} e^{\sum_{j=1}^{D} x_j y^j} &= \Delta(y, \boldsymbol{\delta}) \odot_{y=1} e^{yz - \sum_{i=1}^{D} w_i y^{2i}/2} \odot_{y=1} e^{\sum_{j=1}^{D} x_j y^j} \\ &= \Delta(y, \boldsymbol{\delta}) \odot_{y=1} e^{\sum_{j=1}^{D} x_j y^j - \sum_{j=1}^{D} w_j y^{2j}/2} = P_{\mathrm{LWMG}}(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{\delta}). \end{aligned}$$

□

2 Graphs

Definitions. A weighted graph G is a finite sequence

$$G = (V(G), E_1(G), E_2(G), \dots)$$

where $V(G) = \{1, 2, \dots, n(G)\}$ is the set of $n(G)$ vertices, and $\bigcup_j E_j(G)$ is a finite set of edges, which are unordered pairs of distinct vertices. $E_j(G)$ denotes the set of edges of weight j and its cardinality is $m_j(G)$. Thus, edges are unoriented and unlabeled, loops and multiple edges are forbidden. The degree of a vertex is still defined as the sum of the weights of its adjacent edges.

Unweighted graphs correspond to the case $G = (V(G), E_1(G))$, so $E_j = \varnothing$ for all $j \geqslant 2$.

Generating function. The generating function of weighted graphs $\mathrm{WG}(z, \boldsymbol{w}, \boldsymbol{\delta})$ is defined as a sum over all weighted graphs

$$\mathrm{WG}(z, \boldsymbol{w}, \boldsymbol{\delta}) = \sum_G \left(\prod_{u=1}^{n(G)} \delta_{\deg(u)} \right) \left(\prod_{j \geqslant 1} w_j^{m_j(G)} \right) \frac{z^{n(G)}}{n(G)!}.$$

This generating function is exponential with respect to the variable z marking the vertices, and ordinary with respect to the variables $(w_j)_{j \geqslant 1}$ marking the edges. Note that the generating function of an edge of weight j is w_j (for weighted multigraphs, we used the convention $w_j/2$, linked to the fact that edges were oriented).

Lemma 3. *For $j \geqslant 0$, let $v_j(\boldsymbol{w})$ and $P_{\mathrm{WG}}(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{\delta})$ denote the polynomials*

$$v_j(\boldsymbol{w}) = [y^j] \log\left(1 + \sum_{j \geqslant 1} w_j y^j\right), \qquad P_{\mathrm{WG}}(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{\delta}) = \Delta(y, \boldsymbol{\delta}) \odot_{y=1} \frac{e^{\sum_{j=1}^{D} x_j y^j}}{\sqrt{1 + \sum_{j=1}^{D} w_j y^{2j}}},$$

then the generating function of weighted graphs is equal to

$$\mathrm{WG}(z, \boldsymbol{w}, \boldsymbol{\delta}) = e^{\sum_{j=1}^{D} v_j(\boldsymbol{w}) x_j^2/2} \odot_{x_1=1} \cdots \odot_{x_D=1} e^{z P_{\mathrm{WG}}(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{\delta})}.$$

Proof. Consider a weighted graph G with m_d edge of weight d for each $d \geqslant 1$. Then there exist $\prod_d 2^{m_d} m_d!$ ways to orient and label those edges to turn G into a weighted multigraph. Thus, the generating function of loopless weighted multigraphs that contain no multiple edge is equal to the generating function of weighted graphs. To construct a loopless weighted multigraph G from a loopless weighted multigraph H without multiple edges, one replace each edge of H of weight d with a set of edges linking the same vertices and whose weights sum to d. For each $d \geqslant 1$, let us use the variable v_d to mark edges of weight d in loopless weighted multigraphs, and the variable w_d to mark edges of weight d in weighted graphs. The previous construction corresponds to replacing the variable w_d with

$$w_d(\boldsymbol{v}) = [y^d](e^{\sum_{j \geqslant 1} v_j y^j} - 1).$$

Thus, the generating functions of weighted graphs and of loopless weighted multigraphs are linked by the relation

$$\mathrm{WG}(z, \boldsymbol{w}(\boldsymbol{v}), \boldsymbol{\delta}) = \mathrm{LWMG}(z, \boldsymbol{v}, \boldsymbol{\delta}).$$

Inverting this relation, we obtain

$$v_d(\boldsymbol{w}) = [y^d] \log\left(1 + \sum_{j \geqslant 1} w_j y^j\right)$$

and

$$\mathrm{WG}(z, \boldsymbol{w}, \boldsymbol{\delta}) = \mathrm{LWMG}(z, \boldsymbol{v}(\boldsymbol{w}), \boldsymbol{\delta}).$$

Injecting the expression of $\mathrm{LWMG}(z, \boldsymbol{w}, \boldsymbol{\delta})$ from Lemma 2 concludes the proof. $\square$

In particular, for $\boldsymbol{w} = (w, 0, 0, \dots, 0)$, we recover the case of classical graphs, and more specifically the case of k-regular graphs, by setting $\delta_d = 1$ if $d = k$, and $\delta_d = 0$ otherwise.

Theorem 1. *Let $P_G(\boldsymbol{x}, w, \boldsymbol{\delta})$ denote the polynomial*

$$P_G(\boldsymbol{x}, w, \boldsymbol{\delta}) = \Delta(y, \boldsymbol{\delta}) \odot_{y=1} \frac{e^{\sum_{j=1}^{D} x_j y^j}}{\sqrt{1 + w y^2}},$$

then the generating function of graphs is equal to

$$G(z, w, \boldsymbol{\delta}) = e^{-\sum_{j=1}^{D} (-w)^j x_j^2/(2j)} \odot_{x_1=1} \cdots \odot_{x_D=1} e^{z P_G(\boldsymbol{x}, w, \boldsymbol{\delta})}.$$

In particular, the number of k-regular graphs with n vertices is

$$e^{\sum_{j=1}^{k} (-1)^{j+1} x_j^2/(2j)} \odot_{x_1=1} \cdots \odot_{x_k=1} \left([y^k] \frac{e^{\sum_{j=1}^{k} x_j y^j}}{\sqrt{1 + y^2}}\right)^n.$$

3 Bipartite graphs

Definition. A bipartite graph G is a triplet $(V(G), \tilde{V}(G), E(G))$ with $V(G)$ (resp $\tilde{V}(G)$) the set of labeled left-vertices (resp. right-vertices) and $E(G) \subset V(G) \times \tilde{V}(G)$ the set of edges.

Generating function. The generating function $\mathrm{BG}(z, \tilde{z}, w, \boldsymbol{\delta}, \tilde{\boldsymbol{\delta}})$ of bipartite graphs with degree at most D is defined as a sum over bipartite graphs

$$\mathrm{BG}(z, \tilde{z}, w, \boldsymbol{\delta}, \tilde{\boldsymbol{\delta}}) = \sum_G \left(\prod_{u=1}^{|V(G)|} \delta_{\deg(u)} \right) \left(\prod_{u=1}^{|\tilde{V}(G)|} \tilde{\delta}_{\deg(u)} \right) w^{|E(G)|} \times \frac{z^{|V(G)|}}{|V(G)|!} \times \frac{\tilde{z}^{|\tilde{V}(G)|}}{|\tilde{V}(G)|!}$$

This generating function is exponential with respect to the variable z (resp. $\tilde{z}$) marking the left vertices (resp. right vertices), and ordinary with respect to the variable w marking the edges. For all d, δ_d (resp. $\tilde{\boldsymbol{\delta}}_d$) marks the left vertices (resp. right vertices) of degree d.

Notation. The *multivariate exponential Hadamard product* is defined as

$$\sum_{\boldsymbol{m}} A_{\boldsymbol{m}} \frac{\boldsymbol{z}^{\boldsymbol{m}}}{\boldsymbol{m}!} \odot_{\boldsymbol{m}} \sum_{\boldsymbol{m}} B_{\boldsymbol{m}} \frac{\boldsymbol{z}^{\boldsymbol{m}}}{\boldsymbol{m}!} = \sum_{\boldsymbol{m}} A_{\boldsymbol{m}} B_{\boldsymbol{m}} \frac{\boldsymbol{z}^{\boldsymbol{m}}}{\boldsymbol{m}!}.$$

This extension is compatible with the univariate product in the sense

$$A(z_1) B(z_2) \odot_{z_1 z_2} C(z_1, z_2) = A(z_1) \odot_{z_1} \left(B(z_2) \odot_{z_2} C(z_1, z_2) \right).$$

Theorem 2. *Let* $\boldsymbol{v} = \left((-1)^{d+1} w^d / d \right)_{1 \leqslant d \leqslant D}$, $\Delta(y, \boldsymbol{\delta}) = \sum_{d=0}^{D} \delta_d \frac{y^d}{d!}$ *and let* $P_{\mathrm{BG}}(\boldsymbol{w}, \boldsymbol{\delta})$ *denote the polynomial*

$$P_{\mathrm{BG}}(\boldsymbol{w}, \boldsymbol{\delta}) = \Delta(y, \boldsymbol{\delta}) \odot_{y=1} e^{\sum_{d=1}^{D} w_d y^d}.$$

Then the generating function of bipartite graphs with degree constraints is equal to

$$\mathrm{BG}(z, \tilde{z}, w, \boldsymbol{\delta}, \tilde{\boldsymbol{\delta}}) = e^{z P_{\mathrm{BG}}(\boldsymbol{w}, \boldsymbol{\delta})} \odot_{\boldsymbol{w}=\boldsymbol{v}} e^{\tilde{z} P_{\mathrm{BG}}(\boldsymbol{w}, \tilde{\boldsymbol{\delta}})}.$$

Aknowledgement. We thank Michael Borinsky, Marni Mishna and Clément Requilé for interesting conversations on the topic, suggestions of proof techniques, pointers to the relevant literature, and contagious enthusiasm.

References

[1] E. A. Bender and E. R. Canfield. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory, Series A*, 24(3):296–307, 1978.

[2] F. Bergeron, G. Labelle, and P. Leroux. *Combinatorial Species and Tree-like Structures.* Cambridge University Press, 1997.

[3] B. Bollobás and O. Riordan. An old approach to the giant component problem. *Journal of Combinatorial Theory, Series B*, 113:236–260, 2015.

[4] M. Borinsky. *Graphs in perturbation theory: Algebraic structure and asymptotics.* Springer, 2018.

[5] F. Chyzak, M. Mishna, and B. Salvy. Effective scalar products of d-finite symmetric series. *Journal of Combinatorial Theory, Series A*, 112:1–43, 2005.

[6] E. de Panafieu. Analytic combinatorics of connected graphs. *Random Structures & Algorithms*, 2018.

[7] E. de Panafieu and L. Ramos. Graphs with degree constraints. *Proceedings of the Meeting on Analytic Algorithmics and Combinatorics (Analco16)*, 2016.

[8] S. Dovgal and V. Ravelomanana. Shifting the phase transition threshold for random graphs using degree set constraints. *Proceedings of the American Theoretical Informatics Symposium (LATIN18)*, pages 399–412, 2018.

[9] M. M. Elin. Multidimensional Hadamard composition. *Sibirskii Matematicheskii Zhurnal*, 35(5):1052–1057, 1994.

[10] P. L. Erdős, I. Miklós, and L. Soukup. Towards random uniform sampling of bipartite graphs with given degree sequence. *arXiv preprint arXiv:1004.2612*, 2010.

[11] P. Flajolet and R. Sedgewick. *Analytic Combinatorics.* Cambridge University Press, 2009.

[12] I. M. Gessel. Symmetric functions and p-recursiveness. *Journal of Combinatorial Theory, Series A*, 53:257–285, 1990.

[13] I. P. Goulden, D. M. Jackson, and J. W. Reilly. The Hammond series of a symmetric function and its application to P-recursiveness. *SIAM Journal on Algebraic Discrete Methods*, 4(2):179–193, 1983.

[14] H. Hatami and M. Molloy. The scaling window for a random graph with a given degree sequence. *Random Structures & Algorithms*, 41(1):99–123, 2012.

[15] S. Janson and M. J. Luczak. A new approach to the giant component problem. *Random Structures & Algorithms*, 34(2):197–216, 2009.

[16] F. Joos, G. Perarnau, D. Rautenbach, and B. Reed. How to determine if a random graph with a fixed degree sequence has a giant component. *Probability Theory and Related Fields*, 170(1-2):263–310, 2018.

[17] A. Joyal. Une théorie combinatoire des séries formelles. *Advances in mathematics*, 42(1):1–82, 1981.

[18] M. Kang and T. G. Seierstad. The critical phase for random graphs with a given degree sequence. *Combinatorics, Probability and Computing*, 17(1):67–86, 2008.

[19] A. Kaygun. Enumerating labeled graphs that realize a fixed degree sequence. *arXiv preprint arXiv:2101.02299*, 2021.

[20] E. Leinartas. Multidimensional Hadamard composition and sums with linear constraints on the summation indices. *Siberian Mathematical Journal*, 30(2):250–255, 1989.

[21] L. Lipshitz. The diagonal of a d-finite power series is d-finite. *Journal of algebra*, 113(2):373–378, 1988.

[22] B. D. McKay and N. C. Wormald. Asymptotic enumeration by degree sequence of graphs of high degree. *European Journal of Combinatorics*, 11(6):565–580, 1990.

[23] M. Mishna. Automatic enumeration of regular objects. *arXiv preprint math/0507249*, 2005.

[24] M. Mishna. Regularity in weighted graphs: a symmetric function approach. *arXiv preprint arXiv:1507.05121*, 2015.

[25] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms*, 6(2-3):161–180, 1995.

[26] T. Sadykov. The Hadamard product of hypergeometric series. *Bulletin des Sciences Mathématiques*, 126(1):31–43, 2002.

[27] R. P. Stanley. Differentiably finite power series. *European journal of combinatorics*, 1(2):175–188, 1980.

[28] D. Zeilberger. Sister Celine's technique and its generalizations. *Journal of Mathematical Analysis and Applications*, 85(1):114–145, 1982.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

A PRECISE CONDITION FOR INDEPENDENT TRANSVERSALS IN BIPARTITE COVERS

(EXTENDED ABSTRACT)

Stijn Cambie* Penny Haxell† Ross J. Kang‡ Ronen Wdowinski§

Abstract

Given a bipartite graph $H = (V = V_A \cup V_B, E)$ in which any vertex in V_A (resp. V_B) has degree at most D_A (resp. D_B), suppose there is a partition of V that is a refinement of the bipartition $V_A \cup V_B$ such that the parts in V_A (resp. V_B) have size at least k_A (resp. k_B). We prove that the condition $D_A/k_A + D_B/k_B \leq 1$ is sufficient for the existence of an independent set of vertices of H that is simultaneously transversal to the partition, and show moreover that this condition is sharp.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-036`

1 Introduction

Consider the following question: how much easier is it to colour graphs that are bipartite than to colour graphs in general? Of course, when considered in the context of the usual chromatic number, this is utterly trivial: compared to the general case, for which the chromatic number can be $\Delta(G) + 1$ but no larger (with $\Delta(G)$ denoting the maximum

*Extremal Combinatorics and Probability Group (ECOPRO), Institute for Basic Science (IBS), Daejeon, South Korea. E-mail: `stijn.cambie@hotmail.com` or `stijncambie@ibs.re.kr`. Supported by the Institute for Basic Science (IBS-R029-C4).

†Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON Canada N2L 3G1. E-mail: `pehaxell@uwaterloo.ca`. Partially supported by NSERC.

‡Korteweg–de Vries Institute for Mathematics, University of Amsterdam, PO Box 94248, 1090 GE Amsterdam, Netherlands. E-mail: `r.kang@uva.nl`. Partially supported by a Vidi grant (639.032.614) of the Dutch Research Council (NWO).

§Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON Canada N2L 3G1. E-mail: `ronen.wdowinski@uwaterloo.ca`

degree of G), the factor of reduction in the number of necessary colours is of order $\Delta(G)$. We treat some settings stronger than that of ordinary proper colouring, settings that have both classic and contemporary combinatorial motivation.

Recall the definition of the list chromatic number, a notion introduced nearly half a century ago independently by Erdős, Rubin and Taylor [5] and by Vizing [12]. Let $G = (V, E)$ be a simple, undirected graph. A mapping $L : V(G) \to 2^{\mathbb{Z}^+}$ is called a *list-assignment* of G; if for some positive integer k, the mapping L satisfies $|L(v)| = k$ for all v then it is called a *k-list-assignment*; a colouring $c : V \to \mathbb{Z}^+$ is called an *L-colouring* if $c(v) \in L(v)$ for any $v \in V$. We say G is *k-choosable* if for any k-list-assignment L of G there is a proper L-colouring of G. The *choosability* $\chi_\ell(G)$ (or *choice number* or *list chromatic number*) of G is the least k such that G is k-choosable.

Framing the above question with respect to the list chromatic number, note first that a greedy procedure implies $\chi_\ell(G) \le \Delta(G) + 1$ always, which is exact for G a complete graph. However, for bipartite G, it is a longstanding conjecture that $\chi_\ell(G)$ must be lower than this bound by a factor of order $\Delta(G)/\log \Delta(G)$.

Conjecture 1.1 (Alon and Krivelevich [2])**.** *There is some $C \ge 1$ such that $\chi_\ell(G) \le C \log_2 \Delta(G)$ for any bipartite graph G with $\Delta(G) \ge 2$.*

If true, this statement would be sharp up to the value of C, due to the complete bipartite graphs [5]. For an idea of how stubborn this problem has been, we relate to the reader how the current best progress was essentially *already known* to the conjecture's originators. In particular, a seminal result for triangle-free graphs of Johansson [8] from the mid-1990's implies that $\chi_\ell(G) = O(\Delta(G)/\log \Delta(G))$ as $\Delta(G) \to \infty$ for any bipartite G, so a reduction factor only of order $\log \Delta(G)$.

To stimulate activity, two of the authors with Alon [1, 4] proposed some natural refinements and variations of Conjecture 1.1, and offered modest related progress. Although less directly relevant to Conjecture 1.1, the present work has the momentum of this trajectory. We introduce some definitions needed to properly describe this progression. In particular, we cast the (bipartite) colouring task in a more precise and general way.

Let G and H be simple, undirected graphs. We say that H is a *cover (graph)* of G with respect to a mapping $L : V(G) \to 2^{V(H)}$ if L induces a partition of $V(H)$ and the bipartite subgraph induced between $L(v)$ and $L(v')$ is edgeless whenever $vv' \notin E(G)$. If for some positive integer k, the mapping L satisfies $|L(v)| = k$ for all v, then we call H a *k-fold* cover of G (with respect to L). Moreover, if G and H are bipartite graphs, where G admits a bipartition $V(G) = A_G \cup B_G$ and H admits a bipartition $V(H) = A_H \cup B_H$, then we say that H is a *bipartite cover (graph)* of H with respect to L if $L(A_G)$ induces a partition of A_H and $L(B_G)$ induces a partition of B_H, i.e. the bipartitions of G and H suitably align. We will have reason to be even more specific for this situation by referring to H as an *(A, B)-cover* of G (with respect to L). (Here we regard A as the pair (A_G, A_H) of partitions, and B similarly.)

To connect the notions above to Conjecture 1.1, notice that, for any list-assignment L of some graph G, one may construct a cover graph H as follows. The vertices of H consist of all pairs (v, x) for $v \in G$ and $x \in L(v)$, and $E(H)$ is a subset of the collection of pairs $(v, x)(v', x')$ such that $vv' \in E(G)$ and $x = x'$. By regarding L as a mapping from v to $\{(v, x) \mid x \in L(v)\}$, we can then regard H as a cover graph of G with respect to L. Moreover, if G is bipartite, the corresponding H is a bipartite cover of G with respect to L. We refer to any (bipartite) cover graph constructed as above as a *(bipartite) list-cover*. Moreover, if $E(H)$ is chosen maximally, we may refer to H as the *maximal (bipartite) list-cover* of G with respect to L. Notice that a proper L-colouring of G is equivalent to an independent set in the corresponding maximal list-cover H that is transversal to the partition induced by L (that is, it intersects every part exactly once), or, in short, an *independent transversal (IT)* of H.

Conjecture 1.1 redux. *There is some $C \geq 1$ such that, for any bipartite graph G of maximum degree $\Delta \geq 2$, any bipartite $(C \log_2 \Delta)$-fold list-cover of G admits an independent transversal.*

There are three potential directions to highlight through adoption of the above notation. First, note that list-covers form a proper subclass of all cover graphs, and so we might consider the 'colouring' task under increasingly more general conditions with respect to H. More specifically, we may ask analogous questions about sufficient conditions for the existence of an IT in natural and successively larger superclasses of list-covers (among all cover graphs). Second, note that if G has maximum degree Δ, then any list-cover of G has maximum degree Δ, but the converse is not true in general. And, for instance, we may consider a problem/result about list-colouring in some class of bounded degree graphs and try to generalise it to the analogous class of bounded degree list-covers. This type of 'colour-degree' problem was introduced by Reed [10]. Third, and specific to (A, B)-covers, we may insist on a more refined viewpoint by imposing (degree/list-size/structural) conditions on parts A and B separately. Two of the authors together with Alon [1] introduced this third asymmetric perspective for studying Conjecture 1.1, and in a follow up [4] they furthermore took on the first two perspectives, in particular by generalising the problem to so-called *correspondence-covers*, which we discuss later. Here we concentrate on the most general case for (asymmetric, bipartite) cover graphs with given degree bounds.

The following problem was posed in [4] (see therein the special case of Problem 1.1 with Δ_A, Δ_B infinite).

Problem 1.2. *Let H be an (A, B)-cover of G with respect to L. What conditions on positive integers k_A, k_B, D_A, D_B suffice to ensure the following? If the maximum degrees in A_H and B_H are D_A and D_B, respectively, and $|L(v)| \geq k_A$ for all $v \in A_G$ and $|L(w)| \geq k_B$ for all $w \in B_G$, then there is guaranteed to be an independent transversal of H with respect to L.*

We resolve Problem 1.2 through the following sufficient condition for a bipartite cover graph to admit an IT.

Theorem 1.3. *Let H be an (A, B)-cover of G with respect to L. Let positive integers k_A, k_B, D_A, D_B be such that $\frac{D_B}{k_A} + \frac{D_A}{k_B} \leq 1$. If the maximum degrees in A_H and B_H are D_A and D_B, respectively, and $|L(v)| \geq k_A$ for all $v \in A_G$ and $|L(w)| \geq k_B$ for all $w \in B_G$, then H admits an independent transversal with respect to L.*

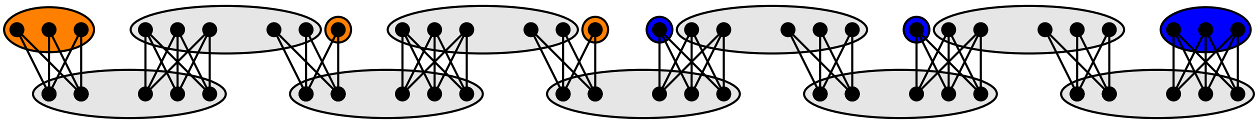


Figure 1: A bipartite graph with maximum degree 3 and partition classes of size 5 with no IT

This result is corollary to a general result for independent transversals found in [6].

In fact, the condition in Theorem 1.3 is best possible, as follows.

Theorem 1.4. *Let positive integers k_A, k_B, D_A, D_B be such that $\frac{D_B}{k_A} + \frac{D_A}{k_B} > 1$. Then there exists an (A, B)-cover H of G with respect to L such that the maximum degrees in A_H and B_H are D_A and D_B, respectively, and $|L(v)| = k_A$ for all $v \in A_G$ and $|L(w)| = k_B$ for all $w \in B_G$, and such that H admits no independent transversal with respect to L.*

It is worth isolating the symmetric situation where we maintain that $D_A = D_B = D$ and $k_A = k_B = k$; in this case the condition in Theorem 1.3 resolves to $k \geq 2D$. In other words, we have the following.

Corollary 1.5. *Any bipartite $(2D)$-fold cover graph of maximum degree D admits an independent transversal. Moreover, the conclusion may fail if the $2D$ part size condition is relaxed to $2D - 1$.*

This condition coincides with that of a well-known, more general result of the second author [7]: that any $(2D)$-fold cover graph of maximum degree D is guaranteed to admit an IT. As such, one may see Theorem 1.4 as simultaneously a strengthening and generalisation of a result of Szabó and Tardos [11] (which in turn built upon a series of results beginning in the original paper of Bollobás, Erdős and Szemerédi [3]): that there exists a $(2D - 1)$-fold cover graph of maximum degree D that does not admit an IT. Recalling the question posed at the beginning, Theorem 1.4 shows in a wider sense how the bipartite assumption does *not* help for the existence of ITs in cover graphs.

We remark that, while the construction of Szabó and Tardos is composed of the union of complete bipartite graphs, its partition classes do not align with a bipartition. Corollary 1.5 affirms that it is possible to achieve such an alignment in some bipartite construction. For an indication of the difference, Figure 1 depicts the $D = 3$ construction in Corollary 1.5, and one can compare it with [11, Fig. 1].

Let us briefly discuss what happens in the special case of correspondence-covers, as explored in [4]. Given a cover graph H of G with respect to L, we say H is a *correspondence-cover* if the bipartite subgraph induced between $L(v)$ and $L(v')$ is a matching for any $vv' \in V(G)$. In other words, the maximum degree induced between two parts of H with respect to L is at most 1. Clearly the class of all correspondence-covers strictly includes that of all list-covers. The next result follows from a 'coupon collector' argument, and this is counterbalanced by a simple probabilistic construction (that was given, for example, in [9]).

Theorem 1.6 ([4]). *For any $\varepsilon > 0$, the following holds for all D sufficiently large. Any bipartite $(1+\varepsilon)\frac{D}{\log D}$-fold correspondence-cover graph of maximum degree D admits an independent transversal. Moreover, the conclusion fails if the $(1+\varepsilon)$ factor is weakened to a $(\frac{1}{2}-\varepsilon)$ factor.*

One reason for highlighting this case is that it could be interesting to gradually tune (between 1 and D) the condition on maximum degree induced between two parts of H with respect to L, in order to gain a better understanding of the transition between the $\Theta(D/\log D)$ (probabilistic) part-size condition in Theorem 1.6 and the $\Theta(D)$ condition in Corollary 1.5 (which was originally established in [7]).

Let us conclude by returning to the original motivation and a related challenge. With Corollary 1.5 and Theorem 1.6 in mind, the following 'colour-degree' generalisation of Conjecture 1.1 seems worth investigating.

Conjecture 1.7. *There is some $C \geq 1$ such that any bipartite $(C\log_2 D)$-fold list-cover graph of maximum degree $D \geq 2$ admits an independent transversal.*

To round out the story, we point out how Conjectures 1.1 and 1.7 are essentially equivalent.

Theorem 1.8. *If Conjecture 1.1 is true for some constant $C \geq 1$, then Conjecture 1.7 is true for some constant $C' \geq 1$. The same implication holds when C and C' are both replaced by $1+o(1)$ (as $\Delta, D \to \infty$).*

Proof. Assume Conjecture 1.1 is true for some $C \geq 1$. We choose $D_0 \geq 2$ such that $\sqrt{D} \geq C\log_2 D$ for every $D \geq D_0$, and take $C' = 2D_0 \geq 2C^2 \geq 2C$. We will prove that any bipartite $(C'\log_2 D)$-fold list-cover graph of maximum degree D admits an independent transversal. Let $k = C'\log_2 D$ and H be a k-fold list-cover of maximum degree $D \geq 2$. If $D \leq D_0$, then $k \geq C' \geq 2D$ and it then follows from Haxell's theorem [7] that H has an independent transversal as desired. We may therefore assume $D > D_0$. Let G be the 'covered' graph of H, i.e. uv is an edge of G if and only if $L(u) \cap L(v)$ is not empty. Then by definition the maximum degree Δ of G satisfies $D_0 \leq D \leq \Delta \leq kD$. By the choice of D_0 it then follows that $\sqrt{\Delta} \geq C\log_2 \Delta$. Suppose now for a contradiction that $k < C\log_2 \Delta$. Then $\Delta < D \cdot C\log_2 \Delta$ and so $D > \frac{\Delta}{C\log_2 \Delta} \geq \sqrt{\Delta}$. But $D > \sqrt{\Delta}$ and $C' \geq 2C$ imply that $k = C'\log_2 D \geq C\log_2 \Delta$, which is a contradiction. Hence $k \geq C\log_2 \Delta$. Now consider the maximal list-cover $H' \supseteq H$ of G with respect to L. By our assumption, H' admits an independent transversal with respect to L, which implies the same conclusion for H, as required.

The proof for the $1+o(1)$ version proceeds analogously. Fix $\varepsilon > 0$. Now one can take D_0 sufficiently large such that $C\log_2 D \leq D^{\varepsilon}$ and Conjecture 1.1 is true with $1+\varepsilon$ whenever $\Delta \geq D_0$. Then for any k-fold list-cover H of maximum degree $D \geq 2$, where $k \geq \frac{1+\varepsilon}{1-\varepsilon}\log_2 D$, we conclude H has an independent transversal by the same strategy. □

In a similar way, nontrivial progress on Conjecture 1.1 may imply nontrivial progress on Conjecture 1.7. Conversely, lower bound constructions related to Conjecture 1.7 may directly yield corresponding constructions related to Conjecture 1.1.

2 A sufficient condition

In this section, we derive Theorem 1.3.

We say that a set U of vertices of a graph G *dominates* the set $W \subseteq V(G)$ if every vertex of W has a neighbour in U. (This is a somewhat nonstandard use of the term since, contrary to the most common usage, here we require each vertex of $U \cap W$ to have a neighbour in U.) Theorem 1.3 is a straightforward consequence of the following result of Haxell (see e.g. [6]), concerning critical graphs with respect to ITs.

Theorem 2.1 ([6]). *Let $H = (V_H, E_H)$ be a cover graph of some graph $G = (V_G, E_G)$ with respect to L. Suppose that H has no IT but $H - e$ does for any $e \in E_H$. Then for any $e \in E_H$, there exists a subset $S \subset V_G$ and a set Z of edges of the subgraph of H induced by $L(S)$ such that $e \in Z$, $|Z| \le |S| - 1$, and $V_H(Z)$ dominates $L(S)$.*

Proof of Theorem 1.3. Suppose H is a counterexample and take it to be edge-minimal. By Theorem 2.1, there exist some a partition classes of A_H and b partition classes of B_H, and a set Z of edges of size at most $a + b - 1$ whose end-vertices dominate the union of these $a + b$ partition classes. The end-vertices of Z dominate at most $(a + b - 1)D_B$ vertices in A_H, while the a partition classes contain at least ak_A vertices. This implies that $ak_A \le (a + b - 1)D_B$. Similarly, considering B_H, we have $bk_B \le (a + b - 1)D_A$. But then

$$\frac{D_A}{k_B} + \frac{D_B}{k_A} \ge \frac{b}{a+b-1} + \frac{a}{a+b-1} > 1,$$

contradicting the hypothesis. □

3 Sharpness of the condition

A proof of Theorem 1.4 is deferred to the full manuscript associated to this extended abstract.

Acknowledgement

This work was initiated during a visit by the second author to Nijmegen in April of 2022. The hospitality of the Department of Mathematics in Nijmegen is gratefully acknowledged.

References

[1] N. Alon, S. Cambie, and R. J. Kang. Asymmetric list sizes in bipartite graphs. *Ann. Comb.*, 25(4):913–933, 2021.

[2] N. Alon and M. Krivelevich. The choice number of random bipartite graphs. *Ann. Comb.*, 2(4):291–297, 1998.

[3] B. Bollobás, P. Erdős, and E. Szemerédi. On complete subgraphs of r-chromatic graphs. *Discrete Math.*, 13(2):97–107, 1975.

[4] S. Cambie and R. J. Kang. Independent transversals in bipartite correspondence-covers. *Canad. Math. Bull.*, 65(4):882–894, 2022.

[5] P. Erdős, A. L. Rubin, and H. Taylor. Choosability in graphs. In *Proceedings of the West Coast Conference on Combinatorics, Graph Theory and Computing (Humboldt State Univ., Arcata, Calif., 1979)*, Congress. Numer., XXVI, pages 125–157, Winnipeg, Man., 1980. Utilitas Math.

[6] P. Haxell. On forming committees. *Amer. Math. Monthly*, 118(9):777–788, 2011.

[7] P. E. Haxell. A note on vertex list colouring. *Combin. Probab. Comput.*, 10(4):345–347, 2001.

[8] A. Johansson. Asymptotic choice number for triangle-free graphs. Technical Report 91-5, DIMACS, 1996.

[9] D. Král, O. Pangrác, and H.-J. Voss. A note on group colorings. *J. Graph Theory*, 50(2):123–129, 2005.

[10] B. Reed. The list colouring constants. *J. Graph Theory*, 31(2):149–153, 1999.

[11] T. Szabó and G. Tardos. Extremal problems for transversals in graphs with bounded degree. *Combinatorica*, 26(3):333–351, 2006.

[12] V. G. Vizing. Coloring the vertices of a graph in prescribed colors. *Diskret. Analiz*, 29 Metody Diskret. Anal. v Teorii Kodov i Shem:3–10, 101, 1976.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Chordal graphs with bounded tree-width

(Extended abstract)

Jordi Castellví* Michael Drmota† Marc Noy‡ Clément Requilé§

Abstract

Given $t \geq 2$ and $0 \leq k \leq t$, we prove that the number of labelled k-connected chordal graphs with n vertices and tree-width at most t is asymptotically $cn^{-5/2}\gamma^n n!$, as $n \to \infty$, for some constants $c, \gamma > 0$ depending on t and k. Additionally, we show that the number of i-cliques ($2 \leq i \leq t$) in a uniform random k-connected chordal graph with tree-width at most t is normally distributed as $n \to \infty$.

The asymptotic enumeration of graphs of tree-width at most t is wide open for $t \geq 3$. To the best of our knowledge, this is the first non-trivial class of graphs with bounded tree-width where the asymptotic counting problem is solved. Our starting point is the work of Wormald [Counting Labelled Chordal Graphs, *Graphs and Combinatorics* (1985)], were an algorithm is developed to obtain the exact number of labelled chordal graphs on n vertices..

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-037`

*Departament de Matemàtiques de la Universitat Politècnica de Catalunya (UPC), Barcelona, Spain. E-mail: `jordi.castellvi@upc.edu`.

†Institute for Discrete Mathematics and Geometry of the Technische Universität Wien, Austria. E-mail: `michael.drmota@tuwien.ac.at`. Supported by the Special Research Program SFB F50-02 "Algorithmic and Enumerative Combinatorics", by the project P35016 "Infinite Singular Systems and Random Discrete Objects" of the FWF, and by the Marie Curie RISE research network "RandNet" MSCA-RISE-2020-101007705.

‡Departament de Matemàtiques and Institut de Matemàtiques de la Universitat Politècnica de Catalunya, and Centre de Recerca Matemàtica, Barcelona, Spain. E-mail: `marc.noy@upc.edu`. Supported by the Spanish State Research Agency through projects MTM2017-82166-P and PID2020-113082GB-I00, by the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence project CEX2020-001084-M, and by the Marie Curie RISE research network "RandNet" MSCA-RISE-2020-101007705.

§Departament de Matemàtiques and Institut de Matemàtiques de la Universitat Politècnica de Catalunya, Barcelona, Spain. E-mail: `clement.requile@upc.edu`. Supported by the Spanish State Research Agency through projects MTM2017-82166-P and PID2020-113082GB-I00, by the grant Beatriu de Pinós BP2019 funded by the H2020 COFUND project No 801370 and AGAUR, and by the Marie Curie RISE research network "RandNet" MSCA-RISE-2020-101007705.

1 Introduction

Tree-width is a fundamental parameter in structural and algorithmic graph theory, as illustrated for instance in [8]. It can be defined in terms of tree-decompositions or equivalently in terms of k-trees. A k-tree is defined recursively as either a complete graph on $k+1$ vertices or a graph obtained by adjoining a new vertex adjacent to a k-clique of a smaller k-tree. The tree-width of a graph Γ is then the minimum k such that Γ is a subgraph of a k-tree. In particular, k-trees are the maximal graphs with tree-width at most k. The number of k-trees on n labelled vertices was independently shown [3, 17, 15] to be

$$\binom{n}{k}(k(n-k)+1)^{n-k-2} = \frac{1}{\sqrt{2\pi}\,k!\,k^{k+2}}\,n^{-5/2}\,(ek)^n\,n!\,(1+o(1)), \tag{1}$$

where the estimate holds for k fixed and $n \to \infty$. However, there are relatively few results on the enumeration of graphs of given tree-width or on properties of random graphs with given tree-width. Graphs of tree-width one are forests (acyclic graphs) and their enumeration is a classical result, while graphs of tree-width at most two are series-parallel graphs and were first counted in [5]. The problem of counting graphs of tree-width three is still open. From now on, we will use t to denote the tree-width while k will denote the connectivity of a graph. All graphs considered in this work will be simple and labelled, that is, with vertex-set $[n]$.

Given that tree-width is non-increasing under taking minors, the class of graphs with tree-width at most t is 'small' when t is fixed, in the sense that the number $g_{n,t}$ of labelled graphs with n vertices and tree-width at most t grows at most like $c^n n!$ for some $c > 0$ depending on t (see [18, 13]). The best known bounds for $g_{n,t}$ are, up to lower order terms,

$$\left(\frac{2^t tn}{\log t}\right)^n \leq g_{n,t} \leq (2^t tn)^n.$$

The upper bound follows by considering all possible subgraphs of t-trees, and the lower bound uses a suitable construction developed in [2]. In the present work we determine the asymptotic number of labelled chordal graphs with tree-width at most t, following the approach in [14] and [11], and based on the analysis of systems of equations satisfied by generating functions.

A graph is chordal if every cycle of length greater than three contains at least one chord, that is, an edge connecting non-consecutive vertices of the cycle. Chordal graphs have been extensively studied in structural graph theory and graph algorithms (see for instance [16]), but not so much from the point of view of enumeration. Wormald [20] used generating functions to develop a method for finding the exact number of chordal graphs with n vertices for a given value of n. It is based on decomposing chordal graphs into k-connected components for each $k \geq 1$. As remarked in [20], it is difficult to define the k-connected components of arbitrary graphs for $k > 3$, but for chordal graphs they are well defined. It is a consequence of Dirac's characterisation [10]: *in a chordal graph every minimal separator is a clique.*

For fixed $n, t \geq 1$ and $0 \leq k \leq t$, let $\mathcal{G}_{t,k,n}$ be the set of k-connected chordal graphs with n labelled vertices and tree-width at most t. Our two main results are the following.

Theorem 1.1. *For $t \geq 1$ and $0 \leq k \leq t$, there exist $c_{t,k} > 0$ and $\gamma_{t,k} > 1$ such that*

$$|\mathcal{G}_{t,k,n}| = c_{t,k}\, n^{-5/2}\, \gamma_{t,k}^n\, n!\, (1 + o(1)) \qquad \textit{as } n \to \infty.$$

We remark that in principle, for fixed t and k the constants $c_{t,k}$ and $\gamma_{t,k}$ can be computed, at least approximately.

Theorem 1.2. *Let $t \geq 1$, $0 \leq k \leq t$. For $i \in \{2, \ldots, t\}$ let $X_{n,i}$ denote the number of i-cliques in a uniform random graph in $\mathcal{G}_{t,k,n}$, and set $\mathbf{X_n} = (X_{n,2}, \ldots, X_{n,t})$. Then $\mathbf{X_n}$ satisfies a multivariate central limit theorem, that is, as $n \to \infty$ we have*

$$\frac{1}{\sqrt{n}}\,(\mathbf{X}_n - \mathbb{E}\,\mathbf{X}_n) \xrightarrow{d} N(0, \Sigma), \qquad \textit{with} \quad \mathbb{E}\,\mathbf{X}_n \sim \alpha n \quad \textit{and} \quad \mathbb{C}\mathrm{ov}\,\mathbf{X}_n \sim \Sigma n,$$

and where α is a $(t-1)$-dimensional vector of positive numbers and Σ is a $(t-1) \times (t-1)$-dimensional positive semi-definite matrix.

Let us point that more structural asymptotic results can be expected. Notably, the class of chordal graphs with tree-width at most t is *subcritical* in the sense of [12]. It follows from [19] that the uniform random connected chordal graph with tree-width at most t with distances rescaled by $1/\sqrt{n}$ admits the *Continuum Random Tree* (CRT) [1] as a scaling limit, multiplied by a constant that depends on t.

A more complete version of the work presented here can be found in [6].

2 Decomposition of chordal graphs

Let $k \geq 1$. A k-separator of a graph Γ is a subset of k vertices whose removal disconnects Γ. And Γ is said to be k-connected if it contains no i-separator for $i \in [k-1]$. With this definition, we consider the complete graph on k vertices to be k-connected, for any $k \geq 1$, contrary to the usual definition of connectivity (see for instance [9]). A k-connected component of Γ is a k-connected subgraph that is maximal, in term of subgraph containment, with that property.

An essential consequence of chordality is that k-connected chordal graphs admit a unique decomposition into $(k+1)$-connected components through its k-separators. This is a generalisation of the well-known decomposition of a connected graph into so-called blocks, that are maximal 2-connected components. And it induces a system of functional equations satisfied by the generating function counting chordal graphs of tree-width at most k.

We now fix some $t \geq 1$ and let $\mathcal{G}$ be the family of chordal graphs with tree-width at most t. For a graph $\Gamma \in \mathcal{G}$ and $j \in [t]$, let us denote by $n_j(\Gamma)$ the number of j-cliques of

Γ. In the rest of the paper, we will write $\mathbf{x}$ as a short-hand for $x_1, \ldots, x_t$, and define the multivariate (exponential) generating function associated to $\mathcal{G}$ to be

$$G(\mathbf{x}) = G(x_1, \ldots, x_t) = \sum_{\Gamma \in \mathcal{G}} \frac{1}{n_1(\Gamma)!} \prod_{j=1}^{t} x_j^{n_j(\Gamma)}.$$

Let g_n denote the number of chordal graphs with n vertices and tree-width at most t. Then,

$$G(x, 1, \ldots, 1) = \sum_{n \geq 1} \frac{g_n}{n!} x^n.$$

For $0 \leq k \leq t+1$, let $\mathcal{G}_k$ be the family of k-connected chordal graphs with tree-width at most t and $G_k(\mathbf{x})$ be the associated generating function. In particular, we have

$$G_{t+1}(\mathbf{x}) = \frac{1}{(t+1)!} \prod_{j \in [t]} x_j^{\binom{t+1}{j}}. \tag{2}$$

For other values of k, we need to consider graphs rooted at a clique. Rooting the graph $\Gamma \in \mathcal{G}_k$ at an i-clique means distinguishing one i-clique K of Γ and choosing an ordering of (the labels of) the vertices of K. In order to avoid over-counting, we will discount the subcliques of K. Let $i \in [k]$ and define $\mathcal{G}_k^{(i)}$ to be the family of k-connected chordal graphs with tree-width at most t and rooted at an i-clique. Let then $G_k^{(i)}(\mathbf{x})$ be the associated generating function, where now for $1 \leq j \leq i$ the variables x_j mark the number of j-cliques that are not subcliques of the root.

Lemma 2.1. *Let $k \in [t]$. Then the following equations hold:*

$$G_{k+1}^{(k)}(\mathbf{x}) = k! \prod_{j=1}^{k-1} x_j^{-\binom{k}{j}} \frac{\partial}{\partial x_k} G_{k+1}(\mathbf{x}), \tag{3}$$

$$G_k^{(k)}(\mathbf{x}) = \exp\left(G_{k+1}^{(k)}\big(x_1, \ldots, x_{k-1}, x_k G_k^{(k)}(\mathbf{x}), x_{k+1}, \ldots, x_t\big) \right), \tag{4}$$

$$G_k(\mathbf{x}) = \frac{1}{k!} \prod_{j=1}^{k-1} x_j^{\binom{k}{j}} \int G_k^{(k)}(\mathbf{x}) \; dx_k. \tag{5}$$

Finally, the fact that a graph is the set of its connected components can be translated as $G(\mathbf{x}) = G_0(\mathbf{x}) = \exp(G_1(\mathbf{x}))$. Then, it is clear that one can derive $G_0(\mathbf{x})$ from $G_{t+1}(\mathbf{x})$ by successively using Identities (3), (4) and (5) from Lemma 2.1.

3 Asymptotic analysis

Fix $t \geq 1$. To prove Theorems 1.1 and 1.2, we use rather classical methods from [14] and [11, Chapter 2] which consist in deriving asymptotic estimates from local expansions of the generating functions from Section 2 at their singularities, typically by applying a *Transfer Theorem* (for instance [11, Lemma 2.18]).

However, the main difficulties here are the multivariate nature of Lemma 2.1, in particular the fact that the local expansions are with respect to different variables from one step to the next, and the fact that local expansions have to be "carried over" from $G_{t+1}(\mathbf{x})$ to $G_0(x_1, 1, \ldots, 1)$. To overcome this, we extend some of the tools and notions present in [11].

Sketch of the proofs of Theorems 1.1 and 1.2. Starting with G_{t+1} which is an explicit monomial, we recursively compute via Lemma 2.1 local representations of $G_t, G_{t-1}, \ldots, G_1$ and finally of $G_0 = \exp(G_1)$.

The first step of the induction amounts to computing a multivariate local representation of the generating function of t-trees. Let $x_2, \ldots, x_t \in \mathbb{R}_+$. Then there exist two functions $h_1(x)$ and $h_2(x)$, that are analytic and non-zero at $x_1 = 1/e$, such that for $x_1 \sim 1/e$ we have

$$G_t(\mathbf{x}) = \frac{\prod_{j=1}^{t} x_j^{\binom{t}{j}}}{t!}\left(h_1(tX) + h_2(tX)(1-etX)^{3/2}\right), \qquad \text{where } X = \prod_{j=1}^{t} x_j^{\binom{t}{j-1}}.$$

From there, one can prove that the above representation for $G_t(\mathbf{x})$ implies corresponding representation for $G_{t-1}(\mathbf{x}), G_{t-2}(\mathbf{x}), \ldots, G_1(\mathbf{x})$, then $G_0(\mathbf{x})$. And the main counting result can be deduced by setting $x_2 = \cdots = x_t = 1$ then applying a *Transfer Theorem*.

Finally, the joint central limit theorem can be obtained in a similar manner: first showing that a local representation of $G_k(\mathbf{x})$ can be extended uniformly in a neighbourhood of $(1, \ldots, 1) \in \mathbb{C}^{t-1}$, then concluding with the *Quasi-Powers Theorem* [11, Theorem 2.22].

4 Concluding remarks

Let us mention a recent result [7] giving an estimate $cn^{-5/2}\gamma^n n!$ for the number of labelled *planar* chordal graphs with $\gamma \approx 11.89$. It turns pout that the class of chordal graphs with tree-width at most three is exactly the same as the class of chordal graphs not containing K_5 as a minor, whose asymptotic growth is, according to Theorem 1.1 and some numerical computations of the form $cn^{-5/2}\delta^n n!$ with $\delta = 1/\rho_{3,1} \approx 12.98$.

Since the number of all chordal graphs grows like $2^{n^2/4}$ (see [4]), the singularity $\rho_t = \rho_{t,1} \to 0$ as $t \to \infty$. Concerning the rate of convergence, since the exponential growth of t-trees is $(etn)^n$, we have $\rho_t = O(1/t)$. And since the growth of all graphs of tree-width at most t is at most $(2^t tn)^n$, we also have $\rho_t = \Omega\left(1/(t2^t)\right)$. A remaining problem is to narrow the gap between the upper and lower bounds. Heuristic arguments suggest that ρ_t decreases exponentially in t.

As a final question, we consider letting $t = t(n)$ grow with n. Recall that a class of labelled graphs is small when the number of graphs in the class grows at most like $c^n n!$ for some $c > 0$, and large otherwise. One can prove that if $t = (1+\epsilon)\log n$ then the class of labelled chordal graphs of tree-width at most t is large for each $\epsilon > 0$. We leave as an open problem to determine at which order of magnitude between $t = O(1)$ and $t = \log n$ the class ceases to be small.

Acknowledgements. We thank the anonymous referees for their comments on an earlier version, and Juanjo Rué and Dimitrios Thilikos for earlier discussions on this topic. Additionally, the authors thank the support from the Marie Curie RISE research network "RandNet" MSCA-RISE-2020-101007705.

References

[1] D. Aldous. The Continuum Random Tree. I. *The Annals of Probability*, 19(1):1 – 28, 1991.

[2] J. Baste, M. Noy, and I. Sau. On the number of labeled graphs of bounded treewidth. *European Journal of Combinatorics*, 71:12–21, 2018.

[3] L. W. Beineke and R. E. Pippert. The number of labeled k-dimensional trees. *Journal of Combinatorial Theory*, 6(2):200–205, 1969.

[4] E. A. Bender, L. B. Richmond, and N. C. Wormald. Almost all chordal graphs split. *Australian Mathematical Society. Journal. Series A. Pure Mathematics and Statistics*, 38(2):214–221, 1985.

[5] M. Bodirsky, O. Giménez, M. Kang, and M. Noy. Enumeration and limit laws for series-parallel graphs. *European Journal of Combinatorics*, 28(8):2091–2105, 2007.

[6] J. Castellví, M. Drmota, M. Noy, and C. Requilé. Chordal graphs with bounded tree-width. 2023. arXiv:2301.00194.

[7] J. Castellví, M. Noy, and C. Requilé. Enumeration of chordal planar graphs and maps. *Discrete Mathematics*, 346(1):113163, 2023.

[8] M. Cygan, F. V. Fomin, L. Kowalik, D. Lokshtanov, D. Marx, M. Pilipczuk, M. Pilipczuk, and S. Saurabh. *Parameterized algorithms.* Springer, 2015.

[9] R. Diestel. *Graph Theory.* Springer Berlin, Heidelberg, Fifth edition, 2016.

[10] G. A. Dirac. On rigid circuit graphs. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 25:71–76, 1961.

[11] M. Drmota. *Random Trees.* Springer-Verlag Wien, 2009.

[12] M. Drmota, E. Fusy, M. Kang, V. Kraus, and J. Rué. Asymptotic study of subcritical graph classes. *SIAM Journal on Discrete Mathematics*, 25(4):1615–1651, 2011.

[13] Z. Dvořák and S. Norine. Small graph classes and bounded expansion. *Journal of Combinatorial Theory. Series B*, 100(2):171–175, 2010.

[14] P. Flajolet and R. Sedgewick. *Analytic Combinatorics.* Cambridge University Press, Cambridge, 2009.

[15] D. Foata. Enumerating k-trees. *Discrete Mathematics*, 1(2):181–186, 1971.

[16] M. C. Golumbic. *Algorithmic Graph Theory and Perfect Graphs*, volume 57 of *Annals of Discrete Mathematics.* Elsevier Science B.V., Amsterdam, second edition, 2004. With a foreword by Claude Berge.

[17] J. W. Moon. The number of labeled k-trees. *Journal of Combinatorial Theory*, 6(2):196–199, 1969.

[18] S. Norine, P. Seymour, R. Thomas, and P. Wollan. Proper minor-closed families are small. *Journal of Combinatorial Theory. Series B*, 96(5):754–757, 2006.

[19] K. Panagiotou, B. Stufler, and K. Weller. Scaling limits of random graphs from subcritical classes. *The Annals of Probability*, 44(5):3291–3334, 2016.

[20] N. C. Wormald. Counting labelled chordal graphs. *Graphs and Combinatorics*, 1(2):193–200, 1985.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

A compactification of the set of sequences of positive real numbers with applications to limits of graphs

(Extended abstract)

David Chodounský* Lluís Vena†

Abstract

We introduce compactification results on the set of sequences of positive real numbers: under the continuum hypothesis, one can find a totally ordered set of sequences whose elements can be used as test sequences to capture every possible asympthotic growth, perhaps along a subsequence; this behaviour mimics the statement that, in a compact set of $\mathbb{R}$, every sequence has a convergent partial subsequence. These compactification results allows us to unify two notions of convergence for graphs into a single graph-convergence notion, while retaining the property that each sequence of graphs have a convergent partial subsequence. These convergent notions are the Benjamini-Schramm convergence for bounded degree graphs, regarding the distribution of r-neighbourhoods of the vertices, and the left-convergence for dense graphs, regarding the existence, for each fixed graph F, of a limiting probability that a random mapping from F to $\{G_i\}$ is a graph homomorphism.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-038`

*Institute of Mathematics of the Czech Academy of Sciences and Charles University, Žitná 25, Praha 1, 115 67, Czech Republic. E-mail: `chodounsky@math.cas.cz`. Supported by the Czech Academy of Sciences (RVO 67985840).

†Departament of Mathematics and Institute of Mathematics (IMTech), Universitat Politècnica de Catalunya, Jordi Girona 1-3,08034 Barcelona, Spain. E-mail: `lluis.vena@upc.edu`. Supported from the grant Beatriu de Pinós BP2018, funded by the H2020 COFUND project No 801370 and AGAUR (the Catalan agency for managment of university and research grants).

1 Introduction

Recently, several authors have considered the study of (large) discrete objects by, after introducing an appropriate limit notion, draw conclusions of the sequence by studying the objects that appear as their limits. Two of the most well known examples is the study of limits of sequences of graphs (e.g. [2, 9, 4]) or of permutations (see [8]). In this work, we focuss our attention to limits of graphs.

In the area of limits of graphs, one of the problems is that the properties of the sequences of graphs are radically different depending on several parameters, one of them being the density of edges. Thus, there have appeared several convergence notions depending on different growth regimes, such as the notion of left-convergence [9, 4] that works well when the density of edges is a positive proportion of the total number (dense case) yet trivializes when the sequence is of sparse graphs (non-dense), or Benjamini-Schramm convergence [2] when the sequence is of bounded degree graphs (very sparse case). Other notions of convergence for limits of graphs have been introduced; these either generalize the previous two in several ways, or consider some strenghthening of them. As some examples we can mention: L^p convergence [5, 3], action convergence [1], log–convergence [11], convergence in fragments of logic [10], for intermediate growth [6], or local-global convergence [7].

In the following, we give compactification results on the set of sequences of positive real numbers Theorem 1 and Theorem 2 that, as far as we know, are new, and we give an application of such results to limits of graphs by considering a graph limit notion that is uniform regardless of the growth regime of the number of edges, thus generalizing both the local convergence [2] and the left-convergence [9, 4]. This notion can be seen as a brute force generalization of the one by Frenkel [6]. First let us present the compactification results.

Theorem 1. *Assume the continuum hypothesis. Let $\mathcal{R}^{>0} = \{f \mid f : \omega \to \mathbb{R}^{>0}\}$ be the set of positive real sequences. Then there exists a set $A \subset \mathcal{R}^{>0}$ such that:*

1. *For each $a, b \in A$,*
$$\lim_{n\to\infty} \frac{a(n)}{b(n)} \qquad \text{is either } 0 \text{ or } \infty.$$

2. *For each $g \in \mathcal{R}^{>0}$ and each ordered embedding $\iota : (\omega, <) \to (\omega, <)$, there exists an ordered embedding $\iota_0 : (\omega, <) \to (\omega, <)$, an $a \in A$, and a $c \in (0, \infty)$ such that*
$$\lim_{n\to\infty} \frac{a(\iota(\iota_0(n)))}{g(\iota(\iota_0(n)))} = c$$

Theorem 1 claims to obtain a totally ordered set A (ordered with the relation $a < b \iff \lim_{n\to\infty} \frac{a(n)}{b(n)} = 0$) with the property that for any partial sequence (given by the pair (g, ι)), there exists a subsequence (given by ι_0), an element of A (given by a), and a constant $c \in (0, \infty)$ such that, up to c, the sequence a gives the asympthotic behaviour of g along a subsequence ι_0. We interpret this result in two ways:

- the set A is a set of test functions that verifies that $\mathcal{R}^{>0}$ satisfies the following "compactification property": "every sequence have a convergent partial subsequence".
- the set of test functions A "captures" every possible asympthotic behaviour.

We can impose additional restrictions on A and on the set of sequences considered.

Theorem 2. *Assume the continuum hypothesis. Let $f_0, f_1 : \omega \to \mathbb{R}^{>0}$ such that $\frac{f_0(i)}{f_1(i)} < \frac{f_0(i-1)}{f_1(i-1)}$ for $i > 0$. Let $\mathcal{R}^{>0}(f_0, f_1) = \{f \mid f : \omega \to \mathbb{R}^{>0}, f(i) \in [f_0(i), f_1(i)]\}$ the set of positive real sequences between f_0 and f_1. There exists a set $A \subset \mathcal{R}^{>0}(f_0, f_1)$ such that 1 and 2 in Theorem 1 are satisfied and, moreover, $f_0, f_1 \in A$.*

Applications to limits of graphs. Let $\mathcal{G}^\circ$ be the set of finite graphs with one loop in each vertex, and $\mathcal{G}$ the set of finite graphs. For each $F \in \mathcal{G}$, let A_F denote a set of sequences of positive real numbers obtained by using Theorem 2 with $f_0(n) = n$ and $f_1(n) = n^{|V(F)|}$, and with both sequences in A_F, where $V(F)$ is the vertex set of F. Note that, for each $G \in \mathcal{G}^\circ$, $|\{h : F \to G \ : \ h \text{ is a graph homomorphism}\}| \in [n, n^{|V(F)|}]$.

Let $\{G_i\}_{i\in I}$ be a sequence of graphs in $\mathcal{G}^\circ$ with strictly increasing number of vertices (not necessarily $|G_i| = i$, just an strictly increasing sequence). We say that (see [12, Definition 2.1])

$$\{G_i\}_{i\in I} \text{ is } q\text{-convergent to } \{(f_F, c_F)\}_{f_F\in A_F, c_F\in(0,\infty)} \iff$$

$$\text{for each } F \in \mathcal{G}, \qquad \lim_{i\to\infty} \frac{|\{h : F \to G_i \ : \ h \text{ is a graph homomorphism}\}|}{f_F(\ |V(G_i)|\ )} = c_F \tag{1}$$

Note that the use of $\mathcal{G}^\circ$ instead of $\mathcal{G}$ is mostly for technical reasons, as we always want to consider sequences of non-zero real numbers. Note also that, by doing inclusion–exclusion arguments, the number of homomorphisms from F to G' (with the loops removed) can be recovered from the number of homomorphisms from F_i to G (graph with one loop on each vertex), where $\{F_i\}$ are the subgraphs of G. Now, a couple of results that gives the application of the "compactification" result to limits of graphs.

Proposition 3. *Assume the continuum hypothesis. Let $\{G_i\}_{i\in I}$ is an infinite sequence of graphs, with strictly increasing number of vertices, then there exists an infinite $I_0 \subseteq I$ such that $\{G_i\}_{i\in I_0}$ is q-convergent.*

Equivalently, any sequence has a partial convergent subsequence.

Proposition 4. *Let $\{G_i\}_{i\in I}$ is an infinite sequence of graphs in $\mathcal{G}^\circ$, with strictly increasing number of vertices, and such that for each $F \in \mathcal{G}$,*

$$\lim_{i\to\infty} |\{h : F \to G_i \ : \ h \textit{ is a graph homomorphism}\}| \,/\, \left[|V(G_i)|^{|V(F)|}\right] = c_F,\ c_F > 0$$

then $\{G_i\}_{i\in I}$ is also q-convergent.

Equivalently, if the sequence is convergent in the dense sense with positive probabilities ([9]), then it is also q-convergent using the same constants and functions $n \to n^{|V(F)|}$ for each $F \in \mathcal{G}$. In this case, the same would be true for the sequences of graphs where the loops have been removed.

Proposition 5. *Let $\{G_i\}_{i\in I}$ be an infinite sequence of graphs each of which has maximum degree upper bounded by d, belong to $\mathcal{G}^\circ$, and the number of vertices is strictly increasing along the sequence. Assume that, for each graph $F \in \mathcal{G}$ we have*

$$\lim_{i\to\infty} |\{h : F \to G_i \,:\, h \text{ is an graph homomorphism}\}| \,/|V(F)| = c_F,\; c_F > 0$$

if and only if $\{G_i\}_{i\in I}$ is q-convergent.

Note that the fact that $\{G_i\}_{i\in I}$ is a sequence of bounded degree graphs we may ask whether it converges in the Benjamini-Schramm sense [2]; in this case, by an inclusion-exclusion argument, the convergence considered in Proposition 5 is equivalent to the local-convergent considered by Benjamini-Schramm [2]. In this case, the loops ensures a bare minimum of homomorphisms for each subgraph. Therefore, Proposition 5 claims that, for bounded-degree graph sequences, Benjamini-Schramm convergence is equivalent to q-convergent.

Altoghether, Proposition 5 and Proposition 4 shows that the notion of q-convergence extends the notion of convergence for the limits of graphs in the dense case [9], and the notion of convergence for in the case of sequences of bounded degree graphs considered by Benjamini and Schramm [2] into a single, uniform framework. Here we should note that asking for the constants $c_F > 0$ in the dense case Proposition 4 is rather natural, as there are many sequences that are convergent in the dense case and that, for instance, have *no* triangles (or copies of K_3), to the same limit, yet there are several subsequences with different growth ratios of triangles, and thus the q-convergence will distinguish between the two subsequences. The q-convergence may distinguish different sequences that the dense notion considers to be equivalent; this is rather a natural behaviour since we want to distinguish between sparse sequences that the dense notion of convergence maps to the zero graphon [9]. The presence of Proposition 3 allows to claim reasonable compactification properties for the set of q-convergent sequences of graphs.

2 Sketch of the arguments

Let us sketch the proof of Theorem 1 (the key difference in Theorem 2 is highlighted below). The (positive) real numbers $\mathbb{R}^{>0}$ have the cardinality of the continuum, the set of sequences of positive reals numbers $\prod_{i\in\omega}\mathbb{R}^{>0}$ has the cardinality of the continuum, since $|\prod_{i\in\omega}\mathbb{R}^{>0}| = (\mathfrak{c})^{\aleph_0} = (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0\aleph_0} = 2^{\aleph_0}$, and the set of ordered injections $\mathcal{I} = \{\iota \mid \iota : \omega \to \omega \text{ ordered injection}\}$ has the cardinality of the continuum as well (note that these injections are a subset of all possible subsets of the natural numbers). Therefore, $[\prod_{i\in\omega}\mathbb{R}^{>0}] \times \mathcal{I}$ the set of infinite subsequences of positive real numbers has the cardinality of the continuum. Assuming the continuum hypothesis and the axiom of choice, we can well-order $[\prod_{i\in\omega}\mathbb{R}^{>0}]\times\mathcal{I}$ and biject it with the countable ordinals (those ordinals $<\omega_1$). Thus we write $[\prod_{i\in\omega}\mathbb{R}^{>0}]\times\mathcal{I}=\{(f_\alpha,\iota_\alpha)\mid\alpha<\omega_1\}$ and use a transfinite induction along ω_1 to find sets $\{A_\alpha\,|\,\alpha\leq\omega_1\}$ that have the desired properties (the sequences of A_α are pairwise comparable, and each subsequence (f_β,ι_β), with $\beta<\alpha$ has a representative in A_α). The sets A_α are built as $A_0=\emptyset$ (or $A_0=\{f_0,f_1\}$ if we want to show Theorem 2), $A_\alpha=\cup_{\beta<\alpha}A_\beta$ for limit ordinals, and where $A_{\alpha+1}$ is build from A_α by adding a new sequence agreeing uppon (f_α,ι_α) along a subsequence and comparable with the others in A_α; to find this new sequence we first examine whether there is already a test sequence in A_α that agrees with (f_α,ι_α) along a subsequence (up to a multiplicative c), if that is the case, then $A_{\alpha+1}=A_\alpha$. If that is not the case, then we can partition the at most countable (here we are using the continuum hypothesis again) elements in A_α in two parts U and D, and find a sequence of subsequences $\{\iota_{\alpha,i}\}_{i<\omega}$ ($\iota_{\alpha,i}$ subsequence of $\iota_{\alpha,i+1}$) for (f_α,ι_α) such that (f_α,ι_α) is below g_i along $\iota_{\alpha,i}$ if $g_i\in A_\alpha$ belongs to U, and is above g_j along $\iota_{\alpha,j}$ if $g_j\in A_\alpha$ belongs to D. Then we find a subsequence $\iota_{\alpha,\omega}$ of (f_α,ι_α) along which (f_α,ι_α) is below each element from A_α in U and above each element in D. Finally, we complete the subsequence along $\iota_{\alpha,\omega}$ into a full sequence between the elements of U and the elments of D by backwards extending the elements along the subsequence $\iota_{\alpha,\omega}$ with elements between the lowest found elements in U and the highest found elements in D (the current element of the subsequence $\iota_{\alpha,\omega}$ ensures that the multiplicative distance to all the previously considered elements in U and D goes to 0 and ∞ respectively). The transfinite induction then gives A_{ω_1}.

Proposition 3 is proven finding an appropriate triple (subsequence, constant, test function) for each finite graph. These test functions capture the asympthotic growth f_F for each subgraph F, consistently along an increasing family of subgraphs F by considering further subsequences of the graphs $\{G_i\}$. Then we use a diagonal argument, similar as before, to find a subsequence of graphs $\{G_i\}$ that, for each F, the number of homomorphisms from F to $\{G_i\}$ has, along the subsequence of graphs $\{G_i\}$, the asympthotic growth f_F up to a multiplicative constant c_F.

Proposition 5 and Proposition 4 follows by observing that, with their respective hypotheses, the test functions at which the subgraph count of a convergent graph sequence grow are, up to a multiplicative factor, the minimum and maximum possible (given that the graphs have loops on each vertex).

References

[1] Á. Backhausz and B. Szegedy. Action convergence of operators and graphs. *Canadian Journal of Mathematics*, 74(1):72–121, 2022.

[2] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13 pp. (electronic), 2001.

[3] C. Borgs, J. Chayes, H. Cohn, and Y. Zhao. An L^p theory of sparse graph convergence I: Limits, sparse random graph models, and power law distributions. *Transactions of the American Mathematical Society*, 372(5):3019–3062, 2019.

[4] C. Borgs, J. Chayes, L. Lovász, V. T. Sós, B. Szegedy, and K. Vesztergombi. Graph limits and parameter testing. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 261–270, 2006.

[5] C. Borgs, J. T. Chayes, H. Cohn, and Y. Zhao. An L^p theory of sparse graph convergence II: LD convergence, quotients and right convergence. *The Annals of Probability*, 46(1):337 – 396, 2018.

[6] P. Frenkel. Convergence of graphs with intermediate density. *Transactions of the American Mathematical Society*, 370(5):3363–3404, 2018.

[7] H. Hatami, L. Lovász, and B. Szegedy. Limits of locally–globally convergent graph sequences. *Geometric and Functional Analysis*, 24(1):269–296, 2014.

[8] C. Hoppen, Y. Kohayakawa, C. G. Moreira, B. Ráth, and R. Menezes Sampaio. Limits of permutation sequences. *Journal of Combinatorial Theory, Series B*, 103(1):93–113, 2013.

[9] L. Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.

[10] J. Nešetřil and P. Ossona De Mendez. A model theory approach to structural limits. *Commentationes Mathematicae Universitatis Carolinae*, 53(4):581–603, 2012.

[11] B. Szegedy. Sparse graph limits, entropy maximization and transitive graphs. *arXiv preprint arXiv:1504.00858*, 2015.

[12] L. Vena. On limits of sparse random graphs. *Electronic Notes in Discrete Mathematics*, 54:343–348, 2016. Discrete Mathematics Days - JMDA16.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

A DIRECT BIJECTION BETWEEN TWO-STACK SORTABLE PERMUTATIONS AND FIGHTING FISH

(EXTENDED ABSTRACT)

Lapo CIONI* Luca FERRARI* Corentin HENRIET*

Abstract

We define a bijection between *two-stack sortable permutations* and *fighting fish*, enriching the garden of bijections linking the numerous combinatorial classes counted by the sequence $A000139$ of the OEIS. Our bijection is (up to symmetry) the non-recursive version of the one of Fang (2018). Along the way, we encounter *labeled sorting trees*, a new class of trees that appear to have nice properties that seem worth to explore.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-039`

1 Introduction

The problem of sorting a permutation through a stack was addressed by Knuth in his seminal work [9] (Section 2.2.1), initiating the development of the field of pattern-avoiding permutations. While one-stack sortable permutations are characterized by their avoidance of the pattern 231 and counted by the Catalan numbers, West [11] established the pattern-avoidance characterization of two-stack sortable permutations, but did not succeed in proving that the number of such permutations of size n is given by the nice formula $\frac{2}{(n+1)(2n+1)}\binom{3n}{n}$, leaving it as a conjecture. This formula was first proved by Zeilberger [12] using generating functions, then refined according to certain statistics by Bousquet-Mélou [1]. Other proofs of this enumeration were later found by Dulucq, Gire and Guibert [6], by Goulden and West [8] establishing bijections between two-stack sortable permutations and nonseparable planar maps using either generating trees or recursive decompositions. Hence

*Dipartimento di Matematica e Informatica 'Ulisse Dini', Università degli Studi di Firenze

two combinatorial explanations of the enumeration formula were found, but they were still not completely satisfying because recursive. More recently, Fang [7] used another recursive decomposition of two-stack sortable permutations to define a bijection with fighting fish, a generalization of parallelogram polyominoes introduced in 2016 by Duchi, Guerrini, Rinaldi and Schaeffer [2]. The main contribution of the present paper is a direct description of the bijection of Fang using a particular class of trees that we call labeled sorting trees, which, combined with the encoding of nonseparable planar maps by fighting fish given by Duchi and the third author [4], leads to a direct bijective path from two-stack sortable permutations to nonseparable planar maps. This is one step further in the bijective understanding of the connections between combinatorial structures enumerated by [10, A000139]: we refer to the first page of [4] for a diagram summarizing known bijections.

The paper is organized as follows: we first present the objects involved in our bijection (two-stack sortable permutations and fighting fish), we then describe our bijection going through labeled sorting trees, and we finish by an overview of open questions arising naturally in our work.

2 Preliminaries

2.1 Stack sorting on permutations

Let τ be a word made of distinct positive integers. If τ is empty, we define $\mathbf{S}(\tau) = \tau$. Else, if we denote by k the largest letter of τ, we can write $\tau = \tau_1 k \tau_2$, and we then define $\mathbf{S}(\tau) = \mathbf{S}(\tau_1)\mathbf{S}(\tau_2)k$. A permutation $\sigma \in \mathfrak{S}_n$ is identified with its one-line notation $\sigma_1...\sigma_n$, where we write σ_i for $\sigma(i)$. Restricted to permutations, $\mathbf{S}$ is said to be the operator of *stack-sorting*. Stack-sorting deserves its name because $\mathbf{S}(\sigma)$ is the word τ obtained from σ via the following procedure on a stack constrained to be decreasing from bottom to top. Initialize the stack to be empty and τ to be the empty word, and at each step, consider the smallest i such that σ_i has not yet been put in the stack : if the stack is empty or if the top element of the stack is greater than σ_i then put σ_i on the top of the stack, otherwise pop the top element of the stack and append it to τ. If all elements of σ have been treated, then pop out the top of the stack and append it to τ. The procedure ends when all elements of σ have been treated and the stack is empty. We can encode this procedure by a Dyck path : we add an up step $u = (1,1)$ to the path each time we put an element on the top of the stack, and a down step $d = (1,-1)$ each time we pop out the top element of the stack. We denote by $\mathrm{D}(\sigma)$ the Dyck path obtained for a permutation $\sigma \in \mathfrak{S}_n$: it is a path from $(0,0)$ to $(2n,0)$ staying weakly above the x-axis. For $k \geq 1$, we define $\mathrm{D}_k(\sigma) = \mathrm{D}(\mathbf{S}^{k-1}(\sigma))$.

For $k \geq 0$, a permutation $\sigma \in \mathfrak{S}_n$ is said to be *k-stack sortable* if $\mathbf{S}^k(\sigma)$ is the identity. A k-stack sortable permutation σ is uniquely determined by the tuple $(\mathrm{D}_1(\sigma), ..., \mathrm{D}_k(\sigma))$ because σ can be recovered from the identity by reverting the stack-sorting process encoded by this tuple of Dyck paths. We denote by $2\mathcal{SS}_n$ the set of two-stack sortable permutations of $\mathfrak{S}_n$, and $2\mathcal{SS} = \bigcup_{n\geq 1} 2\mathcal{SS}_n$.

2.2 Fighting fish

While fighting fish have been introduced in [2] in terms of gluings of square cells, we present them here as words on the alphabet $\{E, N, W, S\}$ (see Figure 1):

Definition 1. *A word $w \in \{E, N, W, S\}^*$ is a* fighting fish *if it can be obtained from the word $ENWS$ using a finite sequence of the following 3 operations:*

- *Upper gluing : replace a subword W by NWS.*
- *Lower gluing : replace a subword N by ENW.*
- *Double gluing : replace a subword WN by NW.*

The size of a fighting fish is half of its length minus 1. We denote by $\mathcal{FF}_n$ the set of fighting fish of size n, and by $\mathcal{FF} = \bigcup\limits_{n\geq 1} \mathcal{FF}_n$.

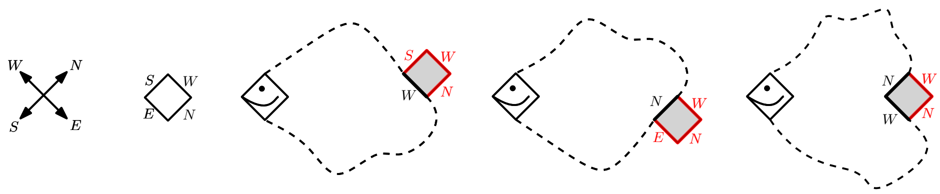


Figure 1: The tilted cardinal steps, a cell and operations of upper, lower and double gluing.

There is a natural notion of symmetry on fighting fish: if F is a fighting fish, its *conjugate*, denoted $\mathrm{Conj}^{\mathcal{FF}}(\mathrm{F})$, is the fighting fish obtained by reversing F and changing its letters with the rules $E \leftrightarrow S$, $W \leftrightarrow N$. Conjugation is an involution on fighting fish, that is $\mathrm{Conj}^{\mathcal{FF}} \circ \mathrm{Conj}^{\mathcal{FF}}$ is the identity on $\mathcal{FF}$, and we can see it as the symmetry with respect to the horizontal axis on our two-dimensional pictures.

3 The bijection from two-stack sortable permutations to fighting fish

3.1 From permutations to labeled sorting trees

Let $\sigma \in \mathfrak{S}_n$ be a permutation, and let us define $\hat{\sigma} \in \mathfrak{S}_{n+1}$ by setting $\hat{\sigma}(1) = n + 1$ and $\hat{\sigma}(i) = \sigma(i-1)$ for $2 \leq i \leq n+1$. We represent $\hat{\sigma}$ as the set of points $\{(i, \hat{\sigma}(i))\}$ in $\mathbb{Z}^2$ (its *grid representation*) and we construct a rooted plane tree on this set of points. The *sorting tree* $\mathrm{ST}(\sigma)$ associated to σ is the rooted plane tree obtained by the following top-to-bottom process :

- Define the root to be $(1, n+1)$.
- At step $j \geq 1$, we insert the point $(k, \hat{\sigma}(k))$ in the tree, where k is such that $\hat{\sigma}(k) = n - j$. To do so, let us consider $0 = i_1 < i_2 < \ldots < i_j$ the x-coordinates of all points already inserted in the tree. There is then a maximal index m such that $i_m < k$. We distinguish two cases :
  - If $m = j$ or $\hat{\sigma}(i_m) < \hat{\sigma}(i_{m+1})$, we define the parent of $(k, \hat{\sigma}(k))$ to be $(i_m, \hat{\sigma}(i_m))$.
  - If $m < j$ and $\hat{\sigma}(i_m) > \hat{\sigma}(i_{m+1})$, we consider the greatest $m + 1 \leq r \leq n$ such that $\hat{\sigma}(i_m) > \hat{\sigma}(i_{m+1}) > \ldots > \hat{\sigma}(i_r)$, and we set the parent of $(k, \hat{\sigma}(k))$ to be $(i_r, \hat{\sigma}(i_r))$.
- The process ends when all points have been inserted, i.e. after the n^{th} step.

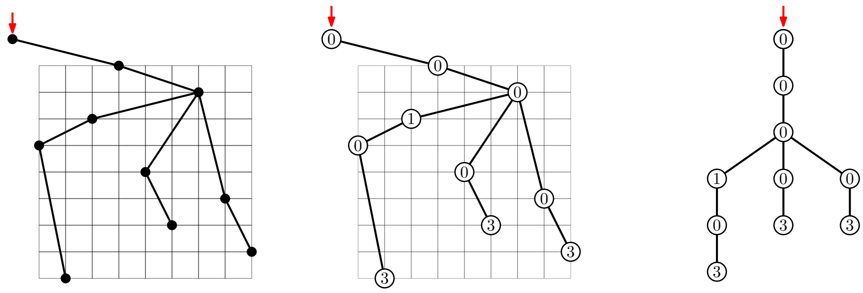


Figure 2: The permutation 617953842, its sorting tree and its labeled sorting tree

This procedure produces a tree since n edges are inserted and no cycles are created, because each non-root vertex has a parent of strictly greater y-coordinate. The permutation $\hat{\sigma}$ can be split into maximal descending runs in a unique way. We associate to every element of $\hat{\sigma}$ its *run label* in the following way: if it is not the last element of its descending run, we label it by 0, else we label it by the number of elements in its descending run. The *labeled sorting tree* $\mathrm{LST}(\sigma)$ associated to σ is the plane rooted tree obtained by labeling each node of $\mathrm{ST}(\sigma)$ with the run label of the element to which it corresponds in the permutation. As an example, we give the (labeled) sorting tree of the permutation 617953842 in Figure 2.

Sorting trees and labeled sorting trees of permutations are intimately linked to the Dyck paths corresponding to their first two passes into a stack:

Proposition 1. *Let $\sigma, \tau \in \mathfrak{S}_n$. Then:*

- $\mathrm{ST}(\sigma) = \mathrm{ST}(\tau)$ *if and only if* $\mathrm{D}_2(\sigma) = \mathrm{D}_2(\tau)$.
- $\mathrm{LST}(\sigma) = \mathrm{LST}(\tau)$ *if and only if* $(\mathrm{D}_1(\sigma), \mathrm{D}_2(\sigma)) = (\mathrm{D}_1(\tau), \mathrm{D}_2(\tau))$.

In particular, when $\sigma, \tau \in 2\mathcal{SS}_n$, $\mathrm{ST}(\sigma) = \mathrm{ST}(\tau)$ *if and only if* $\mathbf{S}(\sigma) = \mathbf{S}(\tau)$.

The proof of the next proposition, that we do not present here, relies mainly on the decomposition of two-stack sortable permutations presented in [7] and on its isomorphic counterpart on labeled sorting trees, transferred by LST:

Proposition 2. *Let T be a rooted labeled plane tree with root r, and having n non-root vertices (we say that it has* size *n). For a given node $v \in T$, we denote by $\lambda(v)$ its nonnegative label,* $\deg(v)$ *its number of children,* $\mathrm{sub}(v)$ *the subtree of T rooted at v and* $\mathrm{anc}(v)$ *the nodes w such that v belongs to* $\mathrm{sub}(w)$ *(the ancestors of v).*
Then there exists a permutation $\sigma \in \mathfrak{S}_n$ such that $\mathrm{LST}(\sigma) = T$ *if and only if:*

$$\begin{cases} \sum_{v \in T} \lambda(v) = n + 1 \\ \forall v \in T, \lambda(v) \leq \sum_{w \in \mathrm{anc}(v)} \big(2 - \deg(w)\big) - 1 \\ \forall v \in T \setminus \{r\}, \sum_{w \in \mathrm{sub}(v)} \big(\lambda(w) - 1\big) \geq 1 \end{cases}$$

Furthermore, denoting by $\mathcal{LST}_n$ the set of trees satisfying these three conditions, the restriction of the map LST *to two-stack sortable permutations is a bijection between $2\mathcal{SS}_n$ and $\mathcal{LST}_n$.*

Since $\mathrm{LST}(\mathfrak{S}_n) = \mathcal{LST}_n$, we can call trees in $\mathcal{LST} = \bigcup_{n \geq 1} \mathcal{LST}_n$ *labeled sorting trees.*

3.2 From labeled sorting trees to fighting fish

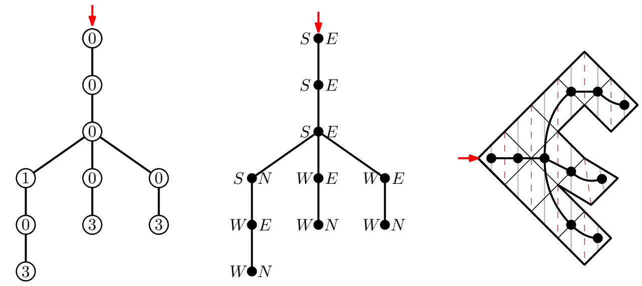


Figure 3: A tree in $\mathcal{LST}_9$, its fish word displayed on the tree and the corresponding fighting fish.

Let T be a tree in $\mathcal{LST}_n$. We build a word w on the alphabet $\{E, N, W, S\}$ with the following algorithm using a stack :

- Set w to be the empty word and the stack to be empty, and run a clockwise tour of the tree T, starting from the root.
- Every time we encounter a vertex v for the first time, we read its label $\lambda(v)$: if $\lambda(v) = 0$, then we put nothing in the stack and append E to w, else $\lambda(v) > 0$ and we insert (in this order) a letter S and $\lambda(v) - 1$ letters W in the stack and append N to w.
- Every time we encounter a vertex for the last time, we pop the top element of the stack and append it to w.

- The algorithm ends after we reach the root vertex for a second (and last) time.

The word w obtained via this procedure is called the *fish word* of T and we denote it by $\mathrm{FW}(T)$. We provide an example in Figure 3. It is not straightforward that the stack is always not empty when we have to pop an element of the stack, but the conditions on labels of labeled sorting trees ensure this property.

Proposition 3. *The map* $\mathrm{FW} : \mathcal{LST} \to \mathcal{FF}$ *is a bijection preserving the size.*

Our proof of the proposition relies on the isomorphic decompositions of labeled sorting trees and of fighting fish (the wasp-waist decomposition presented in [3]) transferred by FW. Combining the maps LST and FW, we then get (up to conjugation) a direct description of the recursive bijection of Fang (see [7]) between two-stack sortable permutations and fighting fish:

Theorem 1. *The map* $\mathrm{FW} \circ \mathrm{LST} : 2\mathcal{SS} \to \mathcal{FF}$ *is a bijection that sends two-stack sortable permutations of size* n *to fighting fish of size* n.

4 Perspectives

- Let $T \in \mathcal{LST}_n$ a labeled sorting tree. Proposition 2 states that there is a unique two-stack sortable permutation σ such that $\mathrm{LST}(\sigma) = T$. It raises the natural question of enumerating the set $\{\tau \in \mathfrak{S}_n, \mathrm{LST}(\tau) = T\}$. An equivalent problem is to determine how many permutations $\tau \in \mathfrak{S}_n$ satisfy $(\mathrm{D}_1(\tau), \mathrm{D}_2(\tau)) = (\mathrm{D}_1(\sigma), \mathrm{D}_2(\sigma))$.
- The conditions of the characterization of trees belonging to $\mathcal{LST}_n$ in Proposition 2 do not depend on the order of the children of a given vertex. In particular, the mirror tree (obtained by recursively reversing the order of the children of the nodes) of a tree in $\mathcal{LST}_n$ is also in $\mathcal{LST}_n$. This symmetry is surprising if we consider that $\mathcal{LST}_n$ is the set of labeled sorting trees of two-stack sortable permutations. On the other hand, it is also not evident that the algorithm defining the bijection FW still produces a valid fighting fish when considering the mirror tree of a tree in $\mathcal{LST}_n$. It might be interesting to investigate the involutions on two-stack sortable permutations and on fighting fish induced by this mirror involution on labeled sorting trees.
- Similarly, conjugation on fighting fish gives rise to involutions on labeled sorting trees and on two-stack sortable permutations. It would be nice to have a direct description of these involutions, since the counterpart of conjugation on nonseparable planar maps is the important notion of duality (see [4]).
- A natural statistic to consider on fighting fish is the *area*: it is the number of square cells composing it (or equivalently the area enclosed by the corresponding quadrant walk). This statistic has an interesting counterpart on synchronized Tamari intervals (which are bijectively linked to fighting fish in [5]), generalizing the notion of area on Dyck paths. We tried to have a nice and direct interpretation on (two-stack sortable) permutations of the area statistic on fighting fish (transferred by $\mathrm{FW} \circ \mathrm{LST}$), without success. Still, we were able to characterize the permutations giving rise to a fighting fish of minimal area in terms

of pattern avoidance. We also conjecture that they are enumerated according to their size by the sequence [10, A131178].

References

[1] M. Bousquet-Mélou. Multi-statistic enumeration of two-stack sortable permutations. *The Electronic Journal of Combinatorics*, 5(1):R21, 1998.

[2] E. Duchi, V. Guerrini, S. Rinaldi, and G. Schaeffer. Fighting Fish. *Journal of Physics A : Mathematical and Theoretical*, 50.2, 2016.

[3] E. Duchi, V. Guerrini, S. Rinaldi, and G. Schaeffer. Fighting Fish: Enumerative Properties. In *29th International Conference on "Formal Power Series and Algebraic Combinatorics" (FPSAC 2017)*. Séminaire Lotharingien de Combinatoire 78B.43, 2017.

[4] E. Duchi and C. Henriet. A bijection between rooted planar maps and generalized fighting fish, 2022.

[5] E. Duchi and C. Henriet. A bijection between Tamari intervals and extended fighting fish. *European Journal of Combinatorics*, 110:103698, 2023.

[6] S. Dulucq, S. Gire, and O. Guibert. A combinatorial proof of J. West's conjecture. *Discrete Mathematics*, 187(1):71–96, 1998.

[7] W. Fang. Fighting Fish and Two-Stack Sortable Permutations. In *30th International Conference on "Formal Power Series and Algebraic Combinatorics" (FPSAC 2018)*. Séminaire Lotharingien de Combinatoire 80B.7, 2018.

[8] I. P. Goulden and J. West. Raney Paths and a Combinatorial Relationship between Rooted Nonseparable Planar Maps and Two-Stack-Sortable Permutations. *Journal of Combinatorial Theory, Series A*, 75(2):220–242, 1996.

[9] D. E. Knuth. *The Art of Computer Programming, Vol. 1: Fundamental Algorithms*. Addison-Wesley, Reading, Mass., third edition, 1997.

[10] OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences. Published electronically at `http://oeis.org`.

[11] J. West. Sorting twice through a stack. *Theoretical Computer Science*, 117(1):303–313, 1993.

[12] D. Zeilberger. A proof of Julian West's conjecture that the number of two-stack-sortable permutations of length n is $2(3n)!/((n+1)!(2n+1)!)$. *Discrete Mathematics*, 102(1):85–93, 1992.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Counting tournament score sequences

(Extended abstract)

Anders Claesson* Mark Dukes† Atli Fannar Franklín‡
Sigurður Örn Stefánsson§

Abstract

The score sequence of a tournament is the sequence of the out-degrees of its vertices arranged in nondecreasing order. The problem of counting score sequences of a tournament with n vertices is more than 100 years old (MacMahon 1920). In 2013 Hanna conjectured a surprising and elegant recursion for these numbers. We settle this conjecture in the affirmative by showing that it is a corollary to our main theorem, which is a factorization of the generating function for score sequences with a distinguished index. We also derive a closed formula and a quadratic time algorithm for counting score sequences.

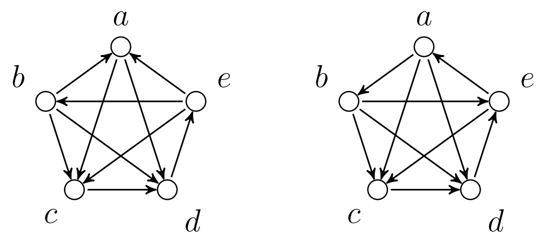
DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-040`

1 Introduction

This extended abstract summarises the results of our paper [4]. In 1953 Landau [9] used oriented complete graphs—also called *tournaments*—to model pecking orders. If the vertices of the complete graph represent players (rather than chickens), then the initial vertex of a directed edge signifies the winner of a game between the two end-point players. The number of wins of a player is equal to the number of outgoing edges from that vertex. A *score sequence* is a sequence of these number of wins given in a nondecreasing order. For instance, with 3 players there are two possible score sequences, namely $(0, 1, 2)$ and

*Science Institute, University of Iceland, Iceland. Email: akc@hi.is.
†School of Mathematics & Statistics, University College Dublin, Ireland. E-mail: mark.dukes@ucd.ie.
‡Science Institute, University of Iceland, Iceland. Email: aff6@hi.is.
§Science Institute, University of Iceland, Iceland. Email: sigurdur@hi.is.

$(1,1,1)$. Note that non-isomorphic tournaments may give rise to the same score sequence. With 5 players there are, up to isomorphism, 12 tournaments but only 9 score sequences. To be even more specific, here are two non-isomorphic tournaments:



The score sequence associated with both is $(1,1,2,3,3)$. The following characterization of score sequences is known as Landau's theorem.

Theorem 1 (Landau [9]). *A sequence of integers $s = (s_0, \dots, s_{n-1})$ is a score sequence if and only if*

(1) $0 \le s_0 \le s_1 \le \cdots \le s_{n-1} \le n-1$,

(2) $s_0 + \cdots + s_{k-1} \ge \binom{k}{2}$ *for* $1 \le k < n$, *and*

(3) $s_0 + \cdots + s_{n-1} = \binom{n}{2}$.

Let S_n be the set of score sequences of length n. There is no known closed formula for the associated cardinalities (A000571 in the OEIS [7])

$$\big(|S_n|\big)_{n\ge 0} = (1,1,1,2,4,9,22,59,167,490,1486,4639,14805,\dots)$$

or their generating function.

It should be noted that Landau was not the first person to study score sequences, or attempt to count them. MacMahon [10] used symmetric functions and hand calculations to determine $|S_n|$ for $n \le 9$ in 1920. Building on Landau's work, Narayana and Bent [11], in 1964, derived a multivariate recursive formula for determining $|S_n|$. They used it to give a table for $n \le 36$. In 1968 Riordan [12] gave a simpler and more efficient recursion, but unfortunately it turned out to be incorrect [13].

Let $[a,b]$ denote the interval of integers $\{a, a+1, \dots, b\}$. We may view a score sequence $s \in S_n$ as an endofunction $s : [0, n-1] \to [0, n-1]$. We now introduce the notion of a *pointed score sequence*. Define $S_n^\bullet$ as the Cartesian product $S_n^\bullet = S_n \times [0, n-1]$. We call the members of $S_n^\bullet$ *pointed score sequences*; e.g. there are 6 pointed score sequences in $S_3^\bullet$:

$$((0,1,2),0),\ ((0,1,2),1),\ ((0,1,2),2),$$
$$((1,1,1),0),\ ((1,1,1),1),\ ((1,1,1),2).$$

Let $(s,i) \in S_n^\bullet$. Depending on the context, the element i will be interpreted as a position (element in the domain) or a value (element in the codomain) of s. If i is a value, then the cardinality of the fiber $s^{-1}(i)$ is the number of times i occurs in s; this number may be zero. Let

$$S_n^\bullet(t) = \sum_{(s,i)\in S_n^\bullet} t^{|s^{-1}(i)|}$$

be the polynomial recording the distribution of the statistic $(s,i) \mapsto |s^{-1}(i)|$ on $S_n^\bullet$. As an example, $S_3^\bullet(t) = 2 + 3t + t^3$. Let

$$S^\bullet(x,t) = \sum_{n\geq 1} S_n^\bullet(t)x^n.$$

To present the bijection that is the main result of this paper, we will first introduce a particular type of multiset that is an essential ingredient in our deconstruction of a pointed score sequence. At first glance it is not obvious what the relevance of these multisets to score sequences is.

We define EGZ_n as the set of multisets of size n with elements in the cyclic group $\mathbb{Z}_n$ whose sum is $\binom{n}{2}$ modulo n. To understand what the elements of EGZ_n look like it may be helpful to note that $\binom{n}{2}$, as an element of $\mathbb{Z}_n$, is 0 if n is odd and $n/2$ if n is even. For instance, EGZ_3 consists of the 4 multisets $\{0,0,0\}$, $\{0,1,2\}$, $\{1,1,1\}$, and $\{2,2,2\}$.

The notation EGZ_n refers to the Erdős-Ginzburg-Ziv Theorem [5], which is stated below. Following it we give a proposition motivating this terminology; its proof gives a simple one-to-one correspondence between EGZ_n and the sets considered by Erdős, Ginzburg, and Ziv.

Theorem 2 (Erdős, Ginzburg, and Ziv [5])**.** *Each set of $2n-1$ integers contains some subset of n elements the sum of which is a multiple of n.*

Proposition 3. *There is a one-to-one correspondence between EGZ_n and n-element subsets of $[1, 2n-1]$ whose sum is a multiple of n.*

Proof. Let $A = \{a_1, \dots, a_n\}$ be a subset of $[1, 2n-1]$ such that $a_1 + \cdots + a_n$ is divisible by n. Without loss of generality we can further assume that $a_1 < a_2 < \cdots < a_n$. Let $b_i = a_i - i$. The mapping $A \mapsto \{b_1, \dots, b_n\}$ is a bijection onto EGZ_n. Further proof details are omitted but can be found in [4], as with other results presented in this abstract. □

The sequence of cardinalities

$$\big(|\mathrm{EGZ}_n|\big)_{n\geq 1} = (1, 1, 4, 9, 26, 76, 246, 809, 2704, 9226, 32066, \dots)$$

is entry A145855 in the OEIS [7]. As recorded in that OEIS entry, Jovović conjectured and Alekseyev [1] proved in 2008 that

$$|\mathrm{EGZ}_n| = \frac{1}{2n}\sum_{d|n}(-1)^{n-d}\varphi(n/d)\binom{2d}{d}, \tag{4}$$

where the sum runs over all positive divisors of n and φ is Euler's totient function. A generalization of this result was given by Chern [3] in 2019.

The zeros in a multiset $M \in \mathrm{EGZ}_n$ play a prominent role in our construction. We now introduce a generating function to record their number. For a multiset $M \in \mathrm{EGZ}_n$ let $|M|_i$ be the number of occurrences of i in M. Furthermore, let

$$\mathrm{EGZ}_n(t) = \sum_{M \in \mathrm{EGZ}_n} t^{|M|_0}$$

be the polynomial recording the distribution of zeros in multisets belonging to EGZ_n. For instance, $\mathrm{EGZ}_3(t) = 2 + t + t^3$ (looking at the distribution of 1s or 2s in EGZ_3 would result in the same polynomial). Define the generating functions

$$\mathrm{EGZ}(x,t) = \sum_{n \geq 1} \mathrm{EGZ}_n(t) x^n \quad \text{and} \quad S(x) = \sum_{n \geq 0} |S_n| x^n.$$

Our main result, Theorem 4, is a factorization of the generating function for pointed score sequences:

$$S^\bullet(x,t) = \mathrm{EGZ}(x,t) S(x). \tag{5}$$

Let $(s,i) \in S_n^\bullet$. Viewing i is an element of the codomain of s we find that $S^\bullet(x,0)$ consists of terms stemming from pairs (s,i) such that $s^{-1}(i)$ is empty; i.e. i is outside the image of s. Thus, $S^\bullet(x,1) - S^\bullet(x,0)$ counts pairs (s,i) for which i is in the image of s. Let

$$S_n^\circ = \{(s,i) \in S_n^\bullet : i \in \mathrm{Im}(s)\} = \{(s,i) \in S_n^\bullet : i = s_j \text{ for some } j \in [n]\}$$

and let $S^\circ(x) = S^\bullet(x,1) - S^\bullet(x,0)$ be the corresponding generating function. For instance, S_3° consists of the 4 elements $((0,1,2),0)$, $((0,1,2),1)$, $((0,1,2),2)$, and $((1,1,1),1)$. We will show (in Corollary 6) that $S^\circ(x) = xC(x)S(x)$, where $C(x) = (1 - \sqrt{1-4x})/(2x)$ is the generating function for the Catalan numbers $C_n = \binom{2n}{n}/(1+n)$. This striking occurrence of the Catalan numbers was in fact the original inspiration for our work. It was in the summer of 2019 that we experimented with score sequences and conjectured the identity. Despite ample attempts we were for the longest time unable to prove it.

By setting $t = 1$ in Equation 5 and noting that $S^\bullet(x,1) = xS'(x)$ it follows that

$$xS'(x) = \mathrm{EGZ}(x,1) S(x), \tag{6}$$

a fact conjectured by Paul D. Hanna as recorded in the OEIS entry A000571 in 2013. Equation 6 may alternatively be written $\big(\log S(x)\big)' = \mathrm{EGZ}(x,1)/x$ and so

$$S(x) \;=\; \exp\!\left(\sum_{n \geq 1} \frac{|\mathrm{EGZ}_n|}{n} x^n \right),$$

which arguably is the most elegant way of expressing the relation between $|S_n|$ and $|\mathrm{EGZ}_n|$. The most efficient way of computing the numbers $|S_n|$ is, however, to use the recursion underlying Equation 6. Namely, $|S_0| = 1$ and, for $n \geq 1$,

$$|S_n| = \frac{1}{n} \sum_{k=1}^{n} |S_{n-k}| |\mathrm{EGZ}_k|.$$

See Corollary 8 and the discussion following it.

2 The main theorem and its bijection

Let the generating functions $S^{\bullet}(x,t)$, $\mathrm{EGZ}(x,t)$ and $S(x)$ be defined as in Section 1.

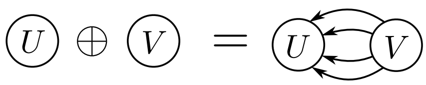
Theorem 4. *We have* $S^{\bullet}(x,t) = \mathrm{EGZ}(x,t)S(x)$.

The proof of Theorem 4 is combinatorial and is achieved by creating a bijection

$$\Phi : S_n^{\bullet} \to \bigcup_{k=1}^{n} \mathrm{EGZ}_k \times S_{n-k}$$

that maps a pointed score sequence to a pair consisting of a multiset and a score sequence. A property of this bijection is that, for $(M,v) = \Phi(s,i)$, the number of occurrences of i in s is equal to the multiplicity of zero in M. Before defining Φ we need to introduce several necessary concepts.

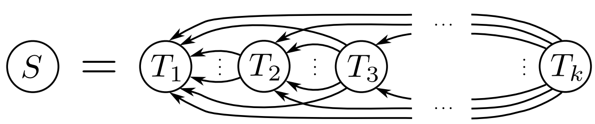
A nonempty directed graph is said to be *strongly connected* if there is a directed path between each pair of vertices of the graph. Note that we do not consider the empty graph to be strongly connected. A *strong score sequence* is one which stems from a strongly connected tournament. Equivalently (see Harary and Moser [6, Theorem 9]), $s = (s_0,\ldots,s_{n-1})$, with $n \geq 1$, is a strong score sequence if the inequality (2) of Theorem 1 is always strict; that is, $s_0 + \cdots + s_{k-1} > \binom{k}{2}$ for $1 \leq k < n$. Let us define the *direct sum* of two score sequences $u \in S_k$ and $v \in S_\ell$ by $u \oplus v = uv'$, where v' is obtained from v by adding k to each of its letters and juxtaposition indicates concatenation. For instance, $(0) \oplus (0) \oplus (1,1,1) = (0,1,3,3,3)$. If U and V are tournaments having score sequences u and v, one may view the direct sum $u \oplus v$ as the score sequence of the tournament where arrows are placed between the vertices of U and V such that they all point towards U:



This may easily be seen to be independent of the choice of tournaments.

Lemma 5. *Let* $s \in S_n$. *If* $s_0 + \cdots + s_{k-1} = \binom{k}{2}$ *for some* $k < n$, *then* $u = (s_0,\ldots,s_{k-1})$ *and* $v = (s_k - k,\ldots,s_{n-1} - k)$ *are both score sequences, and* $s = u \oplus v$.

A direct consequence of Lemma 5 is that every score sequence s can be uniquely written as a direct sum $s = t_1 \oplus t_2 \oplus \cdots \oplus t_k$ of nonempty strong score sequences; in this context, the t_i will be called the *strong summands* of s. In terms of underlying tournaments we have the picture:



We are now almost in a position to define the promised map Φ, but first a couple of definitions. Assume that we are given a score sequence $s = (s_0, s_1, \ldots, s_{n-1}) \in S_n$.

- For any integer j, let $s+j$ denote the sequence obtained by adding j to each element of s, reducing modulo n, and sorting the outcome in nondecreasing order. Note that $s+j$ need not be a score sequence even though s is. E.g. $s=(1,1,1)$ is a score sequence, but $s+1=(2,2,2)$ is not. On the other hand, if $s=(0,1,2)$ then $s+1=s$ is a score sequence. A characterization of when $s+j$ is a score sequence is given in [4, Lemma 7].

- Let $\mu(s+j)$ denote the multiset $\{s_0+j, s_1+j, \ldots, s_{n-1}+j\}$ with elements in the cyclic group $\mathbb{Z}_n$.

Given a pointed score sequence $(s,i) \in S_n^\bullet$, write $s = t_1 \oplus t_2 \oplus \cdots \oplus t_k$ and let j be the smallest index such that $i < |t_1 \oplus \cdots \oplus t_j|$. Another way to define j is as the smallest prefix $t_1 \oplus \cdots \oplus t_j$ of strong summands of s that begins $s_0, s_1, \ldots, s_i$. Define the two score sequences u and v by

$$u = t_1 \oplus \cdots \oplus t_j \quad \text{and} \quad v = t_{j+1} \oplus \cdots \oplus t_k.$$

Finally, we let

$$\Phi(s,i) := \big(\mu(u-i), v\big).$$

As an example, consider the score sequence $s=(0,2,2,3,3,5,7,7,7)$; its decomposition into strong summands is $s=(0)\oplus(1,1,2,2)\oplus(0)\oplus(1,1,1)$. With $i=3$ we get $u = (0)\oplus(1,1,2,2) = (0,2,2,3,3)$, $v=(0)\oplus(1,1,1)=(0,2,2,2)$, $u-3=(0,0,2,4,4)$ and so $\Phi(s,3) = \big(\{0,0,2,4,4\},(0,2,2,2)\big)$.

Corollary 6. *We have $S^\circ(x) = xC(x)S(x)$, where $C(x)$ is the generating function for the Catalan numbers.*

Corollary 7. *We have $S(x) = \exp\left(\sum_{n\geq 1} |\mathrm{EGZ}_n| x^n/n\right)$.*

We end by comparing our result (Corollary 7) with earlier results on the enumeration of the *ordered* score sequences $(s_0, s_1, \ldots, s_{n-1})$, also called *score vectors*. That is, if G is a tournament on the vertex set $\{v_0, v_1, \ldots, v_{n-1}\}$, then s_i is the out-degree of v_i in G. For instance, while there are only two score sequences of length 3, namely $(1,1,1)$ and $(1,2,3)$, there are 7 score vectors of length 3: the vector $(1,1,1)$ together with the 6 permutations of $(1,2,3)$.

Stanley and Zaslavsky [14] have shown that the number of score vectors of length n equals the number of (labeled) forests on n nodes. A combinatorial proof was subsequently given by Kleitman and Winston [8]. Cayley [2] famously gave the formula n^{n-2} for the number of trees on n nodes. From the theory of exponential generating functions it immediately follows that $\exp\big(\sum_{n\geq 1} n^{n-2}x^n/n!\big)$ is the exponential generating function of forests, and thus also of score vectors.

3 The number of score sequences

If two power series $A(x) = 1 + \sum_{n\geq 1} a_n x^n$ and $B(x) = \sum_{n\geq 1} b_n x^n$ satisfy $xA'(x)/A(x) = B(x)$ and hence $\log A(x) = \sum_{n\geq 1} b_n x^n/n$, then one readily obtains a closed formula for a_n by expanding and identifying coefficients in $A(x) = \exp\bigl(b_1 x^1/1\bigr)\exp\bigl(b_2 x^2/2\bigr)\cdots$. Applying this to the equation in Corollary 7 we arrive at

$$|S_n| = \frac{1}{n!}\sum_{\pi\in\mathrm{Sym}(n)}\prod_{\ell\in C(\pi)}|\mathrm{EGZ}_\ell|, \tag{7}$$

where $\mathrm{Sym}(n)$ is the symmetric group of degree n and $C(\pi)$ encodes the cycle type of π; i.e. there is an $\ell \in C(\pi)$ for each ℓ-cycle of π. While having the virtue of being closed, this formula does not lend itself to quickly calculating $|S_n|$. For that purpose the following recursion is better suited.

Corollary 8. *For* $n \geq 1$, $|S_n| = \sum_{k=1}^{n}|S_{n-k}||\mathrm{EGZ}_k| = \dfrac{1}{n}\sum_{k=1}^{n}\dfrac{|S_{n-k}|}{2nk}\sum_{d|k}(-1)^{k-d}\varphi(k/d)\dbinom{2d}{d}$.

This allows us to calculate all values of $|S_k|$ for $k \leq n$ in $\Theta(n^2)$ time, assuming constant time integer operations. This is an improvement on earlier results by Narayana and Bent [11]. Their recursive formula can be implemented to find $|S_n|$ in $\Theta(n^3)$ time, but no faster since their recursive function must always visit $\Theta(n^3)$ states to do so; to get all $|S_k|$ for $k \leq n$ takes $\Theta(n^4)$ time due to lack of overlap in the states recursively visited for different k.

Since $S(x) = (1 - T(x))^{-1}$, where $T(x)$ is the generating function for the number of strong score sequences $|T_k|$ having length k, this recursive computation method can be extended to $|T_k|$. We first calculate the values $|S_k|$ and use this recursion to calculate all the values $|T_k|$ for $k \leq n$ in $\Theta(n^2)$ time. This is the same method as used by Stockmeyer [15], just calculating the underlying $|S_k|$ faster which brings the total time complexity down from $\Theta(n^4)$ to $\Theta(n^2)$.

References

[1] Max Alekseyev. https://oeis.org/A145855/a145855.txt, 2008.

[2] Arthur Cayley. *A theorem on trees*, volume 13 of *Cambridge Library Collection - Mathematics*, pages 26—-28. Cambridge University Press, 2009.

[3] Shane Chern. An extension of a formula of Jovovic. *Integers*, 19:Paper No. A47, 2019.

[4] Anders Claesson, Mark Dukes, Atli Fannar Franklín, and Sigurdur Örn Stefánsson. Counting tournament score sequences. *Proceedings of the American Mathematical Society*, to appear. doi:10.1090/proc/16425. arXiv:2209.03925.

[5] P. Erdős, A. Ginzburg, and A. Ziv. Theorem in the additive number theory. *Bull. Res. Council Israel Sect. F*, 10F(1):41–43, 1961.

[6] Frank Harary and Leo Moser. The theory of round robin tournaments. *Amer. Math. Monthly*, 73:231–246, 1966.

[7] OEIS Foundation Inc. The Online Encyclopedia of Integer Sequences. `https://oeis.org`, 2021.

[8] Daniel J Kleitman and Kenneth J Winston. Forests and score vectors. *Combinatorica*, 1(1):49–54, 1981.

[9] H. G. Landau. On dominance relations and the structure of animal societies: III. The condition for a score structure. *Bull. Math. Biophys.*, 15:143–148, 1953.

[10] P. A. MacMahon. An American tournament treated by the calculus of symmetric functions. *Quart. J. Math.*, 49:1–36, 1920. Reprinted in *Percy Alexander MacMahon Collected Papers, Volume I*, George E. Andrews, ed., MIT Press, 1978.

[11] T. V. Narayana and D. H. Bent. Computation of the number of score sequences in round-robin tournaments. *Can. Math. Bull*, 7(1):133–136, 1964.

[12] John Riordan. The number of score sequences in tournaments. *J. Combinatorial Theory*, 5:87–89, 1968.

[13] John Riordan. Erratum: "The number of score sequences in tournaments". *J. Combinatorial Theory*, 6:226, 1969.

[14] Richard P. Stanley. Decompositions of rational convex polytopes. In J. Srivastava, editor, *Combinatorial Mathematics, Optimal Designs and Their Applications*, volume 6 of *Annals of Discrete Mathematics*, pages 333–342. Elsevier, 1980.

[15] Paul K. Stockmeyer. Counting various classes of tournament score sequences. *arXiv:2202.05238v1*, 2022.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

The k-XORSAT threshold revisited

(Extended abstract)

Amin Coja-Oghlan* Mihyun Kang† Lena Krieg‡ Maurice Rolvien§

Abstract

We provide a simplified proof of the random k-XORSAT satisfiability threshold theorem. As an extension we also determine the full rank threshold for sparse random matrices over finite fields with precisely k non-zero entries per row. This complements a result from [Ayre, Coja-Oghlan, Gao, Müller: Combinatorica 2020]. The proof combines physics-inspired message passing arguments with a surgical moment computation. MSc: 60B20, 15B52

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-041`

1 Introduction

A random k-XORSAT instance consists of a conjunction of random XOR clauses with k literals. The goal of the well-known random k-XORSAT problem is to determine the maximum number of XOR-clauses in a random k-XORSAT formula such that the formula remains satisfiable with high probability (w.h.p. for short). This threshold was derived for the random 3-XORSAT problem ($k = 3$) by Dubois and Mandler [14]. They stated that their proof extends to the general case. But this turned out to be far from straightforward.

*TU Dortmund, Faculty of Computer Science, 12 Otto-Hahn-St, Dortmund 44227, Germany. E-mail: `amin.coja-oghlan@tu-dortmund.de`. Supported by DFG CO 646/3 and DFG CO 646/5

†TU Graz, Institute of Discrete Mathematics, Steyrergasse 30, 8010 Graz, Austria. E-mail: `kang@math.tugraz.at`. Supported in part by a Friedrich Wilhelm Bessel research award of the Alexander von Humboldt Foundation (AUT 1204138 BES)

‡TU Dortmund, Faculty of Computer Science, 12 Otto-Hahn-St, Dortmund 44227, Germany. E-mail: `lena.krieg@tu-dortmund.de`. Supported by DFG CO 646/5.

§TU Dortmund, Faculty of Computer Science, 12 Otto-Hahn-St, Dortmund 44227, Germany. E-mail: `maurice.rolvien@tu-dortmund.de`

Only more than ten years later did Pittel and Sorkin [26] publish the first complete yet complicated proof based on moment computations. Their proof spans well over 30 pages and resorts to computer-assistance. Subsequently, Ayre, Coja-Oghlan, Gao and Müller [4] published a different but still complicated proof based on coupling arguments.

In this work we provide a relatively short proof for the random k-XORSAT satisfiability threshold. Our proof is based on a novel combination of physics-inspired 'quenched' arguments and 'annealed' computations.

We start with a quenched argument. Using a message passing technique called Warning Propagation ('WP') we characterize typical solutions of random k-XORSAT instances. Equipped with this characterization we then carry out a carefully truncated moment calculation ('annealed' computation in physics jargon).

Let $\boldsymbol{F} = \boldsymbol{F}_k(n,m)$ be a random k-XORSAT instance consisting of n Boolean variables and m random XOR-clauses with k literals, where the clauses are drawn independently and uniformly from the set of all possible $2^k\binom{n}{k}$ XOR-clauses of length k on n variables. The k literals of a clause are drawn independently at random. The following theorem, first established in [14] for $k = 3$ and in [26] for $k > 3$, provides the k-XORSAT satisfiablity threshold.

Theorem 1. *For $k \geq 3$ and $d > 0$ let*

$$\Phi_{d,k}(\alpha) = \exp\left(-d\alpha^{k-1}\right) + d\alpha^{k-1} - \frac{d(k-1)}{k}\alpha^k - \frac{d}{k} \qquad \text{and} \tag{1.1}$$

$$d_k = \sup\left\{d > 0 : \max_{\alpha\in[0,1]} \Phi_{d,k}(\alpha) = 1 - d/k\right\}. \tag{1.2}$$

For any $\varepsilon > 0$ w.h.p. the random k-XORSAT formula $\boldsymbol{F}$ is

(i) satisfiable if $m \leq (1-\varepsilon)d_k n/k$, *(ii) unsatisfiable if $m \geq (1+\varepsilon)d_k n/k$.*

A k-XORSAT formula can naturally be translated to a linear system over $\mathbb{F}_2$ and therefore it induces a random matrix over $\mathbb{F}_2$ where each column represents a variable and each row a clause of the formula. Theorem 1 admits a natural generalisation to matrices over finite fields beyond $\mathbb{F}_2$.

Thus, let $q \geq 2$ be a prime power and let $\mathfrak{A} = (\mathfrak{A}_{ij})_{i,j\geq 1}$ be an infinite matrix with non zero entries $\mathfrak{A}_{ij} \in \mathbb{F}_q \setminus \{0\}$. Further, we choose a sequence $(\boldsymbol{e}_i)_{i\geq 1}$ of independent uniformly random subsets of $[n]$ of size $|\boldsymbol{e}_i| = k$. Define the random $m\times n$-matrix $\boldsymbol{A} = \boldsymbol{A}(k,m,n,q,\mathfrak{A})$ over $\mathbb{F}_q$ by letting

$$\boldsymbol{A}_{ij} = \mathfrak{A}_{ij}\mathbb{1}\{j \in \boldsymbol{e}_i\} \qquad (i \in [m],\ j \in [n]).$$

For $q = 2$ we obtain the matrix induced by a k-XORSAT formula.

Theorem 2. *For all $k \geq 3$, all prime powers $q \geq 2$ and all infinite matrices $\mathfrak{A}$ composed of non-zero elements of $\mathbb{F}_q$ the following hold. Let d_k be the threshold from* (1.2)*. Then for any $\varepsilon > 0$,*

(i) if $m \leq (1-\varepsilon)d_k n/k$, then $\boldsymbol{A}$ has full row rank w.h.p.

(ii) if $m \geq (1+\varepsilon)d_k n/k$, then $\boldsymbol{A}$ fails to have full row rank w.h.p.

Theorem 2 complements [4, Theorem 1.1], where only random matrices with identically distributed rows were considered, while in Theorem 2 random matrices may proscribe different non-zero entries for each row. We proceed to outline the proof strategy of Theorem 2.

2 Proof strategy

The main task is to prove the positive statement Theorem 2(i). Assume that for $m < (1-\varepsilon)d_k n/k$ w.h.p. the values of a random kernel vector $\boldsymbol{\sigma} \in \ker \boldsymbol{A}$ are approximately 'balanced' such that each value $s \in \mathbb{F}_q$ appears in $\boldsymbol{\sigma}$ about n/q times. Via a moment calculation we could show that the expected number of such balanced vectors $\sigma \in \ker \boldsymbol{A}$ equals $(1+o(1))q^{n-m}$. Thus $|\ker \boldsymbol{A}| = (1+o(1))q^{n-m}$ w.h.p. and $\boldsymbol{A}$ has full row rank w.h.p. via the second moment method.

Hence, it remains to show that a typical kernel vector $\boldsymbol{\sigma} \in \ker \boldsymbol{A}$ is balanced. However, we are not able to prove directly that a random kernel vector is balanced w.h.p. Instead, we will use a technique called Warning Propagation ('WP') to extract a quantitative picture of the kernels vectors' structure via a 'quenched' argument.

Pinning. We begin with an auxiliary result from [6]. Let A be an $M \times N$ matrix over finite field $\mathbb{F}_q$. We alter the matrix A using a technique called pinning: We add a few rows to the matrix with exactly one non-zero entry at a random position which thus pin the corresponding variables to zero. This randomised pinning operation, devised in this form in [6], mostly removes 'short linear relations' from the matrix and actually works on any arbitrary matrix.

Following [6] we call a set of columns J a *relation of* A if there exists a linear combination of rows with J as the set of non zero entries. Hence J is a relation of A if there exists a vector $y \in \mathbb{F}_q^M$ such that $\mathrm{supp}(y^\top A)$ is a non-empty subset of J. For a k-XORSAT formula these relations can be interpreted as derived XOR-clauses.

Further we call a column or variable *frozen in* A, if the singleton $\{j\}$ is a relation of A. Thus, j is frozen iff every kernel vector is zero on position j. We denote $\mathcal{F}(A)$ as the set of frozen coordinates in A and say that $J \neq \emptyset$ is a *proper relation* of A if $J \setminus \mathcal{F}(A)$ is a relation of A. Finally, we say that A is (δ,ℓ)*-free* if A possesses fewer than $\delta\binom{N}{h}$ proper relations I of size $|I| = h$ for any $2 \leq h \leq \ell$. In other words, a matrix is (δ,ℓ)-free if it contains only few short relations that are not exclusively composed of frozen coordinates.

For an integer $t \geq 0$ let $A[t]$ denote a matrix obtained from A by adding t new rows, each of which contains a single non-zero entry at a random position.

Lemma 1 ([6, Proposition 2.4])**.** *For any $\delta > 0, \ell > 0$ there exists $T_0 = O(\ell^3/\delta^4) > 0$ such that for any $T \geq T_0$ and any matrix A for a random $\boldsymbol{t} \in [T]$ we have $\mathbb{P}\left[A[\boldsymbol{t}] \text{ is } (\delta,\ell)\text{-free}\right] > 1-\delta$.*

Thus, with $T = \lceil \log n \rceil$, the matrix $\boldsymbol{A}^\dagger = \boldsymbol{A}[\boldsymbol{t}]$ is (ω^{-1}, ω)-free with $\omega = \lceil \log \log n \rceil$ w.h.p. This allows us to characterize the set of frozen variables in $\boldsymbol{A}^\dagger$ in terms of the Warning Propagation scheme.

Warning Propagation. We introduce WP for a general $M \times N$ matrix A, not just for $\boldsymbol{A}^\dagger$. The matrix A naturally induces a bipartite graph $G(A)$ called the *Tanner graph* with two different kind of vertices, *variable nodes* and *check nodes.* The set of variable nodes and check nodes coincide with the columns and rows of the matrix.

We define the WP scheme following [5]. The goal is to characterize the set of variables frozen in the matrix A in terms of local interactions between variable nodes and their adjacent checks using WP messages. Each edge $v_j a_i$ is endowed with two messages, one sent by the variable node v_j to the factor node a_i and one from the factor node to the variable node. Each message takes a symbolic value $\{\mathtt{u}, \mathtt{f}\}$ to represent 'unfrozen' and 'frozen'.

The *standard messages* $\mathfrak{m}_{v_j \to a_i}(A)$ encompass the actual effects of adjacent variables and factors emerging of the matrix. Let $A \setminus \{a_i\}$ be the matrix obtained from A by deleting the row a_i. Similarly $A \setminus \{\partial v_j \setminus \{a_i\}\}$ is the matrix where every other row adjacent to v_j except a_i is removed. The standard message $\mathfrak{m}_{v_j \to a_i}(A) = \mathtt{f}$ indicates that the variable v_j is frozen in the matrix $A \setminus \{a_i\}$. Similarly, $\mathfrak{m}_{a_i \to v_j}(A) = \mathtt{f}$ expresses that v_j needs to be frozen in order to satisfy the check a_i and thus is frozen in $A \setminus \{\partial v_j \setminus \{a_i\}\}$.

Warning Propagation update provides a heuristic fixed point equation for these messages:

$$\mathfrak{m}_{v_j \to a_i} = \begin{cases} \mathtt{f} & \text{if } \exists a_h \in \partial v_j \setminus \{a_i\} : \mathfrak{m}_{a_h \to v_j} = \mathtt{f}, \\ \mathtt{u} & \text{otherwise}, \end{cases} \tag{2.1}$$

$$\mathfrak{m}_{a_i \to v_j} = \begin{cases} \mathtt{f} & \text{if } \forall v_h \in \partial a_i \setminus \{v_j\} : \mathfrak{m}_{v_h \to a_i} = \mathtt{f}, \\ \mathtt{u} & \text{otherwise}. \end{cases} \tag{2.2}$$

The idea is that freezing is caused by local effects only. For instance v_j is expected to be frozen in $A \setminus \{a_i\}$ iff some other check a_h freezes v_j via a standard message.

The fixed point equations (2.1),(2.2) are easily verified for matrices with acyclic Tanner graphs. However, they do not hold for general matrices. Nonetheless, we show that for the random matrix $\boldsymbol{A}^\dagger$ (2.1),(2.2) hold for all but $o(n)$ adjacent pairs a_i, v_j w.h.p. and that the messages correctly identify the set of frozen variables. Furthermore, we prove that in most kernel vectors the values of the unfrozen variables are approximately 'balanced'.

Quenched analysis. Recall that our goal is to show that a random kernel vector $\boldsymbol{\sigma}^\dagger \in \ker \boldsymbol{A}^\dagger$ is approximately balanced w.h.p. Since we know that this holds for the unfrozen variables due to the WP-results, we only need to show that the fraction of frozen variables is $\boldsymbol{\alpha} = o(1)$ w.h.p. For this purpose we will extract detailed quantitative information about combinations of messages belonging to an edge as well as the number of certain labels.

Our next goal is to derive this information in terms of the (as of yet) unknown random variable $\boldsymbol{\alpha}$.

We denote by $\ell = (\ell_{\mathtt{uu}}, \ell_{\mathtt{uf}}, \ell_{\mathtt{fu}}, \ell_{\mathtt{ff}}) \in \mathbb{Z}_{\geq 0}^4$ a specification of message combinations, where $\ell_{\mathtt{uf}}$ equals the number of edges with message combination $\mathtt{u}$ (incoming) $\mathtt{f}$ (outgoing), etc. Define Δ_ℓ as the number of variable nodes that receive/send out messages according to ℓ. Analogously, let Γ_ℓ be the number of factor nodes that receive/send according to ℓ.

We are going to estimate $|\Delta_\ell|$ and $|\Gamma_\ell|$ in terms of the fraction $\boldsymbol{\alpha}$ of frozen variables using the hypothesis that the incoming messages at a check node a_i are essentially independent. We can derive predictions $\bar{\Gamma}_\ell(\boldsymbol{\alpha})$ and $\bar{\Delta}_\ell(\boldsymbol{\alpha})$ in terms of the (obvious) Galton Watson tree that mimics the Tanner graph of $\boldsymbol{A}$ and show that these approximations are accurate w.h.p.

Proposition 1. *Let $d > 0, k \geq 3$. Then w.h.p. for all but $o(n)$ adjacent pairs v_j, a_i the fixed point equations* (2.1),(2.2) *hold. Moreover for all ℓ*

$$\mathbb{E}\left||\Delta_\ell| - n\bar{\Delta}_\ell(\boldsymbol{\alpha})\right| + \mathbb{E}\left||\Gamma_\ell| - m\bar{\Gamma}_\ell(\boldsymbol{\alpha})\right| = o(n).$$

Finally, for all but $o(n)$ exceptions variable v_j is frozen iff $\mathfrak{m}_{a_i \to v_j} = \mathtt{f}$ for some $a_i \in \partial v_j$.

The proof of Proposition 1 is based on coupling arguments and does not reveal the likely value of $\boldsymbol{\alpha}$.

Annealed argument. In the next and last step we aim to show that $\boldsymbol{\alpha} = o(1)$ w.h.p. if $d < (1-\varepsilon)d_k$. The present annealed computation differs significantly from the prior works of [14, 26]. These prior works were based on blunt moment computations that generally have the disadvantage that even extremely rare events contribute. These large deviations result in intricate and technically demanding analytical optimisation problems. In contrast, thanks to Proposition 1 we already know the typical shape of kernel vectors and are therefore left with a straightforward and elegant computation.

To elaborate, we proceed in two steps. First, we estimate the expected number of α-WP fixed points with an α-fraction of frozen variables, which turns out to be sub-exponential for any $0 \leq \alpha \leq 1$. In the next step we estimate the number $\boldsymbol{X}_\alpha$ of kernel vectors $\boldsymbol{\sigma}^\dagger \in \ker(\boldsymbol{A}^\dagger)$ that extend a certain α-WP fixed point (frozen variable set to zero, unfrozen variables balanced). Proposition 1 then implies that $|\ker \boldsymbol{A}^\dagger| \sim \boldsymbol{X}_{\boldsymbol{\alpha}}$ w.h.p. Let $\mathfrak{D}$ be the σ-algebra generated by the degree-sequence of the Tanner graph. The following proposition gives a first moment upper bound on $\boldsymbol{X}_\alpha$ for any $0 \leq \alpha \leq 1$ in terms of the function $\Phi_{d,k}$ from (1.1).

Proposition 2. *Let $d > 0, k \geq 3$. W.h.p. for all $\alpha \in [0,1]$ we have*

$$\mathbb{E}[\boldsymbol{X}_\alpha \mid \mathfrak{D}] \leq q^{n\Phi_{d,k}(\alpha)+o(n)}.$$

For $d < d_k$ the function $\Phi_{d,k}$ has its unique maximum at $\alpha = 0$ and $q^{n\Phi_{d,k}(0)} = q^{n-dn/k}$. Thus, we can derive the estimate $\boldsymbol{\alpha} = o(1)$ w.h.p. and finally we can deduce that w.h.p. most kernel vectors $\boldsymbol{\sigma}^\dagger \in \ker(\boldsymbol{A}^\dagger)$ are 'balanced'. This finishes our proof strategy outlined at the beginning.

Acknowledgment Special thanks to Olga Scheftelowitsch for her comments on an earlier version of this work.

References

[1] D. Achlioptas and M. Molloy. The solution space geometry of random linear equations. *Random Struct. Algorithms*, 46:197–231, 2015.

[2] D. Achlioptas, A. Naor, and Y. Peres. Rigorous location of phase transitions in hard optimization problems. *Nature*, 435:759–764, 2005.

[3] M. Aizenman, R. Sims, and S. Starr. An extended variational principle for the sk spin-glass model. *Phys. Rev. B*, 68:214403, 2003.

[4] P. Ayre, A. Coja-Oghlan, P. Gao, and N. Müller. The satisfiability threshold for random linear equations. *Combinatorica*, 40:179–235, 2020.

[5] A. Coja-Oghlan, O. Cooley, M. Kang, J. Lee, and J. Ravelomanana. The sparse parity matrix. *Proc. 33rd SODA*, pages 822–833, 2022.

[6] A. Coja-Oghlan, A. Ergür, P. Gao, S. Hetterich, and M. Rolvien. The rank of sparse random matrices. *Proc. 31st SODA*, pages 579–591, 2020.

[7] A. Coja-Oghlan, P. Gao, M. Hahn-Klimroth, J. Lee, N. Müller, and M. Rolvien. The full rank condition for sparse random matrices. *arXiv:2112.14090*, 2021.

[8] O. Cooley, J. Lee, and J. Ravelomanana. Warning propagation: stability and subcriticality. *arXiv:2111.15577*, 2021.

[9] C. Cooper. The cores of random hypergraphs with a given degree sequence. *Random Struct. Algorithms*, 25:353–375, 2004.

[10] C. Cooper, A. Frieze, and W. Pegden. On the rank of a random binary matrix. *Electron. J. Comb.*, 26, 2019. P4.12.

[11] N. Creignou, H. Daude, and O. Dubois. Approximating the satisfiability threshold for random k-xor-formulas. *arXiv:cs/0106001*, 2001.

[12] M. Dietzfelbinger, A. Goerdt, M. Mitzenmacher, A. Montanari, R. Pagh, and M. Rink. Tight thresholds for cuckoo hashing via xorsat. *Proc. 37th ICALP*, pages 213–225, 2010.

[13] J. Ding, A. Sly, and N. Sun. Proof of the satisfiability conjecture for large k. *Ann. of Math.*, 196:1–388, 2022.

[14] O. Dubois and J. Mandler. The 3-xorsat threshold. *Proc. 43rd FOCS*, pages 769–778, 2002.

[15] D. Fernholz and V. Ramachandran. Cores and connectivity in sparse random graphs. *UTCS Technical Report TR04-13*, 2004.

[16] A. Goerdt and L. Falke. Satisfiability thresholds beyond k-xorsat. *Proc. 7th International Computer Science Symposium in Russia*, pages 148–159, 2012.

[17] M. Ibrahimi, Y. Kanoria, M. Kraning, and A. Montanari. The set of solutions of random xorsat formulae. *Ann. Appl. Probab.*, 25:2743–2808, 2015.

[18] S. Janson and M. Luczak. A simple solution to the k-core problem. *Random Struct. Algorithms*, 30:50–62, 2007.

[19] S. Janson, T. Luczak, and A. Rucinski. Random graphs. *Wiley*, 2000.

[20] J.H. Kim. Poisson cloning model for random graphs. *Proc. International Congress of Mathematicians*, pages 873–897, 2006.

[21] M. Mézard and A. Montanari. Information, physics and computation. *Oxford University Press*, 2009.

[22] M. Mézard, F. Ricci-Tersenghi, and R. Zecchina. Two solutions to diluted p-spin models and xorsat problems. *J. Stat. Phys.*, 111:505–533, 2003.

[23] M. Molloy. Cores in random hypergraphs and boolean formulas. *Random Struct. Algorithms*, 27:124–135, 2005.

[24] A. Montanari. Estimating random variables from random sparse observations. *European Trans. on Telecomm.*, 19(4):385–403, 2008.

[25] D. Panchenko and M. Talagrand. Bounds for diluted mean-field spin glass models. *Probab. Theory Relat. Fields*, 130:319–336, 2004.

[26] B. Pittel and G. Sorkin. The satisfiability threshold for k-xorsat. *Combin. Probab. Comput.*, 25:236–268, 2016.

[27] B. Pittel, J. Spencer, and N. Wormald. Sudden emergence of a giant k-core in a random graph. *J. Combin. Theory Ser. B*, 67:111–151, 1996.

[28] P. Raghavendra and N. Tan. Approximating csps with global cardinality constraints using sdp hierarchies. *Proc. 23rd SODA*, pages 373–387, 2012.

[29] O. Riordan. The k-core and branching processes. *Combin. Probab. Comput.*, 17:111–136, 2008.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Proving a directed analogue of the Gyárfás-Sumner conjecture for orientations of P_4

(Extended abstract)

Linda Cook* Tomáš Masařík† Marcin Pilipczuk† Amadeus Reinald‡ Uéverton S. Souza§¶

Abstract

An oriented graph is a digraph that does not contain a directed cycle of length two. An (oriented) graph D is H*-free* if D does not contain H as an induced sub(di)graph. The Gyárfás-Sumner conjecture is a widely-open conjecture on simple graphs, which states that for any forest F, there is some function f such that every F-free graph G with clique number $\omega(G)$ has chromatic number at most $f(\omega(G))$. Aboulker, Charbit, and Naserasr [Extension of Gyárfás-Sumner Conjecture to Digraphs; E-JC 2021] proposed an analogue of this conjecture to the dichromatic number of oriented graphs. The *dichromatic number* of a digraph D is the minimum number of colors required to color the vertex set of D so that no directed cycle in D is monochromatic.

Aboulker, Charbit, and Naserasr's $\vec{\chi}$-boundedness conjecture states that for every oriented forest F, there is some function f such that every F-free oriented graph D has dichromatic number at most $f(\omega(D))$, where $\omega(D)$ is the size of a maximum

*Discrete Math Group, Institute for Basic Science, Daejeon, Republic of Korea. Email: lindacook@ibs.re.kr

†Institute of Informatics, Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Warsaw, Poland.

‡Université de Montpellier, Montpellier, France.

§Instituto de Computação, Universidade Federal Fluminense, Niterói, Brasil

¶The research of TM, MP, and US is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme Grant Agreement 714704. LC and AR were supported the Institute for Basic Science (IBS-R029-C1). AR was also supported by the ANR project Digraphs (ANR-19- CE48-0013-01).

clique in the graph underlying D. In this paper, we perform the first step towards proving Aboulker, Charbit, and Naserasr's $\overrightarrow{\chi}$-boundedness conjecture by showing that it holds when F is any orientation of a path on four vertices.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-042`

1 Introduction

In a simple graph, the size of a maximum clique gives a lower bound on its chromatic number. But if a graph contains no large cliques, does it necessarily have small chromatic number? This question has been answered in the negative. In 1959, Erdős showed that there exist graphs with arbitrarily high girth and arbitrarily high chromatic number [10]. Hence, if a graph H contains a cycle there a graphs with arbitrarily high chromatic number and no induced copy of H. Around the 1980s, Gyárfás and Sumner independently conjectured [13, 25] that for any forest H, all graphs with bounded clique number and no induced copy of H have bounded chromatic number. The conjecture has been proven for some specific classes of forests but remains largely open; see [23] for a survey of related results. This paper concerns an extension of the Gyárfás-Sumner conjecture to directed graphs proposed by Aboulker, Charbit, and Naserasr [3].

We call a digraph *oriented* if it has no *digon* (directed cycle of length two). This paper will focus on finite, simple, oriented graphs. For a digraph $D = (V, E)$ we define the *underlying graph* of D to be the graph $D^* = (V, E^*)$ where E^* is the set obtained from E by replacing each arc $e \in E$ by an undirected edge between the same two vertices. We say two vertices in D are *adjacent* or *neighbors* if they are adjacent in D^*. We denote the set of neighbors of a vertex $v \in V(D)$ by $N(v)$ and we denote $N(v) \cup \{v\}$ by $N[v]$. For a set of vertices $S \subseteq V(D)$ we let $N(S)$ and $N[S]$ denote the sets $\cup_{v\in S}N(v) \setminus S$ and $\cup_{v\in S}N[v]$. For a subdigraph $H \subseteq D$ we let $N(H)$ denote the set $N(V(H))$. We let P_t denote the path on t vertices and $\overrightarrow{P_t}$ be the path $p_1 \to p_2 \to \cdots \to p_t$. We call an oriented graph whose underlying graph is a clique a *tournament*. Given a (di)graph G and $S \subseteq V$, we denote the sub(di)graph of G induced by S as $G[S]$. We say that a (di)graph G *contains* a (di)graph H if G contains H as an induced sub(di)graph. If G does not contain a (di)graph H we say that G is *H-free*. The *clique number* and the *chromatic number* of a digraph are the chromatic number and clique number of its underlying graph, respectively. We denote the clique number and the chromatic number of a (di)graph G by $\omega(G)$ and $\chi(G)$, respectively. We say that a graph H is *χ-bounding* if there exists a function f with the property that every H-free graph G satisfies $\chi(G) \leq f(\omega(G))$. In this language, the Gyárfás-Sumner conjecture states that every forest is χ-bounding.

How can the Gyárfás-Sumner conjecture be adapted to the directed setting? A first idea is to call an oriented graph H *χ-bounding* if there exists a function f with the property that every H-free oriented graph D satisfies $\chi(D) \leq f(\omega(D))$. Then, once again, by [10], all χ-bounding oriented graphs are oriented forests. Note that if an oriented graph H is χ-bounding, its underlying graph H^* is also χ-bounding. However, the converse does not hold, as, for instance, P_4 is χ-bounding, but there exist orientations of P_4 that are not χ-bounding. There are four different orientations of P_4, up to reversing the order of the vertices on the whole path:

$$\overrightarrow{P_4} : \to\to\to,\ \overrightarrow{A_4} : \to\leftarrow\to,\ \overrightarrow{Q_4} : \to\leftarrow\leftarrow,\ \overrightarrow{Q_4}' : \leftarrow\leftarrow\to$$

Chudnovsky, Scott and Seymour showed $\overrightarrow{Q_4}$ and $\overrightarrow{Q_4}'$ are χ-bounding in [8]. However, $\overrightarrow{P_4}$ and $\overrightarrow{A_4}$ are not $\overrightarrow{\chi}$-bounding as shown by Kierstead and Trotter [18] and Gyárfás [15], respectively. Chudnovsky, Scott and Seymour [8] showed that $\overrightarrow{Q_4}, \overrightarrow{Q_4}'$ are both χ-bounding in 2019. In the same article, the authors show that orientations of stars are also χ-bounding.

Our first attempt at adapting the Gyárfás-Sumner conjecture to oriented graphs failed for oriented paths such as $\overrightarrow{P_4}$ and $\overrightarrow{A_4}$. Hence, we focus on a different approach proposed by Aboulker, Charbit, and Naserasr [3] which uses a concept called "dichromatic number" introduced in [21]. A *dicoloring* of a digraph D is a partition of $V(D)$ into classes, or colors, such that each class induces an acyclic digraph (that is, there is no monochromatic directed cycle). The *dichromatic number* of D, denoted as $\overrightarrow{\chi}(D)$, is the minimum number of colors needed for a dicoloring of D. Notice that every coloring of a directed graph D is also a dicoloring, thus $\overrightarrow{\chi}(D) \leq \chi(D)$. We say a class of digraphs $\mathcal{D}$ is $\overrightarrow{\chi}$*-bounded* if there exists a function f such that every $D \in \mathcal{D}$ satisfies $\overrightarrow{\chi}(D) \leq f(\omega(D))$ and we call such an f a $\overrightarrow{\chi}$*-binding function* for $\mathcal{D}$. We say that a digraph H is $\overrightarrow{\chi}$*-bounding* if the class of H-free oriented graphs is $\overrightarrow{\chi}$-bounded.

We can now state Aboulker, Charbit, and Naserasr's dichromatic analogue to the Gyárfás–Sumner conjecture for digraphs. For brevity, we will call this conjecture the "ACN $\overrightarrow{\chi}$-boundedness" conjecture in the remainder of this extended abstract.

Conjecture 1.1 (The ACN $\overrightarrow{\chi}$-boundedness conjecture [3])**.** *Every oriented forest is $\overrightarrow{\chi}$-bounding.*

The converse of the ACN $\overrightarrow{\chi}$-boundedness conjecture holds; all $\overrightarrow{\chi}$-bounding digraphs must be oriented forests. Indeed, Harutyunyan and Mohar proved that there exist oriented graphs of arbitrarily large undirected girth and dichromatic number [16]. Oriented graphs of sufficiently large undirected girth (and no digon) forbid any fixed digraph that is not an oriented forest.

The ACN $\overrightarrow{\chi}$-boundedness conjecture is still widely open. It is not known whether the conjecture holds for any orientation of any tree T on at least five vertices that is not a star. In particular, it is not known whether the conjecture holds for oriented paths. In contrast, Gyárfás showed that every path is χ-bounding in the 1980's [13, 14] via short and elegant proof. For $t \leq 3$, every orientation of P_t is trivially $\overrightarrow{\chi}$-bounding. However, for $t \geq 4$, the picture gets more complicated. Let T be any fixed orientation of K_3. In [3], Aboulker, Charbit and Naserasr showed that class of $(T,\overrightarrow{P_4})$-free oriented graphs have bounded dichromatic number. The authors also show that $\overrightarrow{P_4}$-free oriented graphs with clique number at most three have bounded dichromatic number. Recently, Aboulker, Aubian, Charbit, and Thomassé showed that $\overrightarrow{P_6}$-free oriented graphs with clique number at most two also have bounded dichromatic number [1]. See [5] for further related results.

Let $\overrightarrow{K_t}$ denote the transitive tournament on t vertices. Steiner showed that the class of $(\overrightarrow{K_3}, \overrightarrow{A_4})$-free oriented graphs has bounded dichromatic number in [24]. In the same paper Steiner asked whether the class of $(H, \overrightarrow{K_t})$-free oriented graphs has bounded dichromatic number for $t \geq 4$ and $H \in \{\overrightarrow{P_4}, \overrightarrow{A_4}\}$. Our main result answers this question in the affirmative as corollary.

1.1 Our contributions

In this paper, we show that every orientation of P_4 is $\overrightarrow{\chi}$-bounding and thus the ACN $\overrightarrow{\chi}$-boundedness conjecture holds for all orientations of P_4. The ACN $\overrightarrow{\chi}$-boundedness conjecture is open for any orientation of P_t for $t \geq 5$. Our main novel result is that $\overrightarrow{P_4}$ and $\overrightarrow{A_4}$ are both $\overrightarrow{\chi}$-bounding. To summarize, our main result is the following:

Theorem 1.2. *Let H be an oriented P_4. Then, the class of H-free oriented graphs is $\overrightarrow{\chi}$-bounded. In particular, for any H-free oriented graph D,*

$$\overrightarrow{\chi}(D) \leq (\omega(D) + 7)^{(\omega(D)+8.5)}.$$

2 Proof Sketch

In this section we will sketch the proof of Theorem 1.2. The full proof is available in the arXiv version of this paper [9]. Our main tool in the proof is an object called a "dipolar set" which was first introduced in [2] as a "nice set".

Definition 2.1. *A* dipolar set *of an oriented graph D is a nonempty subset $S \subseteq V(D)$ that can be partitioned into S^+, S^- such that no vertex in S^+ has an out-neighbor in $V(D \setminus S)$ and no vertex in S^- has an in-neighbor in $V(D \setminus S)$.*

We will use the following lemma from [2] which reduces the problem of bounding the dichromatic number of D to bounding the dichromatic number of a dipolar set in every induced oriented subgraph of D.

Lemma 2.2 (Lemma 17 in [2])**.** *Let $\mathcal{D}$ be a family of oriented graphs closed under taking induced subgraphs. Suppose there exists a constant c such that every $D \in \mathcal{D}$ has a dipolar set S with $\overrightarrow{\chi}(S) \leq c$. Then every $D \in \mathcal{D}$ satisfies $\overrightarrow{\chi}(D) \leq 2c$.*

We will give a way of finding a dipolar set in any oriented graph excluding some orientation of P_4 as an induced subdigraph and show how to bound its dichromatic number. The backbone of our dipolar set is an object we call a closed tournament.

Definition 2.3. *We say K and P* form a closed tournament *$C = K \cup V(P)$ if K is a tournament of maximum order and P is a directed path from a source component to a sink* *component of the directed graph induced by K. We say K and P form* path-minimizing closed tournament *if $|P|$ is minimized amongst all choices of K, P that form a closed tournament.*

Lemma 2.4. *Let H be an orientation of P_4 and D be an H-free oriented graph. Let C be a closed tournament in D and let X be the set of vertices with both an in-neighbor and an out-neighbor in C. Then $N[C \cup X]$ is a dipolar set.*

The proof follows from the fact that every $v \in N(C)$ must have a non-neighbor in C and from the definition of strong connectivity. See Lemma 3.1 in [9] for details.

Our proof that orientations of P_4 are $\overrightarrow{\chi}$-bounding proceeds by induction. Let H be an oriented P_4 and let $\omega > 1$ be an integer. We let γ be the maximum of $\overrightarrow{\chi}(D')$ over every H-free oriented graph D' satisfying $\omega(D') < \omega$. We assume γ is finite. We let D be an H-free oriented graph with clique number ω and assume D is strongly connected.

Observation 2.5. *Every $v \in V(D)$ satisfies $\overrightarrow{\chi}(N(v)) \leq \gamma$.*

Let $C = K \cup V(P)$ be a path-minimizing closed tournament in D. Let X be the set of vertices in $N(C)$ with an in-neighbor and an out-neighbor in C. Then $N[C \cup X]$ is a closed tournament. It remains to show that $\overrightarrow{\chi}(N[C \cup X])$ is bounded by a function of ω and γ.

By Observation 2.5, $\overrightarrow{\chi}(N[K]) \leq \omega \cdot \gamma + \omega$. Let the vertices of P be $p_1 \to p_2 \to \cdots \to p_\ell$, in order. Then since C is path minimizing, P is a shortest directed path from p_1 to p_ℓ. Hence:

Observation 2.6. *For each integer $2 \leq i + 1 < j \leq \ell$, there is no arc from p_i to p_j.*

It follows that $\overrightarrow{\chi}(V(P)) \leq 2$. Hence, it is enough to show that $\overrightarrow{\chi}(N(P) \setminus N[K])$ and $\overrightarrow{\chi}(N(X) \setminus N[C])$ are bounded by a function of ω and γ. We obtain that $\overrightarrow{\chi}(N(X) \setminus N[C]) \leq 2\gamma$ by applying the following lemma with $Q := C$, $R := X$ and $S := N(X) \setminus N[C]$.

Lemma 2.7. *Let H be an oriented P_4 and let D be an H-free oriented graph. Suppose there is a partition of $V(D)$ into sets Q, R, S such that there is no arc between Q and S, every $r \in R$ has both an in-neighbor and an out-neighbor in Q and every $s \in S$ has a neighbor in R. Let γ be an integer such that for every $r \in R$, we have $\overrightarrow{\chi}(N(r)) \leq \gamma$. Then $\overrightarrow{\chi}(S) \leq 2\gamma$.*

Lemma 2.7 follows from an easy inductive argument on $|S|$. (See Lemma 4.3 in [9] for details). By Lemma 2.7, it only remains to show that $\overrightarrow{\chi}(N(P) \setminus N[K])$ is bounded by some function of γ and ω. Note, we cannot simply apply Observation 2.5 because P may be arbitrarily long. For this part of the proof we proceed (slightly) differently for each orientation of P_4. Here, we present a sketch of the case when $H = \overrightarrow{P_4}$. The other cases are similar.

We say the "first" and "last" neighbors of a vertex $v \in N(P)$ are the vertices $p_i \in N(v) \cap V(P)$ minimizing i and maximizing i, respectively. For each integer $1 \leq i \leq \ell$, we let F_i, L_i denote the sets of vertices in $N(P)$ whose first neighbor is p_i and whose last neighbor is p_i, respectively. Let F_i^-, L_i^+ be the sets consisting of all out-neighbors of p_i in F_i and in-neighbors of p_i in L_i, respectively.

Observation 2.8. *Let $2 \leq i < j \leq \ell - 1$. Then there are no arcs from F_i^- to F_j and no arcs from L_i to L_j^+.*

Indeed, as otherwise $D[N[\{p_i, p_j\}]]$ would contain a $\overrightarrow{P_4}$. Let $W = \cup_{i=2}^{\ell-1}(F_i^- \cup L_i^+)$ Then by Observation 2.5 and Observation 2.8, $\overrightarrow{\chi}(W) \leq 2\gamma$. Let $R = N(P) \setminus (W \cup N(\{p_1, p_2, p_\ell\}))$. By Observation 2.5, we need only show showing $\overrightarrow{\chi}(R)$ is at most some function of γ and ω to complete the proof of Theorem 1.2. We will require a technical lemma:

Lemma 2.9. *Let $v, w \in R$, if $(w, v) \in E(D)$, there is a directed path from v to w on at most* $\max\{6, \ell - 1\}$ *vertices.*

Lemma 2.9 follows from Observation 2.6 and a brief case analysis. See Lemma 5.4 from [9] for details.

Lemma 2.10. $\overrightarrow{\chi}(R) \leq 6\gamma$.

Proof. If P contains at most six vertices then $\overrightarrow{\chi}(N(P)) \leq 6\gamma$, hence we may assume this is not the case. We may assume that there is a tournament J of size ω in $D[R]$ for otherwise $\overrightarrow{\chi}(R) \leq \gamma$. Since $P \neq \emptyset$ and $C = K \cup V(P)$ was chosen to be path-minimizing it follows that J cannot be strongly connected. Let v be a vertex in the sink component of J and w be a vertex in the source component of J. Therefore, $(w, v) \in E(D)$. Thus by Lemma 2.9 there is a path Q from v to w of length less than that that of P. Hence, J, P' form a closed tournament. By definition since K, P were chosen to form a path-minimizing closed tournament P' cannot be shorter than P, a contradiction. □

By combining the results from this section, we obtain that $\overrightarrow{\chi}(N[C \cup X])$ is at most some function of γ and ω. Since $N[C \cup X]$ is dipolar, it follows from Lemma 2.2 that $\overrightarrow{\chi}(D)$ is at most some function of γ and ω. Hence, by induction $\overrightarrow{P_4}$ is $\overrightarrow{\chi}$-binding. The proofs that $\overrightarrow{A_4}, \overrightarrow{Q_4}$ and $\overrightarrow{Q_4}'$ are $\overrightarrow{\chi}$-binding are similar (and slightly simpler). Full details can be found in [9].

3 Conclusion

Our result is an initial step towards resolving the ACN $\overrightarrow{\chi}$-boundedness conjecture for orientation of paths in general. However, we think we are still far from this result. It would already be interesting to hear the answer to the easier question: Is it true that for every oriented path H there is a constant c_H such that every oriented graph not containing H or a tournament of size three has dichromatic number at most c_H. By Theorem 1.2 this is known when H is an orientation of a path of length at most four. It is proven in [1] that this is true when $H = \overrightarrow{P_6}$.

Recall that the classes of $\overrightarrow{Q_4}$-free oriented graphs and $\overrightarrow{Q_4}'$-free oriented graphs were already shown to be χ-bounded in [8]. The χ-binding function f' for these two classes from [8] is defined using recurrence $f'(x) := 2(3f'(x-1))^5$ which leads to a double-exponential bound on χ, and cannot guarantee a better bound on $\vec{\chi}$. In this paper,Theorem 1.2 provides an improved $\vec{\chi}$-binding function for these classes. We would like to know whether any orientation of P_4 is *polynomially* $\vec{\chi}$-bounding. In other words, is there some oriented P_4 so that the class of oriented graphs forbidding it has a *polynomial* $\vec{\chi}$-binding function?

References

[1] Pierre Aboulker, Guillaume Aubian, Pierre Charbit, and Stéphan Thomassé. (P6, triangle)-free digraphs have bounded dichromatic number, 2022. URL: https://arxiv.org/abs/2212.02272.

[2] Pierre Aboulker, Nicolas Bousquet, and Rémi de Verclos. Chordal directed graphs are not χ-bounded. *The Electronic Journal of Combinatorics*, 29(2):17, 2022. doi:10.37236/11050.

[3] Pierre Aboulker, Pierre Charbit, and Reza Naserasr. Extension of Gyárfás-Sumner conjecture to digraphs. *The Electronic Journal of Combinatorics*, 28(2):27, 2021. doi:10.37236/9906.

[4] Eli Berger, Krzysztof Choromanski, Maria Chudnovsky, Jacob Fox, Martin Loebl, Alex Scott, Paul Seymour, and Stéphan Thomassé. Tournaments and colouring. *Journal of Combinatorial Theory, Series B*, 103(1):1–20, 2013. doi:10.1016/j.jctb.2012.08.003.

[5] Alvaro Carbonero. On coloring digraphs with forbidden induced subgraphs, 2023. Master's Thesis for University of Waterloo. URL: http://hdl.handle.net/10012/19321.

[6] Alvaro Carbonero, Patrick Hompe, Benjamin Moore, and Sophie Spirkl. Digraphs with all induced directed cycles of the same length are not $\vec{\chi}$-bounded. *The Electronic Journal of Combinatorics*, 29(4), 2022. doi:10.37236/11179.

[7] Alvaro Carbonero, Patrick Hompe, Benjamin Moore, and Sophie Spirkl. A counterexample to a conjecture about triangle-free induced subgraphs of graphs with large chromatic number. *Journal of Combinatorial Theory, Series B*, 158:63–69, 2023. doi:10.1016/j.jctb.2022.09.001.

[8] Maria Chudnovsky, Alex Scott, and Paul Seymour. Induced subgraphs of graphs with large chromatic number. XI. orientations. *European Journal of Combinatorics*, 76:53–61, 2019. doi:10.1016/j.ejc.2018.09.003.

[9] Linda Cook, Tomáš Masařík, Marcin Pilipczuk, Amadeus Reinald, and Uéverton S. Souza. Proving a directed analogue of the Gyárfás-Sumner conjecture for orientations of P_4, 2022. URL: https://arxiv.org/abs/2209.06171.

[10] Paul Erdős. Graph theory and probability. *Canadian Journal of Mathematics*, 11:34–38, 1959. doi:10.4153/CJM-1959-003-9.

[11] Paul Erdős. Problems and results in number theory. In *Proceedings of the Ninth Manitoba Conference on Numerical Mathematics and Computing*, pages 3–21, 1979. URL: https://users.renyi.hu/~p_erdos/1981-21.pdf.

[12] Paul Erdos and Leo Moser. On the representation of directed graphs as unions of orderings. *Math. Inst. Hung. Acad. Sci*, 9:125–132, 1964.

[13] András Gyárfás. On Ramsey covering-numbers. *Infinite and Finite Sets*, 2:801–816, 1975. URL: https://users.renyi.hu/~gyarfas/Cikkek/05_Gyarfas_OnRamseyCoveringNumbers.pdf.

[14] András Gyárfás. Problems from the world surrounding perfect graphs. *Applicationes Mathematicae*, 19(3-4):413–441, 1987. doi:10.4064/am-19-3-4-413-441.

[15] András Gyárfás. Problem 115. *Discrete Mathematics*, pages 109–110, 1989/1990. URL: https://www.renyi.hu/~gyarfas/Cikkek/46_Gyarfas_Problem115.pdf.

[16] Ararat Harutyunyan and Bojan Mohar. Two results on the digraph chromatic number. *Discrete Mathematics*, 312(10):1823–1826, 2012. doi:https://doi.org/10.1016/j.disc.2012.01.028.

[17] Henry A. Kierstead and Stephen G. Penrice. Radius two trees specify χ-bounded classes. *Journal of Graph Theory*, 18(2):119–129, 1994. doi:10.1002/jgt.3190180203.

[18] Henry A. Kierstead and William T. Trotter. Colorful induced subgraphs. *Discrete Mathematics*, 101(1):165–169, 1992. doi:10.1016/0012-365X(92)90600-K.

[19] Henry A. Kierstead and Yingxian Zhu. Radius three trees in graphs with large chromatic number. *SIAM Journal on Discrete Mathematics*, 17(4):571–581, 2004. doi:10.1137/S0895480198339869.

[20] Jan Mycielski. Sur le coloriage des graphs. *Colloquium Mathematicae*, 3(2):161–162, 1955. URL: http://eudml.org/doc/210000.

[21] Víctor Neumann-Lara. The dichromatic number of a digraph. *Journal of Combinatorial Theory, Series B*, 33(3):265–270, 1982. doi:10.1016/0095-8956(82)90046-6.

[22] Alex Scott and Paul Seymour. Induced subgraphs of graphs with large chromatic number. XIII. New brooms. *European Journal of Combinatorics*, 84:103024, 2020. doi:10.1016/j.ejc.2019.103024.

[23] Alex Scott and Paul Seymour. A survey of χ-boundedness. *Journal of Graph Theory*, 95(3):473–504, 2020. doi:https://doi.org/10.1002/jgt.22601.

[24] Raphael Steiner. On coloring digraphs with forbidden induced subgraphs. *Journal of Graph Theory*, 103(2):323–339, 2023. doi:10.1002/jgt.22920.

[25] David P. Sumner. Subtrees of a graph and the chromatic number. In *The Theory and Applications of Graphs: Fourth International Conference, May 6–9, 1980, Western Michigan University, Kalamazoo, Michigan*. Wiley, New York, 1981.

[26] Alexander A. Zykov. On some properties of linear complexes. *Matematicheskii Sbornik Novaya Seriya*, 24(66):163–188, 1949. In Russian. English translation in: American Mathematical Society translations 79 (1952). URL: http://mi.mathnet.ru/msb5974.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Higher degree Erdős distinct evaluations problem

(Extended abstract)

Simone Costa* Stefano Della Fiore† Andrea Ferraguti*

Abstract

Let $\Sigma = \{a_1, ..., a_n\}$ be a set of positive integers with $a_1 < \cdots < a_n$ such that all 2^n subset sums are distinct. A famous conjecture by Erdős states that $a_n > c \cdot 2^n$ for some constant c, while the best result known to date is of the form $a_n > c \cdot 2^n/\sqrt{n}$.

In this paper, we propose a generalization of the Erdős distinct sum problem that is in the same spirit as those of the Davenport and the Erdős-Ginzburg-Ziv constants recently introduced in [7] and in [6]. More precisely, we require that the non-zero evaluations of the m-th degree symmetric polynomial are all distinct over the subsequences of Σ. Even though these evaluations can not be seen as the values assumed by the sum of independent random variables, surprisingly, the variance method works to provide a nontrivial lower bound on a_n. Indeed, the main result here presented is to show that

$$a_n > c_m \cdot 2^{\frac{n}{m}} / n^{1-\frac{1}{2m}}.$$

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-043

1 Introduction

For any $n \geq 1$, consider sets $\{a_1, ..., a_n\}$ of positive integers with $a_1 < \cdots < a_n$ whose subset sums are all distinct. A famous conjecture, due to Paul Erdős, is that $a_n \geq c \cdot 2^n$

*DICATAM - Sez. Matematica, University of Brescia, Via Branze 38, I-25123 Brescia, Italy. E-mails: simone.costa@unibs.it, andrea.ferraguti@unibs.it.

†DI, University of Salerno, Via Giovanni Paolo II 132, Fisciano, Italy. E-mail: s.dellafiore@unisa.it.

for some constant $c > 0$. Using the variance method, Erdős and Moser [10] (see also [1] and [13]) were able to prove that $a_n \geq 1/4 \cdot n^{-1/2} \cdot 2^n$. No advances have been made so far in removing the term $n^{-1/2}$ from this lower bound, but there have been several improvements on the constant factor, including the work of Dubroff, Fox, and Xu [11], Guy [12], Elkies [9], Bae [4], and Aliev [3]. In particular, the best currently known lower bound states that $a_n \geq (1+o(1))\sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{n}}2^n$. Two simple proofs of this result, first obtained unpublished by Elkies and Gleason, are presented in [11]. In the other direction, the best-known construction is due to Bohman [5] (see also [14]), who showed that there exist arbitrarily large such sets with $a_n \leq 0.22002 \cdot 2^n$.

Several variations on the problem appear during the years such as [2] and [8]. In this paper, we generalize the Erdős distinct sum problem by requiring that the non-zero evaluations of the m-th degree symmetric polynomial are all distinct over the sub-sequences of Σ. The problem here considered is inspired by those of the Davenport and the Erdős-Ginzburg-Ziv constants recently introduced in [7] and in [6].

More formally, given a sequence of real numbers $\Sigma = \{a_1, \dots, a_n\}$ and a subset $A \subseteq [1,n]$, we define the m-th (degree) evaluatio $e^m_\Sigma(A) = \sum_{\substack{\{i_1,\dots,i_m\}\subseteq A\\ i_1<\cdots<i_m}} a_{i_1}\cdots a_{i_m}$, where we adopt the convention that $e^m_\Sigma(A) = 0$ if $|A| < m$.

Problem 1.1. *For every positive integer n, find the least positive $M = M(n)$ such that there exists an increasing sequence $\Sigma = (a_1, \dots, a_n)$ of real numbers with $a_i \in [0, M]$ for every i such that for all distinct $A_1, A_2 \subseteq [1,n]$ of size at least m we have that $|e^m_\Sigma(A_1) - e^m_\Sigma(A_2)| \geq 1$.*

A sequence as in Problem 1.1 will be called *M-bounded m-th evaluation distinct.*

In Section 2, we provide lower bounds on the values of M in Problem 1.1 using the variance method proving that

$$M > c_m \cdot 2^{\frac{n}{m}}/n^{1-\frac{1}{2m}}.$$

Then, in Section 3, we derive an upper bound presenting a direct construction.

2 Lower Bounds

One first lower bound to the value of M of Problem 1.1 can be provided using the pigeonhole principle. Indeed, since the number of non-zero evaluations of e^m_Σ is $2^n - \sum_{i=0}^{m-1}\binom{n}{i} = (1+o(1))2^n$, these evaluations are spaced at least by one, and each of these is smaller than $e^m_\Sigma([1,n]) \leq \binom{n}{m}M^m \leq n^m M^m/c_m$, it follows that $M > c_m \cdot 2^{\frac{n}{m}}/n$.

Now we see that using the variance method (see [1], [10] or [12]), it is possible to improve this lower bound.

Theorem 2.1. *Let $\Sigma = (a_1, \dots, a_n)$ be an m-th evaluation distinct sequence in $\mathbb{R}$ (resp. $\mathbb{Z}$) that is M-bounded. Then*

$$M > (1+o(1))\frac{2^{1-\frac{1}{m}}((m-1)!)^{\frac{1}{m}}}{3^{\frac{1}{2m}}}\frac{2^{\frac{n}{m}}}{n^{1-\frac{1}{2m}}}.$$

Proof. Let $\Sigma = (a_1, \ldots, a_n)$ be such a sequence of real (resp. integer) numbers. Pick a subset A uniformly at random from $2^{[1,n]}$ and define the real random variable $X = e^m_\Sigma(A)$. We denote by $\mu := \mathbb{E}[X]$ and $\sigma^2 := \mathbb{E}[X^2] - \mu^2$ respectively the expected value and the variance of the random variable X. Clearly, $\mu = 1/2^n \sum_{A\subseteq[1,n]:\ |A|\geq m} e^m_\Sigma(A)$. Here we have that the monomial $a_{i_1}\ldots a_{i_m}$ appears in the evaluation $e^m_\Sigma(A)$ whenever A contains $i_1, \ldots, i_m$ which happens for 2^{n-m} subsets of $[1,n]$. Therefore, we have that $\mu = e^m_\Sigma([1,n])/2^m$. By definition of variance we have that:

$$2^n\sigma^2 = \sum_{A\subseteq[1,n]} (e^m_\Sigma(A) - \mu)^2 = \sum_{A\subseteq[1,n]} \left(\sum_{\substack{i_1<i_2<\cdots<i_m \\ i_1,\ldots,i_m\in A}} a_{i_1}\ldots a_{i_m} - \sum_{\substack{i_1<i_2<\cdots<i_m \\ i_1,\ldots,i_m\in[1,n]}} \frac{a_{i_1}\ldots a_{i_m}}{2^m} \right)^2 .$$

Due to the symmetry of e^m_Σ, there exist coefficients $C_1, \ldots, C_m$ such that the latter sum can be written as follows:

$$C_0 \sum_{\substack{i_1<i_2<\ldots<i_{2m} \\ i_1,\ldots,i_{2m}\in[1,n]}} a_{i_1}\ldots a_{i_{2m}} + C_1 \sum_{\substack{i_1<i_2<\ldots<i_{2m-1} \\ i_1,\ldots,i_{2m-1}\in[1,n]}} \sum_{\ell\in[1,2m-1]} a_{i_1}a_{i_2}\ldots a^2_{i_\ell}\ldots a_{i_{2m-1}} + \tag{1}$$
$$+ \ldots + C_m \sum_{\substack{i_1<i_2<\ldots<i_m \\ i_1,\ldots,i_m\in[1,n]}} a^2_{i_1}\ldots a^2_{i_m}.$$

One can prove that $C_0 = 0$, $C_1 = 2^{n-2m}\binom{2m-2}{m-1}$ and $C_k = O(2^n)$ for every $k \in \{2, \ldots, m\}$. This can be seen since the coefficient of $a_{i_1}\ldots a_{i_{2m}}$ is $\binom{2m}{m}$ times that obtained by taking the term $a_{i_1}\ldots a_{i_m}$ from the first $(e^m_\Sigma(A) - \mu)$ in the product and $a_{i_{m+1}}\ldots a_{i_{2m}}$ from the second one. Then, the coefficient of $a^2_{i_1}\ldots a_{i_2}\ldots a_{i_{2m-1}}$ is $\binom{2m-2}{m-1}$ times that obtained taking the term $a_{i_1}\ldots a_{i_m}$ from the first $(e^m_\Sigma(A) - \mu)$ in the product and $a_{i_1}a_{i_{m+1}}\ldots a_{i_{2m-1}}$ from the second one. Symmetrically, the same is true for every term $a_{i_1}\ldots a^2_{i_\ell}\ldots a_{i_{2m-1}}$. Finally, the coefficient of $a^2_{i_1}\ldots a^2_{i_k}a_{i_{k+1}}\ldots a_{i_{2m-k}}$ is $\binom{2m-2k}{m-k}$ times that obtained taking the term $a_{i_1}\ldots a_{i_m}$ from the first $(e^m_\Sigma(A) - \mu)$ in the product and $a_{i_1}\ldots a_{i_k}a_{i_{m+1}}\ldots a_{i_{2m-k}}$ from the second one. Summing up, we can rewrite equation (1) as

$$2^n\sigma^2 = C_1 \sum_{\substack{i_1<i_2<\ldots<i_{2m-1} \\ i_1,\ldots,i_{2m-1}\in[1,n]}} \sum_{\ell\in[1,2m-1]} a_{i_1}a_{i_2}\ldots a^2_{i_\ell}\ldots a_{i_{2m-1}} + \tag{2}$$
$$+O(2^n)\left(\sum_{k=2}^{m} C_k \sum_{\substack{i_1<i_2<\ldots<i_{2m-k} \\ i_1,\ldots,i_{2m-k}\in[1,n]}} \sum_{\substack{\ell_1<\ldots<\ell_k \\ \ell_1,\ldots,\ell_k\in[1,2m-k]}} a_{i_1}a_{i_2}\ldots a^2_{i_{\ell_1}}\ldots a^2_{i_{\ell_k}}\ldots a_{i_{2m-k}} \right).$$

In equation (2), each C_k multiplies a sum of $\binom{n}{2m-k}\cdot\binom{2m-k}{k} < \frac{n^{2m-k}}{(2m-2k)!k!}$ terms. Since

a_n is the largest element of the sequence, we get:

$$2^n\sigma^2 < \frac{n^{2m-1}}{(2m-2)!}\binom{2m-2}{m-1}2^{n-2m}a_n^{2m}(1+o(1)) = \left(\frac{n^{2m-1}}{((m-1)!)^2}2^{n-2m}a_n^{2m}\right)(1+o(1)). \tag{3}$$

On the other hand, for $|A| \geq m$, the evaluations $e_\Sigma^m(A)$ are all different and spaced at least by one, and hence we have that $(e_\Sigma^m(A)-\mu)^2$ assumes at least $\frac{1}{2}(2^n - \sum_{i=0}^{m-1}\binom{n}{i})$ different values. Since the sum $\sum_{A\subseteq[1,n]}(e_\Sigma^m(A)-\mu)^2$ is minimized when the values are around μ and are spaced by one, we obtain the lower bound:

$$\frac{1+o(1)}{12}2^{3n} = 2\sum_{i=0}^{\frac{1}{2}(2^n-\sum_{i=0}^{m-1}\binom{n}{i})} i^2 \leq 2^n\sigma^2. \tag{4}$$

To conclude the proof, it is enough to compare (3) and (4). □

3 Upper bounds

In this section we provide an upper bound to the value of M in Problem 1.1 by presenting the following direct construction.

Lemma 3.1. *Let ϵ_1, ϵ_2 be two reals such that $\epsilon_1 > \epsilon_2 > 0$ and let $m \geq 2$ be an integer. Then for every n large enough the sequence $\Sigma = (a_1, a_2, \ldots, a_n)$, where $a_i = (2+\epsilon_1)^n - (2+\epsilon_2)^{i-1}$ for $i = 1, 2, \ldots, n$, is m-evaluation distinct.*

Proof. Suppose by contradiction there exists two distinct subsets $B, C \subseteq [1,n]$ such that

$$|e_\Sigma^m(B) - e_\Sigma^m(C)| < 1\,. \tag{5}$$

For an arbitrary subset $S \subseteq [1,n]$ with $|S| \geq m$, by definition we have:

$$e_\Sigma^m(S) = \sum_{j=0}^{m}(-1)^j(2+\epsilon_1)^{(m-j)n}\binom{|S|-j}{m-j}\sum_{\substack{\{i_1,i_2,\ldots,i_j\}\subseteq S\\ i_1<i_2<\ldots<i_j}}(2+\epsilon_2)^{i_1+i_2+\ldots+i_j-j}\,. \tag{6}$$

We first show that inequality (5) implies $|B| = |C|$. Suppose without loss of generality that $|B| > |C|$. Then (6) implies that:

$$e_\Sigma^m(B) - e_\Sigma^m(C) = (2+\epsilon_1)^{mn}\left[\binom{|B|}{m} - \binom{|C|}{m}\right] + \sum_{j=1}^{m}(-1)^j(2+\epsilon_1)^{(m-j)n}$$
$$\left[\binom{|B|-j}{m-j}\sum_{\substack{\{i_1,i_2,\ldots,i_j\}\subseteq B\\ i_1<i_2<\ldots<i_j}}(2+\epsilon_2)^{i_1+i_2+\ldots+i_j-j} - \binom{|C|-j}{m-j}\sum_{\substack{\{i_1,i_2,\ldots,i_j\}\subseteq C\\ i_1<i_2<\ldots<i_j}}(2+\epsilon_2)^{i_1+i_2+\ldots+i_j-j}\right]. \tag{7}$$

Now it can be seen that each term in the first summation of equation (7) is of order $O\left(n^m(2+\epsilon_1)^{mn}\left(\frac{2+\epsilon_2}{2+\epsilon_1}\right)^{jn}\right)$, for $j=1,2,\ldots,m$ and $n\to\infty$. Hence, asymptotically in n, we can rewrite (7) as $e^m_\Sigma(B)-e^m_\Sigma(C)=(2+\epsilon_1)^{mn}\left[\binom{|B|}{m}-\binom{|C|}{m}\right](1+o(1))$, since $\epsilon_1>\epsilon_2$.This clearly contradicts (5), and hence we must have $|B|=|C|$.

Next, let t be an integer such that $|B|=|C|=t$ and let $B:=\{b_1,b_2,\ldots,b_t\}$ and $C:=\{c_1,c_2,\ldots,c_t\}$, where $b_1<b_2<\ldots<b_t$ and $c_1<c_2<\ldots<c_t$. Since $B\neq C$, there exists an integer $\ell\in[1,t]$ such that $b_\ell\neq c_\ell$ while $b_{\ell+1}=c_{\ell+1}$, $b_{\ell+2}=c_{\ell+2}$,…, $b_t=c_t$. Suppose without loss of generality that $b_\ell>c_\ell$. Then we have:

$$|e^m_\Sigma(B)-e^m_\Sigma(C)|=\Bigg|(2+\epsilon_1)^{(m-1)n}\binom{t-1}{m-1}\left(\sum_{i\le\ell}(2+\epsilon_2)^{b_i-1}-(2+\epsilon_2)^{c_i-1}\right)+\sum_{j=2}^{m}(-1)^{j-1}$$
$$(2+\epsilon_1)^{(m-j)n}\binom{t-j}{m-j}\left(\sum_{\substack{1\le i_1<i_2<\ldots<i_j\le t\\ i_1\le\ell}}(2+\epsilon_2)^{b_{i_1}+b_{i_2}+\ldots+b_{i_j}-j}-(2+\epsilon_2)^{c_{i_1}+c_{i_2}+\ldots+c_{i_j}-j}\right)\Bigg|. \tag{8}$$

To conclude the proof, we need to lower bound equation (8). The summation formula for the geometric series implies that: $\sum_{i\le\ell}(2+\epsilon_2)^{c_i-1}\le\sum_{1\le i\le c_\ell}(2+\epsilon_2)^{i-1}<(2+\epsilon_2)^{c_\ell}/(1+\epsilon_2)\le(2+\epsilon_2)^{b_\ell-1}/(1+\epsilon_2)$, and since each term in the summation over j in equation (8) is, as $n\to\infty$, of order $O\left(n^m(2+\epsilon_1)^{(m-1)n}(2+\epsilon_2)^{b_\ell-1}\left(\frac{2+\epsilon_2}{2+\epsilon_1}\right)^{(j-1)n}\right)$, we obtain the following lower bound:

$$|e^m_\Sigma(B)-e^m_\Sigma(C)|>\left|(2+\epsilon_1)^{(m-1)n}\binom{t-1}{m-1}(2+\epsilon_2)^{b_\ell-1}\left(1-\frac{1}{1+\epsilon_2}\right)\right|(1+o(1)).$$

The theorem now follows since the right hand side of the above inequality is greater than 1 for sufficiently large n's. $\square$

Along the same lines of Lemma 3.1, we can prove the following corollary. We do not report here the proof due to space limitations.

Corollary 3.2. *Let ϵ_1, ϵ_2 be two reals such that $\epsilon_1>\epsilon_2>0$ and let $m\ge2$ be an integer. Then for every n large enough the sequence $\Sigma=(a_1,a_2,\ldots,a_n)$, where $a_i=\lfloor(2+\epsilon_1)^n-(2+\epsilon_2)^{i-1}\rfloor$ for $i=1,2,\ldots,n$, is m-evaluation distinct.*

We observe that Corollary 3.2 holds also for $m=1$ but we obtain a bound that is worse than the ones given in [5] and [14]. As an easy consequence of Corollary 3.2, one can prove the following theorem.

Theorem 3.3. *There exists a sequence $\Sigma=(a_1,a_2,\ldots,a_n)$ of n integers that is m-evaluation distinct and M-bounded such that $M\le2^{n+o(n)}$, for $n\to\infty$.*

References

[1] N. Alon and J. H. Spencer, The probabilistic method, 4th ed. Wiley, Hoboken, NJ, 2016.

[2] M. Axenovich, Y. Caro, R. Yuster, Sum-distinguishing number of sparse hypergraphs, European Journal of Combinatorics 112 (2023), 103712 .

[3] I. Aliev, Siegel's lemma and sum-distinct sets, Discrete Comput. Geom. 39 (2008), 59-66.

[4] J. Bae, On subset-sum-distinct sequences. Analytic number theory, Vol. 1, Progr. Math., 138, Birkhauser, Boston, 1996, 31-37.

[5] T. Bohman, A construction for sets of integers with distinct subset sums, Electron. J. Combin. 5 (1998), Research Paper 3, 14 pages.

[6] Y. Caro and J.R. Schmitt. Higher degree Erdős-Ginzburg-Ziv constants, Integers 22 (2022).

[7] Y. Caro, B. Girard and J.R. Schmitt. Higher degree Davenport constants over finite commutative rings, Integers 21 (2021).

[8] S. Costa, S. Della Fiore, M. Dalai: Variation on the Erdős distinct-sums problem. Discrete Applied Mathematics, 325 (2023), 172–185.

[9] N. D. Elkies, An improved lower bound on the greatest element of a sum-distinct set of fixed order, J. Combin. Theory Ser. A 41 (1986), 89-94.

[10] P. Erdős, Problems and results in additive number theory, Colloque sur la Theorie des Nombres, Bruxelles, 1955, 127-137.

[11] Quentin Dubroff, Jacob Fox and Max Wenqiang Xu, A note on the Erdős distinct subset sums problem, SIAM J. Discret. Math. 35 (2021), 322-324.

[12] R. K. Guy, Sets of integers whose subsets have distinct sums, Theory and practice of combinatorics, 141–154, North-Holland Math. Stud., 60, Ann. Discrete Math., 12, North-Holland, Amsterdam, (1982).

[13] R. K. Guy, Unsolved Problems in Intuitive Mathematics, Vol. I, Number Theory, Problem C8, Springer-Verlag (1981).

[14] W.F. Lunnon, Integer sets with distinct subset sums, Math. Compute, 50 (1988) 297-320.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Monochromatic configurations on a circle

(Extended abstract)

Gábor Damásdi* Nóra Frankl† János Pach‡ Dömötör Pálvölgyi§

Abstract

For $k \geq 3$, call a k-tuple $(d_1, d_2, \ldots, d_k)$ with $d_1 \geq d_2 \geq \cdots \geq d_k > 0$ and $\sum_{i=1}^{k} d_i = 1$ a *Ramsey k-tuple* if the following is true: in every two-colouring of the circle of unit perimeter, there is a monochromatic k-tuple of points in which the distances of cyclically consecutive points, measured along the arcs, are $d_1, d_2, \ldots, d_k$ in some order. By a conjecture of Stromquist, if $d_i = \frac{2^{k-i}}{2^k-1}$, then $(d_1, \ldots, d_k)$ is Ramsey.

Our main result is a proof of the converse of this conjecture. That is, we show that if $(d_1, \ldots, d_k)$ is Ramsey, then $d_i = \frac{2^{k-i}}{2^k-1}$. We do this by finding connections of the problem to certain questions from number theory about partitioning $\mathbb{N}$ into so-called *Beatty sequences*. We also disprove a majority version of Stromquist's conjecture, study a robust version, and discuss a discrete version.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-044`

*Alfréd Rényi Institute of Mathematics and ELTE Eötvös Loránd University, Budapest, Hungary. E-mail: `damasdigabor@caesar.elte.hu`. Partially supported by ERC Advanced Grant GeoScape.

†School of Mathematics and Statistics, The Open University, Milton Keynes UK, and Alfréd Rényi Institute of Mathematics, Budapest, Hungary. E-mail: `nora.frankl@open.ac.uk`. Partially supported by ERC Advanced Grant GeoScape.

‡Alfréd Rényi Institute of Mathematics, Budapest, Hungary and IST, Klosterneuburg, Austria. E-mail: `pach@cims.nyu.edu`. Partially supported by ERC Advanced Grant GeoScape and NKFIH (National Research, Development and Innovation Office) grant K-131529.

§ELTE Eötvös Loránd University and Alfréd Rényi Institute of Mathematics, Budapest, Hungary. E-mail: `domotor.palvolgyi@ttk.elte.hu`. Partially supported by the ERC Advanced Grant "ERMiD" and by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, and by the New National Excellence Program ÚNKP-22-5 and by the Thematic Excellence Program TKP2021-NKTA-62 of the National Research, Development and Innovation Office.

1 Introduction

In the May 2021 issue of the *American Mathematical Monthly*, Robert Tauraso posed the following problem [13]: *If all the points of the plane are arbitrarily coloured blue or red, find an acute pentagon with all vertices the same colour and with prescribed area* 1. A beautiful solution was suggested by Walter Stromquist, which reduced the question to a Ramsey-type problem, interesting on its own right.

Consider 31 points evenly spaced on a circle, and colour each of them arbitrarily blue or red. Then we can always find 5 points with the same colour that divide the circle into arcs proportional to 1:2:4:8:16. (The arcs need not be in the order suggested by the proportion. That is, 1:4:16:2:8 counts as a success.) Notice that no matter in what order 5 points divide the circle into such arcs, their convex hull is a pentagon of the same area. Thus, all we have to do is to start with a circle for which this area is 1. Stromquist managed to verify the above statement by computer, and he formulated the following attractive conjecture.

Conjecture 1.1 (Stromquist's conjecture)**.** *For any $k \geq 3$, consider $2^k - 1$ points evenly spaced on a circle, and colour each of them arbitrarily blue or red.*

Then we can always find k points with the same colour that divide the circle into arcs proportional to $1 : 2 : 4 : \ldots : 2^{k-1}$, but not necessarily in this order.

The case $k = 3$ was settled a long time ago by Bialostocki and Nielsen [4], and it is not hard to verify the case $k = 4$ either. Stromquist kindly informed us that he was able to give a computer assisted proof for $k \leq 6$.

In the present note, we study Stromquist's conjecture. To simplify the presentation, we introduce some notation. For $k \geq 3$, let $\underline{d} = (d_1, d_2, \ldots, d_k)$ be a k-tuple with $d_1 \geq d_2 \geq \cdots \geq d_k > 0$ and $\sum_{i=1}^{k} d_i = 1$. In a two-colouring of the circle S of unit perimeter, we call a k-tuple $(p_1, p_2, \ldots, p_k)$ of points from S *monochromatic* if the colour of every point p_i is the same. The main problem we study is whether for a given $\underline{d}$ it is true that in every two-colouring of S we can find a monochromatic k-tuple in which the distances of consecutive points, measured along the arcs, are exactly $d_1, \ldots, d_k$ in some order. We call a k-tuple $\underline{d}$ with this property a *Ramsey k-tuple*, or simply *Ramsey.*

A *permuted copy* of a k-gon inscribed in S is another k-gon inscribed in S with the same side lengths, but in a possibly different order. If the side lengths of the k-gon, measured along the arcs, are $d_1, \ldots, d_k$, we also call a monochromatic permuted copy of the k-gon a *monochromatic permuted copy*, or simply a *monochromatic copy*, of the k-tuple $\underline{d} = (d_1, d_2, \ldots, d_k)$.

Using this terminology, Stromquist's conjecture is equivalent to that if $k \geq 3$, and $d_i = \frac{2^{k-i}}{2^k-1}$ for every $1 \leq i \leq k$, then $\underline{d} = (d_1, \ldots, d_k)$ is Ramsey. Our main result is proving the converse of the conjecture. That is, we prove that other k tuples are *not* Ramsey.

Theorem 1.2. *If $\underline{d} = (d_1, \ldots, d_k)$ is Ramsey, then $d_i = \frac{2^{k-i}}{2^k-1}$.*

We call the k-tuple $\underline{d} = (d_1, \ldots, d_k)$ with $d_i = \frac{2^{k-i}}{2^k-1}$ the $(k, 2)$*-power.* To prove Theorem 1.2, for every k-tuple $\underline{d}$ that is *not* the $(k, 2)$-power, we construct a two-colouring of

S that does not contain a monochromatic copy of $\underline{d}$. In fact, we show that for any other tuple $\underline{d}$ there exists a $t \in \mathbb{N}$, for which the colouring that consists of $2t$ arcs of equal length, coloured alternating red and blue, does not contain a monochromatic copy of $\underline{d}$. Theorem 1.2 is an immediate corollary of the following lemma, proved in Section 2.

Lemma 1.3. *Let c_t be a uniform colouring of S obtained by dividing it into $2t$ equal circular arcs, and colouring them alternating the two colours. If for every $t \in \mathbb{N}$ the uniform colouring c_t contains a monochromatic copy of $\underline{d} = (d_1, \dots, d_k)$, then $d_i = \frac{2^{k-i}}{2^k-1}$.*

Our proof proceeds by establishing a connection to a conjecture of Fraenkel about *Beatty sequences*, and solving a special case of it, which may be of independent interest.

A Beatty sequence is a sequence of the form $\{\lfloor \alpha n + \beta \rfloor\}_{n=1}^{\infty}$ for some $\alpha, \beta \in \mathbb{R}$. The term Beatty sequence was first used by Connell [5], after a problem proposed by Beatty [3]. Let $\underline{\alpha} = (\alpha_1, \dots, \alpha_k)$ with $0 < \alpha_1 \leq \cdots \leq \alpha_k$ and $\underline{\beta} = (\beta_1, \dots, \beta_k)$ be two k-tuples of real numbers. We say that the pair $(\underline{\alpha}, \underline{\beta})$ *partitions* $\mathbb{N}$, if the Beatty sequences $\{\lfloor \alpha_i n + \beta_i \rfloor\}_{n=1}^{\infty}$ partition $\mathbb{N}$.

Finding a characterisation of those pairs $(\underline{\alpha}, \underline{\beta})$ which partition $\mathbb{N}$ is a well-studied problem, which has connections to a combinatorial game, called *Wythoff's game*, see for example [5, 6, 7, 8, 9, 15]. For $k = 2$, the characterisation is well understood [8, 11]. Fraenkel [8] noted that for $k \geq 3$ and for $\underline{\alpha} = (\alpha_1, \dots, \alpha_k)$ with $\alpha_i = \frac{2^k-1}{2^{k-i}}$ for every $1 \leq i \leq k$, there is a $\underline{\beta}$ such that $(\underline{\alpha}, \underline{\beta})$ partitions $\mathbb{N}$. According to Erdős and Graham,[1] Freankel made the following conjecture.

Conjecture 1.4 (Fraenkel's conjecture)**.** *If for $\underline{\alpha} = (\alpha_1, \dots, \alpha_k)$ with $k \geq 3$ and $0 < \alpha_1 < \cdots < \alpha_k$ the pair $(\underline{\alpha}, \underline{\beta})$ partitions $\mathbb{N}$, then $\alpha_i = \frac{2^k-1}{2^{i-1}}$ for $1 \leq i \leq k$.*

Conjecture 1.4 is confirmed for $k \leq 7$ [1, 2, 10, 16, 14], and is open for $k \geq 8$. To prove Theorem 1.2, we prove Fraenkel's conjecture in a special case.

Theorem 1.5. *If $\alpha_i = \frac{\beta_i}{2}$ for every $1 \leq i \leq k$, and $(\underline{\alpha}, \underline{\beta})$ partitions $\mathbb{N}$, then $\alpha_i = \frac{2^k-1}{2^{k-i}}$ for every $1 \leq i \leq k$.*

We omit the details of the proof of Theorem 1.5 here, due to space restrictions.

In most of our proofs about Ramsey k-tuples, we work with a discrete version of the problem. We can do so because if there is an i for which $\frac{d_i}{\sum_j d_j}$ is irrational, then it is easy to show that $\underline{d}$ is not Ramsey. Indeed, we can to two-colour the points of S with no monochromatic pair of points at a given irrational distance apart.

Assuming $\sum_i d_i = 1$ and that every d_i is rational, then writing $d_i = \frac{p_i}{q_i}$ for every $1 \leq i \leq k$, for $N = \mathrm{lcm}(q_1, \dots, q_k)$ the problem is equivalent to deciding if in any two-colouring of the vertices of a regular N-gon inscribed in S, we can find a monochromatic copy of $\underline{d}$. In other words, the problem is equivalent to deciding if in every two-colouring of $\mathbb{Z}_N$ we can find a monochromatic k-tuple in which the differences of cyclically consecutive elements

[1] This appears at [6, page 19] but the there cited paper [8] of Fraenkel only states a weaker conjecture, asserting that there are i, j with $i \neq j$ such that the ratio α_i/α_j is an integer.

are $N \cdot d_1, \ldots, N \cdot d_k$ in some order. We find connections between certain transformations in the discrete version and avoiding monochromatic copies by using uniform colourings in the original version.

Considering Stromquist's conjecture, we could not answer the more specific question whether every uniform two-colouring of S contains a monochromatic copy of the $(k, 2)$-power, however, we confirmed this for very large values of k by a computer search. This more specific question is related to another problem from number theory, which has connections to vector balancing and combinatorial discrepancy; see Conjecture 5.1.

One might assume that if in a two-colouring one colour class is denser than the other, then it will contain a $(k, 2)$-power. However, this is false. Let $0 < \varepsilon < 1/80$, and divide S into 10 intervals of lengths $1/8-\varepsilon, 1/16+\varepsilon, 1/8-\varepsilon, 1/16+\varepsilon, 1/8-\varepsilon, 1/16+\varepsilon, 1/8-\varepsilon, 1/8+\varepsilon, 1/16-\varepsilon, 1/8+\varepsilon$ in this order, and colour them alternating red and blue, starting with red. Then the set of red points has total length $1/2 + 1/16 - 5\varepsilon > 1/2$, but a straight-forward case analysis shows that there is no red copy of a $(k, 2)$-power for $k \geq 8$.

We also study what happens when instead of a copy of $\underline{d}$, we only want to find a copy ε-close to it. Two k-tuples $(p_1, \ldots, p_k)$ and $(p'_1, \ldots, p'_k)$ in S are *ε-close* if $|p_1 - p'_1|, \ldots, |p_k - p'_k| \leq \varepsilon$. A k-tuple of points $\underline{p} = (p_1, \ldots, p_k)$ in S is an *ε-close copy of $\underline{d}$* if it is ε-close to a copy of $\underline{d}$. We call a k-tuple *nearly-Ramsey*, if for every $\varepsilon > 0$ in every two-colouring of S there is a monochromatic ε-close copy of $\underline{d}$.

We show the following.

Theorem 1.6. *If $d_1 = \frac{1}{2}$, or $\underline{d}$ is $(\frac{4}{7}, \frac{2}{7}, \frac{1}{7})$, $(\frac{5}{8}, \frac{1}{4}, \frac{1}{8})$, $(\frac{3}{4}, \frac{1}{6}, \frac{1}{12})$, $(\frac{7}{12}, \frac{1}{4}, \frac{1}{6})$, then (d_1, d_2, d_3) is nearly-Ramsey.*

We also conjecture that these are the only nearly-Ramsey triples.

2 Proof of Lemma 1.3

Proof. Assume that for every t the colouring c_t contains a monochromatic copy of $\underline{d}$. By symmetry, we may assume that this copy is red. Going around the points corresponding to this monochromatic copy in some cyclic order, we must jump over each blue interval. An arc of distance d_i with red endpoints jumps over $\lfloor td_i \rceil$ blue intervals, where $\lfloor x \rceil$ is the rounding of x to the closest integer. Thus, we must have $\sum_{i=1}^k \lfloor td_i \rceil = t$ for every $t \in \mathbb{N}$. This implies that for every $t > 0$ we have $\sum_{i=1}^k (\lfloor td_i \rceil - \lfloor (t-1)d_i \rceil) = t - (t-1) = 1$.

On the other hand, $\lfloor td_i \rceil - \lfloor (t-1)d_i \rceil$ is either 0 or 1 for each $1 \leq i \leq k$. For a fixed i, we have $\lfloor td_i \rceil - \lfloor (t-1)d_i \rceil = 1$ exactly when t is in the sequence $\{\lfloor (n + \frac{1}{2})\frac{1}{d_i} \rfloor\}_{n=1}^\infty = \{\lfloor n\frac{1}{d_i} + \frac{1}{2d_i} \rfloor\}_{n=1}^\infty$. Thus, the sequences $\{\lfloor n\frac{1}{d_i} + \frac{1}{2d_i} \rfloor\}_{n=1}^\infty$ must partition $\mathbb{N}$, and Theorem 1.5 implies that $d_i = \frac{2^{i-1}}{2^k - 1}$. $\square$

3 Discrete version

Assume that every d_i is rational, $\sum_i d_i = 1$, and write $d_i = \frac{p_i}{q_i}$, and let $N = \mathrm{lcm}(q_1, \dots, q_k)$. In $\mathbb{Z}_N$ a *copy* of $\underline{d} = (d_1, \dots, d_k)$ is a k-tuple in which the distances of cyclically consecutive elements are $N \cdot d_1, \dots, N \cdot d_k$ in some order. A colouring of $\mathbb{Z}_N$ is $\underline{d}$*-free* if it does not contain any monochromatic copy of $(d_1, \dots, d_k)$.

Let $\chi : \mathbb{Z}_N \to \{\text{red}, \text{blue}\}$ be a colouring of $\mathbb{Z}_N$ and let $t \in \mathbb{Z}_N^*$ be such that $\gcd(t, N) = 1$. Let $\chi^t : \mathbb{Z}_N \to \{\text{red, blue}\}$ be defined by $\chi^t(x) = \chi(tx)$. It is a simple fact that χ is $(d_1, \dots, d_k)$-free if and only if χ^t is $(t \cdot d_1, \dots, t \cdot d_k)$-free.

It follows that finding a two-colouring of the vertices of a regular N-gon inscribed in S without a monochromatic copy of a triple $\underline{d} = (d_1, d_2, d_3)$ is equivalent to finding a colouring of it without a monochromatic copy of $\underline{d}^t = (d_1^t, d_2^t, d_3^t)$, where $d_i^t = \frac{t \cdot N \cdot d_i \bmod N}{N}$ if $(t \cdot N \cdot d_1 \bmod N) + (t \cdot N \cdot d_2 \bmod N) + (t \cdot N \cdot d_3 \bmod N) = N$, and $d_i^t = \frac{N - (t \cdot N \cdot d_i \bmod N)}{N}$ if $(t \cdot N \cdot d_1 \bmod N) + (t \cdot N \cdot d_2 \bmod N) + (t \cdot N \cdot d_3 \bmod N) = 2N$.

Notice that if $d_1^t, d_2^t, d_3^t \le \frac{1}{2}$, then colouring a half-arc in S blue, and the other in red, avoids all monochromatic copies of $\underline{d}^t$. Thus, if there is a t such that $\gcd(t, N) = 1$ and $d_1^t, d_2^t, d_3^t \le \frac{1}{2}$, then $\underline{d}$ is not Ramsey. The next claim explains how this property is closely related to avoiding copies by using uniform colourings. Here we omit its proof.

Claim 3.1. *If* $\gcd(t, N) = 1$*, then the uniform colouring* c_t *avoids all monochromatic copies of* $\underline{d}$ *if and only if* $d_1^t, d_2^t, d_3^t \le \frac{1}{2}$.

4 Robust version

Theorem 1.6 states that additionally $(\frac{5}{8}, \frac{1}{4}, \frac{1}{8})$, $(\frac{3}{4}, \frac{1}{6}, \frac{1}{12})$, $(\frac{7}{12}, \frac{1}{4}, \frac{1}{6})$ and any triple with $d_1 = \frac{1}{2}$ are also nearly-Ramsey. We conjecture that there are no other nearly-Ramsey triples.

Conjecture 4.1. (d_1, d_2, d_3) *is nearly-Ramsey if and only if it is* $(\frac{4}{7}, \frac{2}{7}, \frac{1}{7})$, $(\frac{5}{8}, \frac{1}{4}, \frac{1}{8})$, $(\frac{3}{4}, \frac{1}{6}, \frac{1}{12})$, $(\frac{7}{12}, \frac{1}{4}, \frac{1}{6})$ *or a triple with* $d_1 = \frac{1}{2}$.

We sketch the proof of Theorem 1.6 and provide some supporting evidence for Conjecture 4.1. We recolour a point $p \in S$ with *black* if there is a red and a blue point in every neighbourhood of p. If a colouring of S is not monochromatic, then there is at least one black point. If we can find an ε-close copy of $\underline{d}$ such that it only has red and black points (or blue and black), then we can also find a 2ε-close copy of it with only red (or only blue) points, by slightly moving the black points of the corresponding triple in S.

Proof sketch of Theorem 1.6. The proof for $(\frac{5}{8}, \frac{1}{4}, \frac{1}{8})$, $(\frac{3}{4}, \frac{1}{6}, \frac{1}{12})$, $(\frac{7}{12}, \frac{1}{4}, \frac{1}{6})$ is by a case analysis of the possible colourings of a regular 8-gon/12-gon, respectively, with a black point.

For $d_1 = \frac{1}{2}$, we show that for every $\varepsilon > 0$ every red-blue colouring contains a monochromatic ε-close copy of a given triple (d_1, d_2, d_3) with $d_1 = \frac{1}{2}$. We may assume that the colouring is not monochromatic, otherwise the statement is trivial. Thus, we may assume the existence of a black point p. Let p' be the point diametrically opposite to p, and q and q' be two other diametrically opposite points, such that any three of the four points

p, p', q, q' form a copy of (d_1, d_2, d_3). By the pigeonhole principle, without loss of generality, we may assume that at most one of p', q, q' is blue. But then the other three points form a copy of (d_1, d_2, d_3) without a blue point. □

Let $\underline{d} = (d_1, d_2, d_3)$ be a triple that Conjecture 4.1 asserts to be not nearly-Ramsey. We believe that for any such $\underline{d}$, there is a uniform colouring c_t as in Lemma 1.3 that contains no monochromatic ε-close copies of $\underline{d}$. We call $t \in \mathbb{N}$ *suitable* if c_t contains no monochromatic copies of $\underline{d}$, and *nearly-suitable*, if c_t contains no monochromatic ε-close copies of $\underline{d}$. In c_t the black points are exactly the endpoints of the intervals. Thus, t is nearly-suitable if and only if it is suitable and c_t avoids copies of $\underline{d}$ with two points coinciding with endpoints of the segments. As the distance of any two black points (along the circumference) is a multiple of $\frac{1}{2t}$, we obtain the following observation.

Observation 4.2. *A suitable t is nearly-suitable if and only if none of $2td_1, 2td_2, 2td_3$ is an integer.*

If one of d_1, d_2, d_3 is irrational, then any suitable t is also nearly-suitable, thus such (d_1, d_2, d_3) is not nearly-Ramsey. Otherwise, we write $d_i = \frac{p_i}{q_i}$ such that p_i, q_i are integers and $\gcd(p_i, q_i) = 1$ for $i = 1, 2, 3$. To prove Conjecture 4.1, it is thus sufficient to find a suitable t in $T := \{t : q_1, q_2, q_3 \nmid 2t\}$. We can prove that there is such t if q_1, q_2, q_3 are all odd, as well as in several other cases, but here we omit these proofs.

5 $(k, 2)$-powers in uniform colourings

We conjecture that for every t, the uniform colouring of c_t from Lemma 1.3 contains a monochromatic copy of the $(k, 2)$-power for every k. We have seen in the proof of Lemma 1.3 that in this case the sides of a red copy 'jump' over the t blue intervals. However, this is only a necessary condition, and not a sufficient one. Indeed, if a jump starts from a 'bad' part of a red interval, it might end up inside a blue one. More precisely, we can consider the problem as follows.

Let $N = 2t(2^k - 1)$ and colour vertices of a regular N-gon such that t reds are followed by t blues in an alternating manner, so that vertices $0, \dots, 2^k - 1$ are red, $2^k, \dots, 2 \cdot 2^k - 1$ are blue, $2 \cdot 2^k, \dots, 3 \cdot 2^k - 1$ are red etc. If there is a monochromatic copy of the $(k, 2)$-power, there is also a red copy. For each vertex of the red copy of the $(k, 2)$-power, consider its index modulo $2^{k+1} - 2$. Each of these needs to be at most $2^k - 1$. Moreover, the differences among the consecutive vertices need to be $2t \pmod{2^{k+1} - 2}$, $4t \pmod{2^{k+1} - 2}$, …, $2^k t \pmod{2^{k+1} - 2}$, in some order. To have such a k-tuple of indices modulo $2^{k+1} - 2$ is a necessary and sufficient condition for the existence of a red copy.

By computer, we verified this up to a large k. We phrase a problem in a more natural and general form. Interpret the numbers $2^i t \pmod{2^{k+1} - 2}$ that are larger than $2^k - 1$ as $2^{k+1} - 2 - 2^i t$, and denote these k numbers by $v_1, \dots, v_k$. With this, the numbers v_i will determine how one vertex moves compared to the preceding vertex in the $0, \dots, 2^k - 1$ interval. Note that none of these numbers can be equal to $2^k - 1$. Thus, $-2^k + 1 < v_1, \dots, v_k < 2^k - 1$, and $\sum_{i=1}^k v_k = 0$, since the k-gon with these side-distances exists.

We get the following even nicer question if we divide by $2^k - 1$.

Conjecture 5.1. *If a sequence of reals $-1 < x_1, \dots, x_k < 1$ satisfies*

$$x_{i+1} = \begin{cases} 2x_i, & \text{if } 2|x_i| < 1 \\ 2x_i - 2, & \text{if } 2x_i > 1 \\ 2 - 2x_i, & \text{if } 2x_i < -1 \end{cases}$$

for $i = 1, \dots, k$, such that $x_{k+1} = x_1$, then there is a permutation π of $\{1, \dots, k\}$ such that $0 \le \sum_{i=1}^{j} x_\pi(i) < 1$ for every j.

This conjecture is similar to Steinitz's theorem [12], and to other vector balancing problems. Indeed, it can be proved for any x_i's satisfying the conditions of the conjecture, $\sum_{i=1}^{k} x_i = 0$. We note that if the x_i's are any sequence satisfying $\sum_{i=1}^{k} x_i = 0$ and $|x_i| < 1/2$ for every i, then one can easily find a permutation for which $0 \le \sum_{i=1}^{j} x_\pi(i) < 1$ for every j. But without this bound, we have to exploit that $x_{i+1} = 2x_i$, as otherwise there would be counterexamples, i.e., $0.6, 0.6, 0.6, -0.9, -0.9$. Could it be that the conjecture is true because we always have many i's such that $|x_i| < 1/2$, and these can be used somehow to take care of the other x_i's?

References

[1] E. Altman, B. Gaujal, and A. Hordijk. Balanced sequences and optimal routing. *J. ACM*, 47(4):752–775, 2000.

[2] J. Barát and P. P. Varjú. Partitioning the positive integers to seven beatty sequences. *Indag. Math. (N.S.)*, 14(2):149–161, 2003.

[3] S. Beatty, N. Altshiller-Court, O. Dunkel, A. Pelletier, F. Irwin, J. L. Riley, P. Fitch, and D. M. Yost. Problems for solutions: 3173-3180. *Amer. Math. Monthly*, 33(3):159–159, 1926.

[4] A. Bialostocki and M. J. Nielsen. Minimum sets forcing monochromatic triangles. *Ars Combin.*, 81:297–304, 2006.

[5] I. G. Connell. Some properties of beatty sequences I. *Canad. Math. Bull.*, 2(3):190–197, 1959.

[6] P. Erdős and R. L. Graham. *Old and new problems and results in combinatorial number theory*, volume 28. L'Enseignement Mathematiques Un. Geneve, 1980.

[7] A. S. Fraenkel. The bracket function and complementary sets of integers. *Canad. J. Math.*, 21:6–27, 1969.

[8] A. S. Fraenkel. Complementing and exactly covering sequences. *J. Combin. Theory Ser. A*, 14(1):8–20, 1973.

[9] R. L. Graham. Covering the positive integers by disjoint sets of the form $\{[n\alpha + \beta]: n = 1, 2, \ldots\}$. *J. Combin. Theory Ser. A*, 15(3):354–358, 1973.

[10] Ryozo Morikawa. On eventually covering families generated by the bracket function. *Bull. Fac. Liberal Arts, Nagasaki Univ., Natural Science*, 23(1):17–22, 1982.

[11] T. Skolem. Uber einige eigenschaften der zahlenmengen $[\alpha n + \beta]$ bei irrationalem α mit einleitenden bemerkungen über einige kombinatorische probleme. *Norske Vid. Selsk. Forh*, 30:42–49, 1957.

[12] E. Steinitz. Bedingt konvergente Reihen und konvexe Systeme. (Teil I.). *J. Reine Angew. Math.*, 143:128–175, 1913.

[13] R. Tauraso. Problems and solutions, Problem 12251. *Amer. Math. Monthly*, 128(5):467, 2021.

[14] R Tijdeman. On complementary triples of sturmian bisequences. *Indagationes Mathematicae*, 7(3):419–424, 1996.

[15] R. Tijdeman. Exact covers of balanced sequences and Fraenkel's conjecture. *Algebraic number theory and Diophantine analysis (Graz, 1998)*, pages 467–483, 2000.

[16] R. Tijdeman. Fraenkel's conjecture for six sequences. *Discrete Math.*, 222(1-3):223–234, 2000.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Beyond the Erdős-Sós conjecture

(Extended abstract)

Akbar Davoodi* Diana Piguet* Hanka Řada*†
Nicolás Sanhueza-Matamala‡

Abstract

We prove an asymptotic version of a tree-containment conjecture of Klimošová, Piguet and Rozhoň [European J. Combin. 88 (2020), 103106] for graphs with quadratically many edges. The result implies that the asymptotic version of the Erdős-Sós conjecture in the setting of dense graphs is correct.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-045`

1 Introduction

One of the most classical questions in graph theory is to determine the number of edges in a host graph G that forces the existence of a copy of another guest graph H. For H with chromatic number at least 3, this is well understood thanks to the Erdős–Stone–Simonovits theorem [9, 8]. For bipartite H much less is known, and even the case of trees is widely open. A seminal conjecture by Erdős and Sós [7] says that graphs with average degree larger than $k-1$ should contain all k-edge trees.

Conjecture 1 (Erdős-Sós conjecture)**.** Every graph G with average degree $d(G) > k-1$ contains every tree with k edges.

*The Czech Academy of Sciences, Institute of Computer Science, Pod Vodárenskou věží 2, 182 07 Prague, Czech Republic. The research leading to these results has received funding from The Czech Science Foundation, grant number 19-08740S.

†Department of Applied Mathematics, Faculty of Information Technology, Czech Technical University in Prague.

‡Departamento de Ingeniería Matemática, Facultad de Ciencias Físicas y Matemáticas, Universidad de Concepción, Chile. Partly supported by ANID-Chile through the FONDECYT Iniciación N°11220269 grant.

The conjecture has been verified for certain families of trees (see [14]). A solution for all large enough trees was announced by Ajtai, Komlós, Simonovits and Szemerédi [1, 3, 2], but remains unpublished.

It is well-known that any graph that satisfies $d(G) > k-1$ contains a subgraph G' with $\Delta(G') \geq k$ and $\delta(G') \geq k/2$. However, this weaker condition on the host graph is not enough to ensure the containment of all trees. It fails for trees of diameter four, as shown by examples of Havet, Reed, Stein and Wood [10, Section 1]. It is natural to ask how many vertices of degree at least k in G together with $\delta(G) \geq k/2$ would guarantee the containment of all k-edge trees. A conjecture along these lines was proposed by Klimošová, Piguet, and Rozhoň.

Conjecture 2 (Klimošová, Piguet, Rozhoň, [13, Conjecture 1.4])**.** Every n-vertex graph G with $\delta(G) \geq k/2$ and at least $n/(2\sqrt{k})$ vertices of degree at least k contains all k-edge trees.

Our main result is an approximate version of Conjecture 2 for dense graphs.

Theorem 3. *For any $\eta, q > 0$, there exists an $n_0 \in \mathbb{N}$ so that for every $n \geq n_0$ and all $k \geq qn$, any n-vertex graph G with minimal degree $\delta(G) \geq (1+\eta)k/2$ and with at least ηn vertices of degree at least $(1+\eta)k$ contains all k-vertex trees.*

Theorem 3 implies an approximate dense version of the Erdős-Sós conjecture.

Corollary 4. *For any $\eta, q > 0$ there exists an $n_0 \in \mathbb{N}$ so that for every $n \geq n_0$ and all $k \geq qn$ any n-vertex graph with average degree at least $(1+\eta)k$ contains any tree on at most k vertices as its subgraph.*

Corollary 4 strengthens similar results by Rozhoň [13] and Besomi, Pavez-Signé and Stein [4, Theorem 1.3]. They give the same result as our Corollary 4 but only for trees T on k vertices which in addition satisfy $\Delta(T) = o(k)$; in contrast, our result works for all trees. It also gives a proof independent of the one proposed by Ajtai, Komlós, Simonovits, and Szemerédi [1, 3, 2] in the case the host graph is dense with the very mild strengthening that its average degree is required to be slightly larger than k. We also remark that Besomi, Pavez-Signé and Stein [4, Theorem 1.1] proved a version of the Erdős–Sós conjecture for k-edge trees which is sharp in the average degree condition (it only needs $d(G) > k-1$) but works only for large bounded-degree trees (it needs that G is an n-graph, $\Delta(T) \leq \Delta$, and $k \geq qn$, with n large with respect to q and Δ).

1.1 Notation

As is somewhat standard, we write $a \ll b$ in statements to mean "for all $b > 0$, there exists $a > 0$ such that the following is valid". Longer chains of constants are interpreted similarly, choosing the constants from right to left. We always assume those constants are positive, and if $1/n$ appears in such a chain of constants we assume that n is a positive integer.

For two disjoint subsets X and Y of $V(G)$, the *bipartite density* of the pair (X, Y) is given $d(X,Y) = |E(X,Y)|/(|X||Y|)$, where $|E(X,Y)|$ denotes the number of edges between X and Y. For a graph G, we denote by $d(G)$ the *average degree* of G, i.e. $d(G) := 2|E(G)|/|V|$. For a vertex $v \in V(G)$, let $N_G(v)$ denote the set of neighbours of v in G. We will omit G from the notation if the graph is clear from context.

A *digraph* is a graph in which every edge is oriented, meaning that it consists of an ordered pair of vertices. We admit cycles of length 2 (where the pairs of edges $\overrightarrow{uv}$ and $\overrightarrow{vu}$ are both present), but we do not allow for parallel edges in the same direction, and we also forbid loops.

2 Sketch of the proof and Main lemmas

Our proof has three main steps. First, we describe a way to cut the tree to be embedded into suitable chunks. Secondly, we prepare the host graph to embed the tree. For this, we use the Szemerédi's Regularity Lemma, which is somewhat standard in this type of proofs. A crucial definition, and our main innovation, in this step is what we call *skew matching pairs*, which are required to describe the structure which we wish to find in the host graph. The outcome of this step is summarised in what we call the Structural Lemma (Lemma 12), which is the main technical lemma of our work. In the third and final step, we construct an embedding given the structures in both the tree and the host graph, and this process is summarised in the Tree Embedding Lemma (Lemma 13).

The rest of this extended abstract is structured as follows. First, we state the aforementioned lemmas in more detail. Next, assuming the validity of those lemmas, we give short proofs of our main results: Section 3.1 contains the proof of Theorem 3 and Section 3.2 contains the proof of Corollary 4.

2.1 Preparing the tree

To prepare the embedding, we use the following handy concept used by Hladký, Komlós, Piguet, Simonovits, Stein, and Szemerédi [11, Definition 3.3]. It gives a partition of a tree into vertex-disjoint smaller trees which also satisfy several additional useful properties.

If T is a tree rooted at r, and $\widetilde{T} \subseteq T$ is a subtree with $r \notin V(\widetilde{T})$, the *seed of* $\widetilde{T}$ is the unique vertex $x \in V(T) \setminus V(\widetilde{T})$ which is farthest from r and also belongs to every (r, v)-path in T, for every $v \in V(\widetilde{T})$.

Definition 5 (ℓ-fine partition)**.** Let T be a tree on k vertices rooted at a vertex r. An *ℓ-fine partition of* T is a quadruple $(W_A, W_B, \mathcal{F}_A, \mathcal{F}_B)$, where $W_A, W_B \subseteq V(T)$ and $\mathcal{F}_A, \mathcal{F}_B$ are families of subtrees of T such that

(FP1) the three sets W_A, W_B, and $\{V(T^*)\}_{T^* \in \mathcal{F}_A \cup \mathcal{F}_B}$ partition $V(T)$ (in particular, the trees in $\mathcal{F}_A \cup \mathcal{F}_B$ are pairwise vertex-disjoint),

(FP2) $r \in W_A \cup W_B$,

(FP3) $\max\{|W_A|, |W_B|\} \leq 336k/\ell$, (FP4) $|V(T^*)| \leq \ell$ for every $T^* \in \mathcal{F}_A \cup \mathcal{F}_B$,

(FP5) $V(T^*) \cap N(W_B) = \emptyset$ for every $T^* \in \mathcal{F}_A$, and $V(T^*) \cap N(W_A) = \emptyset$ for every $T^* \in \mathcal{F}_B$;

(FP6) each tree of $\mathcal{F}_A \cup \mathcal{F}_B$ has its seeds in $W_A \cup W_B$,

The crucial fact, proven in [11, Lemma 3.5], is that any tree T admits an ℓ-fine partition, for any $1 \leq \ell \leq |V(T)|$. We denote by $\mathcal{T}^{\rho}_{a_1,a_2,b_1,b_2}$ the set of trees T, so that there is a $(\rho|V(T)|)$-fine partition $(W_A, W_B, \mathcal{F}_A, \mathcal{F}_B)$ of T so that $|V_i(\mathcal{F}_A)| = a_i$, $|V_i(\mathcal{F}_B)| = b_i$, for $i \in \{1, 2\}$, where $V_1(\mathcal{F}_A)$ (resp. $V_2(\mathcal{F}_A)$) is the set of vertices of $\mathcal{F}_A$ that are at odd (resp. even) distance from W_A, and $V_i(\mathcal{F}_B)$ are defined analogously with respect to W_B.

2.2 Preparing the host graph

In this step, we find a suitable structure in the host graph to embed the tree, using the information about the fine partition found in the previous step. The description of this step requires Szemerédi's Regularity Lemma. Before stating it, we recall the standard notions involved in its statement.

Definition 6 (Regular pair and regular partitions)**.** A pair (X, Y) with $X, Y \subseteq V(G)$ is said to be *ε-regular*, if for any sets $X' \subseteq X$ and $Y' \subseteq Y$ with $|X'| \geq \varepsilon|X|$ and $|Y'| \geq \varepsilon|Y|$ we have that $|d(G[X', Y']) - d(G[X, Y])| < \varepsilon$.

We say that a partition $\{V_0, \ldots, V_t\}$ of $V(G)$ is an *ε-regular partition* if $|V_0| \leq \varepsilon|V(G)|$, and for every $1 \leq i \leq t$, all but at most εt values of $1 \leq j \leq t$ are such that the pair (V_i, V_j) is not ε-regular.[1] We call the cluster V_0 the *garbage set*. We call a regular partition *equitable* if $|V_i| = |V_j|$ for every $1 \leq i < j \leq t$.

Szemerédi's Regularity Lemma ensures that regular partitions exist for every graph.

Theorem 7 (Szemerédi's Regularity Lemma, [15])**.** *Let $1/n \ll 1/M_0 \ll \varepsilon$. Any n-vertex graph has an ε-regular equitable partition $\{V_0, V_1, \ldots, V_t\}$ with $1/\varepsilon \leq t \leq M_0$.*

We capture the structure of a regular partition in a so-called *reduced graph*.

Definition 8 (Reduced graph)**.** Given a graph G, $d > 0$, and a ε-regular equitable partition $\mathcal{P} = \{V_0, \ldots, V_t\}$ of $V(G)$, we define the *d-reduced graph* Γ as follows. The vertex set of Γ is $\{1, \ldots, t\}$, and there is an edge $ij \in E(\Gamma)$ if and only if the pair (V_i, V_j) is ε-regular and $d(V_i, V_j) \geq d$.

As mentioned before, the Structural Lemma (Lemma 12) will yield a useful structure in the host graph G; more precisely, this structure will be defined in a reduced graph of G. In essence, the structure we want is an "allocation" of the tree in the clusters of the reduced

[1] We point out that ε-regular partitions are most commonly defined in a slightly different way, with the property that at most εt^2 pairs of the partition are not ε-regular. But the version we use is also common, and in fact the existence of such partitions can be deduced from the well-known 'degree form' of Szemerédi's Regularity Lemma, see e.g. [12, Theorem 1.10].

graph which respects the sizes of the various parts of T given by the fine partition. If T is in $\mathcal{T}^{\rho}_{a_1,a_2,b_1,b_2}$, the structure will incorporate the values a_1, a_2, b_1, b_2 to describe where the parts of the tree will be allocated.

In order to describe this structure precisely, we need some definitions. The basic building block is what we call a *skew fractional matching*, whose definition is inspired by the standard fractional matching.

Given a graph G, its associated digraph $G^{\leftrightarrow}$ is the digraph with the same vertex set as G and $\overrightarrow{uv}$ and $\overrightarrow{vu}$ are present for each undirected $uv \in E(G)$.

Definition 9. Let G be a graph and $\gamma \geq 0$.

(i) An *γ-skew fractional matching* is a function $\sigma : E(G^{\leftrightarrow}) \to [0,1]$ such that for any vertex $u \in V(G)$,

$$\frac{1}{1+\gamma} \sum_{v \in N_G(u)} \sigma(\overrightarrow{uv}) + \frac{\gamma}{1+\gamma} \sum_{v \in N_G(u)} \sigma(\overrightarrow{vu}) \leq 1.$$

The *weight* of σ is $W(\sigma) := \sum_{uv \in E(G)} \sigma(\overrightarrow{uv}) + \sigma(\overrightarrow{vu})$.

(ii) Set $\sigma^1(u) := \frac{1}{1+\gamma} \sum_{v \in N(u)} \sigma(\overrightarrow{uv})$ and $\sigma^2(u) := \frac{\gamma}{1+\gamma} \sum_{v \in N(u)} \sigma(\overrightarrow{uv})$. Abusing notation, we shall use the symbol σ to describe the *total charge* of σ on u, defined as $\sigma(u) := \sigma^1(u) + \sigma^2(u)$.

(iii) If σ, σ' are γ-skew and γ'-skew matchings respectively, we say σ, σ' are *disjoint* if, for every $u \in V(G)$, $\sigma(u) + \sigma'(u) \leq 1$.

Intuitively, γ-skew fractional matchings can be understood as fractional matchings in graphs where the weight of the edge is distributed in an unbalanced way, meaning that one end of the edge gets γ times the weight of the other end. Here, the direction of this imbalance is given by the direction of the edge in the digraph.

Definition 10. Let G be a graph, $\gamma \geq 0$, and $u \in V(G)$. We will say a γ-skew fractional matching σ is *anchored in* $N(u)$ if $\sigma^1(v) > 0$ implies that $v \in N(u)$.

Definition 11. Let G be a graph, $\gamma_A, \gamma_B > 0$. Given an edge $cd \in E(G)$, a *(γ_A, γ_B)-skew matching pair anchored in $\overrightarrow{cd}$* is a pair (σ_A, σ_B) such that

(i) σ_A and σ_B are disjoint,

(ii) σ_B is a γ_B-skew fractional matching anchored in $N(d)$, and

(iii) σ_A is a γ_A-skew fractional matching anchored in $N(c)$.

Lemma 12 (Structural Lemma: Simplified version)**.** *Let $k \in \mathbb{N}$ and let H be a graph such that $\delta(H) \geq k/2$ and $\Delta(H) \geq k$. Let $a_1, a_2, b_1, b_2 > 0$ such that $a_1 + a_2 + b_1 + b_2 = k$. Let $\gamma_A := \frac{a_2}{a_1}$ and $\gamma_B := \frac{b_2}{b_1}$. Then $H^{\leftrightarrow}$ admits a (γ_A, γ_B)-skew matching pair (σ_A, σ_B) anchored in some edge $\overrightarrow{cd} \in E(H^{\leftrightarrow})$ such that $W(\sigma_A) = a_1 + a_2$ and $W(\sigma_B) = b_1 + b_2$.*

2.3 Embedding the tree

Based on the structure given by the Structural Lemma, the next lemma ensures that the embedding of the tree T is possible.

Lemma 13 (Tree Embedding Lemma: Simplified version). *Let $1/n \ll \rho \ll 1/M, \varepsilon \ll d \ll \eta, q$. Suppose G is an n-vertex graph, that $\mathcal{P} = \{V_0, V_1, \dots, V_t\}$ is a ε-regular equitable partition of G with $t \le M$, and that Γ is the d-reduced graph obtained from G and $\mathcal{P}$. Suppose $k \ge qn$ and that we have numbers $a_1, b_1 \in \mathbb{N}$ and $\gamma_A, \gamma_B \ge 0$ such that $k = (1+\gamma_A)a_1 + (1+\gamma_B)b_1$. Suppose $\Gamma^{\leftrightarrow}$ admits a (γ_A, γ_B)-skew matching pair (σ_A, σ_B), anchored in $\overrightarrow{cd} \in E(\Gamma^{\leftrightarrow})$, with weights satisfying $W(\sigma_A)n \ge (1+\eta)(1+\gamma_A)a_1 t$ and $W(\sigma_B)n \ge (1+\eta)(1+\gamma_B)b_1 t$. Then G contains any k-vertex tree $T \in \mathcal{T}^{\rho}_{a_1, \gamma_A a_1, b_1, \gamma_B b_1}$.*

3 Proof of the main results

3.1 Proof of Theorem 3

Now we give the proof of Theorem 3, assuming the validity of the main lemmas (Lemma 12, Lemma 13).

Setting up the parameters. Suppose we are given input parameters $\eta > 0$, $q > 0$ and $k \ge qn$. We may assume that $q, \eta \ll 1$, or we just replace them with smaller values. We set the following parameters to satisfy

$$1/n \ll \rho \ll 1/M \ll \varepsilon \ll d \ll \eta, q \ll 1. \tag{1}$$

From now on we fix an arbitrary k-vertex tree T, and the goal is to show that $T \subseteq G$.

Processing the tree. By [11, Lemma 3.5], T has an $(\rho|V(T)|)$-fine partition. Let a_1, a_2, b_1, b_2 such that $T \in \mathcal{T}^{\rho}_{a_1,a_2,b_1,b_2}$. By assumption, G satisfies $\delta(G) \ge (1+\eta)k/2 = (1+\eta)|V(T)|/2$ and at least ηn vertices of G have degree at least $(1+\eta)k = (1+\eta)|V(T)|$.

Preparing the host graph. We apply Theorem 7 on G with parameters ε and $1/M$, and obtain an ε-regular equitable partition $\{V_0, V_1, \dots, V_t\}$, with $t \le M$. Given G and $\mathcal{P}$, let Γ be the d-reduced graph Γ. Standard arguments show that the reduced graph inherits degree properties of the original graph, up to a small loss. In particular, it can be shown that $\delta(\Gamma) \ge (1+\frac{\eta}{40})kt/(2n)$ and $\Delta(\Gamma) \ge (1+\frac{\eta}{40})kt/n$.

We apply Lemma 12 with Γ, $(1+\frac{\eta}{40})kt/n$, $(1+\frac{\eta}{40})a_i t/n$ and $(1+\frac{\eta}{40})b_i t/n$ playing the roles of H, k and a_i, b_i, for $i \in [2]$. This outputs an $(\frac{a_2}{a_1}, \frac{b_2}{b_1})$-skew matching pair (σ_A, σ_B) anchored in some edge $\overrightarrow{cd} \in E(\Gamma^{\leftrightarrow})$ with $W(\sigma_A) = (1+\frac{\eta}{40})(a_1+a_2)t/n$ and $W(\sigma_B) = (1+\frac{\eta}{40})(b_1+b_2)t/n$.

Embedding the tree. Finally, we can apply Lemma 13 with $a_2/a_1, b_2/b_1, \eta/40$ playing the roles of γ_A, γ_B, η respectively. This shows that $T \subseteq G$, as required. □

3.2 Proof of the approximate version of the Erdős–Sós conjecture

Now we derive Corollary 4 from Theorem 3. Let $k = rn$ with $r \geq q > 0$, and let G be a graph on n vertices with average degree at least $(1+\eta)k$. It is well-known [6, Proposition 1.2.2] that G contains an induced subgraph H such that $\delta(H) \geq d(H)/2 \geq d(G)/2 \geq (1+\eta)k/2$. Let m be the number of vertices of H, we clearly have $(1+\eta)k/2 \leq \delta(H) < m \leq n$.

For any $\lambda > 0$, let X_λ be the set of vertices of H whose degree in H is at least $(1+\lambda)k$. Then we have

$$(1+\eta)km \leq md(H) = \sum_{v \in V(H)} \deg_H(v) \leq |X_\lambda|m + (m - |X_\lambda|)(1+\lambda)k,$$

which, by rearranging, gives $|X_\lambda| \geq \frac{(\eta-\lambda)km}{m-(1+\lambda)k} \geq (\eta-\lambda)k$. From now on, fix $\lambda := \eta k/(m+k)$. This choice satisfies $\eta \geq \lambda \geq \eta r/(1+r)$, and from the previous calculations we deduce that H satisfies $\delta(H) > (1+\lambda)k/2$, and has at least $(\eta-\lambda)k = \lambda m$ vertices of degree at least $(1+\lambda)k$. Thus the statement follows by applying Theorem 3 to H, with λ, m playing the role of η and n, respectively. □

4 Final remarks

We stated our main technical lemmas (Lemma 12 and Lemma 13) in simplified versions which are enough to give a faithful version of the main ideas of our proof. In our actual proof, the statements are a bit more complicated since we need to consider *weighted* reduced graphs, where each edge $ij \in \Gamma$ receives a weight $d_{ij} \in [0,1]$ corresponding to the bipartite density $d(V_i, V_j)$ of the pair (V_i, V_j). Further details, and full proofs of the main lemmas, will be found in the full version of the paper [5].

Acknowledgment

We would like to thank Jan Hladký for drawing our attention to the fact that Theorem 3 implies the approximate Erdős-Sós Conjecture for dense graphs (Corollary 4).

References

[1] M. Ajtai, J. Komlós, M. Simonovits, and E. Szemerédi. On the approximative solution of the Erdős-Sós conjecture on trees. In preparation.

[2] M. Ajtai, J. Komlós, M. Simonovits, and E. Szemerédi. The solution of the Erdős-Sós conjecture for large trees. In preparation.

[3] M. Ajtai, J. Komlós, M. Simonovits, and E. Szemerédi. Some elementary lemmas on the Erdős-Sós conjecture on trees. In preparation.

[4] Guido Besomi, Matías Pavez-Signé, and Maya Stein. On the Erdős-Sós conjecture for trees with bounded degree. *Combin. Probab. Comput.*, 30(5):741–761, 2021.

[5] Akbar Davoodi, Diana Piguet, Hanka Řada, and Nicolás Sanhueza-Matamala. The asymptotic version of the Erdős–Sós conjecture and beyond. In preparation.

[6] Reinhard Diestel. *Graph theory*, volume 173 of *Graduate Texts in Mathematics*. Springer, Berlin, fifth edition edition, 2017.

[7] P. Erdős. Extremal problems in graph theory. In *Theory of Graphs and its Applications (Proc. Sympos. Smolenice, 1963)*, pages 29–36. Publ. House Czech. Acad. Sci., Prague, 1964.

[8] P. Erdős and M. Simonovits. A limit theorem in graph theory. *Studia Sci. Math. Hungar.*, 1:51–57, 1966.

[9] P. Erdös and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946.

[10] Frédéric Havet, Bruce Reed, Maya Stein, and David R. Wood. A variant of the Erdös-Sós conjecture. *J. Graph Theory*, 94(1):131–158, 2020.

[11] Jan Hladký, János Komlós, Diana Piguet, Miklós Simonovits, Maya Stein, and Endre Szemerédi. The approximate Loebl-Komlós-Sós conjecture IV: Embedding techniques and the proof of the main result. *SIAM J. Discrete Math.*, 31(2):1072–1148, 2017.

[12] J. Komlós and M. Simonovits. Szemerédi's regularity lemma and its applications in graph theory. In *Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993)*, volume 2 of *Bolyai Soc. Math. Stud.*, pages 295–352. János Bolyai Math. Soc., Budapest, 1996.

[13] Václav Rozhoň. A local approach to the Erdős-Sós conjecture. *SIAM J. Discrete Math.*, 33(2):643–664, 2019.

[14] Maya Stein. Tree containment and degree conditions. In *Discrete Mathematics and Applications*, volume 165 of *Springer Optim. Appl.*, pages 459–486. Springer, Cham, 2020.

[15] Endre Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Dispersion on the Complete Graph

(Extended abstract)

Umberto De Ambroggio* Tamás Makai* Konstantinos Panagiotou*

Abstract

We consider a synchronous process of particles moving on the vertices of a graph G, introduced by Cooper, McDowell, Radzik, Rivera and Shiraga (2018). Initially, M particles are placed on one vertex of G. At the beginning of each time step, for every vertex inhabited by at least two particles, each of these particles moves independently to a neighbour chosen uniformly at random. The process ends at the first step when no vertex is inhabited by more than one particle.

Cooper et al. showed that when the underlying graph is the complete graph on n vertices, then there is a phase transition when the number of particles $M = n/2$. They showed that if $M < (1-\varepsilon)n/2$ for some fixed $\varepsilon > 0$, then the process finishes in a logarithmic number of steps, while if $M > (1+\varepsilon)n/2$, an exponential number of steps are required with high probability. In this paper we provide a thorough analysis of the distribution of the dispersion time in the barely critical regime, where $\varepsilon = o(1)$, and describe the fine details of the transition between logarithmic and exponential time. As a consequence of our results we establish, for example, that the dispersion time is in probability and in expectation $\Theta(n^{1/2})$ when $|\varepsilon| = O(n^{-1/2})$, and provide qualitative bounds for its tail behavior.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-046`

1 Introduction

We consider the synchronous *dispersion process* introduced by Cooper, McDowell, Radzik, Rivera and Shiraga [1]. The process evolves in discrete time. It involves particles that move between vertices of a given graph G. A particle is called *happy*, if there are no other

*University of Munich, Department of Mathematics, Theresienstr. 39, 80333 Munich, Germany. E-mail: `{deambrog,makai,kpanagio}@math.lmu.de`. Supported by ERC Grant Agreement 772606-PTRCSP.

particles on the same vertex, otherwise it is *unhappy*. Initially, M particles are placed on some vertex of G. In every (discrete) time step, all unhappy particles move simultaneously and independently to a neighbouring vertex selected uniformly at random. Happy particles do not move. The process terminates at the first time at which all particles are happy. This (random) time is denoted by $T_{G,M}$ and it is called the *dispersion time*; it constitutes the main object of interest here.

In [1] the authors studied this process on several graphs, and established results concerning $T_{G,M}$ and the *dispersion distance*, which is the maximum distance of any particle from the origin at dispersion (that is, at step $T_{G,M}$). One of the main focus in [1] is the behaviour when the underlying graph is the complete graph with loops, which we will denote by K_n. The most general results come from considering a *lazy* variant of the dispersion process, which was shown to disperse more particles in a smaller number of steps. More precisely, in this lazy version any unhappy particle moves with probability $q \in (0, 1]$ and stays at its current location with probability $1 - q$.

The main result of [1] regarding $T_{K_n,M}$ is that there are constants $c, C > 0$ such that if $M = (1 - q/2 - \alpha)n$ for any $\alpha > 0$ that may depend on n, then

$$T_{K_n,M} \leq C(q\alpha)^{-1} \log(n) \text{ with probability at least } 1 - O(1/n), \tag{1}$$

whereas when $M = n(1 - q/2 + \alpha)$, then

$$T_{K_n,M} \geq e^{cnq^2\alpha^3} \text{ with probability at least } 1 - e^{-cnq^2\alpha^3}. \tag{2}$$

The above statements leave several questions open. Indeed, corresponding bounds for the lower and upper tails of $T_{K_n,M}$ were not provided. It is not clear, moreover, what the actual behavior is when M is close to $n/2$, that is, when $M = (1 + \varepsilon)n/2$ for some $|\varepsilon| = o(1)$ and how the transition from logarithmic to exponential time quantitatively looks like. For example, (2) is not informative when $q = 1$ and $\alpha = o(n^{-1/3})$, as it essentially only states that the number of steps is at least one.

Since we deal exclusively with the complete graph, in the following we will write $T_{n,M} = T_{K_n,M}$. Our main contribution is a thorough and precise analysis of the dispersion process when, as above, we assume that $M = (1 + \varepsilon)n/2$ and $|\varepsilon| = o(1)$. Then we establish that the process exhibits three qualitatively different behaviours based on the asymptotics of ε, where, informally speaking, $T_{n,M}$ smoothly changes from $|\varepsilon|^{-1} \log(\varepsilon^2 n)$ to $n^{1/2}$ and then to $\varepsilon^{-1}e^{\Theta(\varepsilon^2 n)}$; in particular, $T_{n,M} = \Theta(n^{1/2})$ for $M = n/2$. We begin with providing the upper bounds on the distribution of $T_{n,M}$.

Theorem 1.1. *There is a $C > 0$ such that the following is true for sufficiently large n and all $A \geq 1$. Let $\varepsilon = o(1)$ and $M = (1 + \varepsilon)n/2$. If $\varepsilon < -en^{-1/2}$, then*

$$\mathbb{P}\big(T_{n,M} > AC|\varepsilon|^{-1} \log(\varepsilon^2 n)\big) \leq e^{-(A-1)}.$$

Moreover, if $|\varepsilon| \leq en^{-1/2}$, then

$$\mathbb{P}\big(T_{n,M} > ACn^{1/2}\big) \leq e^{-(A-1)}.$$

Finally, if $\varepsilon > en^{-1/2}$*, then*

$$\mathbb{P}\big(T_{n,M} > A\varepsilon^{-1}e^{C\varepsilon^2 n}\big) \le e^{-(A-1)}.$$

The next main result establishes lower bounds for these ranges as well. When $\varepsilon \le en^{-1/2}$ these match, and in the last case we show that the exponential term is of the same order.

Theorem 1.2. *There is a* $c > 0$ *such that the following is true for sufficiently large* n *and all* $A \ge 1$*. Let* $|\varepsilon| = o(1)$ *and* $M = (1+\varepsilon)n/2$*. If* $\varepsilon < -en^{-1/2}$*, then*

$$\mathbb{P}\big(T_{n,M} \le c|\varepsilon|^{-1}\log(\varepsilon^2 n)/A\big) \le A^{-1}.$$

Moreover, if $|\varepsilon| \le en^{-1/2}$*, then*

$$\mathbb{P}\big(T_{n,M} \le cn^{1/2}/A\big) \le A^{-1}.$$

Finally, if $\varepsilon > en^{-1/2}$*, then*

$$\mathbb{P}\big(T_{n,M} \le \max\{e^{c\varepsilon^2 n}, c\varepsilon^{-1}/A\}\big) \le \min\{e^{-c\varepsilon^2 n}, A^{-1}\}.$$

Let us discuss briefly some consequences of our results. First of all, the two theorems combined imply that in probability

$$T_{n,M} = \Theta(|\varepsilon|^{-1}\log(\varepsilon^2 n)) \quad \text{if} \quad \varepsilon < -en^{-1/2},$$

and

$$T_{n,M} = \Theta(n^{1/2}) \quad \text{if} \quad |\varepsilon| = O(n^{-1/2}).$$

In particular, when $M = n/2$ we obtain that $T_{n,M} = \Theta(n^{1/2})$ in probability. For larger ε, we obtain the slightly weaker uniform estimate that in probability

$$\log(T_{n,M}) = \Theta(\varepsilon^2 n + \log n) \quad \text{if} \quad \varepsilon = \omega(n^{-1/2}).$$

This estimate can be improved as soon as $e^{c\varepsilon^2 n} \ge \varepsilon^{-1}$, that is, when $\varepsilon = \Omega((\log n/n)^{1/2})$; after this point the maximum in Theorem 1.2 will be $e^{c\varepsilon^2 n}$ and so, in fact, for such ε we obtain that even $\log(T_{n,M}) = \Theta(\varepsilon^2 n)$ in probability.

Apart from these estimates we can also use our main theorems to obtain information about, for example, the expectation of $T_{n,M}$. In particular, Theorem 1.1 guarantees that $T_{n,M}$ has an exponential(-ly thin) upper tail and so $T_{n,M}$ is integrable; we readily obtain that

$$\mathbb{E}[T_{n,M}] = \Theta(|\varepsilon|^{-1}\log(\varepsilon^2 n)) \text{ if } \varepsilon \le -en^{-1/2}, \qquad \mathbb{E}[T_{n,M}] = \Theta(n^{1/2}) \text{ if } |\varepsilon| = O(n^{-1/2}),$$

and

$$\log \mathbb{E}[T_{n,M}] = \Theta(\varepsilon^2 n + \log n) \text{ if } \varepsilon = \omega(n^{-1/2}).$$

Further Related work The dispersion process was also studied by Frieze and Pegden [3], who sharpened the result on the dispersion distance on L_∞, which denotes the infinite line. In particular, it was shown in [1] that with high probability, the dispersion distance on L_∞ for n particles is $O(n \log n)$; in [3] the logarithmic factor was eliminated. A similar setup was considered by Shang [8], where the author studied the dispersion distance in a non-uniform dispersion process in which an unhappy particle moves at the next time step to the right with probability p_n and to the left with probability $1 - p_n$, independently of other particles.

Processes where particles move on the vertices of a graph have been widely studied over the past decades; we refer the reader to [1] for references. Concerning processes whose scope is to *disperse* particles on a discrete structure, arguably the best known such model is the Internal Diffusion Limited Aggregation (IDLA, for short); see [2] and [4]. In this model, particles sequentially start (one at a time) from a specific vertex designated as the origin. Each particle moves randomly until it finds an unoccupied vertex; then it occupies it forever (meaning that it does not move at subsequent process steps). When a particle stops, the next particle starts moving. We emphasize that whenever a particle jumps to an occupied vertex, it just keeps moving without activating the occupant particle. In the dispersion process, on the other hand, when a (happy) particle standing alone on a node is reached by another particle, it is reactivated and keeps moving until it becomes happy again.

2 Proof Ideas

In the proof we begin with studying the expected change in the number of unhappy particles in every step. Let us write H_t and U_t for the number of happy and unhappy particles at the beginning of step t; in particular, $U_0 = M$ and $H_0 = 0$ and $U_t + H_t = M$ for all $t \in \mathbb{N}_0$. Then it turns out that

$$\mathbb{E}\left[U_{t+1} | U_t\right] = H_t \left(1 - \left(1 - \frac{1}{n}\right)^{U_t}\right) + U_t \left(1 - \frac{n - H_t}{n} \left(1 - \frac{1}{n}\right)^{U_t - 1}\right), \quad t \in \mathbb{N}_0. \tag{3}$$

The two summands correspond to the number of particles counted in H_t that become unhappy and to the number of particles counted in U_t that remain unhappy in step $t+1$. Recall that we write $M = (1+\varepsilon)n/2$ and assume that U_t is not too big, say $U_t \le \delta n$ for some small $\delta > 0$. Then a quick calculation reveals that

$$\mathbb{E}\left[U_{t+1} | U_t\right] = (1+\varepsilon)U_t - \Theta(U_t^2/n). \tag{4}$$

So, as long as U_t is (much) larger than $|\varepsilon|n$, then U_{t+1} will be (much) smaller than U_t in expectation. In other words, when there are 'many' unhappy particles, the expected number of unhappy particles in the next step decreases significantly. However, this is no longer the case when there are only a 'few' unhappy particles, that is, less than $O(|\varepsilon|n)$. In this case the number of unhappy particles is expected to either decrease only by a slight amount, which in particular is problematic when the expected decrease $\mathbb{E}[U_{t+1} - U_t | U_t]$ is smaller than one, or when $\varepsilon > 0$, where we can expect that the number of unhappy particles even *increases*. We will use different methods to analyse the trajectory of U_t depending on the range of ε and whether we are considering an upper or a lower bound. In particular, when $|\varepsilon|$ is not too large we will see that we can compare the situation to a very slightly biased random walk, and so we will end up with an $n^{1/2}$ term, while in the other cases our walk will have a positive/negative drift and the exponential/logarithmic term will emerge.

2.1 Upper tail

Our approach for establishing Theorem 1.1 is to find a lower bound on the probability that when starting with an *arbitrary* number of unhappy particles, the process will stop within a certain number of steps. By splitting the time interval under consideration into a disjoint union of smaller intervals and using the Markov property of the process, we can apply the above bound repeatedly on these smaller sections to achieve an exponentially decreasing upper bound on the probability that the process is still not finished.

When there are many unhappy particles we use drift analysis to analyse the process; we refer the reader to [5] for an excellent introduction and description of the method. Roughly speaking, drift analysis provides an estimate for the expected duration of a homogeneous Markov-process over a discrete state space, when the expected value of the conditional one step change is known for every element in the state space. With (4) at hand we can apply the method to deduce bounds for the probability that dispersion leaves us with many unhappy particles after a certain number of steps.

Once there are only a few unhappy particles left we change our approach. After this point we bound from above the number of unhappy particle with another random process, which we call the *binomial process*. More specifically, beginning with some initial value B_0, we define a random process by setting $B_{t+1} = 2\mathrm{Bin}(B_t, M/n)$, $t \in \mathbb{N}_0$. The quantity B_t provides an upper bound for the number of unhappy particles after t steps, as the probability that an unhappy particle lands on the same vertex as any other particle is at most M/n, and in that case we account for two unhappy particles. As the number of unhappy particles is small, it is rare for two unhappy particles to land on the same vertex, making this coupling relatively tight.

The binomial process is equivalent to B_0 independent copies of a Galton-Watson branching process that have no offspring with probability $1 - M/n$ and two offspring with probability M/n. A simple inductive argument implies that the size of the k-th generation of these B_0 branching processes has the same distribution as B_k.

In the next step we estimate the probability that a single copy of the branching process survives for at least k generations; denote this probability by x_k. Then $x_0 = 1$ and moreover

$$x_{k+1} = \frac{M}{n}\left(2x_k - x_k^2\right),$$

as in order for the branching process to survive for $k+1$ generations, the root has to have 2 children, and at least one of these children has to survive for at least k generations.

Recall the well-known property concerning the survival probability of a Galton-Watson branching process, namely that it is 0 if the expected number of offspring is at most 1, and it is bounded away from 0 when the expected offspring is larger than 1 (see e.g. [7]). Clearly x_k tends to the survival probability as $k \to \infty$, and moreover, $\lim_{k\to\infty} x_k = o(1)$ as $n \to \infty$, as the expected number of children is ~ 1.

So far we have referred to *many* and *few* unhappy particles, without mentioning the level where the change occurs. The exact value is determined by carefully balancing several properties. For the branching process argument, we would like that the number of steps is such that x_k is sufficiently close to its limit. In addition we would like the number of steps that we study using the branching process to match the number of steps we analyse using drift analysis, so as to reach an optimal bound. This leads to two different regimes, namely

- roughly $n^{1/2}$ steps, when $|\varepsilon| \leq en^{-1/2}$;
- roughly $|\varepsilon|^{-1}$ steps, when $|\varepsilon| > en^{-1/2}$,

which coincide with the three regimes in the main theorem.

The branching process behaves similarly in all regimes, as the survival probability is $o(1)$. However, the rate of convergence in n becomes slower as ε increases and thus the probability that all independent copies of this branching process die out (in the required number of steps) goes from almost certain, when $\varepsilon < en^{-1/2}$ to a constant when $|\varepsilon| \leq en^{-1/2}$, to exponentially decreasing when $\varepsilon > en^{-1/2}$. Recall that we use these probabilities as the basis of a geometric distribution, which leads to the upper bounds in Theorem 1.1.

2.2 Lower tail

Now we consider the lower bound. In this case we first show that we can mostly ignore what happens when the number of unhappy particles is large, the only exception is that we have to ensure that it is unlikely that most of the unhappy particles become happy in any single step. In order to achieve this, note that U_{t+1} is a function of U_t and the vertices to which the unhappy particles counted in U_t jump to. Then we consider the Doob martingale induced by exposing the individual destinations (of the unhappy particles) one at a time and show concentration around its expectation with Azuma-Hoeffding, which yields the desired property that the number of unhappy particles does not decrease too quickly.

Having established this, we proceed similarly to the analysis for the upper tail. Let $B'_0 \in \mathbb{N}$ and consider the binomial process with different parameters defined by

$$B'_{t+1} = 2\mathrm{Bin}(B'_t, (M-2K)/n), \quad t \in \mathbb{N},$$

where $K \in \mathbb{N}$ is arbitrary but fixed. Then, as long as $U_t \leq K$, we have that B'_t is a lower coupling for U_t, as there are at least $M - 2K$ happy particles at the end of the corresponding step. We can analyse the binomial process using branching processes as for the upper bound, providing a lower bound with the right order of magnitude when

$\varepsilon \leq en^{-1/2}$. However, when $\varepsilon > en^{-1/2}$, due to the $-2K$ term the associated branching process becomes subcritical, and hence we are only able to show that the process still runs after $\Omega(\varepsilon^{-1})$ steps, which provides the corresponding term in the max in Theorem 1.2.

In order to obtain the term involving the exponential in $\varepsilon^2 n$, we use an alternative approach, adapting the argument of Theorem 2.6 in Lengler and Steger [6]. Note that this term only appears when $\varepsilon > en^{-1/2}$, thereby whenever U_t is small, (4) indicates that in the following step U_t will increase in expectation. In such a case it is unlikely that U_t decreases, and consequently many steps are required before the prcess stops.

References

[1] C. Cooper, A. McDowell, T. Radzik, N. Rivera, and T. Shiraga. Dispersion processes. *Random Structures Algorithms*, 53(4):561–585, 2018.

[2] P. Diaconis and W. Fulton. A growth model, a game, an algebra, Lagrange inversion, and characteristic classes. *Rend. Sem. Mat. Univ. Politec. Torino*, 49(1):95–119, 1991.

[3] A. Frieze and W. Pegden. A note on dispersing particles on a line. *Random Structures Algorithms*, 53(4):586–591, 2018.

[4] G. F. Lawler, M. Bramson, and D. Griffeath. Internal diffusion limited aggregation. *Ann. Probab.*, 20(4):2117–2140, 1992.

[5] J. Lengler. Drift analysis. In B. Doerr and F. Neumann, editors, *Theory of Evolutionary Computation: Recent Developments in Discrete Optimization*, Nat. Comput. Ser., pages 89–131. Springer, 2020.

[6] J. Lengler and A. Steger. Drift analysis and evolutionary algorithms revisited. *Combin. Probab. Comput.*, 27(4):643–666, 2018.

[7] R. Lyons and Y. Peres. *Probability on trees and networks*, volume 42 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, New York, 2016.

[8] Y. Shang. Longest distance of a non-uniform dispersion process on the infinite line. *Inform. Process. Lett.*, 164:106008, 5, 2020.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

The root cluster after percolation on preferential attachment trees

(Extended abstract)

Colin Desmarais* Cecilia Holmgren* Stephan Wagner*,†

Abstract

The class of linear preferential attachment trees includes recursive trees, plane-oriented recursive trees, binary search trees, and increasing d-ary trees. Bond percolation with parameter p is performed by considering every edge in a graph independently, and either keeping the edge with probability p or removing it otherwise. The resulting connected components are called clusters. In this extended abstract, we demonstrate how to use methods from analytic combinatorics to compute limiting distributions, after rescaling, for the size of the cluster containing the root. These results are part of a larger work on broadcasting induced colorings of preferential attachment trees.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-047`

1 Introduction

For a real number α, a sequence of random (linear) preferential attachment trees $(\mathcal{T}_n)_{n=1}^{\infty}$ is constructed in the following manner. The tree $\mathcal{T}_1$ consists of a single vertex labelled 1. For $n \geq 1$, a vertex v is chosen from $\mathcal{T}_n$ with probability

$$\frac{\alpha \deg^+(v) + 1}{\sum_{u \in V(\mathcal{T}_n)} (\alpha \deg^+(u) + 1)} = \frac{\alpha \deg^+(v) + 1}{\alpha(n-1) + n}, \tag{1}$$

*Department of Mathematics, Uppsala University, Uppsala, Sweden. E-mail: `colindesmarais@gmail.com`, `{cecilia.holmgren, stephan.wagner} @math.uu.se`
†Department of Mathematical Sciences, Stellenbosch University, Stellenbosch, South Africa

where $\deg^+(u)$ (called the *outdegree* of u) is the number of children of u. A vertex labelled $n+1$ is then added as a child of v to construct $\mathcal{T}_{n+1}$. The equality above holds since $\sum_{u\in V(\mathcal{T}_n)} \deg^+(u) = n-1$, the number of edges in the tree $\mathcal{T}_n$. The parameter α is restricted to $\{\dots, -\frac{1}{4}, -\frac{1}{3}, -\frac{1}{2}\} \bigcup [0,\infty)$ to avoid the degenerate case $\alpha = -1$ that produces a path, and to avoid the cases where (1) might be negative.

We then perform bond percolation with parameter p on preferential attachment trees. Each edge is considered independently and either kept (or remains *open*) with probability p or removed (*closed*) with probability $1-p$. We then ask how many vertices remain in the component (or *cluster*) $\mathcal{C}_n$ that contains the root. For $\alpha \geq 0$, Baur proved that $|\mathcal{C}_n|/n^{(p+\alpha)/(1+\alpha)}$ converges in distribution to a random variable $\mathcal{C}$ and provided the first two moments [2]. In the special case $\alpha = 0$, it is already known that $\mathcal{C}$ is a Mittag-Leffler random variable [3,4]. We extend these results by showing that $|\mathcal{C}_n|/n^{(p+\alpha)(1+\alpha)}$ converges in distribution for $p > -\alpha$ and provide a recursion for the moments. This recursion uses the partial Bell polynomials (see [5, Chapter 3.3])

$$B_{k,j}(x_1,\dots,x_{k-j+1}) = \sum_{\substack{m_1+\dots+(k-j+1)m_{k-j+1}=k \\ m_1+\dots+m_{k-j+1}=j}} k! \prod_{i=1}^{k-j+1} \frac{x_i^{m_i}}{m_i! i!^{m_i}}.$$

Theorem 1.1. *Let $\alpha > 0$. Then $|\mathcal{C}_n|/n^{(p+\alpha)/(1+\alpha)} \xrightarrow{d} \mathcal{C}$, where $\mathcal{C}$ has integer moments*

$$\mathbb{E}[\mathcal{C}^k] = \frac{C_k(1+\alpha)\Gamma(1/(1+\alpha))}{\alpha\Gamma((kp+\alpha(k-1))/(\alpha+1))},$$

where C_k satisfies the recursion $C_1 = \alpha/(p+\alpha)$ and

$$(k-1)(p/\alpha+1)C_k = \sum_{j=2}^{k} \frac{p^j\Gamma(1/\alpha+j)}{\Gamma(1/\alpha)} B_{k,j}(C_1,\dots,C_{k-j+1}). \tag{2}$$

When $\alpha = 1$ the above recursion is used to express the moments of $\mathcal{C}$ in a closed form

$$\mathbb{E}[\mathcal{C}^k] = \frac{2p^{k-1}\Gamma(kp+k-1)\sqrt{\pi}}{(p+1)^{2k-1}\Gamma(kp)\Gamma((kp+k-1)/2)}.$$

Theorem 1.2. *Let $\alpha = -1/d$, where $d \geq 2$ is a positive integer, and let $p > -\alpha$. Then $|\mathcal{C}_n|/n^{(pd-1)/(d-1)} \xrightarrow{d} \mathcal{C}$, where $\mathcal{C}$ has integer moments*

$$\mathbb{E}[\mathcal{C}^k] = \frac{D_k\Gamma(1/(d-1))}{\Gamma((kpd-k+1)/(d-1))},$$

where D_k satisfies the recursion $D_1 = 1/(pd-1)$ and

$$(k-1)(pd-1)D_k = \sum_{j=2}^{\min\{k,d\}} \frac{p^j d!}{(d-j)!} B_{k,j}(D_1,\dots,D_{k-j+1}).$$

In the special case $\alpha = -1/2$ (so $\mathcal{T}_n$ is a random binary search tree) the recursion above produces the closed form

$$\mathbb{E}[\mathcal{C}^k] = \frac{k! p^{2(k-1)}}{(2p-1)^{2k-1}\Gamma(k(2p-1)+1)}.$$

When $p \leq -\alpha$, we are also able to prove that $|\mathcal{C}_n|$ is almost surely finite and converges to a Galton-Watson tree with binomial $\mathrm{Bin}(d,p)$ offspring distribution.

The results presented in this extended abstract are part of a larger work on broadcasting-induced colorings on preferential attachment trees [6]. A broadcasting process on a tree is performed by assigning the root either the bit 1 or 0 with equal probability. Starting with the children of the root, every other vertex takes the same bit as its parent with probability p and the other bit with probability $1-p$. The reconstruction problem is then to reconstruct the bit value of the root ρ from the bit values of some subset of vertices in T after broadcasting. Addario-Berry, Devroye, Lugosi, and Velona studied the reconstruction problem in random recursive trees and preferential attachment trees [1].

We then colour a vertex red if its bit value is 0 and blue if its bit value is 1 to obtain a broadcasting induced coloring. If we remove edges between vertices with different colours we are left with a forest of trees corresponding to clusters after performing bond percolation with parameter p. Along with results on the size of the root cluster, we also prove in [6] limiting distributions after rescaling for the number of vertices, clusters, and leaves of each colour, as well as the number of fringe subtrees with two-colorings.

2 Proof outline

Consider the function

$$\phi(\delta) = \begin{cases} 1 & \alpha = 0, \\ \frac{\Gamma(\delta+1/\alpha)}{\Gamma(1/\alpha)} & \alpha > 0, \\ \frac{d!}{(d-\delta)!} & \alpha = -\frac{1}{d}, d \in \mathbb{Z}^+. \end{cases} \tag{3}$$

For a particular rooted labelled tree T on n vertices with increasing labels, define the weight of T to be $w(T) = \prod_{v \in V(T)} \phi(\deg^+(v))$. Then the probability that the recursive process described above produces T is given by $\mathbb{P}(\mathcal{T}_n = T) = w(T)/\sum_{T'} w(T')$, where the sum is taken over all rooted labelled trees T on n vertices with increasing labels. Letting b_n be the denominator $\sum_{T'} w(T')$, the exponential generating function for b_n is given by $B(x) = \sum_{n \geq 1} \frac{b_n}{n!} x^n = -\ln(1-x)$ when $\alpha = 0$, $B(x) = 1 - (1-(1+1/\alpha)x)^{\frac{\alpha}{1+\alpha}}$ when $\alpha > 0$, and $B(x) = (1-(d-1)x)^{-\frac{1}{d-1}} - 1$ when $\alpha = -\frac{1}{d}, d = 2, 3, 4, \ldots$. See [7] for detailed derivations of $B(x)$.

For a particular tree T with bond percolation performed, let $\mathcal{C}(T)$ be the cluster that contains the root and let $|\mathcal{C}(T)|$ be the number of vertices in $\mathcal{C}(T)$. Define $r_{n,k} = \sum_{T:|T|=n} w(T)\mathbb{P}(|\mathcal{C}(T)| = k)$. Then $\sum_{k=1}^n r_{n,k} = b_n$, and $\mathbb{P}(|\mathcal{C}_n| = k) = r_{n,k}/b_n$.

We develop a recursion for $r_{n,k}$. Take any tree T on n vertices with increasing labels whose root cluster after percolation has size k. Let δ be the outdegree of the root and let $T_1, \cdots, T_\delta$ be the subtrees of T rooted at the children of the root such that the edges from the root to $T_1, \ldots, T_s$ are open and the edges from the root to $T_{s+1}, \ldots, T_\delta$ are closed. Then

$$w(T)\mathbb{P}(|\mathcal{C}(T)| = k) = \phi(\delta) \prod_{i=1}^{s} pw(T_i)\mathbb{P}(|\mathcal{C}(T_i)| = k_i) \prod_{j=s+1}^{\delta} (1-p)w(T_j),$$

where $k_1 + \cdots + k_s = k-1$. If the trees $T_1, \ldots, T_\delta$ are of size $n_1, \ldots, n_\delta$ then $n_1 + \cdots + n_\delta = n-1$. Summing over all such trees T on n vertices with root cluster of size k, we get the recursion

$$r_{n,k} = \sum_{\delta}^{n-1} \sum_{s=0}^{\delta} \binom{\delta}{s} \frac{\phi(\delta)}{\delta!} \sum_{n_1,\ldots,n_\delta} \binom{n-1}{n_1, \ldots, n_\delta} \sum_{k_1,\ldots,k_s} \prod_{i=1}^{s} pr_{n_i,k_i} \prod_{j=s+1}^{\delta} (1-p)b_{n_j}.$$

Let $R(x,u)$ be the bivariate (exponential) generating function for $r_{n,k}$, so $R(x,u) = \sum_{n,k\geq 1} = \frac{r_{n,k}}{n!} x^n u^k$. To use the moment of methods to prove that $|\mathcal{C}_n|$ converges in distribution (after rescaling) to a random variable $\mathcal{C}$, we first extract the factorial moments of $|\mathcal{C}_n|$ from the generating function $R(x,u)$. The next steps are to show that after proper rescaling, the factorial moments and integer moments conincide asymptotically and converge to the moments of a random variable $\mathcal{C}$, and to prove that $\mathcal{C}$ is uniquely determined by its moments.

Let $R_k(x) := \frac{\partial^k}{\partial u^k} R(x,u)\Big|_{u=1}$. Then from standard methods (see for example [8, Proposition III.2]), the k'th factorial moment of $|\mathcal{C}_n|$ is given by $[x^n]R_k(x)/[x^n]R(x,1)$. To study $R_k(x)$, we first use our recursion to establish the differential equation

$$\frac{\partial}{\partial x} R(x,u) = u \sum_{\delta=0}^{\infty} \frac{\phi(\delta)}{\delta!} \left(pR(x,u) + (1-p)B(x)\right)^\delta. \tag{4}$$

When $\alpha = 0$ the above differential equation simplifies to

$$\frac{\partial}{\partial x} R(x,u) = u \exp\left(pR(x,u) - (1-p)\ln(1-x)\right),$$

and with the initial condition $R(0,u) = 0$, this linear differential equation has the solution $R(x,u) = \frac{-1}{p} \ln\left(1 - u + u(1-x)^p\right)$. Then $R_k(x) = \frac{1}{p}(k-1)!((1-x)^{-p} - 1)^k$, and so the factorial moments of $|\mathcal{C}_n|$ are given by $(k-1)!n^{pk}/(p\Gamma(pk)$. Once divided by n^{pk}, the factorial moments and integer moments conincide asymptotically and

$$\mathbb{E}\left[\frac{|\mathcal{C}_n|^k}{n^{pk}}\right] \sim \frac{(k-1)!n^{pk}}{n^{pk}p\Gamma(pk)} = \frac{k!}{\Gamma(pk+1)},$$

which are the moments of the Mittag-Leffler distribution with paramter p, a distribution uniquely determined by its moments. Therefore, we have that $|\mathcal{C}_n|/n^p$ converges in distribution to a random variable with Mittag-Leffler distribution.

For other cases of α, we were unable to find a closed form for $R(x,u)$. However, we only need to estimate $R_k(x)$ to extract approximations for the factorial moments of $|\mathcal{C}_n|$. When $\alpha > 0$, the differential equation 4 becomes

$$\frac{\partial}{\partial x} R(x,u) = u\left(1 - \left(pR(x,u) + (1-p)\left(1 - (1-(1+1/\alpha)x)^{\alpha/(1+\alpha)}\right)\right)\right)^{-1/\alpha}.$$

By differentiating both sides k times with respect to u and evaluating at $u = 1$, we achieve a differential equation for $R_k(x)$. With the help of induction we prove the following:

Lemma 2.1. *Let $\alpha > 0$. Then $R_k(x)$ is analytic on the cut plane $\mathbb{C}\setminus[1/(1+1/\alpha),\infty)$ and*

$$R_k(x) = C_k(1-(1+1/\alpha)x)^{\frac{-kp-\alpha(k-1)}{1+\alpha}} + O\left((1-(1+1/\alpha)x)^{\frac{-kp-\alpha(k-1)}{1+\alpha}+\varepsilon}\right)$$

for some $\varepsilon > 0$, where C_k satisfies the recursion given in (2).

Using a transfer theorem (see [8, Corollary VI.1]), we can approximate the coefficients of $R_k(x)$. Using [8, Proposition III.2] again, we extract the k'th factorial moments of $|\mathcal{C}_n|$. After rescaling by $n^{(p+\alpha)/(1+\alpha)}$, the factorial and integer moments coincide asymptotically and

$$\mathbb{E}\left[\frac{|\mathcal{C}_n|^k}{n^{(p+\alpha)/(1+\alpha)}}\right] \to \frac{C_k(1+\alpha)\Gamma(1/(1+\alpha))}{\alpha\Gamma((kp+\alpha(k-1))/(\alpha+1))}. \tag{5}$$

All that is left is to prove that the limiting distribution is determined by its moments. We can prove that the exponential generating function for the coefficients C_k above exists for a positive radius around 0. The exponential generating function generated by the limiting integer moments in (5) has a larger radius of convergence. Therefore the distribution $\mathcal{C}$ with integer moments given by (5) has a moment generating function that exists for a positive radius, and so $\mathcal{C}$ is uniquely determined by its moments. The moment of methods can therefore be applied to prove Theorem 1.1. Theorem 1.2 is proved in a similar manner.

References

[1] L. Addario-Berry, L. Devroye, G. Lugosi, and V. Velona, Broadcasting on random recursive trees, *Ann. Appl. Probab.* **32** (2022), 497–528.

[2] E. Baur, On a class of random walks with reinforced memory, *J. Stat. Phys.* **181** (2020), 772–802.

[3] E. Baur and J. Bertoin, The fragmentation process of an infinite recursive tree and Ornstein-Uhlenbeck type processes, *Electron. J. Probab.* **20** (2015), 1–20.

[4] S. Businger, The shark random swim, *J. Stat. Phys.* **172** (2018), 701–717.

[5] L. Comtet, *Advanced combinatorics. The art of finite and infinite expansions*, D. Reidel Publishing Co., 1974.

[6] C. Desmarais, C. Holmgren, and S. Wagner, Broadcasting-induce colorings of preferential attachment trees, *Random Structures & Algorithms* (2023), available online `https://doi.org/10.1002/rsa.21142`, 42pp.

[7] M. Drmota, *Random trees: an interplay between combinatorics and probability*, Springer Science and Business Media, 2009.

[8] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, 2009.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Cycles through two edges in signed graphs

(Extended abstract)

Matt DeVos* Kathryn Nurse†

Abstract

We give a characterization of when a signed graph G with a pair of distinguished edges $e_1, e_2 \in E(G)$ has the property that all cycles containing both e_1 and e_2 have the same sign. This answers a question of Zaslavsky.

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-048

1 Introduction

Throughout we assume (signed) graphs to be finite and loopless (loops add nothing to the problem under consideration), but we permit parallel edges. A *signed* graph is a triple $G = (V, E, \sigma)$ where (V, E) is a graph and $\sigma : E \to \{-1, 1\}$ is a *signature*. We say that the *sign* of a cycle $C \subseteq G$ is *positive (negative)* if $\sigma(C) = \prod_{e \in E(C)} \sigma(e)$ is equal to 1 (-1). If all cycles of G are positive, then we call G *balanced* and otherwise we call G *unbalanced*.

In a 2-connected signed graph G, a single edge e appears in cycles of both signs if and only if $G - e$ is unbalanced. For the "only if" direction, let C_1, C_2 be cycles of opposite sign containing e and note that the symmetric difference of $E(C_1)$ and $E(C_2)$ is a set of edges with negative sign and even degree at every vertex (which can thus be expressed as a disjoint union of edge sets of cycles). For the "if" direction, let $e = uv$, choose a negative cycle C in $G - e$, and apply Menger to choose two vertex disjoint paths from $\{u, v\}$ to $V(C)$; these two paths together with C and e contain the desired cycles.

*Email: mdevos@sfu.ca. Supported by an NSERC Discovery Grant (Canada)
†Email: knurse@sfu.ca. Partially supported by NSERC (Canada).

Our objective in this article is to extend this simple property to a pair of edges. If G is a signed graph and $e_1, e_2 \in E(G)$, then we say that e_1 and e_2 are *untied* if there exist cycles containing e_1 and e_2 of both positive and negative sign, and otherwise we say that e_1 and e_2 are *tied*. Our main result is as follows.

Theorem 1.1. *Let G be a 3-connected signed graph and let $e_1, e_2 \in E(G)$ be distinct and not in parallel with any other edges. Then e_1 and e_2 are tied in G if and only if one of the following holds:*

1. *There exists a parallel class F containing edges of both signs so that $F^+ = F \cup \{e_1, e_2\}$ is an edge-cut and $G - F^+$ is balanced,*
2. *e_1, e_2 are incident with a common vertex v and $G - v$ is balanced,*
3. *$G - \{e_1, e_2\}$ is balanced.*

In Section 2, we provide a reduction that allows us to determine the structure of arbitrary signed graphs that are tied, meaning this result implies a full characterization of when all cycles through two given edges of a signed graph have the same sign. This problem was explicitly asked by Zaslavsky in [13, E2], but let us remark that our motivation for this work is a forthcoming application of these results in the setting of nowhere-zero flows on signed graphs, towards Bouchet's conjecture that every flow-admissible signed graph has a nowhere-zero 6-flow [1]. We apply the results here while finding a decomposition of the edges of a 3-connected signed graph similar to Seymour's decomposition in the first proof of his 6-Flow Theorem [9].

Theorem 1.1 may be viewed as a signed graph generalization of the following result from Lovász's problem book [8, 6.67]. By replacing the edge e_3 of Theorem 1.2 with two parallel edges, one of each sign, forming a signed graph with exactly one negative edge, one observes that Theorem 1.1 does indeed imply Theorem 1.2.

Theorem 1.2. *[Lovász] Let G be a simple 3-connected graph and $e_1, e_2, e_3 \in E(G)$ be distinct. Then there is no cycle containing e_1, e_2, e_3 if and only if one of the following holds:*

1. *$G - \{e_1, e_2, e_3\}$ is disconnected,*
2. *e_1, e_3, e_3 are incident with a common vertex.*

Another generalization of Theorem 1.2 is the following conjecture by Lovász [7] and Woodall [12] (independently): If G is a k-connected graph, and $S \subseteq E(G)$ a set of k independent edges so that either k is even or $G - S$ is connected, then there is a cycle $C \subseteq G$ with $S \subseteq E(C)$. Kawarabayashi [4] showed that S is always contained in either one cycle or two vertex-disjoint cycles. And Thomassen and Häggkvist [3] showed that the conjecture holds if one assumes G is $(k + 1)$-connected. The following well-known conjecture of Lovász also concerns connectivity, paths and cycles.

Conjecture 1.3. *[Lovász] For any natural number k, there exists a least natural number $f(k)$ so that for any $f(k)$-connected graph G and any $x, y \in V(G)$ there exists an induced xy-path P so that $G - V(P)$ is k-connected.*

The above conjecture also has a natural generalization to signed graphs that we state below. To deduce 1.3 from 1.4, simply add a single negative edge xy to the graph (treat all other edges as positive).

Conjecture 1.4. *For any natural number k, there exists a least natural number $f'(k)$ so that for any $f'(k)$-connected, unbalanced, signed graph G there exists an induced negative cycle C so that $G - V(C)$ is k-connected.*

Concerning the two conjectures above, Tutte [10] proved the simplest of these cases, that $f(1) = f'(1) = 3$. Using Tutte's language, a cycle C in a graph G is *peripheral* if C is induced and $G - V(C)$ is connected. Tutte showed that every 3-connected graph has a peripheral cycle through any given edge, so $f(1) = 3$. Moreover, he proves that the peripheral cycles generate the cycle space. That is to say that the peripheral cycles are not contained in any codimension 1 subspace of the cycle space. It follows that every signed graph with a non-trivial signature has a negative peripheral cycle, and $f'(1) = 3$. Kriesell [6] and independently Chen Gould and Yu [2] show that $f(2) = 5$.

And so we have provided two examples of interesting statements about graphs which have a natural and more general interpretation in the setting of signed graphs.

2 Outline of the Proof

2.1 Reduction to 3-connected

A *k-separation* of a graph G is a pair of subgraphs (G_1, G_2) so that $E(G_1) \cap E(G_2) = \emptyset$, $E(G_1) \cup E(G_2) = E(G)$, and $|V(G_1) \cap V(G_2)| = k$. We say that the separation is *proper* if $V(G_1) \setminus V(G_2) \neq \emptyset \neq V(G_2) \setminus V(G_1)$.

Observation 2.1. *Let G be a 2-connected signed graph, let $e_1, e_2 \in E(G)$, and let (G_1, G_2) be a 2-separation of G with $V(G_1) \cap V(G_2) = \{u, v\}$. For $i = 1, 2$ let G_i^+ be obtained from G_i by adding a positive edge f_i with ends u, v.*

1. *If $e_i \in E(G_i)$ for $i = 1, 2$ then e_1 and e_2 are tied in G if and only if e_i and f_i are tied in G_i^+ for $i = 1, 2$.*

2. *If $e_1, e_2 \in E(G_1)$ and every edge in $E(G_2)$ is positive, then e_1 and e_2 are tied in G if and only if they are tied in G_1^+.*

3. *If $e_1, e_2 \in E(G_1)$ and G_2 is unbalanced, then e_1 and e_2 are tied in G if and only if they are tied in the graph obtained from G_1^+ by adding a negative edge f_1' in parallel with f_1.*

The above observation allows us to reduce the problem of two edges being tied to one on smaller graphs. Continuing in this manner, we may reduce the problem to the setting of 3-connected signed graphs. Since the concept of edges e_1, e_2 being tied is vacuous if e_1 and e_2 are in separate blocks, Theorem 1.1 gives a complete answer. All of the steps in our reduction are reversible, so we can turn this around and provide a generic construction of signed graphs where two given edges are tied by taking the three types given in the above theorem and combining them as in the observation. The possible structures of all such graphs can readily be determined but we found no better way of describing them than by way of the decompositions presented here.

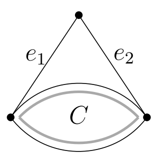
2.2 Some forbidden minors in tied signed graphs

In the setting of signed graphs we are principally focused on the signs of cycles and not those of edges. Accordingly, two signatures σ, σ' of a signed graph G are *equivalent* if every cycle $C \subseteq G$ satisfies $\sigma(C) = \sigma'(C)$. Two signatures are equivalent if and only if one can be obtained from the other by a *switch*, which is changing the sign of every edge in some edge-cut.

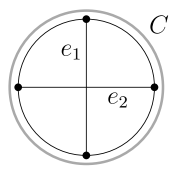
Let $G = (V, E, \sigma)$ be a signed graph and let $e \in E$ ($v \in V$). To *delete* the edge e (vertex v) we remove this edge (vertex and all incident edges) from the graph and adjust the domain of σ accordingly. To *contract* the edge e, first modify σ by switching on an edge-cut (if necessary) so that $\sigma(e) = 1$, and then modify the graph by contracting e and removing e from the domain of σ. If H is a signed graph obtained from G by a (possibly empty) sequence of edge and vertex deletions and edge contractions, we call H a *minor* of G. Note that whenever $C \subseteq H$ is a cycle, there is a corresponding cycle $C^* \subseteq G$ containing all edges in C and having the same sign as C. In particular, this implies the following key property.

Observation 2.2. *Let H be a minor of the signed graph G. If e_1, e_2 are untied edges of H, then they are also untied in G.*

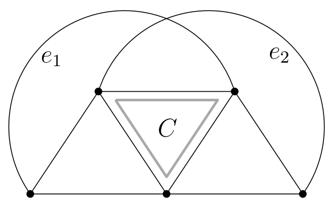
We introduce three families of signed graphs: hat, target, and hedgehog, each of which has a distinguished cycle C that is negative together with distinguished edges e_1, e_2. The edges e_1 and e_2 are untied in all. The heart of our argument is to show that if our graph is not one of the named counterexamples to Theorem 1.1, then it contains a hat, target, or hedgehog graph as a minor.



Hat



Target



Hedgehog

2.3 The main lemma and proof of main result

Our arguments lean on working with a carefully chosen negative cycle C in the graph. For this purpose we adopt Tutte's notation. Let G be a graph and let $H \subseteq G$. A *bridge* of H is a subgraph of $G - E(H)$ of one of the two forms: a single edge uv (and its ends) where $u, v \in V(H)$ and $uv \notin E(H)$, or a component F of $G - V(H)$ together with all edges of G with exactly one end in $V(F)$.

Lemma 2.3. *Let $G = (V, E, \sigma)$ be a simple, signed, 3-connected graph, and let $e_1, e_2 \in E(G)$ be nonadjacent. If there exists a negative cycle in $G - \{e_1, e_2\}$, then e_1 and e_2 are untied.*

Proof sketch. Suppose for contradiction the lemma is false, and let G be a counterexample so that $|V|$ is minimum. Choose a negative cycle $C \subseteq G - \{e_1, e_2\}$ subject to the following constraints: Both e_1 and e_2 are in the same bridge of C if possible, subject to this the bridge of C containing e_1 is maximum, subject to this the bridge of C containing e_2 is maximum, and subject to this the lexicographic ordering of the sizes of the other bridges is maximized. The proof proceeds by establishing the following four claims, whose proofs are omitted. They involve either a rerouting which contradicts the choice of C, or finding one of the minors in Section 2.2.

(1) Every bridge of C must contain e_1 or e_2.
(2) No bridge contains e_1 and e_2.
(3) e_1 is not incident with a vertex of C.
(4) $|V(C)| \geq 4$. □

With this lemma in hand, we prove the main result.

Proof sketch of Theorem 1.1. The "if" direction is straightforward to verify. For the "only if" direction, first suppose that e_1 and e_2 are incident with a common vertex u, say $e_i = uv_i$ for $i = 1, 2$. If $G - u$ is not balanced, then it contains a negative cycle that can be extended to a subgraph with a hat minor. Next, suppose that there exist two parallel edges f, f' of opposite sign. If f, f' are incident with an end of e_1 or e_2, then e_1, e_2 are not tied by 3-connectivity of G. Otherwise, the result follows from Theorem 1.2. This case also follows from an earlier result of Watkins and Mesner [11].

So we may now assume that G does not contain a negative cycle of length 2, and we may assume no parallel edges. If $G - \{e_1, e_2\}$ is balanced, then we have the third structure from the theorem statement. Otherwise, it follows from Lemma 2.3 that e_1 and e_2 are not tied in G, and this completes the proof. □

References

[1] A. Bouchet. Nowhere-zero integral flows on a bidirected graph. *J. Combin. Theory Ser. B*, 34(3):279–292, 1983.

[2] Guantao Chen, Ronald J. Gould, and XingXing Yu. Graph connectivity after path removal. *Combinatorica (Budapest. 1981)*, 23(2):185–203, 2003.

[3] Roland Häggkvist and Carsten Thomassen. Circuits through specified edges. *Discrete mathematics*, 41(1):29–34, 1982.

[4] Kenichi Kawarabayashi. One or two disjoint circuits cover independent edges: Lovász–woodall conjecture. *Journal of combinatorial theory. Series B*, 84(1):1–44, 2002.

[5] Kenichi Kawarabayashi, Orlando Lee, Bruce Reed, and Paul Wollan. A weaker version of lovász' path removal conjecture. *Journal of combinatorial theory. Series B*, 98(5):972–979, 2008.

[6] Matthias Kriesell. Induced paths in 5-connected graphs. *Journal of graph theory*, 36(1):52–58, 2001.

[7] László Lovász. Problem 5. Period. *Math. Hungar*, 4:82, 1974.

[8] László Lovász. *Combinatorial problems and exercises*. North-Holland, 2nd ed. edition, 1993.

[9] P.D Seymour. Nowhere-zero 6-flows. *Journal of combinatorial theory. Series B*, 30(2):130–135, 1981.

[10] W. T. Tutte. How to draw a graph. *Proceedings of the London Mathematical Society*, s3-13(1):743–767, 1963.

[11] M.E. Watkins and D.M. Mesner. Cycles and connectivity in graphs. *Canadian Journal of Mathematics*, 19:1319–1328, 1967.

[12] D.R Woodall. Circuits containing specified edges. *Journal of combinatorial theory. Series B*, 22(3):274–278, 1977.

[13] Thomas Zaslavsky. Negative circles in signed graphs: A problem collection. *Electronic notes in discrete mathematics*, 63:41–47, 2017.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Powers of planar graphs, product structure, and blocking partitions

(Extended abstract)

Marc Distel* Robert Hickingbotham* Michał T. Seweryn† David R. Wood*

Abstract

We show that there exist a constant c and a function f such that the k-power of a planar graph with maximum degree Δ is isomorphic to a subgraph of $H \boxtimes P \boxtimes K_{f(\Delta,k)}$ for some graph H with treewidth at most c and some path P. This is the first product structure theorem for k-powers of planar graphs, where the treewidth of H does not depend on k. We actually prove a stronger result, which implies an analogous product structure theorem for other graph classes, including k-planar graphs (of arbitrary degree).

Our proof uses a new concept of blocking partitions which is of independent interest. An ℓ-blocking partition of a graph G is a partition of the vertex set of G into connected subsets such that every path in G of length greater than ℓ contains two vertices in one set of the partition. The key lemma in our proof states that there exists a positive integer ℓ such that every planar graph of maximum degree Δ has an ℓ-blocking partition with parts of size bounded in terms of Δ.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-049`

*School of Mathematics, Monash University, Melbourne, Australia. E-mails: `{marc.distel,robert.hickingbotham,david.wood}@monash.edu`. Research of Distel and Hickingbotham supported by Australian Government Research Training Program Scholarships. Research of Wood supported by the Australian Research Council.

†Computer Science Department, Université libre de Bruxelles, Brussels, Belgium. E-mail: `michal.seweryn@ulb.be`. Supported by a PDR grant from the Belgian National Fund for Scientific Research (FNRS).

1 Introduction

Given two graphs[1] G and H, their *strong product* $G \boxtimes H$ is defined as the graph on $V(G) \times V(H)$ where distinct vertices $(u_1, v_1), (u_2, v_2) \in V(G) \times V(H)$ are adjacent if $u_1 = u_2$ and $v_1v_2 \in E(H)$, or $u_1u_2 \in E(G)$ and $v_1 = v_2$, or $u_1u_2 \in E(G)$ and $v_1v_2 \in E(H)$.

Graph product structure theory describes complicated graphs as subgraphs of the strong products of graphs with a simple structure like graphs of bounded treewidth[2], paths, or small complete graphs. Arguably the most important result of this theory is the product structure theorem for planar graphs by Dujmović, Joret, Micek, Morin, Ueckerdt, and Wood [7], which has been the key to solving many long-standing problems [1, 2, 4–7, 9, 10]. This theorem states that every planar graph is contained in $H \boxtimes P \boxtimes K_3$ for some graph H with $\mathrm{tw}(H) \leqslant 3$ and some path P. Here a graph G is *contained in* a graph G' if G is isomorphic to a subgraph of G'.

Another product structure for planar graphs by Ueckerdt, Wood, and Yi [14] states that every planar graph is contained in $H \boxtimes P$ for some graph H with $\mathrm{tw}(H) \leqslant 6$ and some path P. Note that $H \boxtimes P$ is isomorphic to $H \boxtimes P \boxtimes K_1$. These two product structure theorems for planar graphs illustrate a trade-off between the treewidth of H and the size of the complete graph involved in the product: If we want to find some fixed planar graph in a graph of the form $H \boxtimes P \boxtimes K_d$ where $\mathrm{tw}(H) \leqslant c$ for some constant c and P is a path, then we can either have $c = 3$ and $d = 3$, or $c = 6$ and $d = 1$.

There are many other graph classes $\mathcal{G}$ for which there exist constants c and d such that every $G \in \mathcal{G}$ is isomorphic to a subgraph of $H \boxtimes P \boxtimes K_d$ for some graph H with $\mathrm{tw}(H) \leqslant c$ and some path P [3, 7, 8, 11, 12]. The strong product $H \boxtimes P \boxtimes K_d$ is isomorphic to $(H \boxtimes K_d) \boxtimes P \boxtimes K_1$ and $\mathrm{tw}(H \boxtimes K_d) \leqslant d(c + 1) - 1$, so it is always possible to drive d down to 1, while minimising c is usually more difficult. Moreover, in many applications of such product structure theorems, the main dependency is on c. Therefore, the primary goal is to minimise c, whereas minimising d is a secondary goal. This paper proves new product structure theorems for k-powers of planar graphs of bounded degree and k-planar graphs. The distinguishing feature of our results is that the bound c on $\mathrm{tw}(H)$ is an absolute constant which does not depend on k.

For an integer $k \geqslant 1$, the *k-power* of a graph G is the graph G^k on $V(G)$ where two distinct vertices u and v are adjacent if and only if the distance between u and v in G is at most k. Dujmović et al. [8] proved that for every planar graph G of maximum degree Δ, and for every integer $k \geqslant 1$, the k-power G^k is contained in $J \boxtimes P \boxtimes K_{6k\Delta^k(k^3+3k)}$ for some graph J of treewidth at most $\binom{k+3}{3} - 1$ and some path P. Note that dependence on Δ is unavoidable since, for example, if G is the complete $(\Delta - 1)$-ary tree of height k, then G^{2k} is a complete graph on roughly $(\Delta - 1)^k$ vertices. Ossona de Mendez [13] asked whether this bound on $\mathrm{tw}(H)$ could be made independent of k. We show that indeed this is the

[1] We consider simple, finite, undirected graphs G with vertex-set $V(G)$ and edge-set $E(G)$.

[2] The *treewidth* $\mathrm{tw}(H)$ of a graph H, is the least integer k such that H is a subgraph of a graph G on a set $\{v_1, \ldots, v_n\}$ such that $n \geqslant k + 1$ and for each $i \in \{k + 1, \ldots, n\}$, the neighbours of v_i in $\{v_1, \ldots, v_{i-1}\}$ form a clique of size k in G.

case.

Theorem 1. *There exist a constant c and a function f such that for every planar graph G of maximum degree Δ and every integer $k \geqslant 1$, the graph G^k is contained in $H \boxtimes P \boxtimes K_{f(\Delta,k)}$ for some graph H with* $\mathrm{tw}(H) \leqslant c$ *and some path P.*

For an integer $k \geqslant 1$, a *k-planar* graph is a graph which has a drawing on the plane such that no three edges cross at a single point and each edge is involved in at most k crossings. Dujmović et al. [8] proved that every k-planar graph is contained in $H \boxtimes P \boxtimes K_{18k^2+48k+30}$, for some graph H with $\mathrm{tw}(H) \leqslant \binom{k+4}{3} - 1$ and some path P. Dujmović et al. [8] asked whether this bound on $\mathrm{tw}(H)$ could be made independent of k. We give an affirmative answer to this question.

Theorem 2. *There exist a constant c and a function f such that every k-planar graph G is contained in $H \boxtimes P \boxtimes K_{f(k)}$ some graph H with* $\mathrm{tw}(H) \leqslant c$ *and some path P.*

Theorems 1 and 2 follow from Theorem 4, which we formulate in Section 2. In particular, in our proof the value of c is the same in both theorems and equal to 15086399. This value is not optimal, but instead of optimising the constant c we chose to simplify the proof.

The proof of Theorem 4 uses a new concept of "blocking partitions". For an integer $\ell \geqslant 1$, an *ℓ-blocking partition* of a graph G is a partition $\mathcal{R}$ of $V(G)$ such that every set in $\mathcal{R}$ induces a connected subgraph of G and every path of length greater than ℓ in G contains two vertices in one part of $\mathcal{R}$. The *width* of $\mathcal{R}$ is the maximum size of a part of $\mathcal{R}$.

The following lemma plays the key role in our proof.

Lemma 3. *There exists a function f such that every planar graph of maximum degree at most Δ has a 222-blocking partition of width at most $f(\Delta)$.*

The construction of a 222-blocking partition is inspired by chordal partitions of triangulations by van den Heuvel et al. [15]. In their construction, a triangulation G is partitioned into paths $P_1, \dots, P_m$ where each path P_j is a shortest path between two distinguished vertices in one component of $G - \bigcup_{i<j} V(P_i)$. In our construction, we partition a planar graph G into trees $T_1, \dots, T_m$ where each T_j is obtained as follows. First, we define T_j^0 as a smallest tree in $G - \bigcup_{i<j} V(T_i)$ which contains some distinguished set of vertices of bounded size, and then, the tree T_j is obtained from T_j^0 by attaching all adjacent vertices as leaves. Finally, we split each tree T_j into subtrees of bounded size by removing an appropriate set of edges. Then, the vertex-sets of these subtrees define the desired 222-blocking partition of G.

Lemma 3 is the most technical part of our proof, and we do not include its proof here. Instead, we sketch the proof of the main theorems assuming Lemma 3. The partition in our proof of Lemma 3, is actually ℓ-blocking for some value ℓ significantly smaller than 222, but we decided to prove a worse bound on ℓ for simplicity's sake.

2 The main result

A *congested model* of a graph G' in a graph G is a set $(B_x : x \in V(G'))$ of connected subgraphs of G such that for every edge $xy \in E(G')$, the subgraphs B_x and B_y *touch* in G, i.e. they share a vertex or there is an edge between $V(B_x)$ and $V(B_y)$ in G. A *rooted congested model* of G' in G is a set $((B_x, v_x) : x \in V(G'))$ such that $(B_x : x \in V(G'))$ is a congested model of G' in G and $v_x \in V(B_x)$ for each $x \in V(G')$. We call a rooted congested model $((B_x, v_x) : x \in V(G'))$ in G an (r, Δ, d)*-model* if

- in each B_x, all vertices are at distance at most r from v_x,
- in each B_x, every vertex distinct from v_x has degree at most Δ, and
- for every $u \in V(G)$, there exist at most d vertices $x \in V(G')$ with $u \in V(B_x)$.

We call a graph G' an (r, Δ, d)*-minor* of G if there exists an (r, Δ, d)-model of G' in G. Note that G' is a minor of G if and only if G' is an $(r, \Delta, 1)$-minor of G for some $r, \Delta \geqslant 0$. A graph G' is an r*-shallow minor* of G if G' is an $(r, \Delta, 1)$-minor of G for some $\Delta \geqslant 0$. Observe that if G' is an (r, Δ, d)-minor of a graph G, then G' is an r-shallow minor of $G \boxtimes K_d$.

If G is a graph with maximum degree at most Δ, then G^k is an (r, Δ, d)-minor of G for $r = \lfloor k/2 \rfloor$ and $d = \sum_{i=0}^{\lfloor k/2 \rfloor} \Delta^i$, as witnessed by the rooted congested model $((B_x, x) : x \in V(G^k))$ where each B_x is the subgraph of G induced by the vertices at distance at most $\lfloor k/2 \rfloor$ from x. Furthermore, it is easy to see that every k-planar graph G' is an (r, Δ, d)-minor of G for $r = \lceil k/2 \rceil$ and $\Delta = d = 2$, where G is the planar graph obtained from G' by adding a dummy vertex at each intersection point. Therefore, Theorems 1 and 2 follow from the following theorem.

Theorem 4. *There exists a function f such that every (r, Δ, d)-minor of a planar graph is contained in $J \boxtimes P \boxtimes K_{f(r,\Delta,d)}$ for some graph J with $\mathrm{tw}(J) \leqslant 15086399$ and some path P.*

Theorem 4 implies a constant-treewidth product structure for other graph classes like δ-string graphs or k-fan-bundle graphs (we refer the reader to [11] for the definitions of these classes).

While it was easy to see that Theorem 4 implies Theorems 1 and 2, it is less obvious why Lemma 3 implies Theorem 4. The main idea behind this implication is captured by the following lemma.

Lemma 5. *There exists a function g such that for any r, Δ, d with $r \geqslant 224$, $\Delta \geqslant 0$ and $d \geqslant 1$, every (r, Δ, d)-minor of a planar graph is an $(r - 1, \Delta', d')$-minor of some planar graph for some $d', \Delta' \in \{1, \dots, g(\Delta, d)\}$.*

Proof. Let f be the function from Lemma 3, and set $g(\Delta, d) = \max\{d, \Delta\} \cdot f(d\Delta)$. Let G be a planar graph, let G' be an (r, Δ, d)-minor of G, and let $((B_x, v_x) : x \in V(G'))$ be an (r, Δ, d)-model of G' in G. Let $G_0 = \bigcup_{x \in V(G')} B_x - v_x$. Note that G_0 is a subgraph of G of maximum degree at most $d\Delta$. Let $\mathcal{R}$ be a 222-blocking partition of G_0 of width at most $f(d\Delta)$, and let us define $\mathcal{R}' = \mathcal{R} \cup \{\{v\} : v \in V(G) \setminus V(G_0)\}$. Let H denote the *quotient* $G/\mathcal{R}'$, i.e., let H be a graph on $\mathcal{R}'$ where two distinct parts are adjacent if G contains an edge with ends in these two parts. Since G is planar and each part of $\mathcal{R}'$ is connected, H is a minor of G, and thus a planar graph. Let $d' = df(d\Delta)$, and let $\Delta' = \Delta f(d\Delta)$. Since the width of $\mathcal{R}$ is at most $f(d\Delta)$, each part of $\mathcal{R}$ has degree at most Δ' in H. Hence, G' is an (r, Δ', d')-minor of H with a corresponding (r, Δ', d')-model $((B'_x, v'_x) : x \in V(G'))$ in H defined as follows. For each $x \in V(H)$, let v'_x be the part of $\mathcal{R}'$ containing v_x, and let B'_x be the "projection" of B_x on G', so that the vertices of B'_x are those parts of $\mathcal{R}'$ which contain at least one vertex of B_x, and two parts are adjacent in B'_x if B_x contains an edge with ends in those parts.

We claim that $((B'_x, v'_x) : x \in V(G'))$ is actually an $(r-1, \Delta', d')$-model. To show that, we need to prove that for any $x \in V(G')$ and $u' \in V(B'_x)$, the distance between v'_x and u' in B'_x is at most $r-1$. Let u be a vertex of B_x which belongs to the part u' of $\mathcal{R}'$. Let $u_0 \cdots u_s$ be a shortest path in B_x with $u_0 = v_x$ and $u_s = u$. Since $((B_x, v_x) : x \in V(G'))$ is an (r, Δ, d)-model, we have $s \leqslant r$. For each $i \in \{0, \dots, s\}$, let u'_i be the part of $\mathcal{R}'$ containing u_i. Hence, $u'_0 = v'_x$, $u'_s = u'$, and for each $i \in \{0, \dots, s-1\}$, either $u'_i = u'_{i+1}$ or $u'_i u'_{i+1} \in E(H)$. Therefore, the distance between v'_x and u' is at most s, and thus at most r. Suppose towards a contradiction that this distance is exactly r. Hence, $s = r$, and the vertices $u'_0, \dots, u'_r$ are pairwise distinct parts of $\mathcal{R}'$. Therefore, $u_1, \dots, u_r$ is a path in G_0, with no two vertices in one part of $\mathcal{R}$. As $r \geqslant 224$, the length of this path is at least 223, which contradicts $\mathcal{R}$ being 222-blocking. This proves that G' is an $(r-1, \Delta', d')$-minor of H. $\square$

The proof of Theorem 4 uses Lemma 5 and the following result by Hickingbotham and Wood [11].

Theorem 6 ([11])**.** *If a graph G is an r-shallow minor of $H \boxtimes P \boxtimes K_d$ where $\mathrm{tw}(H) \leqslant t$, then G is contained in $J \boxtimes P \boxtimes K_{d(2r+1)^2}$ for some graph J with $\mathrm{tw}(J) \leqslant \binom{2r+1+t}{t} - 1$.*

Proof of Theorem 4. Let g be the function from Lemma 5. We may assume that $g(\Delta, d) \leqslant g(\Delta', d')$ whenever $\Delta \leqslant \Delta'$ and $d \leqslant d'$. Define $f(r, \Delta, d)$ recursively:

$$f(r, \Delta, d) = \begin{cases} 3d(2r+1)^2 & \text{if } r \leqslant 223, \\ f(r-1, g(\Delta, d), g(\Delta, d)) & \text{if } r \geqslant 224. \end{cases}$$

We show that this function satisfies the theorem by induction on r. Let G be a planar graph, and let G' be an (r, Δ, d)-minor of G. For the base case, suppose that $r \leqslant 223$. By the product structure theorem for planar graphs, G is contained in $H \boxtimes P \boxtimes K_3$ for some graph H with $\mathrm{tw}(H) \leqslant 3$ and some path P. Hence, G' is an r-shallow minor of $H \boxtimes P \boxtimes K_{3d}$. By Theorem 6, G' is contained in $J \boxtimes P \boxtimes K_{3d(2r+1)^2}$ for some graph J with $\mathrm{tw}(J) \leqslant \binom{2r+1+3}{3} - 1 \leqslant \binom{450}{3} - 1 = 15086399$.

For the induction step, suppose that $r \geqslant 224$. By Lemma 5, there exist a planar graph H and $d', \Delta' \in \{1, \dots, g(\Delta, d)\}$ such that G' is an $(r-1, \Delta', d')$-minor of H. By the induction hypothesis, there exists a graph J with $\mathrm{tw}(J) \leqslant 15086399$ such that G' is contained in $J \boxtimes P \boxtimes K_{f(r-1, \Delta', d')}$. Since $d' \leqslant g(\Delta, d)$ and $\Delta' \leqslant g(\Delta, d)$, we have $f(r-1, \Delta', d') \leqslant f(r, \Delta, d)$, and therefore G' is contained in $J \boxtimes P \boxtimes K_{f(r, \Delta, d)}$. This completes the proof. $\square$

References

[1] Michael A. Bekos, Giordano Da Lozzo, Petr Hliněný, and Michael Kaufmann. Graph product structure for h-framed graphs. In Sang Won Bae and Heejin Park, eds., *33rd International Symposium on Algorithms and Computation (ISAAC '22)*, vol. 248 of *LIPIcs*, pp. 23:1–23. Schloss Dagstuhl, 2022. arXiv:2204.11495.

[2] Édouard Bonnet, O-joung Kwon, and David R. Wood. Reduced bandwidth: a qualitative strengthening of twin-width in minor-closed classes (and beyond). 2022, arXiv:2202.11858.

[3] Marc Distel, Robert Hickingbotham, Tony Huynh, and David R. Wood. Improved product structure for graphs on surfaces. *Discrete Math. Theor. Comput. Sci.*, 24(2):#6, 2022.

[4] Michał Dębski, Stefan Felsner, Piotr Micek, and Felix Schröder. Improved bounds for centered colorings. *Adv. Comb.*, #8, 2021.

[5] Vida Dujmović, Louis Esperet, Cyril Gavoille, Gwenaël Joret, Piotr Micek, and Pat Morin. Adjacency labelling for planar graphs (and beyond). *J. ACM*, 68(6):42, 2021.

[6] Vida Dujmović, Louis Esperet, Gwenaël Joret, Bartosz Walczak, and David R. Wood. Planar graphs have bounded nonrepetitive chromatic number. *Adv. Comb.*, #5, 2020.

[7] Vida Dujmović, Gwenaël Joret, Piotr Micek, Pat Morin, Torsten Ueckerdt, and David R. Wood. Planar graphs have bounded queue-number. *J. ACM*, 67(4):#22, 2020.

[8] Vida Dujmović, Pat Morin, and David R. Wood. Graph product structure for non-minor-closed classes. *J. Combin. Theory Ser. B*, 162:34–67, 2023.

[9] Zdenek Dvorák, Daniel Gonçalves, Abhiruk Lahiri, Jane Tan, and Torsten Ueckerdt. On comparable box dimension. In Xavier Goaoc and Michael Kerber, eds., *Proc. 38th Int'l Symp. on Computat. Geometry* (SoCG 2022), vol. 224 of *LIPIcs*, pp. 38:1–38:14. Schloss Dagstuhl, 2022.

[10] Louis Esperet, Gwenaël Joret, and Pat Morin. Sparse universal graphs for planarity. 2020, arXiv:2010.05779.

[11] Robert Hickingbotham and David R. Wood. Shallow minors, graph products and beyond planar graphs. 2021, arXiv:2111.12412.

[12] Freddie Illingworth, Alex Scott, and David R. Wood. Product structure of graphs with an excluded minor. 2022, arXiv:2104.06627.

[13] Patrice Ossona de Mendez. Product structure for squares of planar graphs. 2021. Open Problems for Workshop on Graph Product Structure Theory, Banff International Research Station (21w5235).

[14] Torsten Ueckerdt, David R. Wood, and Wendy Yi. An improved planar graph product structure theorem. *Electron. J. Combin.*, 29:P2.51, 2022.

[15] Jan van den Heuvel, Patrice Ossona de Mendez, Daniel Quiroz, Roman

Rabinovich, and Sebastian Siebertz. On the generalised colouring numbers of graphs that exclude a fixed minor. *European J. Combin.*, 66:129–144, 2017.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

When is Cartesian product a Cayley graph?

(Extended abstract)

Ted Dobson* Ademir Hujdurović† Wilfried Imrich‡ Ronald Ortner§

Abstract

A graph is said to be *Cayley* graph if its automorphism group admits a regular subgroup. Automorphisms of the Cartesian product of graphs are well understood, and it is known that Cartesian product of Cayley graphs is a Cayley graph. It is natural to ask the reverse question, namely whether all the factors of Cartesian product that is a Cayley graph have to be Cayley graphs. The main purpose of this paper is to initiate the study of this question.

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-050

1 Introduction

Throughout this paper graphs are assumed to be finite, simple, and connected, and groups are finite. Given a graph Γ we let $V(\Gamma)$, $E(\Gamma)$, and $\mathrm{Aut}(\Gamma)$ be the set of vertices, the set of edges, and the automorphism group of Γ, respectively.

Let G be a finite group and $S \subseteq G \setminus \{1\}$ an inverse closed subset of G. Then the *Cayley graph* $Cay(G, S)$ *on* G *with respect to* S is a graph with vertex set G and edge set $\{\{g, gs\} \mid g \in G, s \in S\}$. It is well-known that a graph Γ is a Cayley graph on a group G if there exists a regular subgroup of $\mathrm{Aut}(\Gamma)$ isomorphic to G (see [5]).

*University of Primorska, UP FAMNIT, Glagoljaška 8, 6000 Koper, Slovenia. E-mail: `ted.dobson@upr.si`

†University of Primorska, UP FAMNIT, Glagoljaška 8, 6000 Koper, Slovenia. E-mail: `ademir.hujdurovic@upr.si`

‡Montanuniversität Leoben, 8700 Leoben, Austria. E-mail: `wilfried.imrich@unileoben.ac.at`

§Montanuniversität Leoben, 8700 Leoben, Austria- E-mail: `ronald.ortner@unileoben.ac.at`

Recall that the *Cartesian product* $\Gamma_1 \square \cdots \square \Gamma_k$ of graphs $\Gamma_1, \ldots, \Gamma_k$ has vertex set $V(\Gamma_1) \times \cdots \times V(\Gamma_k)$ with two distinct vertices being adjacent if they are adjacent in one of the coordinates and coincide in all other coordinates. Recall also that two graphs are called *relatively prime* if there exists no non-trivial graph that is a factor – with respect to the Cartesian product – of both of them. A graph is said to be *prime* with respect to the Cartesian product if it cannot be factored as a Cartesian product of two non-trivial graphs. For a graph Γ, the Cartesian product $\underbrace{\Gamma \square \ldots \square \Gamma}_{\text{n times}}$ is denoted with $\Gamma^{\square n}$.

It is well-known that the Cartesian product of Cayley graphs is a Cayley graph. A natural question is to consider whether the converse is true, that is, if the Cartesian product of graphs is a Cayley graph, does each of the factors have to be a Cayley graph? This is the main motivation for the work presented in this article. We provide partial results showing that Cartesian products involving certain vertex-transitive non-Cayley graphs are not Cayley (for example, every graph having a Petersen graph as one of the factors is non-Cayley). We are not aware of any example of a Cayley graph having a non-Cayley factor.

2 Preliminaries

We start by recalling the structure of the automorphism group of the Cartesian products.

Theorem 2.1. *[2, Theorem 6.8] Let Γ be a connected graph with prime factorization $\Gamma = \Gamma_1 \square \Gamma_2 \square \ldots \square \Gamma_k$. Then for any automorphism φ of Γ, there is a permutation π of $\{1, 2, \ldots, k\}$ and isomorphisms $\varphi_i : \Gamma_{\pi(i)} \to \Gamma_i$ for which*

$$\varphi(x_1, x_2, \ldots, x_k) = (\varphi_1(x_{\pi(1)}), \ldots, \varphi_k(x_{\pi(k)})).$$

Let $G_i \leq Sym(V_i)$ for $i \in \{1, \ldots, n\}$. The group $G_1 \times G_2 \times \ldots \times G_n$ acts *canonically* on $V_1 \times V_2 \times \ldots \times V_n$ in such a way that $g_i \in G_i$ is applied to the i-th coordinate. We have the following simple observation.

Lemma 2.2. *Let $G_i \leq Sym(V_i)$ for $i \in \{1, \ldots, n\}$ be transitive groups. If there exists a regular subgroup of $G_1 \times G_2 \ldots \times G_n$ acting canonically on $V_1 \times V_2 \times \ldots \times V_n$, then every G_i admits a regular subgroup.*

Proof. Let H be a regular subgroup of $G_1 \times G_2 \ldots \times G_n$. Since H is regular, it follows that $|H| = |V_1| \cdot \ldots \cdot |V_n|$. Let $j \in \{1, \ldots, n\}$, and let $v_i \in V_i$ be arbitrary for $i \neq j$. Let $K = \{(g_1, \ldots, g_n) \in H \mid g_i(v_i) = v_i \; i \in \{1, \ldots, n\} \setminus \{j\}\}$ and $K(j) = \{g_j \mid (g_1, \ldots, g_j, \ldots, g_n) \in K\}$. It is easy to see that $K(j)$ is a subgroup of G_j, and that the transitivity of H implies that $K(j)$ is transitive subgroup of G_j. Moreover, if $K(j)$ is not semiregular, then H would contain a non-identity element fixing a point of $V_1 \times \ldots \times V_n$, contrary to the assumption that H is regular. We conclude that $K(j)$ is a regular subgroup of G_j. Since j is arbitrary, the result follows. $\square$

The following result follows directly by applying Lemma 2.2 to the fact that the automorphism group of Cartesian product of relatively prime graphs is the direct product of the automorphism groups of the factors, see [2, Corollary 6.12] (see also [3, Theorem 3.1]).

Theorem 2.3. *Let Γ_1 and Γ_2 be two connected relatively prime graphs with respect to the Cartesian product and let $\Gamma = \Gamma_1 \Box \Gamma_2$. Then Γ is a Cayley graph if and only if both Γ_1 and Γ_2 are Cayley graphs.*

In light of Theorem 2.3, the question of which Cartesian products are Cayley is reduced to the question when is a Cartesian power of a graph isomorphic to a Cayley graph. In the following result the automorphism group of a Cartesian power of a graph is given. Let us first recall the definition of a wreath product of permutation groups. Let $G \le Sym(V)$ and $H \le S_n$. The *wreath product of G by H* denoted by $G \wr H$ is the set of all permutations $((g_1, \ldots, g_n), h)$ of V^n (where $g_1, \ldots, g_n \in G$ and $h \in H$) such that $((g_1, \ldots, g_n), h) : (v_1, \ldots, v_n) \mapsto (g_{h(1)}(v_{h(1)}, \ldots, g_{h(n)}(v_{h(n)})$.

Lemma 2.4. *Let Γ be a graph that is prime with respect to the Cartesian product. Then* $\mathrm{Aut}(\Gamma^{\Box n}) \cong \mathrm{Aut}(\Gamma) \wr S_n$.

3 Main results

The following result giving a bound on the order of a Sylow p-subgroup of the symmetric group S_n will be needed later.

Lemma 3.1. *Let $n \ge 1$ be an integer and p a prime divisor of n. A Sylow p-subgroup of S_n has order less than p^n.*

Proof. Let $n = a_0 + a_1 p + \ldots + a_k p^k$ with $0 \le a_i \le p-1$. By [4, pg. 11] a Sylow p-subgroup of S_n has order p^M, where

$$\begin{aligned} M &= \sum_{i=1}^{k} a_i(1 + p + p^2 + \ldots + p^{i-1}) = \sum_{i=1}^{k} a_i \frac{p^i - 1}{p-1} \\ &= \sum_{i=1}^{k} \left[\frac{a_i p^i}{p-1} - \frac{a_i}{p-1} \right] = \frac{\sum_{i=1}^{k} a_i p^i}{p-1} - \sum_{i=1}^{k} \frac{a_i}{p-1} \\ &\le \frac{n}{p-1} - \sum_{i=1}^{k} \frac{a_i}{p-1} < \frac{n}{p-1} \le n. \end{aligned}$$

□

An automorphism of a graph is said to be *semiregular* if all the cycles in its cyclic decomposition have equal lengths.

Theorem 3.2. *Let Γ be a vertex-transitive graph such that* $\mathrm{Aut}(\Gamma)$ *has no semiregular element of order p for some prime p dividing $|V(\Gamma)|$, and Γ is prime with respect to the Cartesian product. Then $\Gamma^{\Box n}$ is not Cayley for every $n \geq 1$.*

Proof. We first observe that Γ is not Cayley, as if it were, it would contain a regular subgroup R of order n. Then R has a subgroup of prime order p, which is necessarily semiregular, contradicting out assumption.

Suppose that R is a regular subgroup of $\mathrm{Aut}(\Gamma^{\Box n})$. As Γ is prime with respect to the Cartesian product, by Lemma 2.4, we have that $\mathrm{Aut}(\Gamma^{\Box n}) = \mathrm{Aut}(\Gamma) \wr S_n$ with the product action. Let P be a Sylow p-subgroup of R. Observe that P has order at least p^n. Define $\varphi : P \to S_n$ with $\varphi((g_1, \ldots, g_n), \sigma) = \sigma$ (where $g_i \in \mathrm{Aut}(\Gamma)$ and $\sigma \in S_n$). Observe that φ is a homomorphism. If the kernel of φ is trivial, then P is isomorphic to a subgroup of S_n. However, by Lemma 3.1, the order of a Sylow p-subgroup of S_n is less than p^n. It follows that the kernel of φ is not trivial. It follows that there exists a non-identity element $\gamma = ((g_1, \ldots, g_n), id) \in P$. Without loss of generality we may assume that the order of γ is p (by taking the p-th powers of γ if necessary), implying that each g_i is identity or of order p. Since by the assumption, $\mathrm{Aut}(\Gamma)$ has no semiregular element of order p, it follows that each g_i of order p fixes a vertex of Γ. Hence γ fixes some point in $V(\Gamma)^n$. But as R is regular, and $\gamma \in P \leq R$, this means $\gamma = 1$, a contradiction. □

Corollary 3.3. *No Cartesian power of the Petersen graph is isomorphic to a Cayley graph.*

Proof. Let P denote the Petersen graph. By Theorem 3.2, we need only show that $\mathrm{Aut}(P)$ has no semiregular element of order 2. The automorphism group of the Petersen graph is isomorphic to the action of S_5 on the 2-subsets of $\{1, 2, 3, 4, 5\}$ by [1, Theorem 2.1.4]. The elements of order 2 in S_5 are a product of two transposition as well as transpositions. It is easy to see that the 2-subset of $\{1, 2, 3, 4, 5\}$ which is permuted in a transposition, is fixed by a transposition, and so no element of order 2 in the automorphism group of the Petersen graph is semiregular. □

Theorem 3.4. *Let Γ be a vertex-transitive graph that is not isomorphic to a Cayley graph, whose automorphism group has order relatively prime to $n!$. Then $\Gamma^{\Box n}$ is not isomorphic to a Cayley graph.*

Proof. Suppose that R is a regular subgroup of $\mathrm{Aut}(\Gamma^{\Box n})$. Then R has order relatively prime to $n!$, in which case every element of R must fix every factor of $V(\Gamma)^n$ (i.e. no element of R can permute factors of $V(\Gamma)^n$). This means that $R \leq \mathrm{Aut}(\Gamma)^n$, hence the result follows by Lemma 2.2. □

Corollary 3.5. *Let Γ be a vertex-transitive graph of odd order that is not a Cayley graph. Then $\Gamma \Box \Gamma$ is not isomorphic to a Cayley graph.*

In the following result we study the structure of a transitive permutation group G admitting no regular subgroup, but such that $A \wr S_2$ in the product action admits a regular subgroup.

Theorem 3.6. *Let $A \leq Sym(V)$ be transitive. If $A \wr S_2$ admits a regular subgroup (in the product action) then A admits a regular subgroup or A admits a semiregular subgroup with two orbits.*

Proof. Let $H \leq A \wr S_2$ be a regular subgroup. If $H \leq A \times A$, then by Lemma 2.2 it follows that A admits a regular subgroup. Suppose that H is not contained in $A \times A$. Let $\overline{H} = H \cap (A \times A)$. Observe that $\overline{H}$ is an index two subgroup of H. Since H is regular, it follows that $\overline{H}$ is semiregular with two orbits.

Let $\overline{H}(v) = \{h_1 \in A \mid \exists h_2 \in A_v \text{ such that } (h_1, h_2) \in \overline{H}\}$. Observe that $\overline{H}(v)$ is a subgroup of A. Moreover, it is semiregular, since $\overline{H}$ is semiregular. If $\overline{H}(v)$ is transitive or has two orbits then we are done. Hence, we may assume that $|Orb_{\overline{H}(v)}(x)| \leq |V|/3$ for every $v \in V$.

Let O be one of the two orbits of $\overline{H}$ on $V \times V$. Let $O(v) = \{y \in V \mid (y, v) \in O\}$. Let $x \in O(v)$ be arbitrary. We claim that $O(v) = Orb_{\overline{H}(v)}(x)$. Let $y \in O(v)$. Then (x, v) and (y, v) belong to the same orbit O of $\overline{H}$, hence there exists $(h_1, h_2) \in \overline{H}$ such that $h_1(x) = y$ and $h_2(v) = v$. It follows that h_1 is an element of $\overline{H}(v)$ mapping x to y. This shows that $O(v)$ is contained in $Orb_{\overline{H}(v)}(x)$.

Let z be an element of $Orb_{\overline{H}(v)}(x)$. There exists $h_1 \in \overline{H}(v)$ such that $h_1(x) = z$. By the definition of $\overline{H}$ it follows that there exists $h_2 \in A$ fixing v such that $(h_1, h_2) \in \overline{H}$. This shows that (h_1, h_2) is an element of $\overline{H}$ mapping (x, v) to (z, v), hence (x, v) and (z, v) belong to the orbit O, implying that $z \in O(v)$. This shows that $O(v) = Orb_{\overline{H}(v)}(x)$.

It is easy to see that $O = \bigcup_{v \in V} O(v)$ is a partition of O, and that $|O| = \sum_{v \in V} |O(v)|$. Since $|Orb_{\overline{H}(v)}(x)| \leq |V|/3$, it follows that $|O| \leq |V|^2/3$, contradicting the assumption that $|O| = |V|^2/2$. The obtained contradiction shows that $\overline{H}(v) \leq A$ must be regular or semiregular with two orbits for some $v \in V$. □

Remark 3.7. *There are examples of transitive groups without regular subgroups such that their wreath product with S_2 in the product action admits regular subgroups. For example, TransitiveGroups(24)[675] (of order 288 and degree 24) is one such group. However, the authors are not aware of any such group which is automorphism group of a graph.*

Corollary 3.8. *Let Γ be a graph that is prime with respect to the Cartesian product such that* $\mathrm{Aut}(\Gamma)$ *admits no semiregular subgroup with two orbits. Then $\Gamma \square \Gamma$ is not isomorphic to a Cayley graph.*

Remark 3.9. *There exist infinitely many vertex-transitive graphs of even order that do not admit a semiregular subgroup with two orbits. One such graph is the Tutte-Coxeter graph, which is a cubic symmetric graph of order 30. Moreover, there exist vertex-transitive graphs of even order admitting a semiregular automorphism of order p, for every prime divisor of the order of the graph, but not admitting a semiregular subgroup with 2 orbits. In particular, any vertex-transitive graph of order 2^n without a semiregular subgroup with 2 orbits is such an example.*

References

[1] Ted Dobson, Aleksander Malnič, and Dragan Marušič, *Symmetry in graphs*, Cambridge Studies in Advanced Mathematics, vol. 198, Cambridge University Press, Cambridge, 2022. MR 4404766

[2] Richard Hammack, Wilfried Imrich, and Sandi Klavžar, *Handbook of product graphs*, second ed., Discrete Mathematics and its Applications (Boca Raton), CRC Press, Boca Raton, FL, 2011, With a foreword by Peter Winkler. MR 2817074

[3] Ademir Hujdurović, Klavdija Kutnar, and Dragan Marušič, *On normality of n-Cayley graphs*, Appl. Math. Comput. **332** (2018), 469–476. MR 3788705

[4] Donald Passman, *Permutation groups*, W. A. Benjamin, Inc., New York-Amsterdam, 1968. MR MR0237627 (38 #5908)

[5] Gert Sabidussi, *Vertex-transitive graphs*, Monatsh. Math. **68** (1964), 426–438. MR MR0175815 (31 #91)

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

A generalization of Bondy's pancyclicity theorem

(Extended abstract)

Nemanja Draganić* David Munhá Correia* Benny Sudakov*

Abstract

The *bipartite independence number* of a graph G, denoted as $\tilde{\alpha}(G)$, is the minimal number k such that there exist positive integers a and b with $a + b = k + 1$ with the property that for any two sets $A, B \subseteq V(G)$ with $|A| = a$ and $|B| = b$, there is an edge between A and B. McDiarmid and Yolov showed that if $\delta(G) \geq \tilde{\alpha}(G)$ then G is Hamiltonian, extending the famous theorem of Dirac which states that if $\delta(G) \geq |G|/2$ then G is Hamiltonian. In 1973, Bondy showed that, unless G is a complete bipartite graph, Dirac's Hamiltonicity condition also implies pancyclicity, i.e., existence of cycles of all the lengths from 3 up to n. In this paper we show that $\delta(G) \geq \tilde{\alpha}(G)$ implies that G is pancyclic or that $G = K_{\frac{n}{2},\frac{n}{2}}$, thus extending the result of McDiarmid and Yolov, and generalizing the classic theorem of Bondy.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-051`

1 Introduction

The notion of Hamiltonicity is one of most central and extensively studied topics in Combinatorics. Since the problem of determining whether a graph is Hamiltonian is NP-complete, a central theme in Combinatorics is to derive sufficient conditions for this property. A classic example is Dirac's theorem [14] which dates back to 1952 and states that every n-vertex graph with minimum degree at least $n/2$ is Hamiltonian. Since then, a plethora of interesting and important results about various aspects of Hamiltonicity have been obtained, see e.g. [1, 11, 12, 13, 18, 24, 26, 27, 32], and the surveys [20, 29].

*Department of Mathematics, ETH Zurich, Switzerland.
E-mails:`{nemanja.draganic,david.munhacanascorreia,benjamin.sudakov}@math.ethz.ch`

Besides finding sufficient conditions for containing a Hamilton cycle, significant attention has been given to conditions which force a graph to have cycles of other lengths. Indeed, *the cycle spectrum of a graph*, which is the set of lengths of cycles contained in that graph, has been the focus of study of numerous papers and in particular gained a lot of attention in recent years [2, 3, 8, 16, 19, 21, 23, 28, 31, 34]. Among other graph parameters, the relation of the cycle spectrum to the minimum degree, number of edges, independence number, chromatic number and expansion of the graph have been studied.

We say that an n-vertex graph is *pancyclic* if the cycle spectrum contains all integers from 3 up to n. Bondy suggested that in the cycle spectrum of a graph, it is usually hardest to guarantee the existence of the longest cycle, i.e. a Hamilton cycle. This intuition was captured by his famous meta-conjecture [5] from 1973, which asserts that any non-trivial condition which implies Hamiltonicity, also implies pancyclicity (up to a small class of exceptional graphs). As a first example, he proved in [6] an extension of Dirac's theorem, showing that minimum degree at least $n/2$ implies that the graph is either pancyclic or that it is the complete bipartite graph $K_{\frac{n}{2},\frac{n}{2}}$. Further, Bauer and Schmeichel [4], relying on previous results of Schmeichel and Hakimi [33], showed that the sufficient conditions for Hamiltonicity given by Bondy [7], Chvátal [10] and Fan [17] all imply pancyclicity, up to a certain small family of exceptional graphs.

Another classic Hamiltonicity result is the Chvátal-Erdős theorem, which states that $\kappa(G) \geq \alpha(G)$ implies that G is Hamiltonian, where $\kappa(G)$ is the connectivity of G, and $\alpha(G)$ its independence number. Motivated by Bondy's meta-conjecture, Jackson and Ordaz [22] thirty years ago suggested that $\kappa(G) > \alpha(G)$ already implies pancyclicity. The first progress towards this problem was obtained by Keevash and Sudakov, who showed pancyclicity when $\kappa(G) \geq 600\alpha(G)$. Recently, in [15] we were able to resolve the Jackson-Ordaz conjecture asymptotically, proving that $\kappa(G) \geq (1 + o(1))\alpha(G)$ is already enough for pancyclicity. It is worth mentioning that, in all the listed work, the proof that the Hamiltonicity condition also implies pancyclicity is usually significantly harder than just proving Hamiltonicity, and requires new ideas and techniques.

An interesting sufficient condition for Hamiltonicity was given by McDiarmid and Yolov [30]. To state their result, we need the following natural graph parameter. For a graph G, its *bipartite independence number* $\tilde{\alpha}(G)$ is the minimal number k, such that there exist positive integers a and b with $a + b = k + 1$, such that between any two sets $A, B \subseteq V(G)$ with $|A| = a$ and $|B| = b$, there is an edge between A and B. Notice that we always have that $\alpha(G) \leq \tilde{\alpha}(G)$. Indeed, if $\tilde{\alpha}(G) = k$, then G does not contain independent sets I of size at least $k + 1$, since evidently for every $a + b = k + 1$, there would exist disjoint sets $A, B \subset I$, so that $|A| = a$ and $|B| = b$ and with no edge between A and B. Let us now state the result of McDiarmid and Yolov.

Theorem 1 ([30]). *If $\delta(G) \geq \tilde{\alpha}(G)$, then G is Hamiltonian.*

This result implies Dirac's theorem, because if $\delta(G) \geq n/2$, then $\lceil n/2 \rceil \geq \tilde{\alpha}(G)$, as for every $|A| = 1$ and $|B| = \lceil n/2 \rceil$ there is an edge between A and B. Hence also $\delta(G) \geq \lceil n/2 \rceil \geq \tilde{\alpha}(G)$, so G is Hamiltonian.

Naturally, the immediate question which arises is whether the McDiarmid-Yolov condition implies that the graph satisfies the stronger property of pancyclicity. As a very preliminary step in this direction, Chen [9] was able to show that for any given positive constant c, for sufficiently large n it holds that if G is an n-vertex graph with $\tilde{\alpha}(G) = cn$ and $\delta(G) \geq \frac{10}{3}cn$, then G is pancyclic. In this paper we completely resolve this problem, showing that $\delta(G) \geq \tilde{\alpha}(G)$ implies that G pancyclic or $G = K_{\frac{n}{2},\frac{n}{2}}$. This generalizes the classical theorem of Bondy [6], and gives additional evidence for his meta-conjecture, mentioned above.

Theorem 2. *If $\delta(G) \geq \tilde{\alpha}(G)$, then G is pancyclic, unless G is complete bipartite $G = K_{\frac{n}{2},\frac{n}{2}}$.*

Our proof is completely self-contained and relies on a novel variant of Pósa's celebrated rotation-extension technique, which is used to extend paths and cycles in expanding graphs (see, e.g., [32]). Define the graph $\tilde{C}_\ell$, to be the cycle of length ℓ together with an additional vertex which is adjacent to two consecutive vertices on the cycle (thus forming a triangle with them). For each $\ell \in [3, n-1]$, our goal is to either find a $\tilde{C}_\ell$ or a $\tilde{C}_{\ell+1}$, which is clearly enough to show pancyclicity. The proof is recursive in nature, as we will derive the existence of a $\tilde{C}_\ell$ or a $\tilde{C}_{\ell+1}$ from the existence of a $\tilde{C}_{\ell-1}$. In our setting, we would like to apply the rotation-extension technique to the $\tilde{C}_{\ell-1}$ with the additional requirement that the extended cycle preserves the attached triangle. However, this is not possible in general and from the existence of a $\tilde{C}_{\ell-1}$ we will in turn derive the existence of a gadget denoted as a *switch*, which is a path with triangles attached to it, to which we can apply our rotation-extension technique. One of the key ideas is to consider the switch which is optimal with respect to how close the triangles are to the beginning of the path. The application of the rotation-extension technique to such an optimal switch will then result in either a $\tilde{C}_\ell$, a $\tilde{C}_{\ell+1}$, or a better switch, contradicting the optimality of the original switch.

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi. First occurrence of Hamilton cycles in random graphs. In Cycles in graphs (Burnaby, B.C., 1982), Vol. 115, North-Holland Mathematical Studies, North-Holland, Amsterdam, 115:173–178, 1985.

[2] N. Alon and M. Krivelevich. Divisible subdivisions. *Journal of Graph Theory*, 98(4):623–629, 2021.

[3] Y. Alon, M. Krivelevich, and E. Lubetzky. Cycle lengths in sparse random graphs. *Random Structures & Algorithms*, 61(3):444–461, 2022.

[4] D. Bauer and E. Schmeichel. Hamiltonian degree conditions which imply a graph is pancyclic. *Journal of Combinatorial Theory, Series B*, 48(1):111–116, 1990.

[5] J. A. Bondy. Pancyclic graphs: recent results, infinite and finite sets. In *Colloq. Math. Soc. János Bolyai*, Keszthely, pp. 181–187, 1973.

[6] J. A. Bondy. Pancyclic graphs I. *Journal of Combinatorial Theory, Series B*, 11(1):80–84, 1971.

[7] J. A. Bondy. *Longest paths and cycles in graphs of high degree*. Department of Combinatorics and Optimization, University of Waterloo, 1980.

[8] M. Bucić, L. Gishboliner, and B. Sudakov. Cycles of many lengths in Hamiltonian graphs. *Forum of Mathematics, Sigma*, 10, E70, 2022.

[9] M. Chen Hamilton-connected, vertex-pancyclic and bipartite holes. *Discrete Mathematics*, 345(12):113158, 2022.

[10] V. Chvátal. On Hamilton's ideals. *Journal of Combinatorial Theory, Series B*, 12(2):163–168, 1972.

[11] V. Chvátal and P. Erdős. A note on Hamiltonian circuits. *Discrete Mathematics*, 2(2):111–113, 1972.

[12] B. Csaba, D. Kühn, A. Lo, D. Osthus, and A. Treglown. Proof of the 1-factorization and Hamilton decomposition conjectures. *Memoirs of the American Mathematical Society*, 244, monograph 1154, 170 pages, 2016.

[13] B. Cuckler and J. Kahn. Hamiltonian cycles in Dirac graphs. *Combinatorica*, 29(3):299–326, 2009.

[14] G. A. Dirac. Some theorems on abstract graphs. *Proceedings of the London Mathematical Society*, 3(1):69–81, 1952.

[15] N. Draganić, D. Munhá Correia and B. Sudakov. Chvátal-Erdős condition for pancyclicity. *arXiv:2301.10190* 2023.

[16] N. Draganić, D. Munhá Correia and B. Sudakov. Pancyclicity of Hamiltonian graphs. *arXiv:2209.03325* 2022.

[17] G. H. Fan. New sufficient conditions for cycles in graphs. *Journal of Combinatorial Theory, Series B*, 37(3):221–227, 1984.

[18] A. Ferber, E. Long, and B. Sudakov. Counting Hamilton decompositions of oriented graphs. *International Mathematics Research Notices*, 2018(22):6908–6933, 2018.

[19] L. Friedman and M. Krivelevich. Cycle lengths in expanding graphs. *Combinatorica*, 41(1):53–74, 2021.

[20] R. J. Gould. Recent advances on the Hamiltonian problem: Survey III. *Graphs and Combinatorics*, 30(1):1–46, 2014.

[21] H. Liu and R. Montgomery. A solution to Erdős and Hajnal's odd cycle problem. *arXiv preprint arXiv:2010.15802*, 2020.

[22] B. Jackson and O. Ordaz. Chvátal-Erdős conditions for paths and cycles in graphs and digraphs. A survey. *Discrete mathematics*, 84(3):241–254, 1990.

[23] P. Keevash and B. Sudakov. Pancyclicity of Hamiltonian and highly connected graphs. *Journal of Combinatorial Theory, Series B*, 100(5):456–467, 2010.

[24] M. Krivelevich. The critical bias for the Hamiltonicity game is $(1 + o(1))n/\ln n$. *Journal of the American Mathematical Society*, 24(1):125–131, 2011.

[25] M. Krivelevich and B. Sudakov. Sparse pseudo-random graphs are Hamiltonian. *J. Graph Theory*, 42(1):17–33, 2003.

[26] M. Krivelevich, C. Lee, and B. Sudakov. Robust Hamiltonicity of Dirac graphs. *Transactions of the American Mathematical Society*, 366(6):3095–3130, 2014.

[27] D. Kühn and D. Osthus. Hamilton decompositions of regular expanders: a proof of Kelly's conjecture for large tournaments. *Advances in Mathematics*, 237:62–146, 2013.

[28] C.-H. Liu and J. Ma. Cycle lengths and minimum degree of graphs. *Journal of Combinatorial*

Theory, Series B, 128:66–95, 2018.

[29] D. Kühn and D. Osthus. Hamilton cycles in graphs and hypergraphs: an extremal perspective. In *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. IV*, pp. 381–406. Kyung Moon Sa, Seoul, 2014.

[30] C. McDiarmid and N. Yolov. Hamilton cycles, minimum degree, and bipartite holes. *Journal of Graph Theory*, 86(3):277–285, 2017.

[31] K. G. Milans, F. Pfender, D. Rautenbach, F. Regen, and D. B. West. Cycle spectra of Hamiltonian graphs. *Journal of Combinatorial Theory, Series B*, 102(4):869–874, 2012.

[32] L. Pósa, Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14(4):359–364, 1976.

[33] E. F. Schmeichel and S. L. Hakimi. A cycle structure theorem for Hamiltonian graphs. *Journal of Combinatorial Theory, Series B*, 45(1):99–107, 1988.

[34] J. Verstraëte. Extremal problems for cycles in graphs. In *Recent trends in combinatorics*, pp. 83–116. Springer, 2016.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Chvátal-Erdős condition for pancyclicity

(Extended abstract)

Nemanja Draganić* David Munhá Correia* Benny Sudakov*

Abstract

An n-vertex graph is *Hamiltonian* if it contains a cycle that covers all of its vertices and it is *pancyclic* if it contains cycles of all lengths from 3 up to n. A celebrated meta-conjecture of Bondy states that every non-trivial condition implying Hamiltonicity also implies pancyclicity (up to possibly a few exceptional graphs). We show that every graph G with $\kappa(G) > (1+o(1))\alpha(G)$ is pancyclic. This extends the famous Chvátal-Erdős condition for Hamiltonicity and proves asymptotically a 30-year old conjecture of Jackson and Ordaz.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-052`

1 Introduction

The notion of Hamiltonicity is one of most central and extensively studied topics in Combinatorics. Since the problem of determining whether a graph is Hamiltonian is NP-complete, a central theme in Combinatorics is to derive sufficient conditions for this property. A classic example is Dirac's theorem [14] which dates back to 1952 and states that every n-vertex graph with minimum degree at least $n/2$ is Hamiltonian. Since then, a plethora of interesting and important results about various aspects of Hamiltonicity have been obtained, see e.g. [1, 11, 12, 13, 19, 26, 27, 28, 32], and the surveys [21, 30].

Besides finding sufficient conditions for containing a Hamilton cycle, significant attention has been given to conditions which force a graph to have cycles of other lengths. Indeed, *the cycle spectrum of a graph*, which is the set of lengths of cycles contained in

*Department of Mathematics, ETH, Zürich, Switzerland. Research supported in part by SNSF grant 200021_196965.
Emails: **{nemanja.draganic,david.munhacanascorreia, benjamin.sudakov}@math.ethz.ch**.

that graph, has been the focus of study of numerous papers and in particular gained a lot of attention in recent years [2, 3, 15, 20, 22, 25, 29, 31, 35]. Among other graph parameters, the relation of the cycle spectrum to the minimum degree, number of edges, independence number, chromatic number and expansion of the graph have been studied.

We say that an n-vertex graph is *pancyclic* if the cycle spectrum contains all integers from 3 up to n. In the cycle spectrum of an n-vertex graph, it is usually hardest to guarantee the existence of the longest cycle, i.e. a Hamilton cycle. This intuition was captured in Bondy's famous meta-conjecture [6] from 1973, which asserts that any non-trivial condition which implies Hamiltonicity, also implies pancyclicity (up to a small class of exceptional graphs). As a first example, he proved in [7] an extension of Dirac's theorem, showing that minimum degree at least $n/2$ implies that the graph is either pancyclic or that it is the complete bipartite graph $K_{\frac{n}{2},\frac{n}{2}}$. Further, Bauer and Schmeichel [5], relying on previous results of Schmeichel and Hakimi [34], showed that the sufficient conditions for Hamiltonicity given by Bondy [8], Chvátal [10] and Fan [18] all imply pancyclicity, up to a certain small family of exceptional graphs.

Another classic condition which implies Hamiltonicity is given by the famous theorem of Chvatál and Erdős [11]. It states that if the connectivity of a graph G is at least as large as its independence number, that is, $\kappa(G) \geq \alpha(G)$, then G is Hamiltonian. The pancyclicity counterpart of this result has also been investigated - see, e.g., [4] and the surveys [23, 33]. In fact, in 1990, Jackson and Ordaz [23] conjectured that G must be pancyclic if $\kappa(G) > \alpha(G)$, which if true would confirm Bondy's meta-conjecture for this classical instance. One can use an old result of Erdős [16] to show pancyclicity if $\kappa(G)$ is large enough function of $\alpha(G)$. A first linear bound on $\kappa(G)$ was given only in 2010 by Keevash and Sudakov [25], who showed that $\kappa(G) \geq 600\alpha(G)$ is enough. In this paper, we resolve the conjecture of Jackson and Ordaz asymptotically, by showing that $\kappa(G) > (1 + o(1))\alpha(G)$ is already enough to guarantee pancyclicity.

Theorem 1. *Let $\varepsilon > 0$ and let n be sufficiently large. Then, every n-vertex graph G for which we have $\kappa(G) \geq (1 + \varepsilon)\alpha(G)$ is pancylic.*

Next we briefly discuss some of the key steps in the proof of this theorem. It will be convenient for us to consider different ranges of cycle lengths whose existence we want to show, and for each range we have a slightly different approach to deal with. But in general, in order to find these different cycle lengths we will combine various tools on shortening/augmenting paths and finding consecutive path lengths between two fixed vertices.

For example, for finding consecutive path lengths we crucially use that since $\kappa(G) > \alpha(G)$, it must be that G contains triangles - moreover, it contains a *path with triangles attached to many of its edges* (see Definition 2), which trivially implies the existence of many consecutive path lengths between the endpoints of such a path. For shortening/augmenting paths, we also introduce new tools. One of them is used to shorten paths using only the minimum degree of the graph (Lemma 6), while another one augments paths using both the independence and connectivity number, and is given in the complete version of the paper. Furthermore, we will also use a novel result proven in [15] using the Gallai-Milgram theorem, in order to shorten paths using the independence number of the graph. In our paper, we present these tools, together with some other useful results of a similar flavour. The general proof strategy is to find a cycle of appropriate length which consists of two paths, one of which has many edges to which triangles are attached. Then we apply our shortening/augmenting results to the other path. Combining the consecutive path lengths from the first path with the path lengths obtained from the second path we get all possible cycle lengths.

2 Cycles with triangles and path shortening

In this section we will give a taste of the methods we use. We will show two simple results – first we show how to obtain a cycle with many triangles, and second, in Lemma 6 we show how to shorten a path between two vertices only using the minimum degree of the graph. We start with the definition of a cycle with many triangles.

Definition 2. *Define the graph C_ℓ^r to be the graph formed by a cycle $v_1v_2\dots v_lv_1$ of length ℓ with the additional edges $v_1v_3, v_3v_5, \dots, v_{2r-1}v_{2r+1}$ (if $r = 0$, then it is just a cycle of length l). We will refer to this as a* cycle of length ℓ with r triangles. *Similarly define P_ℓ^r and refer to it as a* path of length l with r triangles, *where P_0^0 is just a vertex.*

The following is an easy starting point for the existence of the graphs C_ℓ^r with appropriate parameters, as subgraphs in graphs G with $\kappa(G) \geq \alpha(G)$.

Lemma 3. *Every n-vertex graph G with $\kappa(G) \geq \alpha(G)$ contains a C_l^r for all r such that $0 \leq r \leq \frac{\kappa(G)-\alpha(G)}{2}$ and some l with $l - 2(r+1) \leq \max\left(\frac{n}{\kappa(G)-2r+1}, \frac{n}{\kappa(G)-1}\right)$. In particular, it contains a P_{2r}^r for all such r.*

Proof. We will first show that G must always contain a $P_{2r'}^{r'}$ for $r' := \left\lfloor \frac{\kappa(G)-\alpha(G)}{2} \right\rfloor$- we construct such a path greedily. Suppose that we have the vertices $v_1v_2v_3\dots v_{2i+1}$ which form a P_{2i}^i, so that the edges $v_1v_3, \dots, v_{2i-1}v_{2i+1}$ are also present. Provided that $i < r'$, we can augment this path as follows. Consider the set $S := N(v_{2i+1})\backslash\{v_1, \dots, v_{2i}\}$ - by assumption, this has size at least $\delta(G) - 2i > \kappa(G) - 2r' \geq \alpha(G)$. Therefore, it must contain an edge $v_{2i+2}v_{2i+3}$. Clearly, $v_{2i+1}v_{2i+2}v_{2i+3}$ forms a triangle and thus, $v_1v_2v_3\dots v_{2i+1}v_{2i+2}v_{2i+3}$ is a P_{2i+2}^{i+1}. Continuing with this procedure until $i = r'$, gives the desired $P_{2r'}^{r'}$.

Now, fix r with the given condition. If $r = 0$, then take an edge xy in G. By Menger's theorem, there exist at least $\kappa(G)$ internally vertex-disjoint xy-paths in G and thus, at least $\kappa(G) - 1$ of these are not the edge xy. Therefore, there is such a path with at most $\frac{n}{\kappa(G)-1} + 2$ vertices, which together with the edge xy, then creates a cycle of length at most $\frac{n}{\kappa(G)-1} + 2$. If $r \geq 1$, by the previous paragraph, G contains a P_{2r}^r - let x, y be its endpoints. By Menger's theorem, there exist at least $\kappa(G)$ internally vertex-disjoint xy-paths in G. Since at most $2r - 1$ of these intersect $P_{2r}^r \setminus \{x, y\}$, there exists one which is disjoint to $P_{2r}^r \setminus \{x, y\}$ and contains at most $\frac{n}{\kappa(G)-2r+1}$ internal vertices. This produces the desired C_l^r.
□ □

We can also use this type of cycles to extend the celebrated Chvátal-Erdős theorem [11].

Theorem 4 (Chvátal-Erdős [11])**.** *If for a graph G we have that $\kappa(G) \geq \alpha(G)$, then G is Hamiltonian.*

Our resut states that if the Chvátal-Erdős condition is satisfied, then we can find a Hamilton cycle with a certain number of triangles, depending on the discrepancy between the connectivity and the independence number.

Theorem 5. *Every n-vertex graph G such that $\kappa(G) \geq \alpha(G)$ contains a C_n^r with $r = \left\lfloor \frac{\kappa(G)-\alpha(G)}{2} \right\rfloor$.*

Proof. Suppose for contradiction that some $\ell < n$ is maximal such that there exists a copy of C_ℓ^r in G. Note that ℓ exists by Lemma 3. Order the cycle as $v_1v_2\dots v_\ell v_1$ so that the edges $v_1v_3, v_3v_5, \dots, v_{2r-1}v_{2r+1}$ are also present. Since $\ell < n$, there is a vertex v not in C_l^r. Moreover, as $\kappa(G) \geq \alpha(G) + 2r$, there exist $\alpha(G)$ paths contained in $V(G) \setminus \{v_1, \dots, v_{2r}\}$, all of which go from v to C_l^r and are vertex-disjoint apart from the initial vertex v. Let us denote these paths as $P_{i_1}, P_{i_2}, \dots$ so that $v_j = P_j \cap C_l^r$. Consider the set $S := \{v_{i_1+1}, v_{i_2+1}, \dots\}$ with indices taken modulo l, so that $|S| \geq \alpha(G)$. Observe (as illustrated in Figure 1) that then there must be an edge contained in $S \cup \{v\}$ and that any such edge can be used to augment C_l^r to a $C_{l'}^r$ with $l' > l$, contradicting the maximality of l. $\square$ $\square$

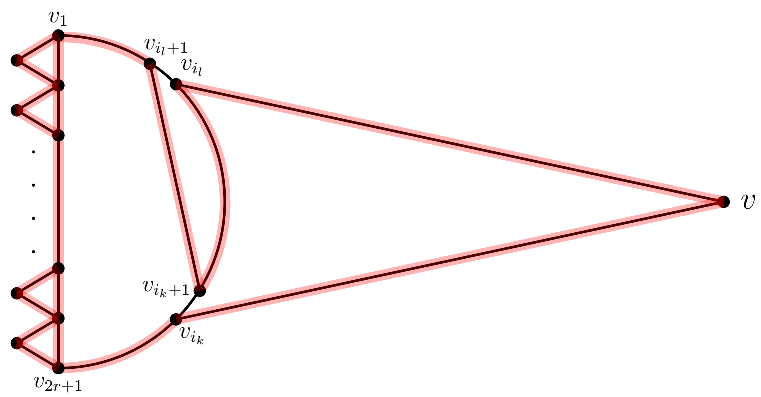


Figure 1: An illustration of how an edge between two elements v_{i_k+1}, v_{i_l+1} of S can be used to construct a new $C_{l'}^r$.

Now we show a result which uses only the minimum degree of the graph to shorten a path between two vertices. Among other shortening/augmenting tools in our paper, this is an important building block for our proof.

Lemma 6. *Let G be an n-vertex graph, $\delta := \delta(G)$ and P a path in G with endpoints x, y such that $|P| > 20n/\delta$. Then there is an xy-path P' such that $|P| - 20n/\delta \leq |P'| < |P|$.*

Proof. Suppose for sake of contradiction that no such path P' exists. Let $P := v_1 v_2 \dots v_{l-1} v_l$ with $v_1 = x, v_l = y$ and let $<_P$ denotes the given ordering of the path P as $v_1 <_P v_2 <_P \dots <_P v_l$. Since $|P| > 10n/\delta$, we can partition P into sub-paths $Q_1, Q_2, \dots, Q_k$ such that $|Q_k| \le 10n/\delta$ and $|Q_i| = 10n/\delta$ for all $i < k$. Moreover, we have $k = \left\lceil \frac{|P|}{10n/\delta} \right\rceil$. Now, consider the vertices in Q_1 and take a subset $Q_1' \subseteq Q_1$ of size $\lfloor |Q_1|/3 \rfloor \ge 3n/\delta$ such that no two vertices in Q_1' are at distance at most 2 in P. Consider then the set of edges incident to Q_1', that is, $E[Q_1', V(G)]$; by the minimum degree condition, there are at least $|Q_1'| \cdot \delta \ge 3n$ such edges.

Now, clearly there cannot exist an edge spanned by Q_1 which does not belong to P since this edge could be used to shorten P by at most $|Q_1| \le 10n/\delta$. Hence, $e(Q_1', Q_1) \le 2|Q_1'|$. Similarly, the following must hold.

Claim. $e(Q_1', V(G) \setminus P) \le n - |P|$.

Proof. Suppose otherwise. Then there is a vertex $v \in V(G) \setminus P$ with at least 2 neighbours in Q_1' - denote these by u, w. Note that since by construction u, w are at distance at least 2 and at most $|Q_1| \le 10n/\delta$ in P, this is a contradiction, since it produces the desired P' by substituting the sub-path of P between u and w by the path uvw. □

To give an upper bound on the total number of edges incident to Q_1' which are contained in $V(P)$, we also use the following claim.

Claim. *For all $i > 1$, we have $e(Q_1', Q_i) < |Q_1'| + |Q_i|$.*

Proof. Suppose otherwise. This implies that there is a cycle in $G[Q_1', Q_i]$ and hence, there must exist two crossing edges in this bipartite graph, that is, edges $a_1 b_1$ and $a_2 b_2$, with $a_1 <_P a_2$ and both in Q_1', and $b_1 <_P b_2$ both in Q_i. These can clearly be used to shorten P (see Figure 2) by at most $|Q_1| + |Q_i| \le 20n/\delta$, which is a contradiction as it produces the desired P'. □

The above claim implies that

$$\sum_{i>1} e(Q_1', Q_i) < \sum_{i>1} (|Q_1'| + |Q_i|) \le (k-1)|Q_1'| + (|P| - |Q_1|) < 2|P| - 2|Q_1'|.$$

Figure 2: Shortening of the path P using the crossing edges $a_1 b_1$ and $a_2 b_2$. The resulting path is P' and is drawn in red.

To conclude, we now must have the following

$$e(Q_1', V(G)) = e(Q_1', Q_1) + e(Q_1', V(G)\backslash P) + \sum_{i>1} e(Q_1', Q_i) < 2|Q_1'| + (n-|P|) + (2|P| - 2|Q_1'|) < 2n.$$

which contradicts the previous observation that $e(Q_1', V(G)) \geq 3n$. □ □

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi. First occurrence of Hamilton cycles in random graphs. In Cycles in graphs (Burnaby, B.C., 1982), Vol. 115, North-Holland Mathematical Studies, North-Holland, Amsterdam, 115:173–178, 1985.

[2] N. Alon and M. Krivelevich. Divisible subdivisions. *Journal of Graph Theory*, 98(4):623–629, 2021.

[3] Y. Alon, M. Krivelevich, and E. Lubetzky. Cycle lengths in sparse random graphs. *Random Structures & Algorithms*, 61(3):444–461, 2022.

[4] D. Amar, I. Fournier, and A. Germa. Pancyclism in Chvátal-Erdős graphs. *Graphs and Combinatorics*, 7(2):101–112, 1991.

[5] D. Bauer and E. Schmeichel. Hamiltonian degree conditions which imply a graph is pancyclic. *Journal of Combinatorial Theory, Series B*, 48(1):111–116, 1990.

[6] J. A. Bondy. Pancyclic graphs: recent results, infinite and finite sets. In *Colloq. Math. Soc. János Bolyai*, Keszthely, pp. 181–187, 1973.

[7] J. A. Bondy. Pancyclic graphs I. *Journal of Combinatorial Theory, Series B*, 11(1):80–84, 1971.

[8] J. A. Bondy. *Longest paths and cycles in graphs of high degree.* Department of Combinatorics and Optimization, University of Waterloo, 1980.

[9] J. A. Bondy and M. Simonovits. *Cycles of even length in graphs.* J. Combinatorial Theory B **16** (1974) 97–105.

[10] V. Chvátal. On Hamilton's ideals. *Journal of Combinatorial Theory, Series B*, 12(2):163–168, 1972.

[11] V. Chvátal and P. Erdős. A note on Hamiltonian circuits. *Discrete Mathematics*, 2(2):111–113, 1972.

[12] B. Csaba, D. Kühn, A. Lo, D. Osthus, and A. Treglown. Proof of the 1-factorization and Hamilton decomposition conjectures. *Memoirs of the American Mathematical Society*, 244, monograph 1154, 170 pages, 2016.

[13] B. Cuckler and J. Kahn. Hamiltonian cycles in Dirac graphs. *Combinatorica*, 29(3):299–326, 2009.

[14] G. A. Dirac. Some theorems on abstract graphs. *Proceedings of the London Mathematical Society*, 3(1):69–81, 1952.

[15] N. Draganić, D. Munhá Correia and B. Sudakov. Pancyclicity of Hamiltonian graphs. *arXiv:2209.03325* 2022.

[16] P. Erdős. Some problems in graph theory. In *Hypergraph Seminar*, pp. 187–190. Springer, 1972.

[17] P. Erdős, R. J. Faudree, C. C. Rousseau, and R. H. Schelp. On cycle-complete graph Ramsey numbers. *Journal of Graph Theory*, 2(1):53–64, 1978.

[18] G. H. Fan. New sufficient conditions for cycles in graphs. *Journal of Combinatorial Theory, Series B*, 37(3):221–227, 1984.

[19] A. Ferber, E. Long, and B. Sudakov. Counting Hamilton decompositions of oriented graphs. *International Mathematics Research Notices*, 2018(22):6908–6933, 2018.

[20] L. Friedman and M. Krivelevich. Cycle lengths in expanding graphs. *Combinatorica*, 41(1):53–74, 2021.

[21] R. J. Gould. Recent advances on the Hamiltonian problem: Survey III. *Graphs and Combinatorics*, 30(1):1–46, 2014.

[22] H. Liu and R. Montgomery. A solution to Erdős and Hajnal's odd cycle problem. *Journal of the American Mathematical Society*, to appear.

[23] B. Jackson and O. Ordaz. Chvátal-Erdős conditions for paths and cycles in graphs and digraphs. A survey. *Discrete mathematics*, 84(3):241–254, 1990.

[24] P. Keevash, E. Long, and J. Skokan. Cycle-complete Ramsey numbers. *International Mathematics Research Notices*, 2021(1):275–300, 2021.

[25] P. Keevash and B. Sudakov. Pancyclicity of Hamiltonian and highly connected graphs. *Journal of Combinatorial Theory, Series B*, 100(5):456–467, 2010.

[26] M. Krivelevich. The critical bias for the Hamiltonicity game is $(1 + o(1))n/\ln n$. *Journal of the American Mathematical Society*, 24(1):125–131, 2011.

[27] M. Krivelevich, C. Lee, and B. Sudakov. Robust Hamiltonicity of Dirac graphs. *Transactions of the American Mathematical Society*, 366(6):3095–3130, 2014.

[28] D. Kühn and D. Osthus. Hamilton decompositions of regular expanders: a proof of Kelly's conjecture for large tournaments. *Advances in Mathematics*, 237:62–146, 2013.

[29] C.-H. Liu and J. Ma. Cycle lengths and minimum degree of graphs. *Journal of Combinatorial Theory, Series B*, 128:66–95, 2018.

[30] D. Kühn and D. Osthus. Hamilton cycles in graphs and hypergraphs: an extremal perspective. In *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. IV*, pp. 381–406. Kyung Moon Sa, Seoul, 2014.

[31] K. G. Milans, F. Pfender, D. Rautenbach, F. Regen, and D. B. West. Cycle spectra of Hamiltonian graphs. *Journal of Combinatorial Theory, Series B*, 102(4):869–874, 2012.

[32] L. Pósa, Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14(4):359–364, 1976.

[33] A. Saito, Chvátal-Erdős theorem: Old theorem with new aspects In *Kyoto International Conference on Computational Geometry and Graph Theory*, pp. 191–200. Springer, 2007.

[34] E. F. Schmeichel and S. L. Hakimi. A cycle structure theorem for Hamiltonian graphs. *Journal of Combinatorial Theory, Series B*, 45(1):99–107, 1988.

[35] J. Verstraëte. Extremal problems for cycles in graphs. In *Recent trends in combinatorics*, pp. 83–116. Springer, 2016.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Tower Gaps in Multicolour Ramsey Numbers

(Extended abstract)

Quentin Dubroff * António Girão † Eoin Hurley ‡ Corrine Yap §

Abstract

Resolving a problem of Conlon, Fox, and Rödl, we construct a family of hypergraphs with arbitrarily large tower height separation between their 2-colour and q-colour Ramsey numbers. The main lemma underlying this construction is a new variant of the Erdős–Hajnal stepping-up lemma for a generalized Ramsey number $r_k(t; q, p)$, which we define as the smallest integer n such that every q-colouring of the k-sets on n vertices contains a set of t vertices spanning fewer than p colours. Our results provide the first tower-type lower bounds on these numbers.

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-053

1 Introduction

Let $K_n^{(k)}$ denote the complete k-uniform hypergraph on n vertices. We define $r_k(G; q)$ for $k, q \in \mathbb{N}$ as the smallest integer n such that in every q-colouring of $K_n^{(k)}$, there is a monochromatic copy of the hypergraph G. For simplicity when G is $K_t^{(k)}$, we write $r_k(G; q) = r_k(t; q)$. Observe that when $q = 2$, $r_k(G; 2)$ and $r_k(t; 2)$ coincide with the classical Ramsey numbers $r_k(G)$ and $r_k(t)$, and we will denote them as such. One of the

*Department of Mathematics, Rutgers University, Piscataway, NJ, 08854, USA. E-mail: qcd2@math.rutgers.edu
†Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK. E-mail: antonio.girao@maths.ox.ac.uk
‡Unaffiliated. E-mail: eoin.hurley@umail.ucc.ie
§Department of Mathematics, Rutgers University, Piscataway, NJ, 08854, USA. E-mail: corrine.yap@rutgers.edu

most central open problems in Ramsey theory is determining the growth rate of the 3-uniform Ramsey number $r_3(t)$. A famous result of Erdős, Hajnal, and Rado [8] from the 60's shows that there exist constants c and c' such that

$$2^{ct^2} \leq r_3(t) \leq 2^{2^{c't}}.$$

Note that the upper bound is essentially exponential in the lower bound. Despite much attention, this remains the state of the art. Perhaps surprisingly, if we allow four colours instead of two, Erdős and Hajnal (see e.g. [10]) showed that the double-exponential upper bound is essentially correct, i.e. there is a $c > 0$ such that $r_3(t;4) \geq 2^{2^{ct}}$. More recently Conlon, Fox, and Sudakov [4] proved a super-exponential bound with three colours, that is, that there exists $c > 0$ such that $r_3(t;3) \geq 2^{t^{c \log t}}$. Erdős conjectured that the double-exponential bound should hold without using extra colours, offering \$500 dollars for a proof that $r_3(t) \geq 2^{2^{ct}}$ for some constant $c > 0$. Raising the stakes for this conjecture is the ingenious *stepping-up* construction of Erdős and Hajnal (see e.g. [10]), which shows that for all q and $k \geq 3$,

$$r_{k+1}(2t + k - 4; q) > 2^{r_k(t;q)-1}. \tag{1}$$

For the past 60 years, we have used (1) to stack our lower bounds for $r_k(t;q)$ upon that of $r_3(t;q)$, yielding that $r_k(t) \geq T_{k-1}(ct^2)$, where $T_k(x)$, *the tower of height k in x*, is defined by $T_1(x) = x$, $T_{i+1}(x) = 2^{T_i(x)}$. The corresponding upper bounds of $r_k(t) \leq T_k(O(t))$ (see [6, 7, 8]) are once again exponential in the lower bounds, and thus a positive resolution of Erdős's conjecture would be the decisive step in showing that $r_k(t) = T_k(\Theta(t))$ for all $k \geq 3$.

Due to the lack of progress on this central conjecture, it is natural to try to understand just how significant a role the number of colours can play in hypergraph Ramsey numbers and whether or not there could really be such a large difference between $r_3(t)$ and $r_3(t;4)$. One argument in favour of the conjecture is that the reliance on extra colours to prove a double-exponential lower bound may be a technical limitation of the stepping-up construction. This is challenged by a stunning discovery of Conlon, Fox, and Rödl [3] who exhibited an infinite family of 3-uniform hypergraphs called *hedgehogs*, whose Ramsey numbers display strong dependence on the number of colours. Namely, they showed that the 2-colour Ramsey number of hedgehogs is polynomial in their order, while the 4-colour Ramsey number is at least exponential. To understand just how significant a role the number of colours could play they asked the following:

Question 1.1. *For any integer $h \geq 3$, do there exist integers k and q and a family of k-uniform hypergraphs for which the 2-colour Ramsey number grows as a polynomial in the number of vertices, while the q-colour Ramsey number grows as a tower of height h?*

Our main contribution is to answer this in the affirmative. Define the *k-uniform balanced hedgehog* $\hat{H}_t^{(k)}$ with body of order t to be the graph constructed as follows: take a set S of t vertices, called the *body*, and for each subset $X \subset S$ of order $\lceil \frac{k}{2} \rceil$ add a k-edge e with $e \cap S = X$ such that for all $e, f \in E(\hat{H}_t^{(k)})$ we have $e \cap f \subset S$. The hedgehog $H_t^{(k)}$ as defined by Conlon, Fox, and Rödl differs only in that they consider every $X \subset S$ of order $k-1$ rather than $\lceil \frac{k}{2} \rceil$. We observe that for $k = 3$ the two definitions coincide. When the uniformity is clear from the context we shall drop the superscript.

Theorem 1.2. *There exist $c > 0$ and $q : \mathbb{N} \to \mathbb{N}$ such that for all $k \in \mathbb{N}$ and sufficiently large t, we have*

(a) $r_{2k+1}(\hat{H}_t) \leq t^{k+3}$*, and*

(b) $r_{2k+1}(\hat{H}_t; q(k)) \geq T_{\lfloor c \log_2 \log_2 k \rfloor}(t)$.

To prove this, we provide new stepping-up lemmas for a more general type of hypergraph Ramsey numbers. Let $r_k(G; q, p)$ for $q \geq p$ be the smallest integer n such that in every q-colouring of $K_n^{(k)}$, there is a copy of the hypergraph G whose edges span fewer than p colours. As before, we use $r_k(t; q, p)$ when $G = K_t^{(k)}$ and suppress p when $p = 2$.

A standard application of the first moment method (see e.g. [1]) shows that for any $k, q \in \mathbb{N}$ there exists $c > 0$ such that $r_k(t; q, q) \geq 2^{ct^{k-1}}$ for all $t \in \mathbb{N}$. We note that in the graph case ($k = 2$) the special case of $q = p$ was already investigated by Erdős and Szemerédi [9] in the 70's; in fact, the more general case when $p < q$ is also indirectly discussed. They showed the following rather precise bounds: for all $q \ll t$, $2^{\Omega(t/q)} \leq r_2(t; q, q) \leq q^{O(t/q)}$.

These generalized hypergraph Ramsey numbers were also considered in a special case by Conlon, Fox, and Rödl [3] who asked if there exist an integer q and number $c > 0$ such that $r_3(t; q, 3) \geq 2^{2^{ct}}$. To date, the only nontrivial improvement on the first moment bound has been made by Mubayi and Suk [11] who proved there exists $c > 0$ such that for $q \geq 9$, we have $r_3(t; q, 3) \geq 2^{t^{2+cq}}$ for $t \in \mathbb{N}$ sufficiently large; for all other values of $k, q, p \geq 3$, the random construction is essentially the state of the art. Our knowledge (or lack thereof) is thus summarised by the following bounds for $k, q, p \geq 3$ and sufficiently large $t \in \mathbb{N}$,

$$2^{t^c} \leq r_k(t; q, p) \leq T_k(O(t)),$$

where $c \geq 1$ is allowed to depend on k, q and p. Note that in this case our upper bounds are a staggering tower of height $k-2$ in the lower bounds.

A related notion called the set-colouring Ramsey number was introduced by Erdős, Hajnal, and Rado in [8] and subsequently studied in [12] and much more recently in [5] and [2]. Borrowing notation from [5], let $R_k(t; q, s)$ denote the minimum number of vertices such that every (q, s)-*set colouring* of $K_n^{(k)}$, that is, a colouring in which each k-set is assigned an element of $\binom{[q]}{s}$, contains a monochromatic $K_t^{(k)}$. Here, monochromatic means the intersection of the colour sets assigned to the edges is nonempty. Observe that certain cases of R_k and r_k coincide. For example, $R_k(t; q, q-1) = r_k(t; q, q)$ and in general, we have the bound

$$r_k(t; q, p) \leq R_k\left(t; \binom{q}{p-1}, \binom{q-1}{p-2}\right).$$

We prove lower bounds on $r_k(t; q, p)$, thus giving lower bounds on certain set-colouring Ramsey numbers. However, we are not able to definitively resolve any questions from [5], due to central gaps in our understanding of hypergraph Ramsey numbers. See Section 2 for more on this.

Our main tool in the proof of Theorem 1.2 is the development of two new stepping-up constructions which yield the first tower-type results of their kind. We show the following three stepping-up statements, listed in order of decreasing strength, with $C_k = \frac{1}{k+1}\binom{2k}{k}$ denoting the k-th Catalan number.

Theorem 1.3. *Let $k, q, p \geq 3$. There exist $c \geq 1$ and t_0 such that for all $t > t_0$,*

(a) if $p \leq C_k - 2$, then $r_{k+1}(t^c; q, p) > 2^{r_k(t;q,p)-1}$,

(b) if $p \leq C_k$, then $r_{k+1}(t^c; 2q + p, p) > 2^{r_k(t;q,p)-1}$, and

(c) if $p \leq k!$, then $r_{2k}(t^c; qp, p) > 2^{r_k(t;q,p)-1}$.

Note that the growth rate in k which is implied by Part (c) (approximately a tower of height $\log_2 k$) of Theorem 1.3 is much smaller than that of Parts (a) and (b) because we can only step up at the cost of doubling the uniformity size. Unfortunately, this does not allow us to answer the question of Conlon, Fox, and Rödl on $r_3(t; q, 3)$, since $C_2 = 2$, but already for $k \geq 4$ we have the following two corollaries:

Corollary 1.4. *For all $k \geq 4$, there is $q \in \mathbb{N}$ and $c > 0$ such that $r_k(t; q, 5) \geq T_{k-1}(t^c)$.*

Corollary 1.5. *For all $k \geq 4$, there is $c > 0$ such that $r_k(t; 3, 3) \geq T_{k-1}(t^c)$.*

Observe that by the second corollary the growth rate of $r_k(t; 3, 3)$ matches the current best lower bounds for $r_k(t)$ up to a polynomial in t. The reason we have an absolute constant c in the exponent is due to the use of an Erdős-Hajnal type result on sequences.

The second main element of our proof connects the problem of avoiding monochromatic balanced hedgehogs to that of avoiding cliques that span few colours. It is a straightforward adaptation of ideas from Conlon, Fox, and Rödl [3].

Lemma 1.6. *Given $k, q, t \in \mathbb{N}$, let $p = \binom{2k+1}{k+1}$ and $q' = \binom{q}{p}$. Then*

$$r_{2k+1}(\hat{H}_t; q', 2) > r_{k+1}(t; q, p+1) - 1.$$

Using this result along with Part (c) of Theorem 1.3 yields the lower bound in Theorem 1.2(b). It is natural to ask whether one can combine the growth rate in k given by Part (a) of Theorem 1.3 with the ability to impose as many colours as in Part (c). Unfortunately, the condition $p \leq C_k$ prevents us from using Part (a) as the right-hand side because $C_k = \frac{1}{k+1}\binom{2k}{k} < \binom{2k+1}{k+1}$. This is tantalisingly close, if not a little curious, as the dependence on C_k comes from our exact solution to a subsequence avoidance problem. We show that C_k presents a natural barrier in this endeavour. This barrier is made concrete by some new and tight results on the Ramsey theory of sequences, including an Erdős-Hajnal-type result.

2 Moving Forward

Both of our new stepping-up constructions rely on a dichotomy: either we can find many suitable substructures within the δ-sequences (which give rise to many colours) or we must have a long monotonic sequence (which allows us to use induction). Since for every k-edge there are at most $k!$ distinct permutations, our methods fail to give good lower bounds for $r_k(t; q, p)$ whenever $k \ll p$. Even in the simplest case $r_3(t; q, 3)$, we were not able to prove a double exponential lower bound, leaving open the following question of Conlon, Fox, and Rödl on $r_3(t; q, 3)$.

Problem 2.1. *[3, Problem 1] Is there an integer q, a positive constant c, and a q-colouring of the 3-uniform hypergraph on $2^{2^{ct}}$ vertices such that every subset of order t receives at least 3 colours?*

We propose here a much weaker problem than Problem 2.1 which we were not able to resolve. We note that a negative answer would uncover a radical new phenomenon in the Ramsey numbers of hypergraphs.

Problem 2.2. *Does there exist $k \in \mathbb{N}$ such that the following holds? For all $p \in \mathbb{N}$ there exist $q \in \mathbb{N}$ and $c > 0$ such that $r_k(t; q, p) \geq 2^{2^{t^c}}$ for all t sufficiently large.*

A similar but much more ambitious problem was posed in [5].

Problem 2.3. *[5, Problem 6.3] Determine the tower height of $R_k(n; r, r-1) = r_k(n; r, r)$ for all $k \geq 3$ and $r \geq 2$.*

The authors of [5] note the apparent difficulty of Problem 2.3 and ask the following weaker question. Is there a fixed integer c such that $R_k(n; r, r-1) \geq T_{k-c}(n)$ for every $k \geq 3$ and $r \geq 2$? We cannot answer this question, but using Theorem 1.3(a), we can prove $R_k(n; r, r-1)$ is at least a tower of height roughly $k - 0.5 \log_2 r$. Any improvement beyond this bound would likely be very interesting.

We make the following conjecture regarding the Ramsey numbers of k-uniform hedgehogs. This would in particular demonstrate that the 2-colour and q-colour Ramsey numbers of these hedgehogs, unlike those of balanced hedgehogs, do not differ by arbitrarily large tower heights.

Conjecture 2.4. *There is $\ell \in \mathbb{N}$ such that for every positive integer k, for every sufficiently large t,*

$$r_k(H_t^{(k)}) \geq T_{k-\ell}(t).$$

Acknowledgments

We thank Bhargav Narayanan for helpful conversations on this topic and the anonymous referees for their detailed and invaluable feedback, especially regarding the organization of the first two sections. The first author was supported in part by Simons Foundation grant

332622. The first and fourth authors were supported in part by NSF Grant CCF-1814409 and NSF Grant DMS-1800521. The second author was supported by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC-2181/1 - 390900948 (the Heidelberg STRUCTURES Cluster of Excellence) and by EPSRC grant EP/V007327/1. The third author was supported in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 428212407.

References

[1] N. Alon and J.H. Spencer. *The Probabilistic Method.* John Wiley & Sons, 2016. 382

[2] Lucas Aragão, Maurício Collares, João Pedro Marciano, Taísa Martins, and Robert Morris. A lower bound for set-colouring Ramsey numbers. *arXiv preprint arXiv:2212.06802*, 2022. 382

[3] D. Conlon, J. Fox, and V. Rödl. Hedgehogs are not color blind. *J. Comb.*, 8(3):475–485, 2017. 381, 382, 383, 384

[4] D. Conlon, J. Fox, and B. Sudakov. Hypergraph Ramsey numbers. *J. Amer. Math. Soc.*, 23(1):247–266, 2010. 381

[5] David Conlon, Jacob Fox, Xiaoyu He, Dhruv Mubayi, Andrew Suk, and Jacques Verstraëte. Set-coloring Ramsey numbers via codes. *arXiv preprint arXiv:2206.11371*, 2022. 382, 384

[6] P. Erdős and A. Hajnal. On Ramsey like theorems, problems and results. In *Combinatorics (Proc. Conf. Combinatorial Math., Math. Inst., Oxford, 1972)*, pages 123–140, 1972. 381

[7] P. Erdős and R. Rado. Combinatorial theorems on classifications of subsets of a given set. *Proc. London Math. Soc.*, 3(1):417–439, 1952. 381

[8] P. Erdős, A. Hajnal, and R. Rado. Partition relations for cardinal numbers. *Acta Math. Acad. Sci. Hung.*, 16:93–196, 1965. 381, 382

[9] P Erdős and A Szemerédi. On a Ramsey type theorem. *Periodica Mathematica Hungarica*, 2(1-4):295–299, 1972. 382

[10] R.L. Graham, B.L. Rothschild, and J.H. Spencer. *Ramsey Theory*, volume 20. John Wiley & Sons, 1990. 381

[11] D. Mubayi and A. Suk. Cliques with many colors in triple systems. *arXiv:2005.03078*, 2020. 382

[12] Xiaodong Xu, Zehui Shao, Wenlong Su, and Zhenchong Li. Set-coloring of edges and multigraph Ramsey numbers. *Graphs and Combinatorics*, 25(6):863–870, 2009. 382

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

On the minimum number of inversions to make a digraph k-(arc-)strong.

(Extended abstract)

Julien Duron[*] Frédéric Havet[†] Florian Hörsch[‡]
Clément Rambaud[†§]

Abstract

The *inversion* of a set X of vertices in a digraph D consists of reversing the direction of all arcs of $D\langle X\rangle$. We study $\mathrm{sinv}'_k(D)$ (resp. $\mathrm{sinv}_k(D)$) which is the minimum number of inversions needed to transform D into a k-arc-strong (resp. k-strong) digraph and $\mathrm{sinv}'_k(n) = \max\{\mathrm{sinv}'_k(D) \mid D \text{ is a } 2k\text{-edge-connected digraph of order } n\}$. We show : (i) $\frac{1}{2}\log(n-k+1) \leq \mathrm{sinv}'_k(n) \leq \log n + 4k - 3$ for all $n \in \mathbb{Z}_{\geq 0}$; (ii) for any fixed positive integers k and t, deciding whether a given oriented graph $\vec{G}$ satisfies $\mathrm{sinv}'_k(\vec{G}) \leq t$ (resp. $\mathrm{sinv}_k(\vec{G}) \leq t$) is NP-complete ; (iii) if T is a tournament of order at least $2k+1$, then $\mathrm{sinv}'_k(T) \leq \mathrm{sinv}_k(T) \leq 2k$, and $\frac{1}{2}\log(2k+1) \leq \mathrm{sinv}'_k(T) \leq \mathrm{sinv}_k(T)$ for some T; (iv) if T is a tournament of order at least $28k-5$ (resp. $14k-3$), then $\mathrm{sinv}_k(T) \leq 1$ (resp. $\mathrm{sinv}_k(T) \leq 6$); (v) for every $\epsilon > 0$, there exists C such that $\mathrm{sinv}_k(T) \leq C$ for every tournament T on at least $2k+1+\epsilon k$ vertices.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-054`

1 Introduction

Notation not given below is consistent with [7]. In particular, a digraph may contain digons but no loops or parallel arcs and an oriented graph is a digraph without digons. We denote by $[k]$ the set $\{1, 2, \ldots, k\}$.

[*]École normale supérieure de Lyon, LIP, France
[†]Université Côte d'Azur, CNRS, Inria, I3S, Sophia Antipolis, France
[‡]CISPA Saarbrücken, Germany
[§]DIENS, École normale supérieure, CNRS, PSL University, Paris, France

A **feedback arc set** in a digraph is a set of arcs whose reversal results in an acyclic digraph. Finding a minimum cardinality feedback arc set is one of the first problems shown to be NP-hard listed by Karp in [18]. Furthermore, it is hard to approximate [17, 12]. For tournaments, the problem remains NP-complete [2, 11], but there is a 3-approximation algorithm [1] and a polynomial-time approximation scheme [19].

To make a digraph D acyclic, one can use a different operation from arc reversal, called inversion. The **inversion** of a set X of vertices consists in reversing the direction of all arcs of $D\langle X\rangle$, the subdigraph induced by X. We say that we **invert** X in D. The resulting digraph is denoted by $\mathrm{Inv}(D; X)$. If $(X_i)_{i\in I}$ is a family of subsets of $V(D)$, then $\mathrm{Inv}(D; (X_i)_{i\in I})$ is the digraph obtained after inverting the X_i one after another. Observe that this is independent of the order in which we invert the X_i : $\mathrm{Inv}(D; (X_i)_{i\in I})$ is obtained from D by reversing the arcs such that an odd number of the X_i contain its two end-vertices. The **inversion number** of an oriented graph D, denoted by $\mathrm{inv}(D)$, is the minimum number of inversions needed to transform D into an acyclic oriented graph. It was first introduced by Belkhechine et al. in [10] and then studied in several papers [6, 23, 4, 3].

The main purpose of this article is to study the possibilities of applying the inversion operation to obtain a different objective than the obtained digraph being acyclic. Instead of making a digraph acyclic, we are interested in making it satisfy a prescribed connectivity property. A digraph D is **strongly connected** or simply **strong** (resp. **k-arc-strong**) for some positive integer k, if for any partition (V_1, V_2) of $V(D)$ with $V_1, V_2 \neq \emptyset$ there is an arc (resp. at least k arcs) with tail in V_1 and head in V_2. For a given digraph D, we denote by $\mathrm{UG}(D)$ the undirected (multi)graph that we obtain by suppressing the orientations of the arcs. A digraph is **k-connected** (resp. **k-edge-connected**) if its underlying (multi)graph is. Clearly, a digraph D can be made k-arc-strong by reversing some arcs if and only if the edges of $\mathrm{UG}(D)$ can be oriented such that the resulting digraph is k-arc-strong. Robbins' Theorem [22] asserts that a graph admits a strong orientation if and only if it is 2-edge-connected, and more generally, Nash–Williams' orientation theorem [21], asserts that a graph admits a k-arc-strong orientation if and only if it is $2k$-edge-connected. It is well known that, by reducing to a minimum-cost submodular flow problem, one can determine, in polynomial time, a minimum set of arcs in D whose reversal gives a k-arc-strong digraph or detect that such a set does not exist, see Section 8.8.4 of [7] for details. The digraphs that contain a linear number of vertices with no outgoing arc show that the number of necessary arc reversals to make a $2k$-edge-connected digraph D k-arc-strong cannot be bounded by a function depending only on k. However this is the case for **tournaments**, which are the orientations of complete graphs: Bang-Jensen and Yeo [5] proved that every tournament on at least $2k+1$ vertices can be made k-arc-strong by reversing at most $\frac{1}{2}k(k-1)$ arcs. This result is tight for transitive tournaments.

We are interested in the problem of using inversions to make a digraph k-arc-strong. The **k-arc-strong inversion number** of a digraph D, denoted by $\mathrm{sinv}'_k(D)$, is the minimum number of inversions needed to transform D into a k-arc-strong digraph. We study $\mathrm{sinv}'_k(n) = \max\{\mathrm{sinv}'_k(D) \mid D \text{ } 2k\text{-edge-connected digraph of order } n\}$. For all $n \in \mathbb{Z}_{\geq 0}$, we show

$$\frac{1}{2}\log(n-k+1) \leq \mathrm{sinv}'_k(n) \leq \log n + 4k - 3.$$

To establish the upper bound, it is enough to consider minimally k-edge-connected digraphs which are k-edge-connected digraphs D such that $D \setminus uv$ is not k-edge connected for any arc uv of D. We show that such a digraph D is d-degenerate for $d = 2k-1$, that is, every subdigraph of D has a vertex of degree at most d. Now a result of a forthcoming paper [16] by a group containing the authors asserts that any orientation $\vec{G}_1$ of an n-vertex d-degenerate graph G can transformed into any other orientation of $\vec{G}_2$ of G by inverting at most $\log n + 2d - 1$ sets. Together with a slight strengthening of Nash-Williams' Theorem, we deduce that $\mathrm{sinv}'_k(D) \leq \log n + 2(2k-1) - 1$.

Then, we prove that, for any fixed positive integers k and t, deciding whether a given oriented graph $\vec{G}$ satisfies $\mathrm{sinv}'_k(\vec{G}) \leq t$ is NP-complete. The case $t = 1$ is proved using a reduction from Monotone Equitable k-SAT. An instance of this problem consists of a set of variables X and a set of clauses $\mathcal{C}$ each of which contains exactly $2k+1$ nonnegated variables and the question is whether there is a truth assignment $\phi : X \to \{\mathrm{true}, \mathrm{false}\}$ such that every clause in $\mathcal{C}$ contains at least k true and k false variables with respect to ϕ. The case $t \geq 2$ is proved using a reduction from k-Cut-Covering. Given a graph G, this problem consists in deciding whether there is a collection $F_1, \dots, F_t$ of cuts such that $\cup_{i=1}^t F_i = E(G)$. We also show that, unless P=NP, sinv'_k cannot be approximated within a factor better than 2.

One may also want to make a digraph k-strong. A digraph D is **k-strong** if $|V(D)| \geq k+1$ and for any set $S \subseteq V(D)$ with less than k vertices $D - S$ is strong. A digraph which can be made k-strong by reversing arcs is **k-strengthenable**. The 1-strengthenable digraphs are the 2-edge-connected ones, because being 1-strong is equivalent to being strong or 1-arc-strong. Thomassen [24] proved that the 2-strengthenable digraphs are the 4-edge-connected digraphs D such that $D - v$ is 2-edge-connected for every vertex $v \in V(D)$, but it is NP-hard to compute the minimum number of arc reversals needed to make a given digraph 2-strong [8]. Furthermore, in contrast to the analogous problem for k-arc-strengthenable digraphs, for $k \geq 3$, it is NP-complete to decide whether a digraph is k-strengthenable. Indeed, for any $k \geq 3$, it is NP-complete to decide whether an undirected graph has a k-strong orientation [13].

It is also natural to use inversions to make a digraph k-strong. A **k-strengthening family** of a digraph D is a family of subsets $(X_i)_{i\in I}$ of subsets of $V(D)$ such that $\mathrm{Inv}(D; (X_i)_{i\in I})$ is k-strong. The **k-strong inversion number** of a k-strengthenable digraph D, denoted by $\mathrm{sinv}_k(D)$, is the minimum number of inversions needed to transform D into a k-strong digraph. We show that for any positive integers k and t, it is NP-complete to decide whether $\mathrm{sinv}_k(D) \leq t$ for a given k-strengthenable oriented graph. We also show that, unless P=NP, sinv_k cannot be approximated within a factor better than 2. The proofs are similar to the ones for sinv'_k.

It is not hard to show that every tournament of order at least $2k+1$ is k-strengthenable and that it can be made k-strong by reversing the orientation of at most $\frac{1}{4}(4k-2)(4k-$

3) arcs, see e.g. [7, p. 379]. In 1994, Bang-Jensen conjectured that every tournament on at least $2k+1$ vertices can be made k-strong by reversing at most $\frac{1}{2}k(k+1)$ arcs. Bang-Jensen, Johansen, and Yeo [9] proved this conjecture for tournaments of order at least $3k-1$. It is then natural to ask whether or not we can make a tournament k-strong or k-arc-strong in a lot less than $\frac{1}{2}k(k+1)$ inversions. This leads to consider $M_k = \max\{\mathrm{sinv}_k(T) \mid T$ tournament of order at least $2k+1\}$ and $M'_k = \max\{\mathrm{sinv}'_k(T) \mid T$ tournament of order at least $2k+1\}$. We show that (for sufficiently large k), we have

$$\frac{1}{2}\log(2k+1) \leq M'_k \leq M_k \leq 2k.$$

The lower bound is obtained for a tournament of order $2k+1$ by using McKay's result [20] on the number of Eulerian tournaments of order $2k+1$, the fact that every k-arc-strong tournament of order $2k+1$ is Eulerian and counting arguments. Let us now prove the upper bound.

Let D be a digraph and u, v two distinct vertices in D. The **strong-connectivity** from u to v in D, denoted by $\kappa_D(u,v)$, is the maximal number α such that $D - X$ contains a (u,v)-path for every $X \subseteq V(D) \setminus \{u,v\}$ with $|X| \leq \alpha - 1$. For some $S \subseteq V(D)$ and positive integer k, we say that S is **k-strong in** D if $\kappa_D(u,v) \geq k$ for all $u, v \in S$. The following statement is well-known.

Lemma 1.1. *Let D be a digraph, S a k-strong set in D and $v \in V(D) \setminus S$. If v has k in-neighbours in S and k out-neighbours in S, then $S \cup \{v\}$ is k-strong in D.*

Theorem 1.2. $M_k \leq 2k$.

Proof. Let D be a tournament with $V(D) = \{v_1, \ldots, v_n\}$ with $n \geq 2k+1$. Further, let T be a k-strong tournament on $\{v_1, \ldots, v_{2k+1}\}$. We now define sets $X_1, \ldots, X_{2k}$. Suppose that the sets $X_1, \ldots, X_{i-1}$ have already been created and let D_{i-1} be the graph obtained from D by inverting $X_1, \ldots, X_{i-1}$. Now let $X_i = \{v_i\} \cup A_i \cup B_i$, where A_i is the set of vertices v_j with $j \in \{i+1, \ldots, 2k+1\}$ for which the edge v_iv_j has a different orientation in T and D_{i-1}, and B_i is, when $i \leq k$ (resp. $i \geq k+1$), the set of vertices v_j with $j \geq 2k+2$ for which D_{i-1} contains the arc v_iv_j (resp. v_jv_i).

Observe that $D_{2k}\langle\{v_1, \ldots, v_{2k+1}\}\rangle = T$ which is k-strong by assumption. Moreover, for any $j \geq 2k+2$, D_{2k} contains the arcs v_jv_i for $i \in [k]$ and the arcs v_iv_j for $i = k+1, \ldots, 2k$. Hence, by Lemma 1.1, D_{2k} is k-strong. □

We also prove that $M_1 = M'_1 = 1$ and $M_2 = M'_2 = 2$ showing that the bound $M_k \leq 2k$ is not tight for $k = 1, 2$. We believe that it is also not tight for larger values of k.

It is not too difficult to prove that every sufficiently large tournament can be made k-strong in one inversion. Hence it is natural to investigate the minimum integer $N_k(i)$ (resp. $N'_k(i)$) such that $\mathrm{sinv}_k(T) \leq i$ (resp. $\mathrm{sinv}'_k(T) \leq i$) for every tournament T of order at least $N_k(i)$. We prove

$$5k-2 \leq N_k(1) \leq 28k-5 \text{ and } N_k(6) \leq 14k-3.$$

The lower bound $N_k(1) \geq 5k-2$ is obtained by considering a tournament T of order $5k-3$ whose vertex set has a partition (A, B, C) such that $T\langle A\rangle$ and $T\langle C\rangle$ are $(k-1)$-diregular tournaments of order $2k-1$, and $A \Rightarrow B \cup C$ and $B \Rightarrow C$, and proving $\mathrm{sinv}'_k(T) > 1$.

The upper bounds $N_k(1) \leq 28k-5$ and $N_k(6) \leq 14k-3$ are obtained using median orders. A **median order** of D is an ordering $(v_1, v_2, \dots, v_n)$ of the vertices of D with the maximum number of forward arcs, that arcs v_iv_j with $j > i$. Our proofs use the two well-known properties (M1) and (M2) in the next lamma (the feedback property in [15]), which allow to prove the third one (M3). We denote by $R^+_D(v)$ (resp. $R^-_D(v)$) the set of vertices which are **reachable** from vertex v (resp. from which v can be reached) in digraph D, that are the vertices w such that there is a directed (v, w)-path (resp. (w, v)-path) in D.

Lemma 1.3. *Let T be a tournament and $(v_1, v_2, \dots, v_n)$ a median order of T. Then, for any two indices i, j with $1 \leq i < j \leq n$:*

(M1) *$(v_i, v_{i+1}, \dots, v_j)$ is a median order of the induced subtournament $T\langle\{v_i, v_{i+1}, \dots, v_j\}\rangle$.*

(M2) *v_i dominates at least half of the vertices $v_{i+1}, v_{i+2}, \dots, v_j$, and v_j is dominated by at least half of the vertices $v_i, v_{i+1}, \dots, v_{j-1}$. In particular, each vertex v_i, $1 \leq i < n$, dominates its successor v_{i+1}.*

(M3) *For any $X \subseteq V(T) \setminus \{v_i\}$, $|R^+_{T-F}(v_i)| \geq n+1-i-2|F|$, and $R^-_{T-F}(v_i) \geq i-2|F|$.*

Let T be a tournament of order $n \geq 28k-5$ and let $(v_1, \dots, v_n)$ be a median order of $V(T)$. Let $A = \{v_{n-6k+1}, \dots, v_n\}$ and $B = \{v_1, \dots, v_{6k}\}$. Using Lemma 1.3, we show that there is a set $X \subseteq A \cup B$ such that in the tournament $T_0 = \mathrm{Inv}(T\langle A \cup B\rangle, X)$, for any $Y \subseteq V(T_0)$ with $|Y| \leq k-1$, there is a directed path from a to $B \setminus Y$ in $T_0 - Y$ for every $a \in A \setminus Y$, and there is a directed path from $A \setminus Y$ to b in $T_0 - Y$ for every $b \in B \setminus Y$. We then show that $\mathrm{Inv}(T, X)$ is k-strong. This proves $N_k(1) \leq 28k-5$.

The fact that there exists a constant $\alpha > 0$ such that every tournament on at least αk vertices can be made k-strong by a single inversion raises the following question: for which $\alpha > 2$, every tournament on at least αk vertices can be made k-strong by a constant number of inversion? We show that every $\alpha > 2$ will do : there is a function f such that for every $\epsilon > 0$ and $k \in \mathbb{N}$, $\mathrm{sinv}_k(T) \leq f(\epsilon)$ for every tournament T on at least $2k+1+\epsilon k$ vertices.

The proof is based on a probabilistic argument: we show that $f(\epsilon)$ inversions drawn uniformly at random, under the constraint that they cover all the vertices, make such a tournament k-strong with high probability.

Finally, the fact that $m_k(n) = 1$ for n sufficiently large (in comparison to k) implies that the set $\mathcal{F}_k$ of tournaments T such that $\mathrm{sinv}_k(T) > 1$ is finite. This implies that for fixed k computing sinv_k and sinv'_k can be done in polynomial time for tournaments.

The proofs of the results announced in this extended abstract can be found in the full version of the paper [14].

References

[1] Nir Ailon, Moses Charikar, and Alantha Newman. Aggregating inconsistent information: Ranking and clustering. *Journal of the ACM*, 55(5), 2008.

[2] Noga Alon. Ranking tournaments. *SIAM Journal on Discrete Mathematics*, 20:137–142, 2006.

[3] Noga Alon, Emil Powierski, Michael Savery, Alex Scott, and Elizabeth Wilmer. Invertibility of digraphs and tournaments, 2022.

[4] Guillaume Aubian, Frédéric Havet, Florian Hörsch, Felix Klingelhoefer, Nicolas Nisse, Clément Rambaud, and Quentin Vermande. Problems, proofs, and disproofs on the inversion number, 2022.

[5] Jœrgen Bang-Jensen and Anders Yeo. Making a tournament k-arc-strong by reversing or deorienting arcs. *Discrete Applied Mathematics*, 136(2):161–171, 2004. The 1st Cologne-Twente Workshop on Graphs and Combinatorial Optimization.

[6] Jørgen Bang-Jensen, Jonas Costa Ferreira da Silva, and Frédéric Havet. On the inversion number of oriented graphs. *Discrete Mathematics & Theoretical Computer Science*, 23(2):# 8, 2022.

[7] Jørgen Bang-Jensen and Gregory Z. Gutin. *Digraphs: theory, algorithms and applications*. Springer-Verlag, London, 2009.

[8] Jørgen Bang-Jensen, Florian Hörsch, and Matthias Kriesell. Complexity of (arc)-connectivity problems involving arc-reversals or deorientations, 2023.

[9] Jørgen Bang-Jensen, Kasper S. Johansen, and Anders Yeo. Making a tournament k-strong. *Journal of Graph Theory*, 103(1):5–11, 2023.

[10] Houmem Belkhechine, Moncef Bouaziz, Imed Boudabbous, and Maurice Pouzet. Inversion dans les tournois. *Comptes Rendus Mathematique*, 348(13-14):703–707, 2010.

[11] Pierre Charbit, Stéphan Thomassé, and Anders Yeo. The minimum feedback arc set problem is NP-hard for tournaments. *Combinatorics, Probability and Computing*, 16:1–4, 2007.

[12] Irit Dinur and Samuel Safra. On the hardness of approximating minimum vertex cover. *Annals of Mathematics, Second Series*, 162(1):439–485, 2005.

[13] Olivier Durand de Gevigney. On Frank's conjecture on k-connected orientations. *Journal of Combinatorial Theory, Series B*, 141:105–114, 2020.

[14] Julien Duron, Frédéric Havet, Florian Hörsch, and Clément Rambaud. On the minimum number of inversions to make a digraph k-(arc-)-strong, 2023.

[15] Frédéric Havet and Stéphan Thomassé. Median orders of tournaments: A tool for the second neighborhood problem and sumner's conjecture. *Journal of Graph Theory*, 35(4):244–256, 2000.

[16] ANR Digraph Inversion Working Group. On the diameter of the inversion graph. unpublished.

[17] Vido Kann. *On the Approximability of NP-complete Optimization Problems*. PhD thesis, Department of Numerical Analysis and Computing Science, Royal Institute of Technology, Stockholm, 1992.

[18] Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, The IBM Research Symposia Series, pages 85–103. Plenum Press, New York, 1972.

[19] Claire Kenyon-Mathieu and Warren Schudy. How to rank with few errors. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '07, page 95–103, 2007.

[20] Brendan D. McKay. The asymptotic numbers of regular tournaments, eulerian digraphs and eulerian oriented graphs. *Combinatorica*, 10(4):367–377, 1990.

[21] Crispin St J.A. Nash-Williams. On orientations, connectivity and odd-vertex-pairings in finite graphs. *Canadian Journal of Mathematics*, 12:555–567, 1960.

[22] Herbert E. Robbins. A theorem on graphs, with an application to a problem on traffic control. *American Mathematical Monthly*, 46(5):281–283, 1939.

[23] Bhalchandra Thatte, Hamza Si Kaddour, and Maurice Pouzet. On the boolean dimension of a graph and other related parameters. *Discrete Mathematics & Theoretical Computer Science*, 23, 2022.

[24] Carsten Thomassen. Strongly 2-connected orientations of graphs. *Journal of Combinatorial Theory, Series B*, 110:67–78, 2015.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Precoloring extension in planar near-Eulerian-triangulations*

(Extended abstract)

Zdeněk Dvořák Benjamin Moore Michaela Seifrtová
Robert Šámal†

Abstract

We consider the 4-precoloring extension problem in *planar near-Eulerian- triangulations*, i.e., plane graphs where all faces except possibly for the outer one have length three, all vertices not incident with the outer face have even degree, and exactly the vertices incident with the outer face are precolored. We give a necessary topological condition for the precoloring to extend, and give a complete characterization when the outer face has length at most five and when all vertices of the outer face have odd degree and are colored using only three colors.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-055`

1 Introduction

Recall that a *k-coloring* of a graph G is a mapping using k colors such that adjacent vertices receive different colors and that a graph is *Eulerian* if all of its vertices have even degree. We study the precoloring extension problem for planar (near) Eulerian triangulations, in particular from an algorithmic perspective.

Famously, the Four Color Theorem states that all planar graphs are 4-colorable [1] and thus from an algorithmic point of view, the problem of determining if a planar graph is 4-colorable is trivial. In contrast, deciding if a planar graph is 3-colorable is a well known

*Supported by project 22-17398S (Flows and cycles in graphs on surfaces) of Czech Science Foundation.
†Computer Science Institute, Charles University, Prague, Czech Republic. E-mail: `{rakdver,brmoore,mikina,samal}@iuuk.mff.cuni.cz`

NP-complete problem [20]. If we move to graphs on surfaces, the situation becomes less clear. Recall that a graph G is $(k+1)$-*critical* if all proper subgraphs of G are k-colorable, but G itself is not. Thus, $(k+1)$-critical graphs are exactly the minimal forbidden subgraphs for k-colorability. A deep result of Thomassen [22] says that for any fixed surface Σ, there are only finitely many $(k+1)$-critical graphs for $k \geq 5$, and combined with a result of Eppstein [7], this implies that there is a polynomial time algorithm for k-coloring graphs on any fixed surface for any $k \geq 5$. Unfortunately, we cannot extend this to 4-colorablity—it is known that there are infinitely many 5-critical graphs on any surface other than the plane [22]. This is a consequence of the folowing result of Fisk [9]: If G is a triangulation of an orientable surface and G has exactly two vertices u and v of odd degree, then u and v must have the same color in any 4-coloring of G, and thus the graph $G + uv$ is not 4-colorable. Even though there are infinitely many 5-critical graphs, it is an important open question if for any fixed surface Σ, there is a polynomial time algorithm to decide if a graph drawn on Σ is 4-colorable.

Let us remark that we have a positive answer to a similar question in the case of 3-coloring triangle-free graphs on surfaces. It is known that there are infinitely many 4-critical triangle-free graphs on all surfaces other than the plane, yet there is a linear time algorithm to decide if a triangle-free graph on any fixed surface is 3-colorable [6]. The algorithm consists of two parts: In the first part, the problem is reduced to (near) quadrangulations [5], and the second part gives a topological criterion for 3-colorability of near quadrangulations [2].

Our hope is that (near) Eulerian triangulations could play the same intermediate role in the case of 4-colorability of graphs on surfaces. Indeed, there is a number of arguments and analogies supporting this idea:

(A) The only constructions of "generic" (e.g., avoiding non-trivial small separations) non-4-colorable graphs drawn on a fixed surface that we are aware of are based on near Eulerian triangulations, such as Fisk's construction [9] or adding vertices to faces of non-3-colorable quadrangulations of the projective plane [13].

(B) As noted above, quadrangulations play key role in the problem of 3-colorability of triangle-free graphs on surfaces, which is a bit surprising at a glance since planar quadrangulations are actually 2-colorable. Analogously, planar Eulerian triangulations are 3-colorable (and in fact, a planar graph is 3-colorable if and only if its a subgraph of a planar Eulerian triangulation).

(C) Many results for quadrangulations of surfaces correspond to results for Eulerian triangulations. As an example, Youngs [25] famously proved that a graph drawn in the projective plane so that all faces have even length is 3-colorable if and only if it does not contain a non-bipartite quadrangulation as a subgraph. For an Eulerian triangulation T of the projective plane, Fisk [8] showed that T has an independent set U such that all faces of $T - U$ have even length, and Mohar [13] proved that T is 4-colorable if and only if $T - U$ is 3-colorable.

As another example, Hutchinson [11] proved that every graph drawn on a fixed orientable surface with only even-length faces and with sufficiently large *edgewidth* (the length of the shortest non-contractible cycle) is 3-colorable, and Nakamoto et al. [16] and Mohar and Seymour [14] have shown that such graphs on non-orientable surfaces are 3-colorable unless they contain a quadrangulation with an odd-length orienting cycle. Analogously, for any orientable surface, any Eulerian triangulation with sufficiently large edgewidth is 4-colorable [12], and for non-orientable surfaces, the only non-4-colorable Eulerian triangulations of large edgewidth are those that have an independent set whose removal results in an even-faced non-3-colorable graph [15].

In this paper, we make the first step towards towards the design of a polynomial-time algorithm to decide whether an Eulerian triangulation of a fixed surface is 4-colorable. In particular, we give the following algorithm. A *planar near-Eulerian-triangulation* is a plane graph where all the faces except possibly for the outer one have length three and all the vertices not incident with the outer face have even degree.

Theorem 1. *There is a linear-time algorithm that given*

- *a planar near-Eulerian-triangulation G with the outer face bounded by a cycle C such that all vertices of C have odd degree in G, and*
- *a precoloring φ of the vertices incident with the outer face of G using only three colors,*

decides whether φ extends to a 4-coloring of G.

The motivation for considering the special case of planar near-Eulerian-triangulations with precolored outer face comes from the general approach towards solving problems for graphs on surfaces, which can be seen e.g. in [2, 3, 4, 18, 19, 21, 23], as well as many other works and is explored systematically in the hyperbolic theory of Postle and Thomas [17]. The general outline of this approach is as follows:

- Generalize the problem to surfaces with boundary, with the boundary vertices precolored (or otherwise constrained).
- Use this generalization to reduce the problem to "generic" instances (e.g., those without short non-contractible cycles, since if an instance contains a short non-contractible cycle, we can cut the surface along the cycle and try to extend all the possible precolorings of the cycle in the resulting graph drawn in a simpler surface).
- The problem is solved in the basic case of graphs drawn in a disk and on a cylinder (plane graphs with one or two precolored faces).
- Finally, the general case is solved with the help of the two basic cases (reducing it to the basic cases by further cutting the surface and carefully selecting the constraints on the boundary vertices [2, 19], using quantitative bounds from the basic cases to show that truly generic cases do not actually arise [17]).

Thus, Theorem 1 is a step towards solving the basic case of graphs drawn in a disk. It unfortunately does not solve this case fully because of the extra assumption that φ only uses three colors (and the assumption that vertices of C have odd degree). Without this assumption, we were able to solve the problem when the precolored outer face has length at most five.

Theorem 2. *There is a linear-time algorithm that given*

- *a planar near-Eulerian-triangulation G with the outer face bounded by a cycle C of length at most five and*
- *a precoloring φ of C*

decides whether φ extends to a 4-coloring of G.

In general, we were only able to find a necessary topological condition for such an extension to exist (which we do not discuss in this extended abstract). We conjecture that using this topological condition in combination with further ideas, it will be possible to resolve the disk case in full.

Conjecture 3. *For every positive integer ℓ, there is a polynomial-time algorithm that given*

- *a planar near-Eulerian-triangulation G with the outer face of length at most ℓ and*
- *a precoloring φ of the vertices incident with the outer face*

decides whether φ extends to a 4-coloring of G.

For the remaining part of the paper, we sketch some of the ideas needed to prove Theorem 1 and Theorem 2.

2 Proving Theorems 1 and 2

Our goal in this section is to show the precoloring extension problem is equivalent to special homomorphisms to the infinite triangular grid equipped with colorings, and then show that in the special cases of Theorems 1 and 2 we can decide if such a homomorphism exists. We need some definitions.

A *hued graph* is a graph G together with a proper 3-coloring $\psi_G : V(G) \to \mathbb{Z}_3$. We will also need to fix a 4-coloring of G, and for notational convenience we use elements of $\mathbb{Z}_2^2$ as colors. With this, a *dappled graph* is a hued graph G together with a proper 4-coloring $\varphi_G : V(G) \to \mathbb{Z}_2^2$. For a vertex $v \in V(G)$, we say that $\psi_G(v)$ is the *hue* and $\varphi_G(v)$ is the *color* of v. *Homomorphisms* of dappled graphs are required to preserve edges and both hue and color, i.e., $f : V(G) \to V(H)$ is a homomorphism if $f(u)f(v) \in E(H)$ for every $uv \in E(G)$ and $\psi_H(f(v)) = \psi_G(v)$ and $\varphi_H(f(v)) = \varphi_G(v)$ for every $v \in V(G)$.

The *dappled triangular grid* is the infinite dappled graph $\mathbf{T}$ with vertex set $\{(i,j) : i,j \in \mathbb{Z}\}$, where vertices (i_1, j_1) and (i_2, j_2) are adjacent if and only if $(i_2 - i_1, j_2 - j_1) \in$

$\{\pm(1,0), \pm(0,1), \pm(1,1)\}$, with vertex hue $\psi_{\mathbf{T}}(i,j) = (i+j) \bmod 3$ for each vertex (i,j), and with vertex color $\varphi_{\mathbf{T}}(i,j) = (i \bmod 2, j \bmod 2)$ for each vertex (i,j).

The following claim follows by an inspection of the definition.

Observation 4. *Let* $\mathbf{T}$ *be the dappled triangular grid. For any vertex* $v \in V(\mathbf{T})$*, if* u_1 *and* u_2 *are distinct neighbors of* v*, then* $(\psi_{\mathbf{T}}(u_1), \varphi_{\mathbf{T}}(u_1)) \neq (\psi_{\mathbf{T}}(u_2), \varphi_{\mathbf{T}}(u_2))$*. Consequently, for any connected dappled graph* G *and any vertex* $x \in V(G)$*, if* $f_1, f_2 : V(G) \to V(\mathbf{T})$ *are homomorphisms and* $f_1(x) = f_2(x)$*, then* $f_1 = f_2$*.*

Let us now give the key property of dappled patches; the proof easily follows from the coloring-flow duality of Tutte [24].

Theorem 5. *Every dappled patch* G *has a homomorphism to the dappled triangular grid* $\mathbf{T}$*.*

A 4-coloring φ of a connected hued graph H is *viable* if and only if the dappled graph H^φ where φ is the associated 4-coloring has a homomorphism to the dappled triangular grid. Let us note that by Observation 4, this condition is easy to verify, as such a homomorphism is unique up to the arbitrary choice of the image of a single vertex of C.

Corollary 6. *Let* G *be a hued patch and* φ *a* 4*-coloring of the boundary of the outer face of* G*. If* φ *extends to a* 4*-coloring of* G*, then* φ *is viable.*

With this, we now sketch how to prove Theorem 1 and 2. The key observation is that in both cases, the homomorphism to the dappled triangular grid $\mathbf{T}$ associated with the precoloring φ maps C to the closed neighborhood of a single vertex of $\mathbf{T}$.

A *hexagon* is a dappled subgraph of $\mathbf{T}$ induced by a vertex and its neighbors. A 4-coloring φ of a connected hued graph C is a *single-hexagon coloring* if it is viable and the corresponding homomorphism f maps C^φ to a subgraph of a hexagon H of $\mathbf{T}$. The *central hue* c and the *central color* k of a single-hexagon coloring is the hue and the color of the central vertex of H. There are two important examples of single-hexagon colorings, corresponding to the assumptions of Theorems 1 and 2, respectively.

- Let us call a patch *odd* if its outer face is bounded by a cycle C and all vertices incident with the outer face have odd degree. Every coloring of C that uses at most three colors is single-hexagon, with the central hue 2 and central color not appearing on C.
- Every viable 4-coloring of a hued (≤ 5)-cycle is single-hexagon.

A *retract* of a hued graph G is an induced subgraph H of G such that there exists a *retraction* f from G to H, i.e., a homomorphism such that $f(v) = v$ for each $v \in V(H)$. The following key observation follows from the fact that each shortest cycle in a bipartite graph is a retract [10]. For a dappled graph H, we let H^- refer to the underlying hued graph.

Lemma 7. *For every hexagon H of the triangular grid $\mathbf{T}$, the hued hexagon H^- is a retract of $\mathbf{T}^-$*

This implies that the precoloring extension problem with four colors in hued patches for single-hexagon precolorings can be reduced to the problem of 3-precoloring extension in bipartite graphs.

Corollary 8. *Let G be a hued patch and let $\varphi : V(C) \to \mathbb{Z}_2^2$ be a single-hexagon 4-coloring of the boundary C of the outer face of G, with central hue c and central color k. Let $K = \mathbb{Z}_2^2 \setminus \{k\}$. Let H be the bipartite subgraph of G induced by the vertices of hue different from c, and let $\varphi' : V(H) \to K$ be the restriction of φ to H. Then φ extends to a 4-coloring of G if and only if φ' extends to a 3-coloring of H (using the colors in K).*

Dvořák, Král' and Thomas [6] gave for every b a linear-time algorithm to decide whether a precoloring of at most b vertices of a planar triangle-free graph extends to a 3-coloring. Hence, we have the following consequence, a common strengthening of Theorems 1 and 2.

Corollary 9. *For every ℓ, there exists an algorithm than, given a hued patch G with the outer face of length at most ℓ and a single-hexagon precoloring φ of the boundary of the outer face, decides in linear time whether φ extends to a 4-coloring of G.*

References

[1] Kenneth Appel and Wolfgang Haken. Every planar map is four colorable. Part I: Discharging. *Illinois Journal of Mathematics*, 21(3):429 – 490, 1977.

[2] Zdeněk Dvořák, Daniel Král', and Robin Thomas. Three-coloring triangle-free graphs on surfaces VI. 3-colorability of quadrangulations. *ArXiv*, 1509.01013, September 2015.

[3] Zdeněk Dvořák, Daniel Král', and Robin Thomas. Three-coloring triangle-free graphs on surfaces II. 4-critical graphs in a disk. *Journal of Combinatorial Theory, Series B*, 132:1–46, 2018.

[4] Zdeněk Dvořák, Daniel Král', and Robin Thomas. Three-coloring triangle-free graphs on surfaces III. Graphs of girth five. *Journal of Combinatorial Theory, Series B*, 145:376–432, 2020.

[5] Zdeněk Dvořák, Daniel Král', and Robin Thomas. Three-coloring triangle-free graphs on surfaces IV. Bounding face sizes of 4-critical graphs. *Journal of Combinatorial Theory, Series B*, 150:270–304, 2021.

[6] Zdeněk Dvořák, Daniel Král', and Robin Thomas. Three-coloring triangle-free graphs on surfaces VII. a linear-time algorithm. *Journal of Combinatorial Theory, Series B*, 152:483–504, 2022.

[7] David Eppstein. Subgraph isomorphism in planar graphs and related problems. In *Proceedings of the Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '95, page 632–640, USA, 1995. Society for Industrial and Applied Mathematics.

[8] Steve Fisk. Geometric coloring theory. *Advances in Mathematics*, 24(3):298–340, 1977.

[9] Steve Fisk. The nonexistence of colorings. *J. Combin. Theory, Ser. B*, 24:247–248, 1978.

[10] Pavol Hell. Absolute retracts in graphs. In Ruth A. Bari and Frank Harary, editors, *Graphs and Combinatorics*, pages 291–301, Berlin, Heidelberg, 1974. Springer Berlin Heidelberg.

[11] Joan Hutchinson. Three-coloring graphs embedded on surfaces with all faces even-sided. *J. Combin. Theory, Ser. B*, 65:139–155, 1995.

[12] Joan Hutchinson, Bruce Richter, and Paul D. Seymour. Colouring Eulerian triangulations. *Journal of Combinatorial Theory, Series B*, 84:225–239, 2002.

[13] Bojan Mohar. Coloring Eulerian triangulations of the projective plane. *Discret. Math.*, 244:339–343, 2002.

[14] Bojan Mohar and Paul D. Seymour. Coloring locally bipartite graphs on surfaces. *J. Combin. Theory, Ser. B*, 84:301–310, 2002.

[15] Atsuhiro Nakamoto. 5-chromatic even triangulations on surfaces. *Discrete Mathematics*, 308(12):2571–2580, 2008.

[16] Atsuhiro Nakamoto, Seiya Negami, and Katsurhiro Ota. Chromatic numbers and cycle parities of quadrangulations on nonorientable closed surfaces. *Discrete Math.*, 285:211–218, 2004.

[17] Luke Postle and Robin Thomas. Hyperbolic families and coloring graphs on surfaces. *Transactions of the American Mathematical Society, Series B*, 5:167–221, 2018.

[18] Neil Robertson and Paul D. Seymour. Graph minors. VI. Disjoint paths across a disc. *Journal of Combinatorial Theory, Series B*, 41:115–138, 1986.

[19] Neil Robertson and Paul D. Seymour. Graph Minors VII. Disjoint paths on a surface. *J. Combin. Theory, Ser. B*, 45:212–254, 1988.

[20] Larry Stockmeyer. Planar 3-colorability is polynomial complete. *SIGACT News*, 5(3):19–25, jul 1973.

[21] Carsten Thomassen. Five-coloring graphs on surfaces. *J. Combin. Theory, Ser. B*, 59:89–105, 1993.

[22] Carsten Thomassen. Color-critical graphs on a fixed surface. *Journal of Combinatorial Theory, Series B*, 70(1):67–100, 1997.

[23] Carsten Thomassen. The chromatic number of a graph of girth 5 on a fixed surface. *J. Combin. Theory, Ser. B*, 87:38–71, 2003.

[24] William Tutte. A contribution to the theory of chromatic polynomials. *Canadian Journal of Mathematics*, 6:80–91, 1954.

[25] Daniel Youngs. 4-chromatic projective graphs. *Journal of Graph Theory*, 21:219–227, 1996.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Maximum genus orientable embeddings from circuit decompositions of dense eulerian graphs and digraphs

(Extended abstract)

M. N. Ellingham* Joanna A. Ellis-Monaghan†

Abstract

Suppose we have an eulerian (di)graph with a (directed) circuit decomposition. We show that if the (di)graph is sufficiently dense, then it has an orientable embedding in which the given circuits are facial walks and there are exactly one or two other faces. This embedding has maximum genus subject to the given circuits being facial walks. When there is only one other face, it is necessarily bounded by an euler circuit. Thus, if the numbers of vertices and edges have the same parity, a sufficiently dense (di)graph D with a given (directed) euler circuit C has an orientable embedding with exactly two faces, each bounded by an euler circuit, one of which is C. The main theorem encompasses several special cases in the literature, for example, when the digraph is a tournament.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-056`

1 Introduction and main results

When can a graph or digraph be cellularly embedding in a orientable surface so that it has exactly two faces, each bounded by an euler circuit, such as shown in Figure 1? Is it possible to specify one of the euler circuits in advance? When is it possible to specify

*Department of Mathematics, Vanderbilt University, 1326 Stevenson Center, Nashville, Tennessee 37240 USA. E-mail: `mark.ellingham@vanderbilt.edu`. Supported by Simons Foundation award 429625.

†Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Science Park 105-107, 1098XH Amsterdam, the Netherlands. E-mail: `j.a.ellismonaghan@uva.nl`.

an arbitrary circuit decomposition of the edges and complete it to an embedding with just one more face, noting that the face is then necessarily bounded by an euler circuit? Finding such a face achieves a maximum genus embedding having the circuits in a given decomposition as facial walks.

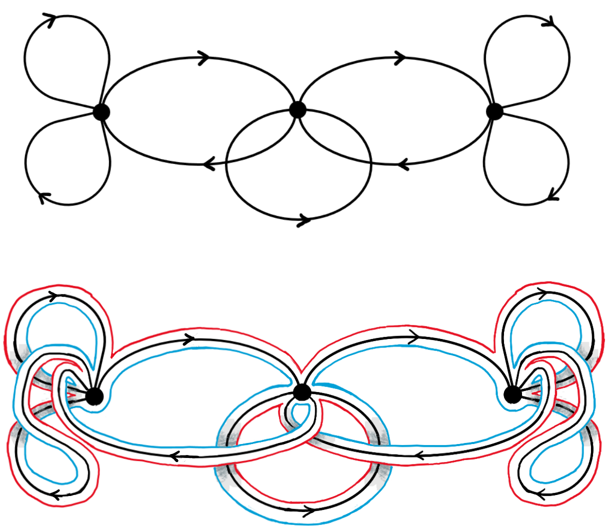


Figure 1: An eulerian digraph (top), and a bi-eulerian orientable embedding of it, showing the two faces bounded by euler circuits, one red and one blue (bottom). The diagram depicts the embedding as a ribbon graph, while sewing a disc into each facial walk gives the surface.

This leads more generally to the question of determining the maximum genus of an embedding relative to a circuit decomposition. Beyond topological graph theory, these questions arise in surprisingly diverse settings, including DNA self-assembly, Steiner triple systems, and latin squares.

Our main result, given in Theorem 1.1, is that if an eulerian (di)graph is sufficiently dense, then it is indeed always possible to achieve these special embeddings of maximum orientable genus.

We allow graphs and digraphs to have loops and multiple edges. A *circuit* in a graph is a closed trail. A graph is *eulerian* if it has a circuit (an *euler circuit*) that uses every edge and every vertex (so the graph is necessarily connected). An embedding of a digraph D is *directed* if every face is bounded by a closed directed walk of D. A *direction-indifferent* embedding of a digraph need not have consistently directed facial walks. In a directed embedding of a digraph, *profaces* are faces whose directed facial walks agree with the clockwise orientation of the surface, while the directed facial walks of the *antifaces* oppose it. We say that an embedding of a (di)graph G is *bi-eulerian* if it has exactly two faces, each bounded by an euler circuit.

Theorem 1.1. *Let D be an n-vertex eulerian digraph where the minimum degree of the underlying simple undirected graph of D is at least $(4n+2)/5$. Let $\mathcal{C}$ be a directed circuit decomposition of D. Then there is an orientable directed embedding of D with the elements of $\mathcal{C}$ as the profaces and with exactly one or two antifaces, depending on whether $|V(D)| + |A(D)| + |\mathcal{C}|$ is odd or even, respectively.*

This embedding has maximum genus among all orientable directed embeddings of D in which all elements of $\mathcal{C}$ are profaces. This embedding also has maximum genus among all orientable direction-indifferent embeddings of D in which all elements of $\mathcal{C}$ are faces.

A number of related results for both graphs and digraphs follow immediately from this theorem, including Corollary 1.2.

Corollary 1.2. *Let G be an n-vertex eulerian graph where the minimum degree of the underlying simple graph of G is greater than or equal to $(4n+2)/5$. Let T be any euler circuit in G. Then there is a 2-face-colorable orientable embedding of G with T as the unique face of one color and with exactly one or two faces of the other color, depending on whether $|V(G)| + |E(G)|$ is even or odd, respectively.*

This embedding has maximum genus among all 2-face-colorable orientable embeddings of G. When $|V(G)|+|E(G)|$ is even this is an orientable bi-eulerian embedding of G with T as a specified face, and the embedding has maximum genus among all orientable embeddings of G.

When G additionally has $|V(G)| + |E(G)|$ even, Corollary 1.2 says there is a maximum orientable genus embedding of G that is bi-eulerian. However, in general not every maximum genus embedding of such a G is bi-eulerian.

We do not know if the bound of $(4n + 2)/5$ in Theorem 1.1 is tight. The original motivation for this problem came from an applied problem in DNA self-assembly posed by Jonoska, Seeman and Wu [8], which required a special closed walk in a graph for a DNA strand to follow. A formalization of the walk requirements in [3] led to *edge-outer embeddings* of a graph, that is, orientable embeddings of a graph in which there is a special face whose boundary uses every edge at least once. While [8] proves the existence of such circuits, and [3] gives short algorithmic proof of existence, and proves the hardness of finding an optimal (shortest outer facial walk) solution, there is no control over the number or sizes of the remaining faces in the embedding.

The startlingly simple (to state!) and intriguing questions in the first paragraph emerged from this application. Although determining the size of optimal edge-outer faces is hard in general, for eulerian graphs any optimal edge-outer face is necessarily bounded by an euler circuit. Thus, we seek to control the remaining faces in an optimal edge-outer embedding of an eulerian graph by specifying them in advance with a circuit decomposition. Of particular interest are bi-eulerian orientable embeddings, particularly when one of the circuits is specified in advance.

While our original motivation was DNA self-assembly, these and some closely related questions have also received considerable attention in various other special settings.

In [2], Edmonds proved that every eulerian graph has an bi-eulerian embedding in some surface, but noted that his main theorem was not sufficient to determine the orientability of the embedding. Restricting to the orientable case makes the existence problem quite challenging. From a different perspective, Andersen, Bouchet, and Jackson in [1] focus specifically on compatible euler circuits (A-trails) in 4-regular graphs and digraphs in low genus surfaces.

Furthermore, a series of papers, [4, 5, 6, 7] authored by Griggs and Širáň and sometimes also Erskine, Grannell, McCourt, or Psomas, discusses upper embeddings relative to a triangular decomposition of a graph or digraph, and more specifically completing such a decomposition to an embedding by adding an euler circuit. They are interested in triangular decompositions which arise from structures in design theory such as Steiner triple systems, configurations, and latin squares. One representative theorem can be stated as follows:

Theorem 1.3 (Griggs, McCourt, and Širáň [6, Theorem 1.1])**.** *Let $\mathcal{C}$ be an oriented Steiner triple system, i.e., a decomposition of a regular tournament D into directed triangles. Then there is an orientable directed embedding of D with the elements of $\mathcal{C}$ as the profaces and with exactly one directed euler circuit antiface.*

The results in these papers are generally specific to triangular decompositions of very special classes of (di)graphs, such as complete or complete tripartite graphs. Our results for arbitrary dense graphs encompass their results that involve complete graphs or tournaments, such as Theorem 1.3.

2 Manipulating facial circuits and proof outline

We first note that we can always form *some* orientable embedding of a digraph with a given circuit decomposition as the profaces. The new faces are necessarily antifaces. We then develop a number of structural results for manipulating the facial circuits in embedded graphs. The following lemmas for example allow us to combine antifaces without altering the given profaces.

A central tool is the following version for directed embeddings of a well-known operation on embeddings of undirected graphs. It allows us to combine three antifaces incident with the same vertex into a single face.

Lemma 2.1. *Let Φ be an orientable directed embedding of an eulerian digraph D, and $v \in V(D)$. If A, B, and C are distinct antifaces that each contain v, then there is an orientable directed embedding Φ' of D that has the same profaces and antifaces as Φ except that A, B and C are merged into a single antiface.*

Interlaced vertices on a circuit, that is, a pair of vertices u and v that appear as $\ldots u \ldots v \ldots u \ldots v$ in the circuit, play a pivotal role here. When we have appropriately interlaced vertices, we can merge faces, as in Lemma 2.2.

Lemma 2.2. *Let Φ be an orientable directed embedding of an eulerian digraph D, and $x, y \in V(D)$. Suppose that there are three distinct antifaces A, B, C where x and y are interlaced on A, x occurs on B, and y occurs on C. Then there is an orientable directed embedding Φ' of D that has the same profaces and antifaces as Φ except that A, B and C are replaced by a single antiface.*

The density of the given (di)graph plays an important role in assuring the necessary interlacements. Given a digraph D, let k be such that for each $v \in V(D)$ there are at most k vertices different from v that are not adjacent to v in D. A *locally irreducible embedding* has at most two antifaces meeting at any vertex (which can be guaranteed by repeated use of Lemma 2.1), and a vertex has type AB if it lies on both of the faces A and B. We write $|AB|$ for the number of vertices of type AB. The following is a representative lemma assuring interlacement.

Lemma 2.3. *Suppose that we have a locally irreducible embedding with distinct antifaces A and B. Suppose that $|AB| \geq \min(3k+4, 4k+3)$ and there exist vertices of type AP and vertices of type BQ with $P, Q \notin \{A, B\}$ (possibly $P = Q$). Then there is a vertex of type AB that is either interlaced on A with a vertex of type AP or interlaced on B with a vertex of type BQ.*

To prove Theorem 1.1, we use results such as those given above to successively reduce the number of antifaces in the orientable directed embedding of the digraph without changing the given profaces.

The proof begins by applying results such as Lemma 2.2 so that no more than two antifaces are incident with any one vertex. We can then create a *touch graph*, K. The vertices of K are the antifaces. Each vertex v of D gives an edge of K: if v is incident with only one antiface the edge is a loop at that vertex of K, and if v is incident with exactly two antifaces the edge joins the corresponding vertices of K.

We use the structure of the touch graph K, in particular the existence and location of its loops, to organize the necessary case work for the proof of Theorem 1.1. In each situation density information allows us to apply results such as Lemma 2.3 to reduce the number of antifaces. By repeated reductions we obtain an orientable embedding with one or two antifaces and the prescribed set of profaces. If there is only one antiface it must be bounded by a directed euler circuit.

Although our focus here has been on dense graphs, there are certainly also sparse graphs that have bi-eulerian orientable embeddings, as for example in Figure 1. Theorem 2.4 below has an constructive proof with an algorithm that, given an eulerian digraph with a directed euler circuit, produces a second eulerian circuit for the desired bi-eulerian embedding. On the other hand, we have examples of 4-edge-connected, 4-regular graphs with no bi-eulerian embeddings.

Theorem 2.4. *Let D be an eulerian digraph with all vertices of degree congruent to* 2 mod 4*, and let T be any directed euler circuit of D. Then D has a bi-eulerian orientable embedding with one of the faces bounded by T.*

References

[1] Lars Døvling Andersen, André Bouchet, and Bill Jackson. Orthogonal A-trails of 4-regular graphs embedded in surfaces of low genus. *J. Combin. Theory Ser. B*, 66(2):232–246, 1996.

[2] Jack Edmonds. On the surface duality of linear graphs. *J. Res. Nat. Bur. Standards Sect. B*, 69B:121–123, 1965.

[3] M. N. Ellingham and Joanna A. Ellis-Monaghan. Edge-outer graph embedding and the complexity of the DNA reporter strand problem. *Theoret. Comput. Sci.*, 785:117–127, 2019.

[4] Grahame Erskine, Terry Griggs, and Jozef Širáň. On the upper embedding of symmetric configurations with block size 3. *Discrete Math.*, 343(4):111774, 6, 2020.

[5] Mike J. Grannell, Terry S. Griggs, and Jozef Širáň. Maximum genus embeddings of Steiner triple systems. *European J. Combin.*, 26(3-4):401–416, 2005.

[6] Terry S. Griggs, Thomas A. McCourt, and Jozef Širáň. On the upper embedding of Steiner triple systems and Latin squares. *Ars Math. Contemp.*, 18(1):127–135, 2020.

[7] Terry S. Griggs, Constantinos Psomas, and Jozef Širáň. Maximum genus embeddings of Latin squares. *Glasg. Math. J.*, 60(2):495–504, 2018.

[8] Nataša Jonoska, Nadrian C. Seeman, and Gang Wu. On existence of reporter strands in DNA-based graph structures. *Theoret. Comput. Sci.*, 410(15):1448–1460, 2009.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Tangled Paths: A Random Graph Model from Mallows Permutations

(Extended abstract)

Jessica Enright* Kitty Meeks† William Pettersson‡ John Sylvester§

Abstract

We introduce the random graph $\mathcal{P}(n, q)$ which results from taking the union of two paths of length $n \geqslant 1$, where the vertices of one of the paths have been relabelled according to a Mallows permutation with real parameter $0 < q(n) \leqslant 1$. This random graph model, the tangled path, goes through an evolution: if q is close to 0 the graph bears resemblance to a path, and as q tends to 1 it becomes an expander. In an effort to understand the evolution of $\mathcal{P}(n, q)$ we determine the treewidth and cutwidth of $\mathcal{P}(n, q)$ up to log factors for all q. We also show that the property of having a separator of size one has a sharp threshold. In addition, we prove bounds on the diameter, and vertex isoperimetric number for specific values of q.

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-057

Introduction

Given two graphs G, H on a common vertex set $[n] = \{1, \dots, n\}$, and a permutation σ on $[n]$, it is natural to consider the following graph

$$\mathsf{layer}(G, \sigma(H)) = ([n],\ E(G) \cup \{\sigma(x)\sigma(y) : xy \in E(H)\}),$$

*School of Computing Science, University of Glasgow, UK. E-mail: jessica.enright@glasgow.ac.uk.
†School of Computing Science, University of Glasgow, UK. E-mail: kitty.meeks@glasgow.ac.uk.
‡School of Computing Science, University of Glasgow, UK. E-mail: william.pettersson@glasgow.ac.uk.
§Dept. of Computing Science, University of Liverpool, UK. E-mail: john.sylvester@liverpool.ac.uk.
Supported by EPSRC project EP/T004878/1: Multilayer Algorithmics to Leverage Graph Structure.

n	v_n	r_n
1	1	1
2	2	12
3	1	312
4	3	3142
5	2	35142
6	5	351462

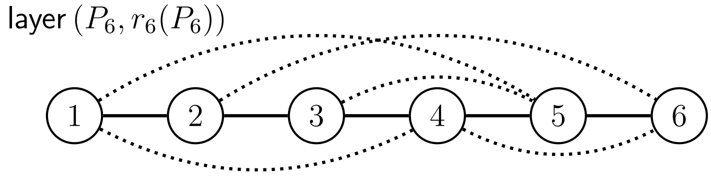


Figure 1: The table on the left gives the sequences of permutations (r_n) generated by the sequence (v_n) for $i = 1, \dots, 6$. On the right we have a tangled path generated by r_6, where the edges of $r_6(P_6)$ are dotted.

which is the union of two graphs where the second graph has been relabelled by a permutation σ. This general and powerful construction has featured previously in the literature in the contexts of constructions and decompositions of graphs [1, 4, 22]. Let P_n be the path on $[n]$ and S_n be the set of all permutations on $[n]$. Consider the following scenario: one must choose a permutation $\sigma \in S_n$ with the goal of making $\mathsf{layer}(P_n, \sigma(P_n))$ as different from a path as possible. There are several parameters one may use to measure the difference between a connected graph G and a path; for example one may look at the diameter $\mathsf{diam}(G)$ or the vertex isoperimetric number $\mathsf{i}(G)$, as the path is extremal for these parameters. The treewidth $\mathsf{tw}(G)$ which, broadly speaking, measures how far (globally) the graph is from being a tree [14], is another natural candidate. It is fairly easy to see that given two or more paths one can build a grid-like graph (see [7, Lemma 8] for more details), and such a graph would have treewidth and diameter $\Theta(\sqrt{n})$. If we choose a permutation uniformly at random, then as a consequence of a result of Kim & Wormald [12, Theorem 1], with high probability the resulting graph is a bounded degree expander. Thus, in this case, the graph $\mathsf{layer}(P_n, \sigma(P_n))$ has treewidth $\Theta(n)$ and diameter $\Theta(\log n)$, so by these parameters it is essentially as far from a path as a sparse graph can be.

The example above shows that even restricting the input graphs to paths can produce rich classes of graphs. Having seen what happens for a uniformly random permutation, one may ask about the structure of $\mathsf{layer}(P_n, \sigma(P_n))$ when σ is drawn from a distribution on S_n that is not uniform. One of the most well known non-uniform distributions on S_n is the Mallows distribution, introduced by Mallows [17] in the late 1950's in the context of statistical ranking theory. Recently it has been the subject of renewed interest for other settings [6, 3, 11], and as an interesting and natural model to study in its own right [2, 10, 20]. The distribution has a parameter q which, roughly speaking, controls the amount of disorder in the permutation.

Mallows Permutations. For real $q > 0$ and integer $n \geqslant 1$, the (n, q)-*Mallows measure* $\mu_{n,q}$ on S_n is given by

$$\mu_{n,q}(\sigma) = \frac{q^{\mathrm{Inv}(\sigma)}}{Z_{n,q}} \qquad \text{for any } \sigma \in S_n,$$

where $\mathrm{Inv}(\sigma) = |\{(i,j) : i < j \text{ and } \sigma(i) > \sigma(j)\}|$ is the number of inversions in the permutation σ and $Z_{n,q}$ is given explicitly by the following formula [2, Equation (2)]:

$$Z_{n,q} = \prod_{i=1}^{n} \left(1 + q + \cdots + q^{i-1}\right) = \prod_{i=1}^{n} \frac{1-q^i}{1-q}.$$

When $q \to 0$, the distribution $\mu_{n,q}$ converges weakly to the degenerate distribution on the identity permutation. We extend $\mu_{n,q}$ to $q = 0$ by setting $\mu_{n,0}$ to be the probability measure assigning 1 to the identity permutation. On the other hand if $q = 1$ then $\mu_{n,1}$ is the uniform measure on S_n. One can see that $\sigma \sim \mu_{n,q}$ has distribution $\mu_{n,1/q}$ when reversed.

A key feature of Mallows permutations is that they can be constructed by a simple procedure from a sequence of independent random variables. For any $q > 0$, the *Mallows Process* gives a sequence (r_n) such that $r_n \sim \mu_{n,1/q}$, for each $n \geqslant 1$. Furthermore each r_n is constructed from r_{n-1} by inserting n at a position in r_{n-1} sampled via a simple distribution that is independent from $r_1, \ldots, r_{n-1}$. Several desirable properties of Mallows permutations can be deduced from this construction, see [10, Sec. 2] for more details.

The Tangled Paths Model. We study the random graph distribution induced by $\mathsf{layer}(P_n, \sigma(P_n))$, where $\sigma \sim \mu_{n,q}$ and $0 \leqslant q := q(n) \leqslant 1$. From now on we call this graph distribution the *tangled path* model and denote it by $\mathcal{P}(n,q)$. Thus a random graph $\mathcal{P}(n,q)$ has vertex set $[n]$ and (random) edge set $E(P_n) \cup \{\sigma(i)\sigma(i+1) : i \in [n-1]\}$, where $\sigma \sim \mu_{n,q}$. We restrict to $q \in [0,1]$ as reversing the permutation does not affect our construction (up to a relabelling). We also identify any multi-edges created as one edge, however this detail is not important for any of our results. This paper will focus on $\mathcal{P}(n,q)$; as we have seen already combining paths can give rise to interesting and varied graphs, and the Mallows permutation gives our model a parameter q which, roughly speaking, increases the 'tangled-ness' of the graph.

We see, from above, that $\mathcal{P}(n,0)$ is a path and $\mathcal{P}(n,1)$ is an expander with high probability by [12, Theorem 1]. Our ultimate aim is to understand the structure of $\mathcal{P}(n,q)$ for intermediate values of q, and this paper takes the first steps in this direction. Informally, one aspect of this is knowing when $\mathcal{P}(n,q)$ stops looking 'path-like'; we show that if $q < 1$ is fixed the diameter is linear (Theorem 3), and there is a sharp threshold for having a single vertex cut at $q_c = 1 - \frac{\pi^2}{6 \log n}$ (Theorem 2). For $q \to 1$ sufficiently fast, it makes more sense to measure the complexity of the internal structure of $\mathcal{P}(n,q)$ by how much it differs from a tree. In this direction we show that, up to logarithmic factors, the treewidth [14] of $\mathcal{P}(n,q)$ grows at rate $(1-q)^{-1}$ (Theorem 4) until the graph becomes an expander at around $q = 1 - \frac{1}{n \log n}$ (Theorem 1), indicating that, in the sense of treewidth, the complexity of the structure grows smoothly with q. This behaviour contrasts with the binomial/Erdős-Rényi random graph [8] where the treewidth increases rapidly from being bounded by a constant, to $\Theta(n)$ as the average degree rises from below one to above one [15].

Aside from this model being natural, motivation for this line of study comes from practical algorithmic applications. Many real-world systems – including social, biological

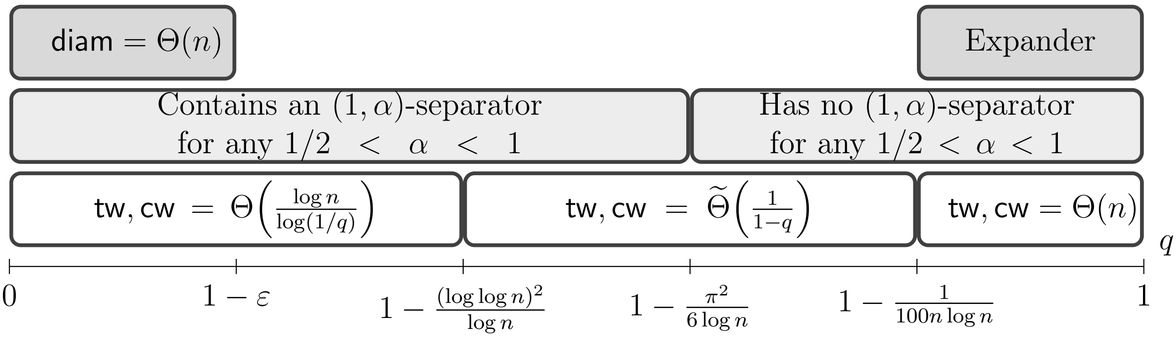


Figure 2: The diagram above gives a pictorial representation of our results. All results above hold with high probability, and we say that $f(n) = \widetilde{\Theta}(g(n))$ if $f(n) = \mathcal{O}(g(n)\log g(n))$ and $f(n) = \Omega(g(n)/\log g(n))$.

and transport networks – involve qualitatively different types of edges, where each type of edge generates a "layer" with specific structural properties [13, 19]. For example, when modelling the spread of disease in livestock, one layer of interest arises from physical adjacency of farms, and so is determined entirely by geography. A second epidemiologically-relevant layer could describe the pairs of farms which share equipment: this is no longer fully determined by geography, but will nevertheless be influenced by the location of farms, as those that are geographically close are more likely to cooperate in this way. It is known that algorithmically useful structure in individual layers of a graph is typically lost when the layers are combined adversarially [7]. The present work can be seen as an attempt to understand the structure of graphs generated from two simple layers which are both influenced to some extent by a shared underlying "geography". In this setting the treewidth tw is a natural parameter as many NP-hard problems become tractable when parametrised by tw [5, Ch. 7].

Our Results

In what follows, the integer $n \geqslant 1$ denotes the number of vertices in the graph (or elements in a permutation) and $q := q(n)$, the parameter of the Mallows permutation (or related tangled path), is a real valued function of n taking values in $[0, 1]$. We say a sequence of events $\mathcal{E}_n$ occurs *with high probability* (w.h.p.) if $\mathbb{P}(\mathcal{E}_n) \to 1$ as $n \to \infty$. Throughout log is base e. See Figure 2 for a summary of our results.

A graph G is a *vertex-expander* if there exists a fixed real $c > 0$ such that any set $S \subseteq V$ with $|S| \leqslant \lceil n/2 \rceil$ is adjacent to at least $c|S|$ vertices in $V\backslash S$. As mentioned above, when $q = 1$ the permutation is uniform, and so the fact that $\mathcal{P}(n, 1)$ is an expander follows from [12, Theorem 1]. Our first result shows that for q sufficiently close to 1, this still holds.

Theorem 1. *If $q \geqslant 1 - \frac{1}{100n\log n}$, then w.h.p. $\mathcal{P}(n,q)$ is a bounded degree vertex-expander.*

For an integer $s \geqslant 1$ and real $1/2 \leqslant \alpha < 1$ we say G has an (s,α)-*separator* if there is a vertex subset S with $|S| \leqslant s$ such that $G \backslash S$ can be partitioned into two disjoint sets of at most $\alpha|V|$ vertices with no crossing edges. Balanced separators (e.g. $\alpha = 2/3$) are useful for designing divide and conquer algorithms, in particular for problems on planar graphs [16]. Balanced separators have intimate connections to notions of sparsity for graphs [18].

Observe that, for any fixed $1/2 < \alpha < 1$, if G is a vertex expander then there exists a $c > 0$ such that G has no (cn, α)-separator. At the other extreme, the path has an $(1,\alpha)$-*separator*. We show that for $\mathcal{P}(n,q)$ this 'path-like' property disappears around $q_c = 1 - \frac{\pi^2}{6 \log n}$.

Theorem 2. *For any fixed real* $1/2 < \alpha < 1$ *we have*

$$\lim_{n \to \infty} \mathbb{P}(\,\mathcal{P}(n,q) \text{ has a } (1,\alpha)\text{-separator}\,) = \begin{cases} 0 & \text{if } \frac{\pi^2}{6(1-q)} - \log n + \frac{\log\log n}{2} \to \infty \\ 1 & \text{if } \frac{\pi^2}{6(1-q)} - \log n + \frac{5\log\log n}{2} \to -\infty \end{cases}.$$

We say that q_0 is *sharp threshold* for a graph property $\mathfrak{P}$ if for any $\varepsilon > 0$ w.h.p. $\mathcal{P}(n,p) \notin \mathfrak{P}$ for any $p \leqslant q_0(1-\varepsilon)$, and $\mathcal{P}(n,r) \in \mathfrak{P}$ for any $r \geqslant q_0(1+\varepsilon)$, see [9]. Theorem 2 is quite precise as it determines the second order term in the threshold up to a constant, showing that the property of having a cut vertex has a sharp threshold of width $\mathcal{O}\left(\frac{\log\log n}{(\log n)^2}\right)$. Theorem 2 is established by finding first and second moment thresholds for the property. Positive correlation between cuts suggests this result cannot be significantly improved using standard methods alone.

The diameter $\mathsf{diam}(G)$ of a graph G is the greatest distance between any pair of vertices. Theorem 1 implies that $\mathsf{diam}(\mathcal{P}(n,q)) = \mathcal{O}(\log n)$ when q is close to 1. On the other hand $\mathsf{diam}(\mathcal{P}(n,0)) = n-1$ as it is a path; by applying bounds on the number of cut vertices, we show this holds (up to a constant) for any fixed $q < 1$.

Theorem 3. *Let* $0 \leqslant q < 1$ *be any function of* n *bounded away from* 1. *Then there exists a constant* $c > 0$ *such that w.h.p.,* $\mathsf{diam}(\mathcal{P}(n,q)) \geqslant cn$.

The treewidth $\mathsf{tw}(G)$ of a graph G is the minimum size (minus one) of the largest vertex subset (i.e. bag) in a tree decomposition of G, minimised over all such decompositions [14]. The cutwidth $\mathsf{cw}(G)$ is the greatest number of edges crossing any real point under an injective function $\varphi : V \to \mathbb{Z}$, minimised over all φ. It is known that for any graph G we have $\mathsf{tw}(G) \leqslant \mathsf{cw}(G)$, however there may a multiplicative discrepancy of order up to n. We show there is at most only a constant factor discrepancy for $\mathcal{P}(n,q)$ for certain ranges of q and give bounds for all q which are tight up to log factors.

Theorem 4. *If there exists a real constant* $\kappa > 0$ *such that* $0 \leqslant q \leqslant 1 - \kappa \cdot \frac{(\log\log n)^2}{\log n}$, *then there exist constants* $0 < c_1, C_2 < \infty$ *such that w.h.p.*

$$c_1 \cdot \left(\sqrt{\frac{\log n}{\log(1/q)}} + 1\right) \leqslant \mathsf{tw}(\mathcal{P}(n,q)) \leqslant C_2 \cdot \left(\sqrt{\frac{\log n}{\log(1/q)}} + 1\right). \tag{1}$$

Furthermore, there exists some $c_3 > 0$ such that if $1 - \frac{(\log\log n)^2}{\log n} \leqslant q \leqslant 1 - \frac{1}{100n\log n}$, then w.h.p.

$$\frac{c_3}{1-q} \cdot \log\left(\frac{1}{1-q}\right)^{-1} \leqslant \mathsf{tw}(\mathcal{P}(n,q)) \leqslant \min\left\{\frac{5}{1-q}\cdot\log\left(\frac{1}{1-q}\right),\ 2n\right\}. \tag{2}$$

In addition, if $q \geqslant 1 - \frac{1}{100n\log n}$ then w.h.p.

$$\frac{n}{50} \leqslant \mathsf{tw}(\mathcal{P}(n,q)) \leqslant 2n. \tag{3}$$

The same upper and lower bounds in (1)*,* (2)*, and* (3) *also hold for the cutwidth* $\mathsf{cw}(\mathcal{P}(n,q))$*.*

Observe that if $q \to 1$ then $\log(1/q) \approx 1-q$ and so when $q = 1 - \Theta\left((\log\log n)^2/\log n\right)$ we have $\sqrt{\log(n)/\log(1/q)} \approx -\log(1-q)/(1-q)$. Therefore, the first two upper bounds for cutwidth are equal up to constants for this value of q. Thus, for this q, the upper bound for the cutwidth given in (2) is tight and the lower bound for treewidth is off by a multiplicative factor of order $(\log\log n)^2$.

Proof Sketch of Theorem 4. The rough strategy for the lower bounds in (1) and (2) is as follows:

(i) relate containing a k-vertex expander as a minor in $\mathcal{P}(n,q)$, and thus $\Omega(k)$-treewidth, to a property of the underlying Mallows permutation or the random sequence generating it,

(ii) show this property holds, for a suitable k, with high probability by utilising the (asymptotic) independence of elements in a Mallows permutation or the sequence generating it.

However, the properties sought and method for controlling the probabilities in (1) and (2) differ slightly.

For Step (i), of the lower bound in (1), we show that if r_k and r_n are generated by sequences $x = x_1,\dots,x_k$ and $y = y_1,\dots,y_n$ respectively via the Mallows process, and x is contained in y as a consecutive sub-sequence, then $P_n \cup r_n(P_n)$ contains $P_k \cup r_k(P_k)$ as a minor. To prove (2) we instead show that if a permutation $\pi \in S_n$ contains a permutation $\sigma \in S_k$ as a consecutive pattern then $P_n \cup \pi(P_n)$ contains $P_k \cup \sigma(P_k)$ as a minor. In particular, both relations hold in the case where $P_k \cup \sigma(P_k)$ and $P_k \cup r_k(P_k)$ are expanders.

For Step (ii), the lower bound in (1) is shown using the second moment method to a given consecutive sub-sequence of k inputs occur w.h.p., whereas for (2) we use independence of permutations induced by disjoint intervals of elements in a Mallows permutation to show a given consecutive pattern occurs.

We now give a proof sketch for the upper bounds on $\mathsf{cw}(\mathcal{P}(n,q))$ in (1) and (2). To begin, we fix the ordering $\varphi : [n] \to [n]$ in the definition of cutwidth to be the identity map.

That is, we order the vertices of $\mathcal{P}(n,q)$ along the line with respect to the ordering of the vertices given by the un-permuted path P_n. We then bound the number of edges crossing any vertex i by showing that not too many elements with values $j > i$ are inserted next to elements with values less than i by the Mallows process. To do this we show that, for $b = \Theta(\frac{\log n}{1-q})$ and some suitable $L \geqslant \ell$, within any consecutive sequence of L steps of the Mallows process (with generating inputs v_i) the following events hold with high probability:

(i) no insert position v_i has value greater than b,

(ii) after L steps the leftmost b places will each contain an element added added at most L steps ago,

(iii) there are at most ℓ values of v_i greater than ℓ.

The events (i) and (iii) ensure that not too many long edges are created from new entries being added far away from the left-hand end of the process. The event (ii) is a little bit more subtle but key to the success of our approach as it ensures that the left-hand end of the permutation grown by the Mallows process cannot retain entries that were inserted long ago, again preventing long edges caused by new elements lying next to old ones. If these three events hold then we can show that the number of edges crossing any vertex under the identity map is $\mathcal{O}(\ell)$. Optimising the choice of L and ℓ then gives the upper bounds in (1) and (2).

Open Problems

One could study the effect of q in $\mathcal{P}(n,q)$ on almost any graph property of interest for sparse graphs. One fundamental problem is to determine the number of edges in $\mathcal{P}(n,q)$ (recall that we disregard multi-edges). This deceptively non-trivial problem is related to clustering of consecutive numbers in Mallows permutations [21]. It would also be interesting to close the gap for treewidth by obtaining tight bounds for all q.

Theorem 2 proves that $q = 1 - \pi^2/(6\log n)$ is a sharp threshold for containing a single vertex whose removal separates the graph into two macroscopic components. A key open problem is to determine if there is a notion of monotone property in the setting of tangled paths which guarantees the existence of a threshold (or even a sharp threshold). One candidate feature (for a property to be monotone with respect to) is the number of inversions in the permutation generating $\mathcal{P}(n,q)$. However, one issue with parameterizing by the number of inversions is the fact that the tangled paths generated by $\sigma = (\sigma_1,\ldots,\sigma_n)$ and its reverse $\sigma^R = (\sigma_n,\ldots,\sigma_1)$ are isomorphic, but the number of inversions may differ greatly as $\mathrm{Inv}(\sigma^R) = \binom{n}{2} - \mathrm{Inv}(\sigma)$.

References

[1] Bogdan Alecu, Vadim Lozin, Daniel A. Quiroz, Roman Rabinovich, Igor Razgon, and Viktor Zamaraev. The treewidth and pathwidth of graph unions, 2022.

[2] Nayantara Bhatnagar and Ron Peled. Lengths of monotone subsequences in a Mallows permutation. *Probab. Theory Related Fields*, 161(3-4):719–780, 2015.

[3] Mark Braverman and Elchanan Mossel. Sorting from noisy information, 2009.

[4] Yijia Chen and Jörg Flum. On the ordered conjecture. In *Proceedings of the 27th Annual IEEE Symposium on Logic in Computer Science, LICS 2012*, pages 225–234. IEEE Computer Society, 2012.

[5] Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer, 2015.

[6] Persi Diaconis and Arun Ram. Analysis of systematic scan Metropolis algorithms using Iwahori-Hecke algebra techniques. *Michigan Math. J.*, 48:157–190, 2000.

[7] Jessica Enright and Kitty Meeks. Two dichotomies for model-checking in multi-layer structures. *CoRR*, abs/1710.08758, 2017.

[8] P. Erdős and A. Rényi. On random graphs. I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[9] Ehud Friedgut and Gil Kalai. Every monotone graph property has a sharp threshold. *Proc. Amer. Math. Soc.*, 124(10):2993–3002, 1996.

[10] Alexey Gladkich and Ron Peled. On the cycle structure of Mallows permutations. *Ann. Probab.*, 46(2):1114–1169, 2018.

[11] Alexander E. Holroyd, Tom Hutchcroft, and Avi Levy. Mallows permutations and finite dependence. *Ann. Probab.*, 48(1):343–379, 2020.

[12] Jeong Han Kim and Nicholas C. Wormald. Random matchings which induce Hamilton cycles and hamiltonian decompositions of random regular graphs. *J. Comb. Theory, Ser. B*, 81(1):20–44, 2001.

[13] Mikko Kivelä, Alex Arenas, Marc Barthelemy, James P. Gleeson, Yamir Moreno, and Mason A. Porter. Multilayer networks. *J. Complex Networks*, 2(3):203–271, 2014.

[14] Ton Kloks. *Treewidth, Computations and Approximations*, volume 842 of *Lecture Notes in Computer Science*. Springer, 1994.

[15] Choongbum Lee, Joonkyung Lee, and Sang-il Oum. Rank-width of random graphs. *J. Graph Theory*, 70(3):339–347, 2012.

[16] Richard J. Lipton and Robert Endre Tarjan. A separator theorem for planar graphs. *SIAM Journal on Applied Mathematics*, 36(2):177–189, 1979.

[17] Colin L. Mallows. Non-null ranking models. I. *Biometrika*, 44:114–130, 1957.

[18] Jaroslav Nesetril and Patrice Ossona de Mendez. *Sparsity - Graphs, Structures, and Algorithms*, volume 28 of *Algorithms and combinatorics*. Springer, 2012.

[19] Shai Pilosof, Mason A Porter, Mercedes Pascual, and Sonia Kéfi. The multilayer nature of ecological networks. *Nature Ecology & Evolution*, 1(4):1–9, 2017.

[20] Ross G. Pinsky. Permutations avoiding a pattern of length three under mallows distributions. *Random Struct. Algorithms*, 58(4):676–690, 2021.

[21] Ross G. Pinsky. Clustering of consecutive numbers in permutations under Mallows distributions and super-clustering under general p-shifted distributions. *Electronic Journal of Probability*, 27:1 – 20, 2022.

[22] Daniel A Quiroz. Chromatic and structural properties of sparse graph classes. *PhD thesis, The London School of Economics and Political Science (LSE)*, 2017.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Cop number of random k-uniform hypergraphs

(Extended abstract)

Joshua Erde* Mihyun Kang† Florian Lehner‡ Bojan Mohar§
Dominik Schmid¶

Abstract

The game of *Cops and Robber* is usually played on a graph, in which a group of cops attempt to catch a robber moving along the edges of the graph. The *cop number* of a graph is the minimum number of cops required to win the game. An important conjecture in this area, due to Meyniel, states that the cop number of an n-vertex connected graph is $O(\sqrt{n})$. In 2016, Prałat and Wormald [Meyniel's conjecture holds for random graphs, Random Structures Algorithms. 48 (2016), no. 2, 396–421. MR3449604] showed that this conjecture holds with high probability for random graphs above the connectedness threshold. Moreoever, Łuczak and Prałat [Chasing robbers on random graphs: Zigzag theorem, Random Structures Algorithms. 37 (2010), no. 4, 516–524. MR2760362] showed that on a log-scale the cop number demonstrates a surprising *zigzag* behaviour in dense regimes of the binomial random

*Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria. E-mail: erde@math.tugraz.at. Supported in part by the Austrian Science Fund (FWF): P 36131.

†Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria. E-mail: kang@math.tugraz.at. Supported in part by the Austrian Science Fund (FWF): W 1230.

‡Department of Mathematics, University of Auckland, 38 Princes Street, 1010, Auckland, New Zealand. E-mail: florian.lehner@auckland.ac.nz.

§Department of Mathematics, Simon Fraser University, 8888 University Drive, Burnaby, BC, Canada. E-mail: mohar@sfu.ca. Supported in part by the NSERC Discovery Grant R611450 (Canada). This work is the result of author's visit of TU Graz under Oberwolfach's Simons Visiting Professors program in December 2021.

¶Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria. E-mail: schmid@math.tugraz.at. Supported in part by the Austrian Science Fund (FWF): W 1230.

graph $G(n,p)$. In this paper, we consider the game of Cops and Robber on a hypergraph, where the players move along hyperedges instead of edges. We show that with high probability the cop number of the k-uniform binomial random hypergraph $G^k(n,p)$ is $O\left(\sqrt{\frac{n}{k}}\log n\right)$ for a broad range of parameters p and k. As opposed to the case of $G(n,p)$, on a log-scale our upper bound on the cop number arises as the minimum of *two* complementary zigzag curves. Furthermore, we conjecture that the cop number of a connected k-uniform hypergraph on n vertices is $O\left(\sqrt{\frac{n}{k}}\right)$.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-058`

1 Introduction and results

1.1 Motivation

The game of *Cops and Robber* was introduced by Quilliot [18] in 1978 and independently by Nowakowski and Winkler [16] in 1983. It is a two-player game played on a simple connected graph $G = (V, E)$, with one player controlling a set of m cops and the other player controlling a single robber. For convenience, we will sometimes refer to the cops and the robber as *pieces*. At the start of the game, the first player chooses a starting vertex for each of the cops, then the second player chooses a starting vertex for the robber. Subsequently, the players take alternating turns and in each turn a player can move each of their pieces to an adjacent vertex (i.e., the pieces move along the edges of G). Note that more than one cop can simultaneously occupy a single vertex and that not every piece must be moved in every turn. The position of all the pieces is known to both players throughout the game. The cops win if at some point in the game a cop occupies the same vertex as the robber, otherwise the robber wins. As this is a game with full information, for each graph G and each number of initial cops m, one of the two players has a winning strategy. The *cop number* $c(G)$ of a graph G is defined as the minimum number $m \in \mathbb{N}$, such that m cops have a winning strategy on G. The cop number has been extensively studied since the introduction of this game.

Whilst there is a structural characterisation of the graphs with cop number one [16], in general the problem of determining the cop number of a graph is EXPTIME-complete [12], and so research in this area has been focused on bounding the cop number of particular graph classes. For example, a classic result of Aigner and Fromme [2] shows that the cop number of a connected planar graph is at most three. More generally, it is known that the cop number is bounded for any proper minor-closed class of graphs [3], and there has been much research into determining the largest cop number of a graph that can be embedded in a fixed surface [7, 9, 13, 19, 20].

Perhaps the most well-known conjecture in this area is Meyniel's conjecture (communicated by Frankl [10]).

Conjecture 1.1. *Let G be a connected graph on n vertices. Then $c(G) = O(\sqrt{n})$.*

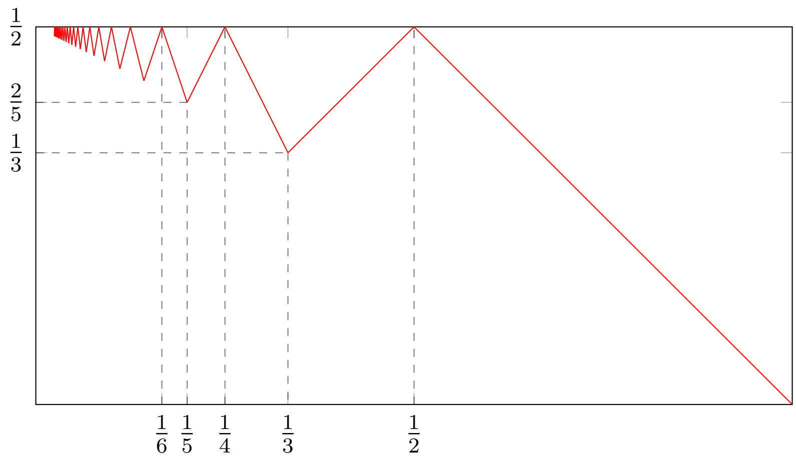


Figure 1: Zigzag shape of the function f

Despite much interest in this conjecture, there has been relatively little improvement to the trivial bound of $O(n)$. Frankl [10] gave the first non-trivial upper bound on the cop number of $O\left(\frac{n \log \log n}{\log n}\right)$, and this bound was improved to $O\left(\frac{n}{\log n}\right)$ by Chiniforooshan [8]. As of today, the best known general upper bound on the cop number is $n2^{-(1+o(1))\sqrt{\log n}}$, given independently by Lu and Peng [14] and by Scott and Sudakov [21]. We note that this bound is still $\Omega\left(n^{1-o(1)}\right)$, and it remains an open question as to whether the cop number can be bounded by $O\left(n^{1-\epsilon}\right)$ for any fixed $\epsilon > 0$ [4].

A natural step towards understanding Conjecture 1.1 is to consider the cop number of the *random* graph $G(n,p)$. For p constant, it was shown by Bonato, Hahn and Wang [6] that with high probability (*whp* for short), meaning with probability tending to one as n tends to infinity, the cop number of $G(n,p)$ is logarithmic in n, and hence Conjecture 1.1 holds for almost all graphs. However, if we let p vary as a function of n, then more interesting behaviour can be seen to develop. Indeed, Łuczak and Prałat [15] showed that whp the function $f\colon (0,1) \to \mathbb{R}$, defined as

$$f(x) = \frac{\log\left(\bar{c}(G\left(n, n^{x-1}\right)\right)}{\log n}, \tag{1.1}$$

where $\bar{c}$ denotes the median of the cop number, has a characteristic zigzag shape (see Figure 1).

Taking the worst case bounds of this function, their result already implies that $c\left(G(n,p)\right) = \tilde{O}\left(\sqrt{n}\right)$ for a broad range of parameters, and that conversely there are choices of p where whp $c(G(n,p)) = \tilde{\Theta}(\sqrt{n})$, where we use $\tilde{\Theta}(\cdot)$ to indicate a bound which holds up to logarithmic factors.

Theorem 1.2 ([15], $\sim$ Theorem 1.1)**.** *Let $\epsilon > 0$, and let $p = \Omega\left(n^{\epsilon-1}\right)$. Then whp*

$$c\left(G(n,p)\right) = \tilde{O}\left(\sqrt{n}\right).$$

In particular, Theorem 1.2 indicates that Conjecture 1.1 holds up to log-factors for this range of p. Bollobás, Kun and Leader [5] gave a similar bound which holds also for sparser regimes of p.

Meyniel's conjecture was finally resolved for all random graphs above the connectedness threshold by Prałat and Wormald [17]. In fact, their result holds for all random graphs with density above $\frac{1}{2}\log n$.

Theorem 1.3 ([17], Theorem 1.2)**.** *Let $\epsilon > 0$, and let $p(n-1) \geq \left(\frac{1}{2}+\epsilon\right)\log n$. Then whp*

$$c\,(G(n,p)) = O(\sqrt{n}).$$

In this paper we consider a generalisation of the Cops and Robber game to hypergraphs, and in particular k-uniform hypergraphs, which are called *k-graphs*. The game is defined analogously to the 2-graph case, with the only difference being that the pieces move along hyperedges instead of edges. For the sake of brevity, when it is clear from the context that we are talking about a hypergraph, we will refer to hyperedges as simply edges. Similarly as for 2-graphs, we define for a hypergraph H

$$c\,(H) \coloneqq \min\left\{m \in \mathbb{N} \colon m \text{ cops have a winning strategy to catch a robber on } H\right\}.$$

This game was first considered by Gottlob, Leone and Scarcello [11] and by Adler [1]. For more recent results on the hypergraph game we refer the reader to [22], where some classic results on the cop number of 2-graphs are generalised to this setting.

Note that by replacing every edge in the hypergraph by a clique, we arrive at an equivalent 2-graph game. Thus, the game of Cops and Robber on hypergraphs is equivalent to the 2 -graph game played on a restricted class of graphs. On the other hand, we can transform a graph G into a $2k$-uniform hypergraph H with $c\,(G) = c\,(H)$ via a simple blow-up construction: We replace each vertex v in G by k vertices $\{v_1, v_2, \ldots v_k\}$ and form a hypergraph $H(G)$ on $\{v_i \colon v \in V(G), i \in [k]\}$ by taking an edge of the form $\{u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_k\}$ for each edge $e = \{u, v\}$ of G (see Figure 2). It is then easy to check that $c\,(G) = c\,(H(G))$, and moreover $|V(H)| = k|V(G)|$.

From these two observations, it is easy to see that the following holds

$$\begin{aligned}\max\left\{c\,(G) : G \text{ a graph }, |V(G)| = \frac{2n}{k}\right\} &\leq \max\left\{c\,(H) : H \text{ a } k\text{-graph }, |V(H)| = n\right\}\\ &\leq \max\left\{c\,(G) : G \text{ a graph }, |V(G)| = n\right\}.\end{aligned}$$

In particular, as there are graphs with $c\,(G) = \Omega\left(\sqrt{n}\right)$, there are also k-graphs with $c\,(H) = \Omega\left(\sqrt{\frac{n}{k}}\right)$. It would seem surprising that such a simple construction, which is essentially graphical in nature, could capture the worst case behaviour for the cop number in hypergraphs of higher uniformity, but we conjecture that this bound is in fact tight.

Conjecture 1.4. *Let H be a connected k-graph on n vertices. Then $c\,(H) = O\left(\sqrt{\frac{n}{k}}\right)$.*

As with Meyniel's Conjecture, a first step towards Conjecture 1.4 would be to consider the behaviour of the cop number of *random* k-graphs.

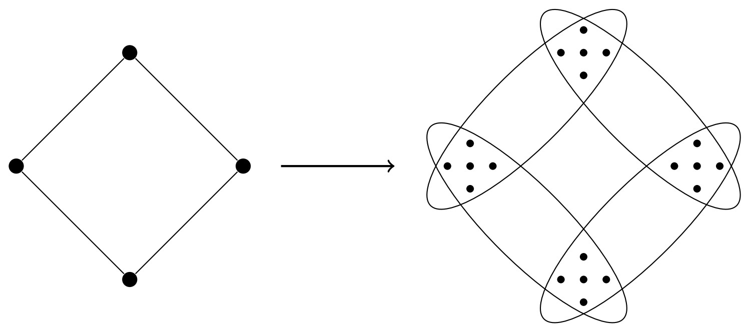


Figure 2: An example of the blow-up construction to generate a $2k$-graph H from a 2-graph that has the same cop number. In this case, $k = 5$, $|V(H)| = 20$ and $c(H) = 2$.

1.2 Main results

The *k-uniform binomial random hypergraph*, which we denote by $G^k(n, p)$, is a random k-graph with vertex set $[n]$ in which each edge, that is, each subset of $[n]$ of size k, appears independently with probability p. Although the main focus of this paper is $G^k(n, p)$, the strategies we develop for the cops will in fact work in a more general class of k-graphs, those satisfying certain expansion properties.

Very roughly, if we denote by $N_V^r(v)$ the vertices that are at most at a fixed distance r from v, then in $G^k(n, p)$ we expect this set to be growing exponentially quickly in r, with its size tightly concentrated around its expectation. Furthermore, for different vertices v and w we do not expect the neighbourhoods $N_V^r(v)$ and $N_V^r(w)$ to have a large intersection, and so, for small subsets $A \subseteq [n]$ we expect the number of vertices at most at a fixed distance r from A to be around $|A|$ times the size of $N_V^r(v)$. Similarly, we expect the set of edges $N_E^r(v)$ at most at a fixed distance $r - 1$ from v to be growing at some uniform exponential rate, and for ranges of p where the random hypergraph is sparse enough, and so few pairs of edges have a large intersection, this rate of growth should be roughly $\frac{1}{k}$ times that of the vertex-neighbourhoods.

Informally, given $\xi > 0$ we will say that a graph is ξ-expanding if the sizes of its vertex and edge-neighbourhoods have this uniform exponential growth, up to some multiplicative error in terms of ξ.

Our first result supports Conjecture 1.4 up to a log-factor for k-graphs that are ξ-expanding for a fixed expansion constant ξ.

Theorem 1.5. *Let $k \geq 2$, let $\xi > 0$ and let G be a ξ-expanding k-graph on n vertices. Then*

$$c(G) \leq 20\xi^{-2}\sqrt{\frac{n}{k}}\log n.$$

Next, we show that whp $G^k(n, p)$ satisfies the desired expansion properties for a broad range of parameters.

Theorem 1.6. *There exists a universal constant ξ such that if $k(n), p(n) > 0$ are such that $k = \omega(\log n)$ and $\frac{n}{k} \geq p\binom{n-1}{k-1} = \omega(\log^3 n)$, then whp $G^k(n,p)$ is ξ-expanding .*

More specifically, after taking a sensible parameterisation, our upper bound on $c(G^k(n,p))$ shows somewhat interesting behaviour, similar to Figure 1. Let us define $\hat{d} = pk\binom{n-1}{k-1}$, which can be thought of as the expected size of the neighbourhood of a vertex in $G^k(n,p)$ and let $\hat{d} = n^\alpha$ and $k = n^\beta$ for some $0 < \beta \leq \alpha \leq 1$. Let us consider the function $f_\beta \colon (\beta, 1) \to \mathbb{R}$ defined as

$$f_\beta(\alpha) = \frac{\log\left(\bar{c}\left(G^k(n,p)\right)\right)}{\log n},$$

with $\bar{c}$ being the upper bound of the cop number as obtained by our strategies (see Section 1.3). Then, f_β again has a characteristic zigzag shape, see Figure 3. In contrast to the case of $G(n,p)$ [15] (see Figure 1), the zigzag shape in the hypergraph case arises as the intersection of two complementary zigzags, coming from two different strategies, and so has twice as many peaks and troughs. We note that under the reasonable assumption that $G^k(n,p)$ is connected, it follows that $d \geq k$, which is the reason as to why the graph in Figure 3 is cut off at $x = \beta$.

We note that, perhaps surprisingly, if we fix k and n and vary d, in certain regimes increasing the average degree, and hence the number of edges, can help the cops, and in other regimes increasing the number of edges can hinder the cops.

Moreover, note that Figure 3 is ignoring log-factors, and in particular, that the worst case bounds, attained when both zigzag lines meet, are of order $O\left(\sqrt{\frac{n}{k}}\log n\right)$. It would be interesting to see, if the log-factor could be removed using similar strategies, and thus show that whp Conjecture 1.4 holds for $G^k(n,p)$.

1.3 Techniques

To give a lower bound for the cop number we need to exhibit a strategy for the cops. As in the work of Łuczak and Prałat [15] we show the existence of a strategy for the cops to *surround* the robber using a probabilistic argument. Whilst in [15] the strategies focused solely on surrounding a small *vertex*-neighbourhood of the robber, we also consider a second type of strategy which aims to surround a small *edge*-neighbourhood, and utilise both these strategies in our result.

Assuming the robber starts on a vertex v, after his first r moves the robber has to be in the r-th vertex-neighbourhood $N_V^r(v)$, and specifically in some edge of the r-th edge-neighbourhood $N_E^r(v)$. The cops aim to occupy each edge in $N_E^r(v)$ before the robber has had time to leave this set. Since the cops move first and a cop can catch the robber in a single move once they occupy the same edge, the cops need to occupy each edge in $N_E^r(v)$ within their first r moves. The strategy of surrounding via vertices, which was used in the 2-graph case by Łuczak and Prałat [15], works quite similarly with the only difference being that the cops surround the r-th vertex-neighbourhood and have $r+1$ moves before the robber can escape. The pay-off in choosing to surround via vertices or edges can be seen as

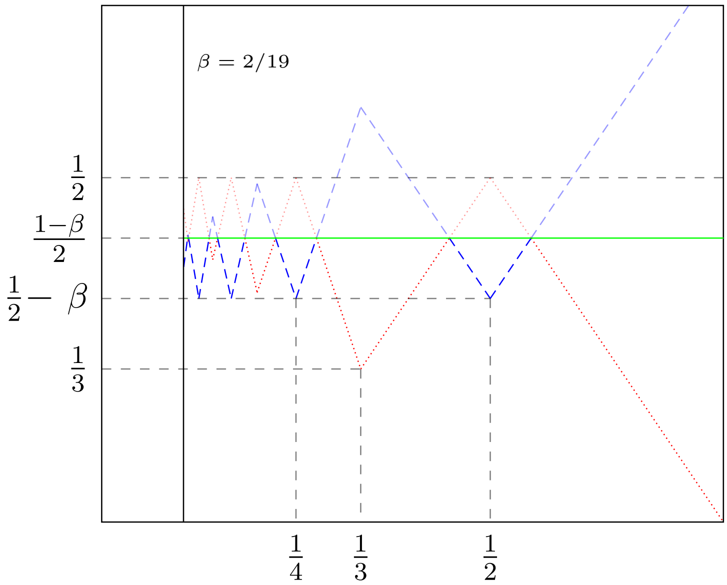


Figure 3: The blue (dashed) line is the upper bound coming from the edge strategy, the red (dotted) line is the upper bound coming from the vertex strategy. As can be seen, the two strategies give rise to two alternating zigzag shapes, that together make up the single zigzag with increased frequency. We note the worst bounds occur at the intersection points of the two lines, which all lie on the green (solid) line at $\frac{1-\beta}{2}$, where β here is equal to $\frac{2}{19}$.

follows – in the former we can use cops at a larger distance, and so in general we will have more cops to work with, whereas in the latter, since each edge contains many vertices, we will not have to occupy as many edges as we would have vertices, and so perhaps we can make do with fewer cops.

For a fixed vertex v and a fixed distance r, the existence of such a strategy can then be reduced to a *matching problem* – for instance in the case of the edge strategy, for each edge e at distance at most r from v we need to assign a unique cop at distance at most r from e, whose strategy will be to occupy e within the first r turns of the game. We aim to show that such an assignment of cops can be found with *positive* probability if we choose a *random* set of cops, assigning a cop to each vertex in the graph independently with some probability $q = q(r)$.

Assuming that our k-graph G is ξ-expanding for some constant $\xi > 0$, we have quite good control over the sizes of $N_V^r(v)$ and $N_E^r(v)$, and also over the number of vertices at a fixed distance from each vertex and edge contained in these sets. Using some standard probabilistic and combinatorial tools, we can show that for an appropriate choice of $q(r)$ with positive probability we can find an appropriate assignment of cops for *each* possible starting vertex v, and bound the number of cops $m(r)$ we use in such a strategy,which in general will depend not only on r, but also on the uniformity k and average degree d of G.

This leads to a family of bounds on the cop number, one for each $r \in \mathbb{N}$, for both the vertex and edge surrounding strategy. For a fixed choice of parameters k and d, we then have to solve an integer optimisation problem to find which choice of r (and of a vertex or edge surrounding strategy) leads to the best bound on the cop number, from which we can derive the bounds leading to Figure 3.

References

[1] I. Adler. Marshals, monotone marshals, and hypertree-width. *J. Graph Theory*, 47(4):275–296, 2004.

[2] M. Aigner and M. Fromme. A game of cops and robbers. *Discrete Applied Mathematics*, 8(1):1–11, 1984.

[3] T. Andreae. On a pursuit game played on graphs for which a minor is excluded. *Journal of Combinatorial Theory, Series B*, 41(1):37–47, 1986.

[4] W. Baird and A. Bonato. Meyniel's conjecture on the cop number: a survey. *J. Comb.*, 3(2):225–238, 2012.

[5] B. Bollobás, G. Kun, and I. Leader. Cops and robbers in a random graph. *Journal of Combinatorial Theory, Series B*, 103(2):226–236, 2013.

[6] A. Bonato, G. Hahn, and C. Wang. The cop density of a graph. *Contrib. Discrete Math.*, 2(2):133–144, 2007.

[7] N. Bowler, J. Erde, F. Lehner, and M. Pitz. Bounding the cop number of a graph by its genus. *SIAM J. Discrete Math.*, 35(4):2459–2489, 2021.

[8] E. Chiniforooshan. A better bound for the cop number of general graphs. *J. Graph Theory*, 58(1):45–48, 2008.

[9] N. E. Clarke, S. Fiorini, G. Joret, and D. O. Theis. A note on the cops and robber game on graphs embedded in non-orientable surfaces. *Graphs Comb.*, 30(1):119–124, 2014.

[10] P. Frankl. Cops and robbers in graphs with large girth and Cayley graphs. *Discrete Applied Mathematics*, 17(3):301–305, 1987.

[11] G. Gottlob, N. Leone, and F. Scarcello. Robbers, marshals, and guards: game theoretic and logical characterizations of hypertree width. *Journal of Computer and System Sciences*, 66(4):775–808, 2003. Special Issue on PODS 2001.

[12] W. B. Kinnersley. Cops and robbers is EXPTIME-complete. *J. Combin. Theory Ser. B*, 111:201–220, 2015.

[13] F. Lehner. On the cop number of toroidal graphs. *J. Combin. Theory Ser. B*, 151:250–262, 2021.

[14] L. Lu and X. Peng. On Meyniel's conjecture of the cop number. *J. Graph Theory*, 71(2):192–205, 2012.

[15] T. Łuczak and P. Prałat. Chasing robbers on random graphs: zigzag theorem. *Random Structures Algorithms*, 37(4):516–524, 2010.

[16] R. Nowakowski and P. Winkler. Vertex-to-vertex pursuit in a graph. *Discrete Mathematics*, 43(2):235–239, 1983.

[17] P. Prałat and N. Wormald. Meyniel's conjecture holds for random graphs. *Random Structures Algorithms*, 48(2):396–421, 2016.

[18] A. Quilliot. *Jeux et pointes fixes sur les graphes*. PhD thesis, Université de Paris VI, 1978.

[19] A. Quilliot. A short note about pursuit games played on a graph with a given genus. *Journal of Combinatorial Theory, Series B*, 38(1):89–92, 1985.

[20] B. S. W. Schroeder. The copnumber of a graph is bounded by $\lfloor \frac{3}{2}$ genus $(G) \rfloor + 3$. In *Categorical perspectives (Kent, OH, 1998)*, Trends Math., pages 243–263. Birkhäuser Boston, Boston, MA, 2001.

[21] A. Scott and B. Sudakov. A bound for the cops and robbers problem. *SIAM J. Discrete Math.*, 25(3):1438–1442, 2011.

[22] P. Siriwong, R. Boonklurb, H. Liu, and S. Singhun. Cops and robber on cartesian products and some classes of hypergraphs. *arXiv preprint arXiv:2107.11741*, 2021.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

The structure of quasi-transitive graphs avoiding a minor with applications to the domino problem

(Extended abstract)

Louis Esperet* Ugo Giocanti† Clément Legrand-Duchesne‡

Abstract

An infinite graph is quasi-transitive if its vertex set has finitely many orbits under the action of its automorphism group. In this paper we obtain a structure theorem for locally finite quasi-transitive graphs avoiding a minor, which is reminiscent of the Robertson-Seymour Graph Minor Structure Theorem. We prove that every locally finite quasi-transitive graph G avoiding a minor has a tree-decomposition whose torsos are finite or planar; moreover the tree-decomposition is canonical, i.e. invariant under the action of the automorphism group of G. As applications of this result, we prove the following.

- Every locally finite quasi-transitive graph attains its Hadwiger number, that is, if such a graph contains arbitrarily large clique minors, then it contains an infinite clique minor. This extends a result of Thomassen (1992) who proved it in the 4-connected case and suggested that this assumption could be omitted.

*Univ. Grenoble Alpes, CNRS, Laboratoire G-SCOP, Grenoble, France. E-mail: louis.esperet@grenoble-inp.fr. Partially supported by the French ANR Projects GrR (ANR-18-CE40-0032), TWIN-WIDTH (ANR-21-CE48-0014-01), and by LabEx PERSYVAL-lab (ANR-11-LABX-0025).

†Univ. Grenoble Alpes, CNRS, Laboratoire G-SCOP, Grenoble, France. E-mail: ugo.giocanti@grenoble-inp.fr. Partially supported by the French ANR Projects GrR (ANR-18-CE40-0032), TWIN-WIDTH (ANR-21-CE48-0014-01), and by LabEx PERSYVAL-lab (ANR-11-LABX-0025).

‡CNRS, LaBRI, Université de Bordeaux, Bordeaux, France. E-mail: clement.legrand@u-bordeaux.fr. Partially supported by the French ANR Projects TWIN-WIDTH (ANR-21-CE48-0014-01), and by LabEx PERSYVAL-lab (ANR-11-LABX-0025).

- Locally finite quasi-transitive graphs avoiding a minor are accessible (in the sense of Thomassen and Woess), which extends known results on planar graphs to any proper minor-closed family.
- Minor-excluded finitely generated groups are accessible (in the group-theoretic sense) and finitely presented, which extends classical results on planar groups.
- The domino problem is decidable in a minor-excluded finitely generated group if and only if the group is virtually free, which proves the minor-excluded case of a conjecture of Ballier and Stein (2018).

.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-059`

1 Introduction

1.1 A structure theorem

A central result in modern graph theory is the Graph Minor Structure Theorem of Robertson and Seymour [16], later extended to infinite graphs by Diestel and Thomas [8]. This theorem states that any graph G avoiding a fixed minor has a *tree-decomposition*, such that each piece of the decomposition, called a *torso*, is close to being embeddable on a surface of bounded genus. A natural question is the following: if the graph G has non trivial symmetries, can we make these symmetries apparent in the tree-decomposition? In other worlds, do graph avoiding a fixed minor have a tree-decomposition as above, but with the additional constraint that the decomposition is *canonical*, i.e., invariant under the automorphism group of G? In this paper we answer this question positively for infinite, locally finite graphs G that are *quasi-transitive*, i.e., the vertex set of G has finitely orbits under the action of the automorphism group of G. This additional restriction has the advantage of making the structure theorem much cleaner: instead of being almost embeddable on a surface of bounded genus, each piece of the tree-decomposition is now simply finite or planar.

Theorem 1.1. *Every locally finite quasi-transitive graph avoiding the countable clique as a minor has a canonical tree-decomposition whose torsos are finite or planar.*

We also give the following more precise version of this result, which might be useful for applications.

Theorem 1.2. *Every locally finite quasi-transitive graph G avoiding the countable clique as a minor has a canonical tree-decomposition with adhesion at most 3 in which each torso is a minor of G, and is planar or has bounded treewidth.*

Interestingly, the proof does not use the original structure theorem of Robertson and Seymour [16] and its extension to infinite graphs by Diestel and Thomas [8]. Instead, we rely mainly on a series of results and tools introduced by Grohe [13] to study decompositions of finite 3-connected graphs into quasi-4-connected components, together with a result of Thomassen [18] on locally finite quasi-4-connected graphs. The main technical contribution of our work consists in extending the results of Grohe to infinite, locally finite graphs and in addition, making sure that the decompositions we obtain are canonical (in a certain weak sense). Our proof crucially relies on a recent result of Carmesin, Hamann, and Miraftab [7], which shows that there exists a canonical tree-decomposition that distinguishes all tangles of a given order (in our case, of order 4).

We now discuss some applications of Theorem 1.1.

1.2 Hadwiger number

As a consequence of Theorem 1.1, we obtain a result on the Hadwiger number of locally finite quasi-transitive graphs. The *Hadwiger number* of a graph G is the supremum of the sizes of all complete minors in G. We say that a graph *attains its Hadwiger number* if the supremum above is attained, that is if it is either finite, or G contains an infinite clique minor. Thomassen [18] proved that every locally finite quasi-transitive 4-connected graph attains its Hadwiger number, and suggested that the 4-connectedness assumption might be unnecessary. We prove that this is indeed the case.

Theorem 1.3. *Every locally finite quasi-transitive graph attains its Hadwiger number.*

1.3 Accessibility in graphs

We now introduce the notion of accessibility in graphs considered by Thomassen and Woess [19]. To distinguish it from the related notion in groups (see below), we will call it vertex-accessibility in the remainder of the paper. A *ray* in an infinite graph G is an infinite one-way path in G. Two rays of G are equivalent if there are infinitely many disjoint paths between them in G (note that this is indeed an equivalence relation). An *end* of G is an equivalence class of rays in G. When there is a set X of vertices of G, two distinct components C_1, C_2 of $G - X$, and two distinct ends ω_1, ω_2 of G such that for each $i = 1, 2$, all rays of ω_i have infinitely many vertices in C_i, we say that X *separates* ω_1 and ω_2. A graph G is *vertex-accessible* if there is an integer k such that for any two ends ω_1, ω_2 in G, there is a set of at most k vertices that separates ω_1 and ω_2.

It was proved by Dunwoody [12] (see also [14, 15] for a more combinatorial approach) that locally finite quasi-transitive planar graphs are vertex-accessible. Here we extend the result to graphs excluding a countable clique K_∞ as a minor, and in particular to any proper minor-closed family.

Theorem 1.4. *Every locally finite quasi-transitive K_∞-minor-free graph is vertex-accessible.*

1.4 Accessibility in groups

The notion of vertex-accessibility introduced above is related to the notion of accessibility in groups. Given a finitely generated group Γ, and a finite set of generators S, the *Cayley* *graph* of Γ with respect to the set of generators S is the edge-labelled graph $\mathrm{Cay}(\Gamma, S)$ whose vertex set is the set of elements of Γ and where for every two $g, h \in \Gamma$ we put an edge (g, h) labelled with $a \in S$ when $h = a \cdot g$. It is known that the number of ends of a Cayley graph of a finitely generated group does not depend of the choice of generators, so we can talk about the number of ends of a finitely generated group. A classical theorem of Stallings [17] states that if a finitely generated group Γ has more than one end, it can be split as a non-trivial free product with finite amalgamation, or as an HNN-extension over a finite subgroup. If any group produced by the splitting still has more than one end we can keep splitting it using Stallings theorem. If the process eventually finishes (with Γ being obtained from finitely many 0-ended or 1-ended groups using free products with amalgamation and HNN-extensions), then Γ is said to be *accessible*. Thomassen and Woess [19] proved that a finitely generated group is accessible if and only if at least one of its Cayley graphs is vertex-accessible, if and only if all of its Cayley graphs are vertex-accessible.

A finitely generated group is *minor-excluded* if at least one of its Cayley graphs avoids a finite minor. Similarly a finitely generated group is K_∞*-minor-free* if one of its Cayley graphs avoids the countable clique as a minor, and *planar* if one of its Cayley graphs is planar. Note that planar groups are minor-excluded and Theorem 1.3 immediately implies that a finitely generated group is minor-excluded if and only if it is K_∞-minor-free.

Droms [9] proved that finitely generated planar groups are finitely presented, while Dunwoody [11] proved that finitely presented groups are accessible, which implies that finitely generated planar groups are accessible. Theorem 1.4 immediately implies the following, which extends this result to all minor-excluded finitely generated groups, and equivalently to all finitely generated K_∞-minor-free groups.

Corollary 1.5. *Every finitely generated K_∞-minor-free group is accessible.*

In fact, combining Theorem 1.1 with techniques introduced by Hamann [14, 15] in the planar case, we prove the following stronger result which also implies Corollary 1.5 using the result of Dunwoody [11] that all finitely presented groups are accessible.

Theorem 1.6. *Every finitely generated K_∞-minor-free group is finitely presented.*

1.5 The domino problem

We refer to [2] for a detailed introduction to the domino problem. A *coloring* of a graph G with colors from a set Σ is simply a map $V(G) \to \Sigma$. The domino problem for a finitely generated group Γ together with a finite generating set S is defined as follows. The input is a finite alphabet Σ and a finite set $\mathcal{F} = \{F_1, \ldots, F_p\}$ of forbidden *patterns*, which are colorings with colors from Σ of the closed neighborhood of the neutral element 1_Γ in the Cayley graph $\mathrm{Cay}(\Gamma, S)$, viewed as an edge-labelled subgraph of $\mathrm{Cay}(\Gamma, S)$. The problem then asks if there is a coloring of $\mathrm{Cay}(\Gamma, S)$ with colors from Σ, such that for each $v \in \Gamma$, the coloring of the closed neighborhood of v in $\mathrm{Cay}(\Gamma, S)$ (viewed as an edge-labelled subgraph of Cay(Γ, S)), is not isomorphic to any of the colorings $F_1, \ldots, F_p$, where we consider isomorphisms preserving the edge-labels.

It turns out that the decidability of the domino problem for (Γ, S) is independent of the choice of the finite generating set S, hence we can talk of the decidability of the domino problem for a finitely generated group Γ. If we consider $\Gamma = (\mathbb{Z}^2, +)$, then the domino problem corresponds exactly to the well-known Wang tiling problem, which was shown to be undecidable by Berger in [6]. On the other hand, there is a simple greedy procedure to solve the domino problem in free groups, which admit trees as Cayley graphs. More generally, the domino problem is decidable in virtually free groups, which can equivalently be defined as finitely generated groups having a Cayley graph of bounded treewidth [2, 1]. A remarkable conjecture of Ballier and Stein [5] asserts that these groups are the only one having decidable domino problem.

Conjecture 1.7 (Domino problem conjecture [5])**.** *A finitely generated group has decidable domino problem if and only if it is virtually free.*

Recall that virtually free groups are precisely the groups having a Cayley graph of bounded treewidth, which is a property that is closed under taking minor. It is therefore natural to ask whether Conjecture 1.7 holds for minor-excluded groups (or equivalently, using Theorem 1.3 to K_∞-minor-free groups). Using Theorem 1.5, together with classical results on planar groups and recent results on fundamental groups of surfaces [3], we prove that this is indeed the case.

Theorem 1.8. *A finitely generated K_∞-minor-free group has decidable domino problem if and only if it is virtually free.*

2 Proof overview

2.1 Sketch of the proof of Theorems 1.1 and 1.2

Consider a locally finite quasi-transitive graph G. First note that if G is quasi-4-connected, the following result of Thomassen immediately implies that the trivial tree-decomposition has the desired properties, hence Theorem 1.2 can be seen as a generalization of it.

Theorem 2.1 ([18])**.** *Let G be a quasi-transitive, quasi-4-connected, locally finite graph which excludes the countable clique as a minor. Then G is planar or has finite treewidth.*

To deal with the more general case, the first step is to obtain a canonical tree-decomposition of G of adhesion at most 2 in which all torsos are minors of G that are 3-connected graphs, cycles, or complete graphs on at most 2 vertices. The existence of such a decomposition in the finite case is a well-known result of Tutte [20] and was proved in the locally finite case in [10]. For our proof we need to go one step further. A graph is said to be *quasi-4-connected* if it is 3-connected and for every set $S \subseteq V(G)$ of size 3 such that $G - S$ is not connected, $G - S$ has exactly two connected components and one of them consists of a single vertex.

A crucial step for us would be to prove a version of the following result of Grohe in which the tree-decomposition would be canonical, and which would hold for locally finite graphs.

Theorem 2.2 ([13]). *Every finite graph G has a tree-decomposition of adhesion at most 3 whose torsos are minors of G and are complete graphs on at most 4 vertices or quasi-4-connected graphs.*

However, as observed by Grohe, even in the finite case the decomposition he obtains is not canonical in general. Our main technical contribution is to extend Theorem 2.2 to locally finite graphs, while making sure that most of the construction (except the very end) is canonical. For this, we proceed in two steps. First, we use a result of [7] to find a canonical tree-decomposition of any 3-connected graph G that distinguishes all its tangles of order 4. Using this result, we show that we can assume that the graph under consideration admits a unique tangle $\mathcal{T}$ of order 4. We then follow the main arguments from [13] and show that G has a canonical tree-decomposition of adhesion 3 which is a star and whose torsos are all minors of G and finite, except for the torso H associated to the center of the star, which has the following property: there exists a matching $M \subseteq E(H)$ which is invariant under the automorphism group of G and such that the graph $H' := H/M$ obtained after the contraction of the edges of M is quasi-transitive, locally finite, and quasi-4-connected. In particular, Theorem 2.1 implies that H' is planar or has bounded treewidth. We then prove that even if H itself is not necessarily quasi-4-connected, it is still planar or has bounded treewidth, which is enough to conclude the proof of Theorem 1.2. The final step to prove Theorem 1.1 consists in refining the tree-decomposition to make sure that torsos of bounded treewidth are replaced by torsos of finite size (moreover, this refinement has to be done in a canonical way). $\square$

2.2 Sketch of the proof of Theorem 1.8

Let Γ be a finitely generated K_∞-minor-free group. Let S be a finite set of generators such that $G := \mathrm{Cay}(\Gamma, S)$ excludes the countable clique as a minor. If Γ has 0 or 2 ends, then the domino problem is known to be decidable in Γ. Assume now that Γ (and thus G) is one-ended. As G is vertex-transitive, a result of Thomassen [18] implies that G is planar. It is known that one-ended planar groups contain fundamental groups of surfaces as subgroups of finite index. For such groups the domino problem is known to be undecidable [3], and this directly implies that the domino problem is undecidable in Γ. Finally, assume that Γ has an infinite number of ends. Theorem 1.5 implies that Γ can be described as the fundamental group of a finite graph of group H, whose vertex-groups Γ_v ($v \in V(H)$) all have at most one end and are subgroups of Γ. If all the vertex-groups Γ_v have 0 ends (or equivalently, are finite), then Γ is virtually free and the domino problem is known to be decidable in Γ. Otherwise at least one vertex-group Γ_v is one-ended. A result of Babai [4] implies that since Γ_v is a subgroup of Γ, Γ_v is also K_∞-minor-free. The proof above then shows that in this case the domino problem in undecidable in Γ_v, and thus also in Γ. $\square$

Acknowledgement We thank Marthe Bonamy, Vincent Delecroix, Reinhard Diestel, Matthias Hamann, Raphael Jacobs, Paul Knappe, Johannes Carmesin and Jan Kurkofka for the helpful discussions.

References

[1] Yago Antolín. On Cayley graphs of virtually free groups. *Groups Complexity Cryptology*, 3(2):301–327, 2011.

[2] Nathalie Aubrun, Sebastián Barbieri, and Emmanuel Jeandel. About the Domino Problem for Subshifts on Groups. In V. Berthé and M. Rigo, editors, *Sequences, Groups, and Number Theory*, Trends in Mathematics, pages 331–389. Birkhäuser, Cham, 2018.

[3] Nathalie Aubrun, Sebastián Barbieri, and Etienne Moutot. The domino problem is undecidable on surface groups. In Peter Rossmanith, Pinar Heggernes, and Joost-Pieter Katoen, editors, *44th International Symposium on Mathematical Foundations of Computer Science, MFCS 2019, August 26-30, 2019, Aachen, Germany*, volume 138 of *LIPIcs*, pages 46:1–46:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

[4] László Babai. Some applications of graph contractions. *Journal of Graph Theory*, 1(2):125–130, 1977.

[5] Alexis Ballier and Maya Stein. The domino problem on groups of polynomial growth. *Groups, Geometry, and Dynamics*, 12(1):93–105, 2018.

[6] Robert L. Berger. The undecidability of the domino problem. *Memoirs of the American Mathematical Society*, 1966.

[7] Johannes Carmesin, Matthias Hamann, and Babak Miraftab. Canonical trees of tree-decompositions. *Journal of Combinatorial Theory, Series B*, 152:1–26, 2022.

[8] Reinhard Diestel and Robin Thomas. Excluding a countable clique. *Journal of Combinatorial Theory, Series B*, 76(1):41–67, 1999.

[9] Carl Droms. Infinite-ended groups with planar Cayley graphs. *Journal of Group Theory*, 9(4):487–496, 2006.

[10] Carl Droms, Brigitte Servatius, and Herman Servatius. The structure of locally finite two-connected graphs. *Electronic Journal of Combinatorics*, 2, 01 1998.

[11] Martin J. Dunwoody. The accessibility of finitely presented groups. *Inventiones Mathematicae*, 81:449–458, 1985.

[12] Martin J. Dunwoody. Planar graphs and covers. *arXiv preprint arXiv:0708.0920*, 2007.

[13] Martin Grohe. Quasi-4-connected components. In Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi, editors, *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11-15, 2016, Rome, Italy*, volume 55 of *LIPIcs*, pages 8:1–8:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016.

[14] Matthias Hamann. Planar transitive graphs. *Electronic Journal of Combinatorics*, 25:4.8, 2015.

[15] Matthias Hamann. Accessibility in transitive graphs. *Combinatorica*, 38(4):847–859, 2018.

[16] Neil Robertson and Paul Seymour. Graph minors. XVI. Excluding a non-planar graph. *Journal of Combinatorial Theory, Series B*, 89(1):43–76, 2003.

[17] John Stallings. *Group theory and three-dimensional manifolds.* Yale University Press, 1972.

[18] Carsten Thomassen. The Hadwiger number of infinite vertex-transitive graphs. *Combinatorica*, 12:481–491, 1992.

[19] Carsten Thomassen and Wolfgang Woess. Vertex-transitive graphs and accessibility. *Journal of Combinatorial Theory, Series B*, 58(2):248–268, 1993.

[20] W.T. Tutte. *Graph Theory.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1984.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Sharp threshold for embedding balanced spanning trees in random geometric graphs

(Extended abstract)

Alberto Espuny Díaz* Lyuben Lichev† Dieter Mitsche‡
Alexandra Wesolek§

Abstract

Consider the random geometric graph $\mathcal{G}(n, r)$ obtained by independently assigning a uniformly random position in $[0, 1]^2$ to each of the n vertices of the graph and connecting two vertices by an edge whenever their Euclidean distance is at most r. We study the event that $\mathcal{G}(n, r)$ contains a spanning copy of a balanced tree T and obtain sharp thresholds for these events. Our methods provide a polynomial-time algorithm for finding a copy of such trees T above the threshold.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-060`

1 Introduction

The *random geometric graph* $\mathcal{G}(n, r)$ is a classic model of random graphs defined as follows. Let n be a positive integer, and let r be a positive real number. The vertices of the graph

*Institut für Mathematik, Technische Universität Ilmenau, 98684 Ilmenau, Germany. E-mail: `alberto.espuny-diaz@tu-ilmenau.de`. Supported by the Carl-Zeiss-Foundation and by DFG grant PE 2299/3-1.

†Institut Camille Jordan, University Jean Monnet, Saint-Etienne, France. E-mail: `lyuben.lichev@univ-st-etienne.fr`.

‡IMC, Pont. Univ. Católica, Chile and Institut Camille Jordan, Univ. Jean Monnet, Saint-Etienne, France. E-mail: `dieter.mitsche@mat.uc.cl`. Supported by grant GrHyDy ANR-20-CE40-0002 and by Fondecyt grant 1220174.

§Université Lyon, CNRS, INSA Lyon, UCBL, LIRIS, UMR5205, F-69622 Villeurbanne, France. E-mail: `agwesole@sfu.ca`.

are n points sampled uniformly at random and independently from $[0,1]^2$, and two vertices are connected by an edge if their Euclidean distance is at most r. Since their introduction by Gilbert [7] as a model for telecommunication networks, random geometric graphs have received a lot of attention from both applied and theoretical points of view [1, 6, 11, 13, 14].

1.1 Thresholds in random geometric graphs

Random geometric graphs are known to exhibit *threshold* behavior for many graph properties, meaning that there are some special values of the parameters of the model around which a drastic change in the behavior of the graph with respect to these properties takes place. Formally, a function $r^* = r^*(n)$ is a *threshold* for some monotone increasing property $\mathcal{P}$ in $\mathcal{G}(n,r)$ if

$$\lim_{n\to\infty} \mathbb{P}[\mathcal{G}(n,r) \in \mathcal{P}] = \begin{cases} 0 & \text{if } r = o(r^*), \\ 1 & \text{if } r = \omega(r^*). \end{cases}$$

Moreover, we say that a function $r^* = r^*(n)$ is a *sharp threshold* for $\mathcal{P}$ if, for every $\varepsilon \in (0,1)$,

$$\lim_{n\to\infty} \mathbb{P}[\mathcal{G}(n,r) \in \mathcal{P}] = \begin{cases} 0 & \text{if } r \le (1-\varepsilon) r^*, \\ 1 & \text{if } r \ge (1+\varepsilon) r^*. \end{cases}$$

Goel, Rai and Krishnamachari [8] gave a general upper bound for the threshold width in $\mathcal{G}(n,r)$, and Bradonjić and Perkins [2] characterized vertex-monotone properties which exhibit a sharp threshold. While the results in [2, 8] serve to prove the existence of (sharp) thresholds, they give no indication of where these thresholds actually are. Determining the (sharp) thresholds for different properties of interest is one of the main problems in the area, and it has received much attention. In this extended abstract, our aim is to determine the (sharp) threshold for the appearance of different spanning trees.

There are some results about thresholds which are closely related to our problem. The sharp threshold for connectivity was determined independently by Gupta and Kumar [9] and Penrose [12], who proved that it is $\sqrt{\log n / \pi n}$. Díaz, Mitsche and Pérez-Giménez [4] obtained the sharp threshold for $\mathcal{G}(n,r)$ to contain a Hamilton cycle, which coincides with the sharp threshold for connectivity. In particular, this gives the sharp threshold for the containment of a *spanning path*.

The question of trying to find the (sharp) threshold for the appearance of different families of spanning trees pops up naturally. The results about the sharp threshold for Hamiltonicity can be used to deduce thresholds for path-like trees. Indeed, using the triangle inequality, for any fixed $k \ge 2$, one may show that the threshold for Hamiltonicity is also a threshold for the property that $\mathcal{G}(n,r)$ contains the k-th power of a Hamilton cycle. It immediately follows that every spanning tree which can be embedded into the k-th power of a Hamilton cycle has this same threshold. This is the case, for instance, of spanning caterpillars with constant maximum degree.

One may naturally wonder whether all spanning trees with bounded maximum degree have the same threshold. Incidentally, in the model of binomial random graphs $\mathcal{G}(n,p)$ where each of the $\binom{n}{2}$ possible edges appears independently with probability p, Montgomery [10] proved that this is the case: the threshold for all bounded degree spanning trees is $\log n/n$. However, this turns out to be very far from the truth in random geometric graphs. Indeed, there are bounded degree trees T whose diameter is logarithmic in the number of vertices, and this diameter directly imposes a much higher lower bound on the threshold r^* for the property of containing a copy of T: since a spanning subgraph of $\mathcal{G}(n, r)$ cannot have smaller diameter than $\mathcal{G}(n, r)$ itself, the threshold for trees with diameter $O(\log n)$ must satisfy $r^* = \Omega(1/\log n)$, which is far larger than the connectivity threshold mentioned above. The results of Goel, Rai and Krishnamachari [8] imply that, for any such tree, there is a sharp threshold. Our goal is to determine the value of this threshold.

Out of all trees with logarithmic diameter, binary trees are especially interesting due to their many applications as data structures (see, e.g., [3]). Identifying the sharp threshold for embedding these trees into $\mathcal{G}(n, r)$ is thus an important particular case of our study.

1.2 Main results

An *s-ary tree* is a rooted tree where every node has at most s children. We say that an s-ary tree is *balanced* if there is an integer h such that all vertices at (graph) distance at most $h - 1$ from the root have exactly s children, and all vertices at distance h from the root are leaves. Our first result determines the sharp threshold for $\mathcal{G}(n, r)$ to contain a spanning copy of the balanced s-ary tree for any fixed integer $s \geq 2$.

Theorem 1. *Fix an integer $s \geq 2$. Let h be a positive integer, and set $n := \sum_{i=0}^{h} s^i$. Let T_h be the balanced s-ary tree of height h (and on n vertices). Then, $r^* := 1/\sqrt{2}h$ is the sharp threshold for $\mathcal{G}(n, r)$ to contain a copy of T_h.*

In fact, Theorem 1 is a particular case of a similar result for a larger class of trees. We may think of the vertices of the rooted tree T as being partitioned into *layers* $V_0, V_1, \ldots$, where V_i contains all vertices at (graph) distance i from the root. For any vertex $v \in V(T)$, if $v \in V_i$, we refer to its neighbors in V_{i+1} as its *children*, and to all vertices which can be reached by a path from v without going through V_{i-1} as its *descendants*. The *height* of the rooted tree is the maximum (graph) distance between the root and another vertex of T. Given positive integers h and $(s_i)_{i=1}^h$, we say that a tree T is the *balanced tree over the sequence* $(s_i)_{i=1}^h$ if it has height h and, for each $i \in \{1, \ldots, h\}$, every vertex of T in V_{i-1} has exactly s_i children. In particular, such a tree T contains exactly $\sum_{i=0}^{h} \prod_{j=1}^{i} s_j$ vertices (where, by convention, the empty product equals 1). If $2 \leq s_i \leq M$ for all $i \in \{1, \ldots, h\}$ and for some positive integer M, we say that T is a *balanced M-tree*.

For a graph G, we denote by $|G|$ the size of the vertex set of G. As typical in random graphs literature, we focus on asymptotic statements. Our asymptotic notations will be taken with respect to the height of the trees, which then also yields asymptotic results with respect to n. Our next result extends Theorem 1 to all balanced M-trees as long as M is not too large compared to h.

Theorem 2. *Let $2 \leq M = M(h) = o(h/\log h)$. Let $(T_h)_{h\geq 1}$ be a sequence of trees where T_h is a balanced M-tree of height h. Then, $r^* := 1/\sqrt{2}h$ is the sharp threshold for $\mathcal{G}(|T_h|, r)$ to contain a copy of T_h.*

We first provide here the (much easier) proof of the lower bound of $1/\sqrt{2}h$ in Theorem 2. Note that a.a.s. (that is, with probability tending to 1 as $h \to \infty$) there are vertices $u, v \in V(\mathcal{G}(|T_h|, r))$ at Euclidean distance $(1 - o(1))\sqrt{2}$ from each other. Indeed, a.a.s. the squares $c_0, c_1 \subseteq [0,1]^2$ of side length $1/h$ containing the corners $(0,0)$ and $(1,1)$, respectively, each contain at least a vertex. Conditioning on this event, assume that G admits T_h as a spanning tree. Since T_h has diameter $2h$, u and v must be at distance at most $2h$ in G. Thus, by the triangle inequality, the Euclidean distance between them must be at most $2hr$, so we must have $2hr \geq \mathrm{dist}(c_0, c_1) = (1 - o(1))\sqrt{2}$.

The proof of the upper bound is much more involved. In order to simplify our exposition, in the next section we provide a sketch of the proof only for the case of balanced binary trees. This particular case already contains most ideas of the more general theorem. For all the details of the proof, we refer the reader to the full version of our paper [5].

2 Embedding algorithm for the upper bound

Fix $\varepsilon \in (0,1)$. Let T be a balanced binary tree of height h, and set $n := 2^{h+1} - 1$. Let $G = \mathcal{G}(n, r)$. Suppose that $r \geq (1+\varepsilon)r^*$. We embed the layers $V_0, V_1, \ldots, V_h$ of T into G one at a time, starting from the root. Throughout, we refer to the embedding of one of the layers of T as a *step* of the algorithm (for simplicity, we assume that V_0 is embedded at the 0-th step). For each $i \in \{0, \ldots, h\}$, at the end of the i-th step, we call a vertex of G into which a vertex on layer V_i has been embedded *active*. Moreover, we refer to the vertices of G into which no vertex of T has been embedded as *unseen*. For simplicity of notation, once a vertex of T has been embedded into G, we often interchangeably use the same notation to refer to either of the two vertices.

Let k be the smallest integer which satisfies that

$$2^{1/2-k} < \frac{\varepsilon r^*}{8}. \tag{1}$$

Let $\mathcal{S}$ be the tessellation of $[0,1]^2$ into 2^{2k} congruent closed axis-parallel squares. We first focus on embedding the layers $V_0, \ldots, V_{m+2}$ into G, where m will be defined below, in such a way that the vertices in V_{m+2} are distributed "sufficiently uniformly" in $[0,1]^2$. To be more precise, we ensure that each square of $\mathcal{S}$ contains 2^{m+2-2k} vertices from V_{m+2}. We then finish the embedding of the remaining layers with a suitable application of Hall's theorem.

For each $\ell \in \{0, \ldots, k\}$, let $\mathcal{S}_\ell$ be the tessellation of $[0,1]^2$ into $2^{2\ell}$ congruent closed axis-parallel squares obtained by combining the squares of $\mathcal{S}$ into groups of size $2^{2(k-\ell)}$. In particular, $\mathcal{S}_0 = \{[0,1]^2\}$ and $\mathcal{S}_k = \mathcal{S}$. For a square $q \subseteq [0,1]^2$, we denote its center by $c(q)$. Moreover, for every $i, j \in \{0, \ldots, k\}$ with $i < j$ and any square $q \in \mathcal{S}_i$, let $\sigma_j(q)$ denote the set of four subsquares of q in $\mathcal{S}_j$ that form the axis-parallel square of side length 2^{1-j}

and center $c(q)$. In particular, $\sigma_{i+1}(q)$ is a tessellation of q into four subsquares in $\mathcal{S}_{i+1}$, and for each $j \in \{i+2, \ldots, k\}$, the set $\sigma_j(q)$ is obtained from $\sigma_{j-1}(q)$ by homothety with center $c(q)$ and ratio $1/2$. Note that we sometimes abuse notation and identify $\sigma_j(q)$ with its geometric realization; in particular, we identify $\bigcup_{p\in\sigma_j(q)} p$ with $\sigma_j(q)$ itself.

To begin with, we embed the layers V_0 and V_1 of T into an arbitrary square $q_0 \in \sigma_k([0,1]^2)$, and then the four vertices in V_2 are evenly distributed among the four squares in $\sigma_k([0,1]^2)$. Our algorithm for embedding $V_3, \ldots, V_h$ has two main parts, which we call *subroutines*. The first subroutine is used to embed layers $V_3, \ldots, V_{m+2}$, while the second subroutine deals with the remaining layers.

2.1 The first subroutine

The m steps of this subroutine are grouped into $k-1$ different blocks. For each $\ell \in \{1, \ldots, k-1\}$, we proceed iteratively as follows. Suppose that at the beginning of the ℓ-th block we have a configuration in which every square $q \in \mathcal{S}_{\ell-1}$ contains the same number of active vertices, and that these are equally distributed among all subsquares in $\sigma_k(q)$ (note that this is verified in the case $\ell = 1$). Then, for each $q \in \mathcal{S}_{\ell-1}$, we proceed as follows.

Iteration: We proceed to distributing the descendants of the currently active vertices in a way that we embed them at increasing distances from $\sigma_k(q)$ as follows.

Define $\phi_\ell \colon \sigma_k(q) \to \sigma_\ell(q)$ as the bijection obtained by homothety with center $c(q)$ and ratio $2^{k-\ell}$. To each square $p \in \sigma_k(q)$ we associate a sequence of squares $(p_1, \ldots, p_t)$ in $\mathcal{S}_k$, for some appropriately chosen t which does not depend on p, which satisfies that $p_1 = p$, $p_t \in \sigma_k(\phi_\ell(p))$, and for all $i \in \{1, \ldots, t-1\}$ we have $\|c(p_{i+1}) - c(p_i)\| \le (1 + 7\varepsilon/8)\, r^*$. Note that the last condition together with the triangle inequality and (1) ensures that, for every $i \in \{1, \ldots, t-1\}$ and every choice of points $x \in p_i$ and $y \in p_{i+1}$, we have

$$\|x-y\| \le \|x-c(p_i)\|+\|c(p_i)-c(p_{i+1})\|+\|c(p_{i+1})-y\| \le \frac{\varepsilon}{16}r^*+\left(1+\frac{7\varepsilon}{8}\right)r^*+\frac{\varepsilon}{16}r^* = r. \quad (2)$$

Hence, if there is an active vertex v in p_i, it is possible to embed all children of v in p_{i+1}.

Let us prove that such sequences of squares can indeed be constructed. Fix $p \in \sigma_k(q)$ as above. We begin by setting $p_1 := p$. For each $i \ge 1$, while $\|c(p_i) - c(\phi_\ell(p))\| > (1 + 7\varepsilon/8)r^*$, we choose a square $p_{i+1} \in \mathcal{S}_k$ such that

$$\|c(p_{i+1})-c(\phi_\ell(p))\| \le \|c(p_i)-c(\phi_\ell(p))\|-\left(1+\frac{5\varepsilon}{8}\right)r^* \text{ and } \|c(p_{i+1})-c(p_i)\| \le \left(1+\frac{7\varepsilon}{8}\right)r^*. \quad (3)$$

Finally, when we reach a point where $\|c(p_i) - c(\phi_\ell(p))\| \le (1 + 7\varepsilon/8)r^*$, we set $t := i + 1$ and choose p_t as a square in $\sigma_k(\phi_\ell(p))$ at smallest distance to $c(p_i)$. Note that (2) holds for $i = t-1$ and this choice of p_t. It remains to prove that a choice as prescribed by (3) can indeed be made.

Assume we have already defined p_i so that it satisfies $\|c(p_i) - c(\phi_\ell(p))\| > (1 + 7\varepsilon/8)r^*$. Then, choose $p_{i+1} \in \mathcal{S}_k$ to be a square containing the point w on the segment $c(p_i)c(\phi_\ell(p))$

at distance exactly $(1+3\varepsilon/4)r^*$ from $c(p_i)$ (if more than one such square exists, we choose one arbitrarily). Then, the triangle inequality and (1) imply that

$$\begin{aligned}\|c(p_{i+1}) - c(\phi_\ell(p))\| &\le \|w - c(\phi_\ell(p))\| + \|c(p_{i+1}) - w\| \\ &= \|c(p_i) - c(\phi_\ell(p))\| - \|c(p_i) - w\| + \|c(p_{i+1}) - w\| \\ &\le \|c(p_i) - c(\phi_\ell(p))\| - \left(1 + \frac{5\varepsilon}{8}\right) r^*\end{aligned}$$

and

$$\|c(p_i) - c(p_{i+1})\| \le \|c(p_i) - w\| + \|w - c(p_{i+1})\| \le \left(1 + \frac{3\varepsilon}{4}\right) r^* + \frac{\varepsilon}{16} r^* \le \left(1 + \frac{7\varepsilon}{8}\right) r^*,$$

so (3) is verified.

Now, once the sequence $(p_1, \ldots, p_t)$ is constructed, we can describe the distribution of the vertices of the tree in each step of the ℓ-th block of the algorithm. Recall that, at the beginning of the ℓ-th block of the algorithm, p_1 contains some active vertices. Then, for the next $t-1$ steps, we simply embed the children of all currently active vertices in p_i into p_{i+1} arbitrarily (which is possible thanks to (2)). Lastly, in one more step of the algorithm, we distribute the children of all currently active vertices in p_t equally among the four subsquares of $\sigma_k(\phi_\ell(p))$. Finally, if $\ell = k-1$ we terminate the subroutine, and otherwise we increment the value of ℓ by 1 and proceed to the next block of the first subroutine.

2.2 The second subroutine

We say that two squares of $\mathcal{S}$ are *adjacent* if they share at least one corner vertex. Our next goal is to embed the vertices in the remaining layers. In particular, if $v \in V_{m+2}$ lies in $q \in \mathcal{S}$, then we embed all descendants of v into q or a square adjacent to q. Note that, for every such v, this is possible by our choice of k since the vertices in any square of side length 2^{1-k} form a clique. To show that this can be done simultaneously for all $v \in V_{m+2}$, we use Hall's theorem. The application of Hall's theorem is standard, hence we omit it, but it can be found in the full version of our paper [5].

3 Proof sketch for the feasibility of the embedding algorithm

We show that our embedding algorithm succeeds a.a.s. We make use of the following distribution property of the vertices of G, which follows directly from Chernoff bounds.

Claim 1. *A.a.s., for every square $q \in \mathcal{S}$, the number of vertices of G in q is in $[2^{-2k}n - n^{2/3}, 2^{-2k}n + n^{2/3}]$.*

Following the description of the first subroutine, we must prove that it reaches a desired configuration in a suitable number of steps. A bound on the number of steps is given by the following claim.

Claim 2. *The first subroutine runs for at most* $(1-\varepsilon/4)h$ *steps.*

Proof. For every $\ell \in \{1,\dots,k-1\}$, we have that

$$\|c(p)-c(\phi_\ell(p))\| \le \|c(q)-c(\phi_\ell(p))\| = \frac{(2^{1-\ell}-2^{-\ell})}{\sqrt{2}},$$

where the inequality holds since $c(p)$ belongs to the segment $c(q)c(\phi_\ell(p))$ (recall that $\phi_\ell(p)$ is obtained from p by homothety with center $c(q)$ and ratio $2^{k-\ell}>1$). Hence, by (3), the total number of steps performed by the first subroutine is at most

$$\sum_{\ell=1}^{k-1}\left(\frac{(2^{1-\ell}-2^{-\ell})/\sqrt{2}}{(1+5\varepsilon/8)r^*}+1\right) \le k+\frac{1/\sqrt{2}}{(1+5\varepsilon/8)r^*} \le \left(1-\frac{\varepsilon}{4}\right)h. \qquad \blacktriangleleft$$

It follows that the total number of vertices embedded during the first subroutine is

$$\sum_{i=0}^{m+2} 2^i \le 2^{-(h-m-2)}n \le 2^{-\varepsilon h/5}n = o(2^{-2k}n). \tag{4}$$

Therefore, by Claim 1, there are sufficiently many vertices in each $q\in\mathcal{S}_k$ to guarantee that the choices we made in the first subroutine can indeed be carried out.

Claims 1 and 2 guarantee that a.a.s. the algorithm succeeds. That is, a.a.s. we have an embedding of the layers $V_0,\dots,V_{m+2}$ of T into G and, moreover, each $q\in\mathcal{S}_k$ contains the same number of vertices of V_{m+2}.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, Wireless sensor networks: a survey. *Comput. Netw.* 38.4 (2002), 393—422, DOI: 10.1016/S1389-1286(01)00302-4.

[2] M. Bradonjić and W. Perkins, On sharp thresholds in random geometric graphs. *Approximation, randomization, and combinatorial optimization. Algorithms and techniques. Proceedings of the 17th international workshop on approximation algorithms for combinatorial optimization problems (APPROX 2014) and the 18th international workshop on randomization and computation (RANDOM 2014), Universitat Politècnica de Catalunya, Barcelona, Spain, September 4–6, 2014*, 500–514, Wadern: Schloss Dagstuhl – Leibniz Zentrum für Informatik, ISBN 978-3-939897-74-3 (2014), DOI: 10.4230/LIPIcs.APPROX-RANDOM.2014.500.

[3] T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein, *Introduction to algorithms.* MIT Press, Cambridge, MA, 4th ed. (2022), ISBN 978-0-262-04630-5.

[4] J. Díaz, D. Mitsche and X. Pérez, Sharp Threshold for Hamiltonicity of Random Geometric Graphs. *SIAM J. Discrete Math.* 21 (2007), 57–65, DOI: 10.1137/060665300.

[5] A. Espuny Díaz, L. Lichev, D. Mitsche and A. Wesolek, Sharp threshold for embedding balanced spanning trees in random geometric graphs. *arXiv e-prints* (2023). arXiv: 2303.14229.

[6] D. Funke, S. Lamm, U. Meyer, M. Penschuck, P. Sanders, C. Schulz, D. Strash and M. von Looz, Communication-free massively distributed graph generation. *J. Parallel Distrib. Comput.* 131 (2019), 200–217, DOI: 10.1016/j.jpdc.2019.03.011.

[7] E. N. Gilbert, Random Plane Networks. *J. Soc. Ind. Appl. Math.* 9 (1961), 533–543, DOI: 10.1137/0109045.

[8] A. Goel, S. Rai and B. Krishnamachari, Monotone properties of random geometric graphs have sharp thresholds. *Ann. Appl. Probab.* 15.4 (2005), 2535–2552, DOI: 10.1214/105051605000000575.

[9] P. Gupta and P. R. Kumar, Critical power for asymptotic connectivity in wireless networks. W. M. McEneaney, G. G. Yin and Q. Zhang (eds.), *Stochastic analysis, control, optimization and applications. A volume in honor of Wendell H. Fleming, on the occasion of his 70th birthday*, 547–566, Birkhäuser, Boston, ISBN 0-8176-4078-9 (1999), DOI: 10.1007/978-1-4612-1784-8_33.

[10] R. Montgomery, Spanning trees in random graphs. *Adv. Math.* 356 (2019), 92, DOI: 10.1016/j.aim.2019.106793. Id/No 106793.

[11] M. Nekovee, Worm epidemics in wireless ad-hoc networks. *New J. Phys.* 9.6 (2007), paper n. 189, DOI: 10.1088/1367-2630/9/6/189.

[12] M. Penrose, The longest edge of the random minimal spanning tree. *Ann. Appl. Probab.* 7.2 (1997), 340–361, DOI: 10.1214/aoap/1034625335.

[13] ———, *Random geometric graphs*, *Oxf. Stud. Probab.*, vol. 5. Oxford: Oxford University Press (2003).

[14] ———, Lectures on random geometric graphs. *Random graphs, geometry and asymptotic structure*, 67–101, Cambridge: Cambridge University Press (2016), DOI: 10.1017/CBO9781316479988.004.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Odd-Sunflowers

(Extended abstract)

Peter Frankl* János Pach† Dömötör Pálvölgyi‡

Abstract

Extending the notion of sunflowers, we call a family of at least two sets an *odd-sunflower* if every element of the underlying set is contained in an odd number of sets or in none of them. It follows from the Erdős–Szemerédi conjecture, recently proved by Naslund and Sawin, that there is a constant $\mu < 2$ such that every family of subsets of an n-element set that contains no odd-sunflower consists of at most μ^n sets. We construct such families of size at least 1.5021^n.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-061`

1 Introduction

A family of at least 3 sets is a *sunflower* (or a Δ-*system*) if every element is contained either in all of the sets, or in at most one. If a family of sets contains no sets that form a sunflower, it is called *sunflower-free*. This notion was introduced by Erdős and Rado [10] in 1960, and it has become one of the standard tools in extremal combinatorics [14]. Erdős and Rado conjectured that the maximum size of any sunflower-free family of k-element

*Alfréd Rényi Institute of Mathematics, Budapest, Hungary. E-mail: `peter.frankl@gmail.com`. Partially supported by ERC Advanced Grant GeoScape.

†Alfréd Rényi Institute of Mathematics, Budapest, Hungary and IST, Klosterneuburg, Austria. E-mail: `pach@cims.nyu.edu`. Partially supported by ERC Advanced Grant GeoScape and NKFIH (National Research, Development and Innovation Office) grant K-131529.

‡ELTE Eötvös Loránd University and Alfréd Rényi Institute of Mathematics, Budapest, Hungary. E-mail: `domotor.palvolgyi@ttk.elte.hu`. Partially supported by the ERC Advanced Grant "ERMiD" and by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, and by the New National Excellence Program ÚNKP-22-5 and by the Thematic Excellence Program TKP2021-NKTA-62 of the National Research, Development and Innovation Office.

sets is at most c^k, for a suitable constant $c > 0$. This conjecture is still open; for recent progress, see [4].

Erdős and Szemerédi [11] studied the maximum possible size of a sunflower-free family of subsets of $\{1, \ldots, n\}$. Denote this quantity by $f(n)$ and let $\mu = \lim f(n)^{1/n}$. Erdős and Szemerédi conjectured that $\mu < 2$, and this was proved by Naslund and Sawin [18], using the methods of Croot, Lev, P. Pach [6], Ellenberg and Gijswijt [8], and Tao [19]. They showed that $\mu < 1.89$, while the best currently known lower bound, $\mu > 1.551$, follows from a construction of Deuber *et al.* [7].

Several variants of the above notion have also been considered. Erdős, Milner and Rado [9] called a family of at least 3 sets a *weak sunflower* if the intersection of any pair of them has the same size. For a survey, see Kostochka [16]. In the literature, we can also find pseudo-sunflowers [13] and near-sunflowers [3]. By restricting the parities of the sets, other interesting questions can be asked, some of which can be answered by the so-called linear algebra method (even-town, odd-town theorems; see [5]).

We introduce the following new variants of sunflowers.

Definition 1. *A nonempty family of nonempty sets forms an* even-degree sunflower *or, simply, an* even-sunflower*, if every element of the underlying set is contained in an even number of sets (or in none). Analogously, a family of at least two nonempty sets forms an* odd-degree sunflower *or, simply, an* odd-sunflower*, if every element of the underlying set is contained in an odd number of sets, or in none.*

Note that any family of pairwise disjoint sets is an odd-sunflower but not an even-sunflower. A (classical) sunflower is an odd-sunflower if and only if it consists of an odd number of sets. In particular, an odd-sunflower-free family is also sunflower-free, as any sunflower contains a sunflower that consists of 3 sets. on the other hand, there exist many odd-sunflowers that contain no sunflower of size 3. For example, $\{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$ is a minimal odd-sunflower. This example can be generalized as follows.

Let $\mathcal{C}_n$ denote the $(n-1)$-uniform family consisting of all $(n-1)$-element subsets of $\{1, \ldots, n\}$. Let $\mathcal{C}_n^+$ denote the same family completed with the set $\{1, \ldots, n\}$. Obviously, $\mathcal{C}_n$ is an odd-sunflower if and only if n is even, and it is an even-sunflower if and only if n is odd. The family $\mathcal{C}_n^+$ is an odd-sunflower if and only if n is odd, and it is an even-sunflower if and only if n is even. Notice that in any subfamily of these families the nonzero degrees of the vertices differ by at most one. Therefore, in every subfamily of $\mathcal{C}_n$ and $\mathcal{C}_n^+$ which is odd- or even-suflower, all nonzero degrees need to be the same, showing that $\mathcal{C}_n$ and $\mathcal{C}_n^+$ are minimal odd- or even-sunflowers. There are many other examples; e.g., all graphs in which every degree is odd/even are 2-uniform odd/even-sunflowers. In fact, we can show that it is NP-complete to decide whether an input family is odd-sunflower-free or not, so there is no hope of a characterization. This is in contrast with (classic) sunflowers, where the problem is trivially in P.

The main goal of this paper is to raise the following questions: What is the maximum size of a family $\mathcal{F}$ of subsets of $\{1, \ldots, n\}$ that contains no even-sunflower or no odd-sunflower? We denote these maximums by $f_{even}(n)$ and by $f_{odd}(n)$, respectively. As in the case of the even-town and odd-town theorems, the answers to these questions are quite different.

Theorem 2. *$f_{even}(n) = n$, i.e., for any even-sunflower-free family $\mathcal{F} \subset 2^{\{1,\ldots,n\}}$ we have $|\mathcal{F}| \leq n$.*

Theorem 3. *$f_{odd}(n) > 1.502148^n$ if $n > n_0$, i.e., there are odd-sunflower-free families $\mathcal{F} \subset 2^{\{1,\ldots,n\}}$, for any large enough n with $|\mathcal{F}| > 1.502148^n$.*

Let $\mu_{odd} = \lim f_{odd}(n)^{1/n}$. (The existence of the limit easily follows from our Lemma 5 and Fekete's lemma, just like for classical sunflowers; see [1].) Using the fact that any odd-sunflower-free family $\mathcal{F}$ is also sunflower-free, the result of Naslund and Savin [18] mentioned above implies that $f_{odd}(n) \leq 1.89^n$. Thus, we have

$$1.502148 < \mu_{odd} \leq \mu < 1.89.$$

It would be interesting to decide whether μ_{odd} is strictly smaller than μ, and to find a direct proof for $\mu_{odd} < 2$. Is the new slice rank method required?

Some of the ideas used in the proof of Theorem 3 originate, in a slightly different form, in [2]; see Lemmas 5 and 6, and also the discussion on the MathOverflow website [17]. Here, we use a similar approach to recursively construct large odd-sunflower-free families of subsets of $\{1, \ldots, n\}$.

In Section 3, we will also establish a (negative) structural result: If n is large enough, the largest odd-sunflower-free families on the underlying set $\{1, \ldots, n\}$ cannot be obtained by repeatedly adding a small construction to itself, in a simple way (to be described in Lemma 5). We will refer to this method as the "brick construction."

We end this section with a definition. A family $\mathcal{F}$ is called an *antichain*, or *Sperner*, if it is containment-free, i.e., $F, G \in \mathcal{F}$ and $F \subset G$ imply that $F = G$. Let $f_{oa}(n)$ denote the maximum size of an antichain $\mathcal{F}$ on the underlying set $\{1, \ldots, n\}$ that contains no odd-sunflower. Note that any *slice* of $\mathcal{F}$, i.e., any subfamily of $\mathcal{F}$ whose sets are of the same size, form an antichain. Obviously, we have $f_{odd}(n)/n \leq f_{oa}(n) \leq f_{odd}(n)$ and, therefore,

$$\lim f_{oa}(n)^{1/n} = \mu_{odd}.$$

2 Proof of Theorem 2

The lower bound $f_{even}(n) \geq n$ follows from taking n singleton sets. For the upper bound $f_{even}(n) \leq n$, we sketch the argument in two different forms: using linear algebra (as in the usual proof of the odd-town theorem) and by a parity argument (which does not work there).

First proof. Represent each set by its characteristic vector over $\mathbb{F}_2^n$. If $|\mathcal{F}| > n$, these vectors have a nontrivial linear combination that gives zero. The sets whose coefficients are one in this combination yield an even-sunflower. □

Second proof. There are $2^{|\mathcal{F}|} - 1$ nonempty subfamilies of $\mathcal{F}$. If $|\mathcal{F}| > n$, by the pigeonhole principle, there are two different subfamilies that contain precisely the same elements of $\{1, \dots, n\}$ an odd number of times. But then their symmetric difference is an even-sunflower. $\square$

3 Brick Constructions

We start with the following simple construction.

Construction 1: Let $k = \lfloor n/3 \rfloor$. Make k disjoint groups of size 3 from $\{1, \dots, n\}$. Define $\mathcal{F}$ as the family of all sets that intersect each group in exactly 2 elements. Then we have $|\mathcal{F}| = 3^k$, i.e., $\sqrt[3]{3}^n$, whenever n is divisible by 3. This shows that

$$\mu_{odd} \geq \sqrt[3]{3} > 1.44. \tag{1}$$

To prove that this construction is odd-sunflower-free, we need some simple lemmas.

In a *multifamily* of sets, every set F can occur a positive integer number of times. This number is called the *multiplicity* of F. A multifamily of at least two nonempty sets is an *odd-sunflower* if the degree of every element of the underlying set is odd or zero. Note that, similarly to sunflowers, restricting an odd-sunflower multifamily to a smaller underlying set also gives an odd-sunflower multifamily, unless fewer than two nonempty sets remain.

Lemma 4. *If $\mathcal{F}$ is odd-sunflower-free family, and $\mathcal{H}$ is a multifamily of size at least two, comprised of elements $\mathcal{F}$, then $\mathcal{H}$ is an odd-sunflower multifamily if only if it consists of an odd number of copies of a single member $F \in \mathcal{F}$, and an even number of copies of some subsets of F.*

In particular, if $|\mathcal{H}|$ is even, it cannot be an odd-sunflower.

Remark. If $\mathcal{F}$ is an *antichain*, that is, if no $F \in \mathcal{F}$ has a proper subset that belongs to $\mathcal{F}$, then the multifamily $\mathcal{H}$ is an odd-sunflower if and only if it consists of an odd number of copies of the same set $F \in \mathcal{F}$.

Proof. The "if" part of the statement is obvious.

Assume that $\mathcal{H}$ is an odd-sunflower. Reduce the multifamily $\mathcal{H}$ to a family $\mathcal{H}'$ by deleting all sets of even multiplicity and keeping only one copy of each set whose multiplicity is odd. This does not change the parity of the degree of any vertex.

Suppose that $\mathcal{H}' \subseteq \mathcal{F}$ consists of at least two sets. Since $\mathcal{H}' \subseteq \mathcal{F}$ is odd-sunflower-free, there is an element which is contained in a nonzero even number of sets of $\mathcal{H}'$ and, therefore, in a nonzero even number of sets in the multifamily $\mathcal{H}$. This contradicts our assumption that $\mathcal{H}$ was an odd-sunflower.

If $\mathcal{H}'$ is empty, then any element covered by $\mathcal{H}$ is contained in an even number of sets from $\mathcal{H}'$, thus $\mathcal{H}$ again cannot be an odd-sunflower.

Finally, consider the case when the reduced family $\mathcal{H}'$ consists of a single set $F \in \mathcal{F}$. If all sets in the multifamily $\mathcal{H}$ are copies of F, we are done. Otherwise, there are some other sets $F' \neq F$ participating in $\mathcal{H}$ with even multiplicity. If any such F' has an element that does not belong to F, then this element is covered by a nonzero even number of sets of the multifamily $\mathcal{H}$, contradicting the assumption that $\mathcal{H}$ is an odd-sunflower. Therefore, all such F' are subsets of F, as claimed. □

Lemma 5. *If $\mathcal{F}$ and $\mathcal{G}$ are odd-sunflower-free families, and at least one of them is an antichain, then $\mathcal{F}+\mathcal{G}$ is also odd-sunflower-free. Moreover, if both $\mathcal{F}$ and $\mathcal{G}$ are antichains, then so is $\mathcal{F}+\mathcal{G}$.*

Remark. If none of $\mathcal{F}$ and $\mathcal{G}$ are antichains, then it can happen that $\mathcal{F}+\mathcal{G}$ contains an odd-sunflower. For example, if $\mathcal{F} = \{\{1\}, \{1,2\}\}$ and $\mathcal{G} = \{\{3\}, \{3,4\}\}$, then $\{\{1,3\}, \{1,2,3\}, \{1,3,4\}\}$ is an odd-sunflower.

Proof. The "moreover" part of the statement, according to which $\mathcal{F}+\mathcal{G}$ is an antichain, is trivial.

Suppose for contradiction that $\mathcal{F}+\mathcal{G}$ has a subfamily $\mathcal{H}$ consisting of at least two sets that form an odd-sunflower. Without loss of generality, $\mathcal{G}$ is an antichain.

Assume first that the parts of the sets of $\mathcal{H}$ that come from $\mathcal{G}$ are not all the same. These parts are the restriction of $\mathcal{H}$ to the underlying set of $\mathcal{G}$, so they form a multifamily which is an odd-sunflower. Applying Lemma 4 to this subfamily, it follows that the parts of the sets in $\mathcal{H}$ that come from $\mathcal{F}$ all coincide, contradicting our assumption.

Otherwise, the parts of the sets of $\mathcal{H}$ that come from $\mathcal{G}$ are all the same, in which case the parts that come from $\mathcal{F}$ are all different. But then we can use that $\mathcal{F}$ is sunflower-free. □

Corollary 6. *For any integers $n, m, t > 0$, we have $f_{oa}(n)+f_{oa}(m) \geq f_{oa}(n+m)$, $f_{oa}(tn) \geq t f_{oa}(n)$, $\mu_{odd} \geq f_{oa}(n)^{1/n}$, where f_{oa} is the function introduced at the end of Section 1.*

This follows by repeated application of Lemma 5. We can think of $\mathcal{F}$ as the "building block" of $\mathcal{F}+\mathcal{F}+\cdots+\mathcal{F}$, so such constructions may be referred to as *brick constructions* [7]. When $\mathcal{F} = \mathcal{C}_3$ consists of the two-element subsets of $\{1,2,3\}$, we recover Construction 1. This proves (1).

4 Wreath Product Constructions

In this section, we describe another construction that uses the *wreath product* of two families. This is a common notion in group theory [15], but less common in set theory. It was introduced in the PhD thesis of the first author [12]; see also [17].

Let n, m be positive integers, $\mathcal{F} \subseteq 2^{\{1,\dots,n\}}$, $\mathcal{G} \subseteq 2^{\{1,\dots,m\}}$ families of subsets of $N = \{1,\dots,n\}$ and $M = \{1,\dots,m\}$, respectively. Take n isomorphic copies $\mathcal{G}_1,\dots,\mathcal{G}_n$ of $\mathcal{G}$ with pairwise disjoint underlying sets $M_1,\dots,M_n$. Define the *wreath product* of $\mathcal{F}$ and $\mathcal{G}$, denoted by $\mathcal{F} \wr \mathcal{G}$, on the underlying set $\cup_{i=1}^n M_i$, as follows.

$$\mathcal{F} \wr \mathcal{G} = \{\bigcup_{i \in F} G_i \mid F \in \mathcal{F}, G_i \in \mathcal{G}_i\}.$$

That is, for each $F \in \mathcal{F}$, for every $i \in F$, for every $G_i \in \mathcal{G}_i$, we take the set $\cup_{i \in F} G_i$. We obviously have $|\mathcal{F} \wr \mathcal{G}| = \sum_{F \in \mathcal{F}} |\mathcal{G}|^{|F|}$. Thus, $|\mathcal{F} \wr \mathcal{G}| = |\mathcal{F}||\mathcal{G}|^k$ holds, provided that $\mathcal{F}$ is *k-uniform*, i.e., $|F| = k$ for every $F \in \mathcal{F}$.

Lemma 7. *If $\mathcal{F}$ and $\mathcal{G}$ are odd-sunflower-free families and $\mathcal{G}$ is an antichain, then $\mathcal{F} \wr \mathcal{G}$ is also odd-sunflower-free. Moreover, if $\mathcal{F}$ is also an antichain, then so is $\mathcal{F} \wr \mathcal{G}$.*

Remark. If $\mathcal{G}$ is not an antichain, then it may happen that $\mathcal{F} \wr \mathcal{G}$ contains an odd-sunflower, even if $\mathcal{F}$ was an antichain. For example, if $\mathcal{F} = \{\{1, 2\}\}$ and $\mathcal{G} = \{\{3\}, \{3, 4\}\}$, then the three sets $\{3_1, 3_2\}$,$\{3_1, 3_2, 4_1\}$,$\{3_1, 3_2, 4_2\}$ form an odd-sunflower.

Proof. The "moreover" part of the statement, according to which $\mathcal{F} \wr \mathcal{G}$ is an antichain, is trivial.

We need to show that in any family $\mathcal{H}$ of at least two sets from $\mathcal{F} \wr \mathcal{G}$, there is an element contained in a nonzero even number of sets from $\mathcal{H}$. Consider the multifamily $\mathcal{H}'$ of sets from $\mathcal{F}$, in which the multiplicity of a set F is as large, as many sets of the form $\cup_{i \in F} G_i$ belong to $\mathcal{H}$.

Since $\mathcal{F}$ is sunflower-free, there are two possibilities.

Case A: Some set in the multifamily $\mathcal{H}'$ has multiplicity greater than one.

In this case there exists an element $i \in F$ such that the multifamily of sets from $\mathcal{G}_i$, consisting of the intersections of the sets from $\mathcal{H}$ with M_i, has at least two *distinct* sets. Otherwise, the sets of $\mathcal{H}$ that correspond to the repeated set of $\mathcal{H}'$ would coincide, and $\mathcal{H}$ has no repeated sets. Applying Lemma 4 to the multifamily of sets from $\mathcal{G}_i$ for such an i, we find an element of M_i contained in a nonzero even number of sets from $\mathcal{H}$, as required.

Case B: The multifamily $\mathcal{H}'$ is not an odd-sunflower. That is, there exists an element $i \in \{1, \dots, n\}$ which is covered by an even number of sets in $\mathcal{H}'$.

This means that $\mathcal{H}$ has a nonzero even number of sets with nonempty intersections with M_i. Thus, applying Lemma 4 to the multifamily of sets from $\mathcal{G}_i$ formed by these nonempty intersections, again we find an element of M_i contained in a nonzero even number of sets from $\mathcal{H}$. This completes the proof. $\square$

Corollary 8. *Let $\mathcal{F}$ is a k-uniform odd-sunflower-free antichain on n elements. Then we have*

$$f_{oa}(nm) \geq |\mathcal{F}|(f_{oa}(m))^k.$$

In particular, $f_{oa}(nm) \geq n(f_{oa}(m))^{n-1}$, for odd n.

The second part of the corollary follows by choosing $\mathcal{F} = \mathcal{C}_n$, the family of all $(n-1)$-element subsets of $\{1, \dots, n\}$. These families have high uniformity, so they are natural candidates to increase the size of the family fast, because the uniformity k appears in the exponent in Corollary 8. As a simple, concrete application, consider the following.

Construction 2: The family $\mathcal{C}_9 \wr \mathcal{C}_3$ consists of $|\mathcal{C}_9||\mathcal{C}_3|^8 = 9 \cdot 3^8 = 3^{10}$ subsets of a $9 \cdot 3 = 27$-element set. Thus, we have

$$\mu_{odd} \geq |\mathcal{C}_9 \wr \mathcal{C}_3|^{1/27} = 3^{10/27} > 1.502144. \tag{2}$$

Lemma 7 implies that $\mathcal{C}_9 \wr \mathcal{C}_3$ contains no odd-sunflower. Thus, $f_{oa}(27) \geq 3^{10}$, and by Corollary 6, $\mu_{odd} \geq f_{oa}(27)^{1/27}$.

By Corollaries 6 and 8, we get $\mu_{odd} \geq f_{oa}(mn)^{1/mn} \geq (n|\mathcal{G}|^{n-1})^{1/mn}$. Here, to get the best bound, we need to choose n so as to maximize the last expression. Letting $n = x|\mathcal{G}|$, we obtain

$$\mu_{odd} \geq (n|\mathcal{G}|^{n-1})^{1/mn} = (x|\mathcal{G}|^n)^{1/mn} = |\mathcal{G}|^{1/m} x^{1/xm|\mathcal{G}|}.$$

Since $|\mathcal{G}|$ and m are independent of n, this is equivalent to maximizing $x^{1/x}$. A simple derivation shows that the optimal choice is $x = e$, so we need n to be the largest odd integer smaller than $e|\mathcal{G}|$, or the smallest odd integer greater than $e|\mathcal{G}|$. In the case of Construction 2, $3e$ is closest to 9.

The above reasoning also shows that any lower bound $|\mathcal{G}|^{1/m} \leq \mu_{odd}$ that comes from the brick construction using $\mathcal{G}$ as a brick, can be slightly improved by taking $\mathcal{C}_n \wr \mathcal{G}$ for some odd n close to $e|\mathcal{G}|$. For example, if $\mathcal{G} = \mathcal{C}_9 \wr \mathcal{C}_3$ is the 16-uniform family of 3^{10} sets on 27 elements obtained in Construction 2, then we can choose n to be $160511 \approx e3^{10}$.

Construction 3: The family $\mathcal{C}_{160511} \wr (\mathcal{C}_9 \wr \mathcal{C}_3)$ consists of $|\mathcal{C}_{160511}||\mathcal{C}_9 \wr \mathcal{C}_3|^{160510} = 160511 \cdot 3^{1605100}$ subsets of a $160511 \cdot 27 = 4333797$-element set. Thus, we have

$$\mu_{odd} \geq (160511 \cdot 3^{1605100})^{1/4333797} > 1.502148. \tag{3}$$

Of course, the improvement on the lower bound for μ_{odd} is extremely small as the families grow.

Concluding remarks

Here we studied the Erdős–Szemerédi-type sunflower problem for odd-sunflowers. We want to remark that our structural result is also true for (classical) sunflowers, using essentially the same proof. That is, if n is large enough, brick constructions will never be optimal. As far as we know, this result is new. The best currently known examples of Deuber *et al.* [7] use a combination of a brick construction and some other *ad hoc* tricks that do not work for odd-sunflowers.

What about the Erdős–Rado-type sunflower problem, i.e., what is the maximum possible size of an odd-sunflower-free k-uniform set system? We pose the following weakening of Erdős and Rado's conjecture:

Conjecture 9. *The maximum size of any odd-sunflower-free family of k-element sets is at most c^k, for a suitable constant $c > 0$.*

Note that the respective problem does not make sense for even-sunflowers, as any number of disjoint sets is even-sunflower-free.

Acknowledgment

We are grateful to Balázs Keszegh, and to the members of the Miklós Schweitzer Competition committee of 2022 for valuable discussions.

References

[1] H. L. Abbott, D. Hanson: On finite Δ-systems II, *Discrete Math.,* **17** (1977), 121–126.

[2] H. L. Abbott, D. Hanson, and N. Sauer: Intersection theorems for systems of sets, *J. Combin. Theory Ser. A,* **12** (3) (1972), 381–389.

[3] N. Alon and R. Holzman: Near-sunflowers and focal families, preprint, `https://arxiv.org/abs/2010.05992`.

[4] R. Alweiss, S. Lovett, K. Wu, and J. Zhang: Improved bounds for the sunflower lemma, *Ann. of Math. (2)* **194** (2021), no. 3, 795–815.

[5] L. Babai and P. Frankl: *Linear Algebra Methods in Combinatorics (Preliminary Version 2*, Dept. of Computer Science, The University of Chicago, 1992.

[6] E. Croot, V. Lev, and P. Pach: Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small, *Ann. of Math. (2)* **185** (2017), 331–337.

[7] W. A. Deuber, P. Erdős, D. S. Gunderson, A. V. Kostochka, and A. G. Meyer, Intersection statements for systems of sets, *J. Combin. Theory, Ser. A*, **79** (1997), 118–132.

[8] J. S. Ellenberg and D. Gijswijt: On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression, *Ann. of Math. (2)* **185** (2017), 339–343.

[9] P. Erdős, E. C. Milner, and R. Rado: Intersection theorems for systems of sets (III), *J. Austral. Math. Soc.* **18** (1974), 22–40.

[10] P. Erdős and R. Rado: Intersection theorems for systems of sets, *J. London Math. Soc., Second Series,* **35** (1) (1960), 85–90.

[11] P. Erdős and E. Szemerédi: Combinatorial properties of systems of sets, *J. Combin. Theory, Ser. A* **24** (1978), 308–313.

[12] P. Frankl: Extremalis halmazrendszerek (in Hungarian), kandidátusi értekezes, MTA, Budapest (1977).

[13] P. Frankl: Pseudo sunflowers, *European J. Combin.* **104** (2022), Paper No. 103553, 6 pp.

[14] P. Frankl and N. Tokushige: *Extremal Problems for Finite Sets, Student Math. Library,* vol. **86**, Amer. Math. Soc., Providence, R.I., 2018.

[15] B. Huppert: *Endliche Gruppen I*, Grundlehren der mathematischen Wissenschaften (GL, volume 134), Springer, 1967.

[16] A. V. Kostochka: Extremal problems on Δ-systems, in: *Numbers, Information and Complexity, Bielefeld, 1998,*, Kluwer Acad. Publ., Boston, MA (2000), 143-150

[17] `https://mathoverflow.net/questions/163689`

[18] E. Naslund and W. Sawin: Upper bounds for sunflower-free sets, *Forum Math. Sigma* **5** (2017), Paper No. e15, 10 pp.

[19] T. Tao: A symmetric formulation of the Croot-Lev-Pach-Ellenberg-Gijswijt capset bound, blog post, 2016, *http://terrytao.wordpress.com/2016/05/18/a*.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Upper bounds on Ramsey numbers for vector spaces over finite fields

(Extended abstract)

Bryce Frederickson * Liana Yepremyan †

Abstract

For $B \subseteq \mathbb{F}_q^m$, let $\mathrm{ex}_{\mathrm{aff}}(n, B)$ denote the maximum cardinality of a set $A \subseteq \mathbb{F}_q^n$ with no subset which is affinely isomorphic to B. Furstenberg and Katznelson proved that for any $B \subseteq \mathbb{F}_q^m$, $\mathrm{ex}_{\mathrm{aff}}(n, B) = o(q^n)$ as $n \to \infty$. For certain q and B, some more precise bounds are known. We connect some of these problems to certain Ramsey-type problems, and obtain some new bounds for the latter. For $s, t \geq 1$, let $R_q(s, t)$ denote the minimum n such that in every red-blue coloring of one-dimensional subspaces of $\mathbb{F}_q^n$, there is either a red s-dimensional subspace of $\mathbb{F}_q^n$ or a blue t-dimensional subspace of $\mathbb{F}_q^n$. The existence of these numbers is implied by the celebrated theorem of Graham, Leeb, Rothschild. We improve the best known upper bounds on $R_2(2, t)$, $R_3(2, t)$, $R_2(t, t)$, and $R_3(t, t)$.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-062`

1 Main Results

We consider bounds for Ramsey-type and Turán-type problems in the setting of vector-spaces over finite fields. In this paper, we use $\left[{V \atop t} \right]$ to denote the collection of all t-dimensional linear subspaces of a vector space V. The following theorem is a special case of a classical theorem of Graham, Leeb, and Rothschild [12], which proves the existence of the Ramsey numbers we consider.

*Department of Mathematics, Emory University, Atlanta, Georgia 30322. Email: `bfrede4@emory.edu`.
†Department of Mathematics, Emory University, Atlanta, Georgia 30322. Email: `lyeprem@emory.edu`.

Theorem 1 (Graham, Leeb, Rothschild)**.** *Let $\mathbb{F}_q$ be any finite field. For any positive integers $t_1, \ldots, t_k$, there exists a minimum $n =: R_q(t_1, \ldots, t_k)$ such that for every k-coloring $f : \left[\begin{smallmatrix}\mathbb{F}_q^n\\ 1\end{smallmatrix}\right] \to [k]$ of the 1-dimensional linear subspaces of $\mathbb{F}_q^n$, there exist $i \in [k]$ and a linear subspace $U \subseteq \mathbb{F}_q^n$ of dimension t_i, such that $\left[\begin{smallmatrix}U\\ 1\end{smallmatrix}\right]$ is monochromatic in color i.*

In the case $t_1 = \cdots = t_k = t$, we write $R_q(t_1, \ldots, t_k) = R_q(t; k)$. The bounds for $R_q(t_1, \ldots, t_k)$ implied by early proofs of Theorem 1 (see [12], [18]) are quite large due to repeated use of the Hales-Jewett Theorem [13]. In the case $q = 2$, the problem can be reduced to the disjoint unions problem for finite sets, considered by Taylor [19], which gives the following bound.

Theorem 2 (Taylor)**.** *The number $R_2(t; k)$ is at most a tower of height $2k(t-1)$ of the form*

$$R_2(t; k) \le k^{3^{k^{\cdot^{\cdot^{\cdot^{3}}}}}}.$$

For comparison, lower bounds attained from applying the techniques from [1] such as the Lovász Local Lemma to a uniform random coloring are only on the order of

$$R_2(t; k) = \Omega\left(\frac{2^t}{t}\log_2 k\right).$$

We improve the bound of Theorem 2 by removing the 3's from the tower.

Theorem 3. *For any t, k, $R_2(t; k)$ is at most a tower of height $k(t-1)$ of the form*

$$R_2(t; k) \le k^{k^{\cdot^{\cdot^{\cdot^{k}}}}}.$$

More recently, Nelson and Nomoto considered the off-diagonal version of this problem over $\mathbb{F}_2$ with two colors, and they proved the following bound.

Theorem 4 (Nelson, Nomoto)**.** *For every $t \ge 2$,*

$$R_2(2, t) \le (t+1)2^t.$$

Similar probabilistic arguments to those mentioned after Theorem 2 only give lower bounds linear in t for $R_2(2, t)$. Nelson and Nomoto asked if a subexponential upper bound is possible. While the answer to that question remains to be seen, we provide the following exponential improvement.

Theorem 5. *There exists a constant C such that for all $t \ge 2$,*

$$R_2(2, t) \le Ct6^{t/4}.$$

We obtain the following analogous results over $\mathbb{F}_3$, using the same methodology.

Theorem 6. *There exists an absolute polynomial $p(x)$ such that for any t, k, $R_3(t; k)$ is at most a tower of height $k(t-1)$ of the form*

$$R_3(t; k) \le p(k)^{p(k)^{\cdot^{\cdot^{\cdot^{p(k)}}}}}.$$

Theorem 7. *There exist constants C and A, with $A \approx 13.901$ such that for all $t \ge 2$,*

$$R_3(2, t) \le CtA^t.$$

2 Background and Methodology

Before we discuss the proofs of these results, we give a brief introduction to affine extremal numbers, which are our principal tool. We say that a subset $A \subseteq \mathbb{F}_q^n$ contains an *affine copy* of $B \subseteq \mathbb{F}_q^m$ if there is an injective affine map $f : \mathbb{F}_q^m \to \mathbb{F}_q^n$ with $f(B) \subseteq A$. If $\mathcal{B} = \{B_i\}_{i\in I}$ is a family of subsets $B_i \subseteq \mathbb{F}_q^{m_i}$, we say that A is *$\mathcal{B}$-free* if A has no affine copy of any B_i. The largest size $\mathrm{ex}_{\mathrm{aff}}(n, \mathcal{B})$ of a $\mathcal{B}$-free subset of $\mathbb{F}_q^n$ is called the nth *affine extremal number* of $\mathcal{B}$. If $\mathcal{B} = \{B\}$, we write $\mathrm{ex}_{\mathrm{aff}}(n, \{B\}) = \mathrm{ex}_{\mathrm{aff}}(n, B)$. Determining these affine extremal numbers dates back at least to the following theorem of Furstenberg and Katznelson [8].

Theorem 8 (Furstenberg, Katznelson)**.** *Let $\mathbb{F}_q$ be any finite field. For any positive integer t,*

$$\mathrm{ex}_{\mathrm{aff}}(n, \mathbb{F}_q^t) = o(q^n).$$

Since any $\mathcal{B}$-free set is $\mathbb{F}_q^t$-free for some t, Theorem 8 says that affine extremal numbers are always $o(q^n)$. Furstenberg and Katznelson went on to prove a density version of the Hales-Jewett Theorem [9], from which Theorem 8 is immediate. Alternative proofs of these results can be found in [16] and [15], respectively.

The projective version of this problem is even older, beginning with the following result of Bose and Burton [3].

Theorem 9 (Bose, Burton)**.** *Let $\mathbb{F}_q$ be a finite field, and let $t \geq 1$. Let A be a subset of $\left[\begin{smallmatrix}\mathbb{F}_q^n\\ 1\end{smallmatrix}\right]$ for which there is no linear t-dimensional subspace $U \subseteq \mathbb{F}_q^n$ with $\left[\begin{smallmatrix}U\\ 1\end{smallmatrix}\right] \subseteq A$. Then*

$$|A| \leq \frac{q^n - q^{n-t+1}}{q-1},$$

with equality if and only if $\left[\begin{smallmatrix}\mathbb{F}_q^n\\ 1\end{smallmatrix}\right] \setminus A = \left[\begin{smallmatrix}W\\ 1\end{smallmatrix}\right]$ for some linear $(n-t+1)$-dimensional linear subspace $W \subseteq \mathbb{F}_q^n$.

The problem of determining projective extremal numbers asymptotically for general projective configurations over $\mathbb{F}_q$ was solved by Geelen and Nelson [10], who proved a theorem analogous to the Erdős-Stone-Simonivits Theorem for graphs.

Returning to the affine context, it is unknown in general (see [11], Open Problem 32) whether the $o(q^n)$ term in Theorem 8 can be taken to be of the form $(q^{1-\varepsilon})^n$ for some $\varepsilon = \varepsilon(q, t) > 0$. However, for $q = 2$ and $q = 3$, we have the following respective results of Bonin and Qin [2], and of Fox and Pham [7].

Theorem 10 (Bonin, Qin)**.** *There exists an absolute constant c such that for every $t \geq 1$, every subset of $\mathbb{F}_2^n$ of size at least $(2^{1-c2^{-t}})^n$ contains an affine t-space.*

Later on, we will use a more precise bound implied by their argument, namely

$$\mathrm{ex}_{\mathrm{aff}}(n, \mathbb{F}_2^t) < 2^{(1-2^{1-t})n+1}.$$

Theorem 11 (Fox, Pham)**.** *There exist absolute constants c and C_0, with $C_0 \approx 13.901$ such that for every $t \geq 1$, every subset of $\mathbb{F}_3^n$ of size at least $\left(3^{1-cC_0^{-t}}\right)^n$ contains an affine t-space.*

The proof of Theorem 10 is entirely self-contained and is no more than a page. Theorem 11, on the other hand, is the culmination of several breakthroughs related to the Cap Set Problem, starting with the advances in polynomial methods from Croot, Lev, and Pach [4] and the subsequent proof of the Cap Set Theorem by Ellenberg and Gijswijt [5], which says that $\mathrm{ex}_{\mathrm{aff}}(n, \mathbb{F}_3^1) \leq (3^{1-\varepsilon})^n$ for some $\varepsilon > 0$. Fox and Lovász [6] then proved a supersaturation version of this result, from which Fox and Pham derived Theorem 11, which is a multidimensional extension of the Cap Set Theorem. It is unknown whether the constant C_0 given in the theorem is tight, as probabilistic lower bounds for $\mathrm{ex}_{\mathrm{aff}}(n, \mathbb{F}_3^t)$ are on the order of $\left(3^{1-3^{-(1+o(1))t}}\right)^n$ [7].

3 Proof Outlines

We now show that Theorems 3, 6, and 7 are easy consequences of the affine extremal results Theorem 10 and Theorem 11. To prove Theorem 5, we prove an additional extremal result over $\mathbb{F}_2$ by way of supersaturation and some observations about sumsets and products of affine structures.

To begin, we show how Theorem 3 follows from Bonin and Qin's result, Theorem 10.

Proof of Theorem 3. Since $R_2(1, \ldots, 1) = 1$, and we can reasonably define $R_2(t_1, \ldots, t_k) = 0$ if some $t_i = 0$, it suffices to show that

$$R_2(t_1, t_2, \ldots, t_k) \leq (\log_2 k)2^r,$$

where $r = \max_{i \leq k} R_2(t_1, \ldots, t_i - 1, \ldots, t_k)$. In this case, we get by induction that

$$R_2(t_1, \ldots, t_k) \leq (\log_2 k)k^{k^{\cdot^{\cdot^{\cdot^{k^2}}}}} \leq k^{k^{\cdot^{\cdot^{\cdot^{k^k}}}}},$$

where the height of the tower is $\sum_{i \leq k}(t_i - 1)$. Let $n = (\log_2 k)2^r$, and consider a k-coloring of $\left[\begin{smallmatrix}\mathbb{F}_2^n \\ 1\end{smallmatrix}\right]$, which we view as a k-coloring of $\mathbb{F}_2^n \setminus \{0\}$. Without loss of generality, assume that at least $2^n/k = 2^{n-\log_2 k}$ points are given color 1. By our choice of n and Theorem 10, we have

$$\mathrm{ex}_{\mathrm{aff}}(n, \mathbb{F}_2^r) < 2^{(1-2^{1-r})n+1} = 2^{n-2\log_2 k+1} \leq 2^{n-\log_2 k},$$

so there is an affine r-dimensional subspace A which is monochromatic in color 1. Note that $0 \notin A$ since 0 was not given a color. Let W be the translate of A containing 0, which is a linear r-space. Suppose that there is no linear t_i-space U_i with $U_i \setminus \{0\}$ monochromatic in color i for any $i \geq 2$. Then by our choice of r, there exists a linear $(t_1 - 1)$-space $U_1' \subseteq W$ with $U_1' \setminus \{0\}$ monochromatic in color 1. Let $u \in A$, and take $U_1 = \mathrm{span}\{U_1', u\}$, which is a linear t_1-space contained in $U_1' \cup A$, with $U_1 \setminus \{0\}$ monochromatic in color 1.

□

The proof of Theorem 6 is essentially the same, except that we use Theorem 11 in place of Theorem 10.

We now reformulate the off-diagonal Ramsey problem as an affine extremal problem. For a subset $A \subseteq \mathbb{F}_2^n$, let $\omega(A)$ be the maximum t such that $A \cup \{0\}$ contains a linear t-space. Define the *sumset* of A to be the set $A + A := \{x + y : x, y \in A\}$, and let $\mathcal{B}_t = \{B \subseteq \mathbb{F}_2^m : m \geq 1, \omega(B + B) \geq t\}$ for $t \geq 1$. Define $m(t)$ to be the minimum n such that $\mathrm{ex}_{\mathrm{aff}}(\mathcal{B}_t) < 2^{n-t+1}$. The following observation is implicit in the work of Nelson and Nomoto [14] on the structural characterization of claw-free binary matroids.

Lemma 12 (Nelson, Nomoto)**.** *For any $t \geq 2$, $R_2(2, t) \leq m(t)$.*

Nelson and Nomoto used the following result of Sanders [17] to prove Theorem 4.

Theorem 13 (Sanders)**.** *Let A be a subset of $\mathbb{F}_2^n$ of density $\alpha < 1/2$. Then*

$$\omega(A + A) \geq n - \left\lceil n / \log_2 \frac{2 - 2\alpha}{1 - 2\alpha} \right\rceil.$$

The proof of Theorem 4 from [14] is simply an application of Theorem 13 with $\alpha = 2^{1-t}$ and $n = (t+1)2^t$, noting that $n - \lceil n / \log_2 \frac{2-2\alpha}{1-2\alpha} \rceil \geq \alpha n/2 - 1 = t$ for this choice of parameters, so $m(t) \leq n$. By Lemma 12, $R_2(2, t) \leq n$ as well.

We observe that the same bound can be obtained by simply applying Theorem 10 instead of Sanders' result, noting that any set A which properly contains an affine $(t-1)$-space has $\omega(A + A) \geq t$, and hence

$$\mathrm{ex}_{\mathrm{aff}}(n, \mathcal{B}_t) \leq \mathrm{ex}_{\mathrm{aff}}(n, \mathbb{F}_2^{t-1}) < 2^{(1-2^{2-t})n+1}.$$

This implies by Lemma 12 that $R_2(2, t) \leq m(t) \leq t2^{t-2}$.

The same argument, together with Theorem 11, gives Theorem 7 for $\mathbb{R}_3(2, t)$. In place of the sumset $A + A$, we consider a set of the form

$$A_{\rightarrow} := \{d \in \mathbb{F}_3^n : \text{there exists } x \text{ such that } x + \lambda d \in A \text{ for all } \lambda \in \mathbb{F}_3\}.$$

To improve on this initial bound for $m(t)$, we consider additional affine structures beyond $\mathbb{F}_2^{t-1}$ that belong to the family $\mathcal{B}_t$. By taking products of smaller structures which have a certain supersaturation property, we construct a sequence $(B_t)_{t \geq 4}$ with $B_t \in \mathcal{B}_t$ and $\mathrm{ex}_{\mathrm{aff}}(n, B_t) < (2^{1-c6^{-t/4}})^n$ for some absolute constant c. This implies Theorem 5, as we have

$$R_2(2, t) \leq m(t) \leq \frac{1}{c}(t-1)6^{t/4}.$$

We leave out most of the details of our argument for the sake of brevity, but we outline our methods. We construct $B_t \in \mathcal{B}_t$ as follows. For $k \geq 2$, define $C_{2k} = \{e_1, \ldots, e_{2k-1}, \sum_{i=1}^{2k-1} e_i\} \subseteq \mathbb{F}_2^{2k-1}$, where e_i is the ith standard basis vector. We observe that $\omega(C_6 + C_6) = 4$. We further observe that for any $A \subseteq \mathbb{F}_2^n$ and $B \subseteq \mathbb{F}_2^m$, the Cartesian product $\{(x, y) \in \mathbb{F}_2^{n+m} : x \in A, y \in B\}$ satisfies

$$\omega((A \times B) + (A \times B)) = \omega(A + A) + \omega(B + B).$$

Thus taking B_t to be $C_6^{\lceil t/4 \rceil}$ gives $B_t \in \mathcal{B}_t$. We also obtain $\mathrm{ex}_{\mathrm{aff}}(n, B_t) < \left(2^{1-c6^{-t/4}}\right)^n$, as desired, via an iterative process that makes use of supersaturation of C_6, in the spirit of [7].

We believe our bounds on $\mathrm{ex}_{\mathrm{aff}}(n, \mathcal{B}_t)$, and hence on $m(t) \geq R_2(2, t)$, to be far from the truth. It remains an open problem to improve these bounds.

References

[1] N. Alon and J. H Spencer, *The probabilistic method*, John Wiley & Sons, 2016. ↑451

[2] J. E Bonin and H. Qin, *Size functions of subgeometry-closed classes of representable combinatorial geometries*, Discrete Mathematics **224** (2000), no. 1-3, 37–60. ↑452

[3] R. C. Bose and R. Burton, *A characterization of flat spaces in a finite geometry and the uniqueness of the Hamming and the MacDonald codes*, Journal of Combinatorial Theory **1** (1966), no. 1, 96–104. ↑452

[4] E. Croot, V. F Lev, and P. P. Pach, *Progression-free sets in are exponentially small*, Annals of Mathematics (2017), 331–337. ↑453

[5] J. S Ellenberg and D. Gijswijt, *On large subsets of with no three-term arithmetic progression*, Annals of Mathematics (2017), 339–343. ↑453

[6] J. Fox and L. Lovász, *A tight bound for Green's arithmetic triangle removal lemma in vector spaces*, Proceedings of the twenty-eighth annual acm-siam symposium on discrete algorithms, 2017, pp. 1612–1617. ↑453

[7] J. Fox and H. T. Pham, *Popular progression differences in vector spaces II* (2017). ↑452, 453, 455

[8] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for IP-systems and combinatorial theory*, Journal d'Analyse Mathématique **45** (1985), no. 1, 117–168. ↑452

[9] ———, *A density version of the Hales-Jewett theorem*, Journal d'Analyse Mathématique **57** (1991), no. 1, 64–119. ↑452

[10] J. Geelen and P. Nelson, *An analogue of the Erdős-Stone theorem for finite geometries*, Combinatorica **35** (2015), 209–214. ↑452

[11] D. Gijswijt, *Excluding affine configurations over a finite field*, arXiv preprint arXiv:2112.12620 (2021). ↑452

[12] R. L. Graham, K. Leeb, and B. L. Rothschild, *Ramsey's theorem for a class of categories*, Advances in Math. **8** (1972), 417–433. MR306010 ↑450, 451

[13] A. W. Hales and R. I. Jewett, *Regularity and positional games*, Trans. Amer. Math. Soc. **106** (1963), 222–229. MR143712 ↑451

[14] P. Nelson and K. Nomoto, *The structure of claw-free binary matroids*, Journal of Combinatorial Theory, Series B **150** (2021), 76–118. ↑454

[15] D. Polymath, *A new proof of the density Hales-Jewett theorem*, Annals of Mathematics (2012), 1283–1327. ↑452

[16] V Rödl, E Tengan, M Schacht, and N Tokushige, *Density theorems and extremal hypergraph problems*, Israel Journal of Mathematics **152** (2006), no. 1, 371–380. ↑452

[17] T. Sanders, *Green's sumset problem at density one half*, Acta Arithmetica **146** (2011), no. 1, 91–101 (eng). ↑454

[18] J. H. Spencer, *Ramsey's theorem for spaces*, Trans. Amer. Math. Soc. **249** (1979), no. 2, 363–371. MR525678 ↑451

[19] A. D Taylor, *Bounds for the disjoint unions theorem*, Journal of Combinatorial Theory, Series A **30** (1981), no. 3, 339–344. ↑451

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

A GENERAL BOUND FOR THE INDUCED POSET SATURATION PROBLEM

(EXTENDED ABSTRACT)

Andrea Freschi* Simón Piga† Maryam Sharifzadeh‡
Andrew Treglown§

Abstract

For a fixed poset P, a family $\mathcal{F}$ of subsets of $[n]$ is induced P-saturated if $\mathcal{F}$ does not contain an induced copy of P, but for every subset S of $[n]$ such that $S \notin \mathcal{F}$, then P is an induced subposet of $\mathcal{F} \cup \{S\}$. The size of the smallest such family $\mathcal{F}$ is denoted by $\mathrm{sat}^*(n, P)$. Keszegh, Lemons, Martin, Pálvölgyi and Patkós [Journal of Combinatorial Theory Series A, 2021] proved that there is a dichotomy of behaviour for this parameter: given any poset P, either $\mathrm{sat}^*(n, P) = O(1)$ or $\mathrm{sat}^*(n, P) \geq \log_2 n$. We improve this general result showing that either $\mathrm{sat}^*(n, P) = O(1)$ or $\mathrm{sat}^*(n, P) \geq 2\sqrt{n-2}$. Our proof makes use of a Turán-type result for digraphs.

Curiously, it remains open as to whether our result is essentially best possible or not. On the one hand, a conjecture of Ivan states that for the so-called diamond poset $\Diamond$ we have $\mathrm{sat}^*(n, \Diamond) = \Theta(\sqrt{n})$; so if true this conjecture implies our result is tight up to a multiplicative constant. On the other hand, a conjecture of Keszegh, Lemons, Martin, Pálvölgyi and Patkós states that given any poset P, either $\mathrm{sat}^*(n, P) = O(1)$ or $\mathrm{sat}^*(n, P) \geq n + 1$. We prove that this latter conjecture is true for a certain class of posets P.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-063`

*University of Birmingham, United Kingdom. E-mail: `axf079@bham.ac.uk`.
†University of Birmingham, United Kingdom. E-mail: `s.piga@bham.ac.uk`. Research supported by EPSRC grant EP/V002279/1.
‡Umeå Universitet, Sweden. E-mail: `maryam.sharifzadeh@umu.se`.
§University of Birmingham, United Kingdom. E-mail: `a.c.treglown@bham.ac.uk`. Research supported by EPSRC grant EP/V002279/1.

1 Introduction

Saturation problems have been well studied in graph theory. A graph G is H*-saturated* if it does not contain a copy of the graph H, but adding any edge to G from its complement creates a copy of H. Turán's celebrated theorem [15] can be stated in the language of saturation: it determines the maximum number of edges in a K_r-saturated n-vertex graph. In contrast, Erdős, Hajnal and Moon [5] determined the minimum number of edges in a K_r-saturated n-vertex graph; see the survey [3] for further results in this direction.

In recent years there has been an emphasis on developing the theory of saturation for posets. Turán-type problems have been extensively studied in this setting (see, e.g., the survey [9]). In this paper we are interested in *minimum* saturation questions à la Erdős–Hajnal–Moon. In particular, we consider *induced saturation* problems.

All posets we consider will be (implicitly) viewed as finite collections of finite subsets of $\mathbb{N}$. In particular, we say that P is a *poset on* $[p] := \{1, 2, \dots, p\}$ if P consists of subsets of $[p]$. Let P, Q be posets. A *poset homomorphism* from P to Q is a function $\phi : P \to Q$ such that for every $A, B \in P$, if $A \subseteq B$ then $\phi(A) \subseteq \phi(B)$. We say that P is a *subposet* of Q if there is an injective poset homomorphism from P to Q; otherwise, Q is said to be P*-free.* Further we say P is an *induced subposet* of Q if there is an injective poset homomorphism ϕ from P to Q such that for every $A, B \in P$, $\phi(A) \subseteq \phi(B)$ if and only if $A \subseteq B$; otherwise, Q is said to be *induced* P*-free.*

For a fixed poset P, we say that a family $\mathcal{F} \subseteq 2^{[n]}$ of subsets of $[n]$ is P*-saturated* if $\mathcal{F}$ is P-free, but for every subset S of $[n]$ such that $S \notin \mathcal{F}$, then P is a subposet of $\mathcal{F} \cup \{S\}$. A family $\mathcal{F} \subseteq 2^{[n]}$ of subsets of $[n]$ is *induced* P*-saturated* if $\mathcal{F}$ is induced P-free, but for every subset S of $[n]$ such that $S \notin \mathcal{F}$, then P is an induced subposet of $\mathcal{F} \cup \{S\}$.

The study of minimum saturated posets was initiated by Gerbner, Keszegh, Lemons, Palmer, Pálvölgyi and Patkós [8] in 2013. In their work the relevant parameter is $\mathrm{sat}(n, P)$, which is defined to be the size of the smallest P-saturated family of subsets of $[n]$. See, e.g., [8, 12, 14] for various results on $\mathrm{sat}(n, P)$.

The induced analogue of $\mathrm{sat}(n, P)$ – denoted by $\mathrm{sat}^*(n, P)$ – was first considered by Ferrara, Kay, Kramer, Martin, Reiniger, Smith and Sullivan [6]. Thus, $\mathrm{sat}^*(n, P)$ is defined to be the size of the smallest induced P-saturated family of subsets of $[n]$. The following result from [12] (and implicit in [6]) shows that the parameter $\mathrm{sat}^*(n, P)$ has a dichotomy of behaviour.

Theorem 1.1. *[6, 12] For any poset P, either there exists a constant K_P with $sat^*(n, P) \leq K_P$ or $sat^*(n, P) \geq \log_2 n$, for all $n \in \mathbb{N}$.*

Probably the most important open problem in the area is to obtain a tight version of Theorem 1.1; that is, to replace the $\log_2 n$ in Theorem 1.1 with a term that is as large as possible. In fact, Keszegh, Lemons, Martin, Pálvölgyi and Patkós [12] made the following conjecture in this direction.

Conjecture 1.2. *[12] For any poset P, either there exists a constant K_P with $sat^*(n, P) \leq K_P$ or $sat^*(n, P) \geq n + 1$, for all $n \in \mathbb{N}$.*

Note that the lower bound of $n+1$ is rather natural here. For example, it is the size of the largest *chain* in $2^{[n]}$ as well as the smallest possible size of the union of two consecutive 'layers' in $2^{[n]}$, namely the layer containing $[n]$ and the layer containing all subsets of $[n]$ of size exactly $n-1$. Furthermore, such structures form minimum induced saturated families for the so-called fork poset $\vee$, i.e., $\mathrm{sat}^*(n, \vee) = n+1$ [6]; so the lower bound in Conjecture 1.2 cannot be increased. There are also no known examples of posets P for which $\mathrm{sat}^*(n, P) = \omega(n)$.

In contrast, Ivan [11, Section 3] presented evidence that led her to conjecture a rather different picture for the *diamond* poset $\Diamond$ (see Figure 1 for the Hasse diagram of $\Diamond$).

Conjecture 1.3. *[11] sat$^*(n, \Diamond) = \Theta(\sqrt{n})$.*

Our main result is the following improvement of Theorem 1.1.

Theorem 1.4. *For any poset P, either there exists a constant K_P with sat$^*(n, P) \leq K_P$ or sat$^*(n, \mathcal{P}) \geq 2\sqrt{n-2}$, for all $n \in \mathbb{N}$.*

Thus, if Conjecture 1.3 is true, the lower bound in Theorem 1.4 would be tight up to a multiplicative constant.

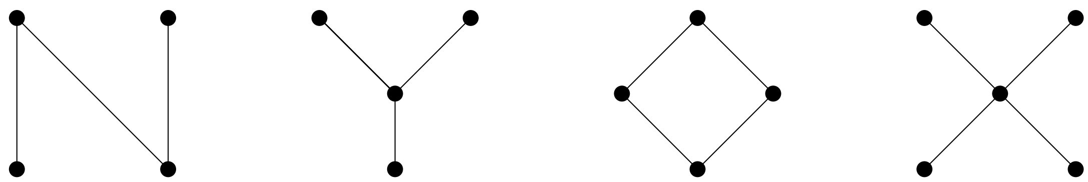


Figure 1: Hasse diagrams for the posets N, Y, $\Diamond$ and X.

On the other hand, we prove that Conjecture 1.2 does hold for a class of posets (that does not include $\Diamond$). Given $p \in \mathbb{N}$ and a poset P on $[p]$ we define the *dual* $\overline{P}$ of P as $\overline{P} := \{[p] \backslash F \colon F \in P\}$. We say a poset P *has legs* if there are distinct elements $L_1, L_2, H \in P$ such that L_1, L_2 are incomparable, $L_1, L_2 \subseteq H$ and for any other element $A \in P \setminus \{L_1, L_2, H\}$ we have $A \supseteq H$. The elements L_1 and L_2 are called *legs* and H is called a *hip*.

Theorem 1.5. *Let P be a poset with legs and $n \geq 3$. Then sat$^*(n, P) \geq n+1$. Moreover, if both P and $\overline{P}$ have legs, then sat$^*(n, P) \geq 2n+2$.*

Our results still leave both Conjecture 1.2 and Conjecture 1.3 open, and it is unclear to us which of these conjectures is true. However, if Conjecture 1.3 is true we believe it highly likely that there will be other posets P for which $\mathrm{sat}^*(n, P) = \Theta(\sqrt{n})$.

It is also natural to seek exact results on $\mathrm{sat}^*(n, P)$. However, despite there already being several papers concerning $\mathrm{sat}^*(n, P)$ [1, 4, 6, 10, 11, 12, 13], there are relatively few posets P for which $\mathrm{sat}^*(n, P)$ is known precisely (see Table 1 in [12] for a summary of most of the known results). Our next result extends this limited pool of posets, determining $\mathrm{sat}^*(n, X)$ and $\mathrm{sat}^*(n, Y)$ (see Figure 1 for the Hasse diagrams of X and Y).

Theorem 1.6. *Given any $n \in \mathbb{N}$ with $n \geq 3$,*

(i) $sat^*(n, Y) = n + 2$ *and*

(ii) $sat^*(n, X) = 2n + 2$.

Note that Theorem 1.6(*ii*) easily follows via Theorem 1.5 and an extremal construction. An application of Theorem 1.5 to Y only yields that $\mathrm{sat}^*(n, Y) \geq n + 1$, so we require an extra idea to obtain Theorem 1.6(*i*).

It is natural to consider induced saturation problems for families of posets. Given a family of posets $\mathcal{P}$, we say that $\mathcal{F} \subseteq 2^{[n]}$ is *induced $\mathcal{P}$-saturated* if $\mathcal{F}$ contains no induced copy of any poset $P \in \mathcal{P}$ and for every $S \in 2^{[n]} \setminus \mathcal{F}$ there exists an induced copy of some poset $P \in \mathcal{P}$ in $\mathcal{F} \cup \{S\}$. We denote the size of the smallest such family by $\mathrm{sat}^*(n, \mathcal{P})$. By following the proof of Theorem 1.4 precisely, one obtains the following result.

Theorem 1.7. *For any family of posets $\mathcal{P}$, either there exists a constant $K_{\mathcal{P}}$ with $sat^*(n, \mathcal{P}) \leq K_{\mathcal{P}}$ or $sat^*(n, \mathcal{P}) \geq 2\sqrt{n-2}$, for all $n \in \mathbb{N}$.*

In light of Theorem 1.7 it is natural to ask whether an analogue of Conjecture 1.2 is true in this more general setting, or whether (for example) the lower bound on $\mathrm{sat}^*(n, \mathcal{P})$ in Theorem 1.7 is best possible up to a multiplicative constant.

The proofs of Theorems 1.4–1.7 appear in [7]. In the next section we describe how we make use of a Turán-type result for digraphs in the proof of Theorem 1.4.

2 A connection to a Turán problem for digraphs

In [13] a trick was introduced which can be used to prove lower bounds on $\mathrm{sat}^*(n, P)$ for some posets P. The idea is to construct a certain auxiliary digraph D whose vertex set consists of the elements in an induced P-saturated family $\mathcal{F}$; one then argues that how this digraph is defined forces D to contain many edges, which in turn forces a bound on the size of the vertex set of D (i.e., lower bounds $|\mathcal{F}|$). This trick has been used to prove that $\mathrm{sat}^*(n, \Diamond) \geq \sqrt{n}$ [13, Theorem 6] and $\mathrm{sat}^*(n, N) \geq \sqrt{n}$ [10, Proposition 4] (see Figure 1 for the Hasse diagram of N).

Our proof of Theorem 1.4 utilises a variant of this digraph trick. In particular, by introducing an appropriate modification to the auxiliary digraph D used in [13], we are able to deduce certain Turán-type properties of D. Turán problems in digraphs are classical in extremal combinatorics and their study can be traced back to the work of Brown and Harary [2]. In [7] we prove a Turán-type result concerning *transitive cycles*, stated as Theorem 2.1 below.

Given $k \geq 3$, the *transitive cycle on k vertices* $\overrightarrow{TC}_k$ is a digraph with vertex set $[k]$ and every directed edge from i to $i+1$ for every $i \in [k-1]$, as well as the directed edge from 1 to k. We establish an upper bound on the number of edges of a digraph not containing any transitive cycle.

Theorem 2.1. *Let $n \in \mathbb{N}$ and let D be a digraph on n vertices. If D is $\overrightarrow{TC}_k$-free for all $k \geq 3$, then*

$$e(D) \leq \left\lfloor \frac{n^2}{4} \right\rfloor + 2.$$

Note that the bound in Theorem 2.1 is best possible up to an additive constant. Indeed, consider the n-vertex digraph D with vertex classes A, B of size $\lfloor n/2 \rfloor$ and $\lceil n/2 \rceil$ respectively and all possible directed edges from A to B. So D has $\lfloor n^2/4 \rfloor$ edges and contains no transitive cycle.

Data availability statement. A full paper containing the proofs of our results can be found on arXiv [7].

References

[1] Paul Bastide, Carla Groenland, Hugo Jacob, and Tom Johnston. Exact antichain saturation numbers via a generalisation of a result of Lehman–Ron. *arXiv:2207.07391*, 2022.

[2] William G. Brown and Frank Harary. Extremal digraphs. *Colloquia Mathematica Societatis János Bolyai*, 4:135–198, 1969.

[3] Bryan L. Currie, Jill R. Faudree, Ralph J. Faudree, and John R. Schmitt. A survey of minimum saturated graphs. *The Electronic Journal of Combinatorics*, DS19, 2011.

[4] Irina Đanković and Maria-Romina Ivan. Saturation for small antichains. *The Electronic Journal of Combinatorics*, 30(1):P1.3, 2023.

[5] Paul Erdős, András Hajnal, and John W. Moon. A problem in graph theory. *The American Mathematical Monthly*, 71(10):1107–1110, 1964.

[6] Michael Ferrara, Bill Kay, Lucas Kramer, Ryan R. Martin, Benjamin Reiniger, Heather C. Smith, and Eric Sullivan. The saturation number of induced subposets of the Boolean lattice. *Discrete Mathematics*, 340(10):2479–2487, 2017.

[7] Andrea Freschi, Simón Piga, Maryam Sharifzadeh, and Andrew Treglown. The induced saturation problem for posets. *arXiv:2207.03974*, 2022.

[8] Dániel Gerbner, Balázs Keszegh, Nathan Lemons, Cory Palmer, Dömötör Pálvölgyi, and Balázs Patkós. Saturating Sperner families. *Graphs and Combinatorics*, 29(5):1355–1364, 2013.

[9] Jerrold R. Griggs and Wei-Tian Li. Progress on poset-free families of subsets. *Recent Trends in Combinatorics*, The IMA Volumes in Mathematics and its Applications, vol. 159, 317–338. Springer, New York, 2016.

[10] Maria-Romina Ivan. Saturation for the butterfly poset. *Mathematika*, 66(3):806–817, 2020.

[11] Maria-Romina Ivan. Minimal diamond-saturated families. *Contemporary Mathematics*, 3(2):81–88, 2022.

[12] Balázs Keszegh, Nathan Lemons, Ryan R. Martin, Dömötör Pálvölgyi, and Balázs Patkós. Induced and non-induced poset saturation problems. *Journal of Combinatorial Theory, Series A*, 184:105497, 2021.

[13] Ryan R. Martin, Heather C. Smith, and Shanise Walker. Improved bounds for induced poset saturation. *The Electronic Journal of Combinatorics*, 27(2):P2.31, 2020.

[14] Natasha Morrison, Jonathan A. Noel, and Alex Scott. On saturated k-Sperner systems. *The Electronic Journal of Combinatorics*, 21(3):P3.22, 2014.

[15] Paul Turán. On an extremal problem in graph theory. *Matematikai és Fizikai Lapok*, 48:436–452, 1941.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Extremal number of graphs from geometric shapes

(Extended abstract)

Jun Gao* Oliver Janzer† Hong Liu* Zixiang Xu*

Abstract

We study the Turán problem for highly symmetric bipartite graphs arising from geometric shapes and periodic tilings commonly found in nature.

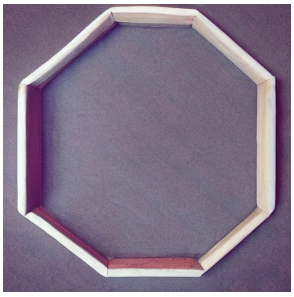
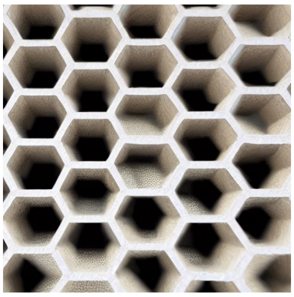
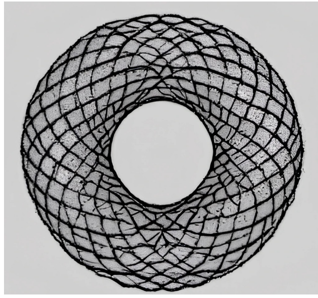
1. The prism $C_{2\ell}^{\square} := C_{2\ell} \square K_2$ is the graph consisting of two vertex disjoint 2ℓ-cycles and a matching pairing the corresponding vertices of these two cycles. We show that for every $\ell \geqslant 4$, $\mathrm{ex}(n, C_{2\ell}^{\square}) = \Theta(n^{3/2})$. This resolves a conjecture of He, Li and Feng.
2. The hexagonal tiling in honeycomb is one of the most natural structures in the real world. We show that the extremal number of honeycomb graphs has the same order of magnitude as their basic building unit 6-cycles.
3. We also consider bipartite graphs from quadrangulations of the cylinder and the torus. We prove near optimal bounds for both configurations. In particular, our method gives a very short proof of a tight upper bound for the extremal number of the 2-dimensional grid, improving a recent result of Bradač, Janzer, Sudakov and Tomon.

Our proofs mix several ideas, including shifting embedding schemes, weighted homomorphism and subgraph counts and asymmetric dependent random choice.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-064`

*Extremal Combinatorics and Probability Group (ECOPRO), Institute for Basic Science (IBS), Daejeon, South Korea. Emails: {jungao, hongliu, zixiangxu}@ibs.re.kr. Supported by IBS-R029-C4.

†Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, United Kingdom. Email:oj224@cam.ac.uk.

1 Introduction

The Turán problem, one of the most central topics in extremal combinatorics, is concerned with determining the maximum density of graphs without containing a given graph as a subgraph. Formally, for a graph F, the *extremal number of* F, denoted by $\mathrm{ex}(n, F)$, is the maximum number of edges in an n-vertex graph not containing F as a subgraph. The celebrated Erdős-Stone-Simonovits Theorem [9, 12] asymptotically solves the problem when $\chi(F) \geqslant 3$. However, for bipartite graphs, not even the order of magnitude is known in general. Turán [21] in 1941 proposed the study of the five graphs from platonic solids, and his result covers the tetrahedron graph K_4. The problem of octahedron, dodecahedron and icosahedron graphs were later resolved by Erdős and Simonovits [11] and by Simonovits [19, 20] respectively; while the innocent looking cube graph remains elusive. Two basic classes of bipartite graphs with high symmetry are even cycles and complete bipartite graphs; both of them have been widely studied for several decades [2, 3, 5, 6, 8, 15, 17, 22]. For more on the bipartite Turán problem, we refer the reader to the comprehensive survey of Füredi and Simonovits [13].

In this paper, we continue this line of study and determine the order of magnitude of the extremal number for several highly symmetric bipartite graphs stemming from certain geometric shapes and periodic tilings, including the prism, the grid, the honeycomb and certain quadrangulations of the cylinder and the torus.

1.1 The prisms

The 2ℓ-*prism* $C_{2\ell}^{\square} := C_{2\ell} \square K_2$ is the Cartesian product of 2ℓ-cycle with an edge, consisting of two vertex disjoint $C_{2\ell}$ and a matching joining the corresponding vertices on these two cycles. As $C_{2\ell}^{\square}$ contains many 4-cycles, we have a lower bound $\mathrm{ex}(n, C_{2\ell}^{\square}) \geqslant \mathrm{ex}(n, C_4) = \Omega(n^{3/2})$. Note that $C_4^{\square}$ is the notorious cube graph, for which the best known bounds are $\Omega(n^{3/2}) \leqslant \mathrm{ex}(n, C_4^{\square}) \leqslant O(n^{8/5})$ [10, 18]. Studying the 2ℓ-prism $C_{2\ell}^{\square}$ could shed some light on the cube problem. An upper bound $\mathrm{ex}(n, C_{2\ell}^{\square}) = O(n^{5/3})$ can be easily obtained via the celebrated dependent random choice method [1].

Very recently, He, Li and Feng [14] studied the odd prisms, determined $\mathrm{ex}(n, C_{2k+1}^{\square})$ for any $k \geqslant 1$ for large n and characterized the extremal graphs. They proposed the following conjecture to break the 5/3 barrier for 2ℓ-prism.

Conjecture 1 ([14]). *For every $\ell \geqslant 2$, there exists $c = c(\ell) > 0$ such that* $\mathrm{ex}(n, C_{2\ell}^{\square}) = O(n^{5/3-c})$.

Our first result provides an optimal upper bound for $C_{2\ell}^{\square}$ for every $\ell \geqslant 4$.

Theorem 1.1. *For any integer $\ell \geqslant 4$, we have*

$$\mathrm{ex}(n, C_{2\ell}^{\square}) = \Theta_\ell(n^{3/2}).$$

We remark that larger prisms are easier to handle. We can provide a shorter and different proof of $\mathrm{ex}(n, C_{2\ell}^{\square}) = O_\ell(n^{3/2})$ for $\ell \geqslant 7$, which can also be used to show that $\mathrm{ex}(n, C_6^{\square}) = O(n^{21/13}(\log n)^{24/13})$. This, together with the known bound for the cube and Theorem 1.1, proves Conjecture 1.

It is worth mentioning a closely related conjecture of Erdős. A graph is *r-degenerate* if each of its subgraphs has minimum degree at most r. Erdős [7] conjectured that for a bipartite H, $\mathrm{ex}(n, H) = O(n^{3/2})$ if and only if H is 2-degenerate. This conjecture was recently disproved by Janzer [16], who constructed, for each $\varepsilon > 0$, a 3-regular bipartite graph H with girth 6 such that $\mathrm{ex}(n, H) = O(n^{4/3+\varepsilon})$. Theorem 1.1 provides a family of 3-regular *girth-4* counterexamples.

1.2 The honeycomb

The hexagonal tiling in honeycomb is one of the most common geometric structures, appearing in nature in many crystals. It is also the densest way to pack circles in the plane. As the honeycomb graph H of any size contains C_6 as a subgraph, we have a lower bound $\mathrm{ex}(n, H) \geqslant \mathrm{ex}(n, C_6) = \Omega(n^{4/3})$.

Our second result is a matching upper bound $O(n^{4/3})$, showing that the hexagonal tiling appears soon after the appearance of a single hexagon. In particular, we consider the following graph $H_{k,\ell}$ (see Figure 1), which contains any (finite truncation of a) honeycomb graph as a subgraph when k and ℓ are sufficiently large.

Definition. For an odd integer $k \geqslant 1$ and even integer $\ell \geqslant 2$, let $H_{k,\ell}$ be the graph with vertex set $V(H_{k,\ell}) = \{x_{i,j} : 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant \ell\}$, where $x_{k,1} = x_{k,3} = \cdots = x_{k,\ell-1} = u$ and $x_{1,2} = x_{1,4} = \cdots = x_{1,\ell} = v$ (but all the other vertices are distinct) and edge set

$$\begin{aligned} E(H_{k,\ell}) = \{x_{i,j}x_{i,j+1} : 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant \ell - 1\} &\cup \{x_{2i-1,j}x_{2i,j} : 1 \leqslant i \leqslant k/2, 1 \leqslant j \leqslant \ell,\ j \text{ is odd}\} \\ &\cup \{x_{2i,j}x_{2i+1,j} : 1 \leqslant i \leqslant k/2, 1 \leqslant j \leqslant \ell,\ j \text{ is even}\}. \end{aligned}$$

Theorem 1.2. *For positive odd integers $k \geqslant 1$ and $\ell \geqslant 2$,*

$$\mathrm{ex}(n, H_{k,\ell}) = \Theta_{k,\ell}(n^{4/3}).$$

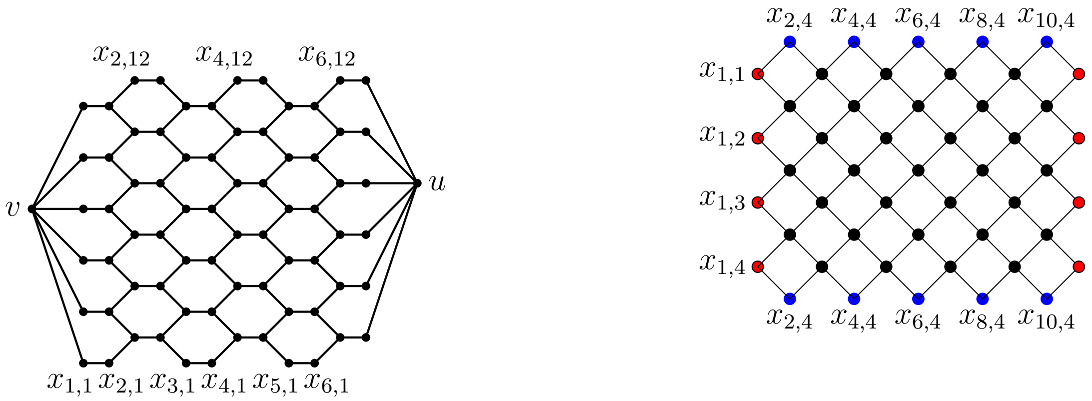


Figure 1: The first graph is $H_{7,12}$. In the second graph, identifying the blue vertices (in the same column) yields a copy of $P_{11,4}$; if additionally the red vertices (in the same row) are identified, then we obtain a copy of $T_{10,4}$.

1.3 The grid

We will also give an improved bound for the extremal number of the grid. For a positive integer t, $F_{t,t}$ is the graph with vertex set $[t] \times [t]$ in which two vertices are joined by an edge if they differ in exactly one coordinate and in that coordinate they differ by one. Bradač, Janzer, Sudakov and Tomon [4] determined the extremal number of $F_{t,t}$ up to a multiplicative constant which depends on t, showing that for any $t \geqslant 2$,

$$\Omega(t^{1/2}n^{3/2}) \leqslant \mathrm{ex}(n, F_{t,t}) \leqslant e^{O(t^5)}n^{3/2}.$$

They have asked to determine the correct dependence on t. We make substantial progress on this question by giving a very short proof of the following bound, which shows that the dependence on t is polynomial.

Theorem 1.3. *For any positive integer t, if n is sufficiently large in terms of t, then*

$$\mathrm{ex}(n, F_{t,t}) \leqslant 5t^{3/2}n^{3/2}.$$

It would be interesting to determine the correct power of t in $\mathrm{ex}(n, F_{t,t})$.

1.4 Quadrangulations of cylinder and torus

Next, we consider certain quadrangulations of the cylinder and the torus.

Definition (Quadrangulation of a cylinder)**.** For integers $k, \ell \geqslant 2$, let $P_{k,\ell}$ be the graph with vertex set $V(P_{k,\ell}) = \{x_{i,j} : 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant \ell\}$, and edge set

$$\begin{aligned} E(P_{k,\ell}) = \{x_{i,j}x_{i+1,j} : 1 \leqslant i \leqslant k-1, 1 \leqslant j \leqslant \ell\} &\cup \{x_{i,j+1}x_{i+1,j} : 1 \leqslant i \leqslant k-1, 1 \leqslant j \leqslant \ell,\ i \text{ is odd}\} \\ &\cup \{x_{i,j}x_{i+1,j+1} : 1 \leqslant i \leqslant k-1, 1 \leqslant j \leqslant \ell,\ i \text{ is even}\}, \end{aligned}$$

where $x_{i,\ell+1} = x_{i,1}$ for all $i \in [k]$.

Clearly, the extremal number of such a quadrangulated cylinder is at least that of the 4-cycle. Our next result infers that in fact they are of the same order of magnitude.

Theorem 1.4. *Let k and ℓ be positive integers. Then we have*

$$\mathrm{ex}(n, P_{k,\ell}) = \Theta_{k,\ell}(n^{3/2}).$$

If k is even and we glue the two sides of the cylinder $P_{k+1,\ell}$, then we obtain a torus, see Figure 1.

Definition (Quadrangulation of a torus)**.** For an even integer $k \geqslant 4$ and integer $\ell \geqslant 2$, let $T_{k,\ell}$ be the graph with vertex set $V(T_{k,\ell}) = \{x_{i,j} : 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant \ell\}$, and edge set

$$\begin{aligned} E(T_{k,\ell}) = \{x_{i,j}x_{i+1,j} : 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant \ell\} &\cup \{x_{i,j+1}x_{i+1,j} : 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant \ell,\ i \text{ is odd}\} \\ &\cup \{x_{i,j}x_{i+1,j+1} : 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant \ell,\ i \text{ is even}\}, \end{aligned}$$

where $x_{k+1,j} = x_{1,j}$ for all $j \in [\ell]$ and $x_{i,\ell+1} = x_{i,1}$ for all $i \in [k]$.

For the quadrangulated torus, we provide a general upper bound as follows.

Theorem 1.5. *For an even integer $k \geqslant 4$ and an integer $\ell \geqslant 2$, we have*

$$\mathrm{ex}(n, T_{k,\ell}) = O_{k,\ell}(n^{\frac{3}{2}+\frac{\ell}{k}}(\log n)^2).$$

Thus, when k is sufficiently large compared to ℓ, the exponent can be arbitrarily close to 3/2. On the other hand, the exponent is always strictly greater than 3/2 as the probabilistic deletion method yields the lower bound $\mathrm{ex}(n, T_{k,\ell}) = \Omega_{k,\ell}(n^{\frac{3}{2}+\frac{3}{4k\ell-2}})$.

2 Ideas of proofs

In this section, we briefly discuss some key ideas in our proofs.

2.1 Shifting embedding schemes: Grid, quadrangulated cylinder, torus and honeycomb

For Theorems 1.2, 1.3 and 1.4, our embedding strategy is based on the observation that, if we can find a large collection of paths or cycles with a certain nice property, then we can repeatedly replace the vertices (or edges) of the chosen paths or cycles with vertices (or edges) from a new one in the collection to build the desired tilings. Formally, the definition of an α-rich collection of paths in a graph is as follows.

Definition 2.1. Let $\alpha > 0$ and $k \in \mathbb{N}$. We say that a collection $\mathcal{P}$ of (labelled) paths P_k is *α-rich* if for any member $x_1x_2 \cdots x_k \in \mathcal{P}$ and any $2 \leqslant i \leqslant k-1$, there exist at least α distinct vertices x_i' such that $x_1x_2 \cdots x_{i-1}x_i'x_{i+1} \cdots x_k \in \mathcal{P}$.

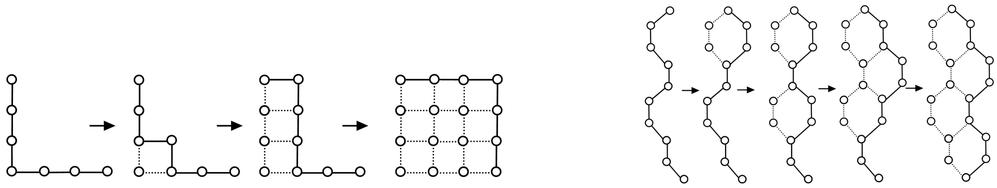


Figure 2: The process to build grids and honeycomb graphs.

Finding a $t \times t$ grid then boils down to constructing an α-rich collection of paths of length $2t-2$ with sufficiently large α; see the left side of Figure 2 for an illustration. In order to find the desired quadrangulations of the cylinder and the torus, we use an analogous definition of rich *cycles*.

To deal with honeycomb graphs, we introduce the following definition.

Definition 2.2. Let $\alpha > 0$ and $k \in \mathbb{N}$. A collection $\mathcal{P}$ of paths P_k is *α-good* if the following holds. For any $x_1x_2\cdots x_k \in \mathcal{P}$ and $2 \leqslant i \leqslant k-2$, there are at least α pairwise disjoint edges $x'_i x'_{i+1}$ such that $x_1x_2\cdots x_{i-1}x'_i x'_{i+1}x_{i+2}\cdots x_k \in \mathcal{P}$.

Finding a honeycomb graph boils down to constructing an α-good collection of paths for a sufficiently large α; see the right side of Figure 2 for an illustration.

While it is not too hard to find a collection of rich paths and even cycles via supersaturation, it is a lot more challenging to construct a collection of good paths. In order to accomplish the latter, rather than doing a direct counting using supersaturation, we carry out a weighted count.

2.2 Weighted count of homomorphisms for Theorem 1.1

In this subsection, we give a brief outline of the proof of Theorem 1.1. Let us call an n-vertex graph H with average degree d *clean* if for any $uv \in E(H)$, u has at least $d/16$ neighbours w in H such that $d_H(v,w) \geqslant \frac{d^2}{128n}$. It can be shown that any graph with average degree at least $2d$ contains a clean subgraph with average degree at least d.

Let G be a graph of average degree d and let distinct vertices x_i, y_i for $0 \leqslant i \leqslant \ell$ form a copy of $P^{\square}_{\ell+1} := K_2 \square P_{\ell+1}$, where $x_iy_i \in E(G)$ for every i and $x_{i-1}x_i, y_{i-1}y_i \in E(G)$ for every $1 \leqslant i \leqslant \ell$. Now the *weight* of this copy is defined to be $1/\prod_{i=1}^{\ell} \max(d_G(x_{i-1},y_i), \frac{d^2}{n})$. For distinct vertices u,v,w,z, we call the 4-tuple (u,v,w,z) *rich* if $uv, wz \in E(G)$, and moreover there are at least 4ℓ pairwise vertex-disjoint edges $xy \in E(G)$ such that $ux, xw, vy, yz \in E(G)$. We say that vertices x_i, y_i (for $0 \leqslant i \leqslant \ell$) form a *nice* copy of $P^{\square}_{\ell+1}$ if they form a copy of $P^{\square}_{\ell+1}$, for every $1 \leqslant i \leqslant \ell$ the codegrees satisfy $d(x_{i-1},y_i), d(x_i,y_{i-1}) \leqslant C_0 d^{1/2}$ (for some suitably defined constant C_0), and for every $2 \leqslant i \leqslant \ell$, the 4-tuple $(x_{i-2}, y_{i-2}, x_i, y_i)$ is not rich. We also say that vertices x_i, y_i, x'_i, y'_i (for $0 \leqslant i \leqslant \ell$) form a *nice* homomorphic copy of $C^{\square}_{2\ell}$ if $x_0 = x'_0, y_0 = y'_0, x_\ell = x'_\ell, y_\ell = y'_\ell$, both $\{x_i, y_i : 0 \leqslant i \leqslant \ell\}$ and $\{x'_i, y'_i : 0 \leqslant i \leqslant \ell\}$ form a nice copy of $P^{\square}_{\ell+1}$, each x_i is distinct from all other vertices except possibly x'_i and each y_i is distinct from all other vertices except possibly y_i'. We define the weight of a homomorphic copy of $C_{2\ell}^{\square}$ to be $\left(\prod_{i=1}^{\ell}\max(d(x_{i-1},y_i),\frac{d^2}{n})\cdot\prod_{i=1}^{\ell}\max(d(x_{i-1}',y_i'),\frac{d^2}{n})\right)^{-1}$.

Let G be a clean, bipartite, n-vertex graph with average degree $d \geqslant Cn^{1/2}$ and maximum degree at most Kd, where K is some absolute constant and C is a sufficiently large constant (which can depend on ℓ). Our proof consists of the following steps.

1. We first prove that the total weight of nice copies of $P_{\ell+1}^{\square}$ in G is at least $\Omega_\ell(nd^{\ell+1})$.
2. Noting that by gluing together two nice copies of $P_{\ell+1}^{\square}$, we get a nice homomorphic copy of $C_{2\ell}^{\square}$, one can easily deduce from step 1 that the total weight of nice homomorphic copies of $C_{2\ell}^{\square}$ in G is $\Omega_\ell(d^{2\ell})$.
3. By carefully analyzing different types of degenerate homomorphic copies of $C_{2\ell}^{\square}$, we can show that for $\ell \geqslant 4$, the total weight of degenerate nice homomorphic copies of $C_{2\ell}^{\square}$ in G is at most $O(nd^{2\ell-2})$. This is negligible compared to $\Omega_\ell(d^{2\ell})$, showing that G contains a genuine copy of $C_{2\ell}^{\square}$.

3 Open problem

An open problem left in this paper is determining the extremal number of $C_6^{\square}$. We conjecture that $\mathrm{ex}(n, C_6^{\square}) = \Theta(n^{3/2})$.

References

[1] Noga Alon, Michael Krivelevich, and Benny Sudakov. Turán numbers of bipartite graphs and related Ramsey-type questions. *Combinatorics, Probability and Computing*, 12(5-6):477–494, 2003.

[2] Noga Alon, Lajos Rónyai, and Tibor Szabó. Norm-graphs: variations and applications. *J. Combin. Theory Ser. B*, 76(2):280–290, 1999.

[3] J. A. Bondy and M. Simonovits. Cycles of even length in graphs. *J. Combin. Theory Ser. B*, 16:97–105, 1974.

[4] Domagoj Bradač, Oliver Janzer, Benny Sudakov, and István Tomon. The Turán number of the grid. *Bull. Lond. Math. Soc.*, 55:194–204, 2023.

[5] W. G. Brown. On graphs that do not contain a Thomsen graph. *Canadian Mathematical Bulletin*, 9(3):281–285, 1966.

[6] Boris Bukh. Extremal graphs without exponentially-small bicliques. *arXiv preprint*, arXiv: 2109.04167, 2021.

[7] P. Erdős. Problems and results in graph theory. In *The theory and applications of graphs (Kalamazoo, Mich., 1980)*, pages 331–341. Wiley, New York, 1981.

[8] P. Erdős, A. Rényi, and V. T. Sós. On a problem of graph theory. *Studia Sci. Math. Hungar.*, 1:215–235, 1966.

[9] P. Erdős and M. Simonovits. A limit theorem in graph theory. *Studia Sci. Math. Hungar.*, 1:51–57, 1966.

[10] P. Erdős and M. Simonovits. Some extremal problems in graph theory. In *Combinatorial theory and its applications, I (Proc. Colloq., Balatonfüred, 1969)*, pages 377–390. North-Holland, Amsterdam, 1970.

[11] P. Erdős and M. Simonovits. An extremal graph problem. *Acta Math. Acad. Sci. Hungar.*, 22:275–282, 1971/72.

[12] P. Erdős and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946.

[13] Zoltán Füredi and Miklós Simonovits. The history of degenerate (bipartite) extremal graph problems. In *Erdős centennial*, volume 25 of *Bolyai Soc. Math. Stud.*, pages 169–264. János Bolyai Math. Soc., Budapest, 2013.

[14] Xiaocong He, Yongtao Li, and Lihua Feng. Extremal graphs for the odd prism. *arXiv preprint*, arXiv: 2302.03278, 2023.

[15] Zhiyang He. A new upper bound on the Turán number of even cycles. *Electron. J. Combin.*, 28(2):Paper No. 2.41, 18, 2021.

[16] Oliver Janzer. Disproof of a conjecture of Erdős and Simonovits on the Turán number of graphs with minimum degree 3. *Int. Math. Res. Not. IMRN*, 2022, to appear.

[17] János Kollár, Lajos Rónyai, and Tibor Szabó. Norm-graphs and bipartite Turán numbers. *Combinatorica*, 16(3):399–406, 1996.

[18] Rom Pinchasi and Micha Sharir. On graphs that do not contain the cube and related problems. *Combinatorica*, 25(5):615–623, 2005.

[19] M. Simonovits. Extermal graph problems with symmetrical extremal graphs. Additional chromatic conditions. *Discrete Math.*, 7:349–376, 1974.

[20] Miklós Simonovits. The extremal graph problem of the icosahedron. *J. Combinatorial Theory Ser. B*, 17:69–79, 1974.

[21] P. Turán. Eine extremalaufgabe aus der graphentheorie. *Fiz Lapok*, pages 436–452, 1941.

[22] R. Wenger. Extremal graphs with no C^4's, C^6's, or C^{10}'s. *J. Combin. Theory Ser. B*, 52(1):113–116, 1991.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

The dimension of the feasible region of pattern densities

(Extended abstract)

Frederik Garbe* Daniel Král'* Alexandru Malekshahian† Raul Penaguiao‡

Abstract

A classical result of Erdős, Lovász and Spencer from the late 1970s asserts that the dimension of the feasible region of homomorphic densities of graphs with at most k vertices in large graphs is equal to the number of connected graphs with at most k vertices. Glebov et al. showed that pattern densities of indecomposable permutations are independent, i.e., the dimension of the feasible region of densities of k-patterns is at least the number of non-trivial indecomposable permutations of size at most k. We identify a larger set of permutations, which are called Lyndon permutations, whose pattern densities are independent, and show that the dimension of the feasible region of densities of k-patterns is equal to the number of non-trivial Lyndon permutations of size at most k.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-065`

1 Introduction

A classical result of Erdős, Lovász and Spencer [8] describes the independence of homomorphic densities of graphs in large graphs. Informally speaking, they showed that homomorphic densities of connected graphs are independent and actually determine the densities of

*Faculty of Informatics, Masaryk University, Botanická 68A, 602 00 Brno, Czech Republic. E-mail: `{garbe,dkral}@fi.muni.cz`. Supported by the MUNI Award in Science and Humanities (MUNI/I/1677/2018) of the Grant Agency of Masaryk University.

†Department of Mathematics, King's College London. E-mail: `alexandru.malekshahian@kcl.ac.uk`.

‡Max Planck Institute for the Sciences, Inselstraße 22, 04103 Leipzig, Germany. E-mail: `raul.penaguiao@mis.mpg.de`.

all graphs. We now state their result formally using the language of the theory of graph limits (referring to Section 2 for definitions). Let $\mathcal{G}_k$ be the set of all graphs with at most k vertices and $\mathcal{G}_k^c$ be the set of all connected graphs with at most k vertices; $t(G, W)$ denotes the homomorphism density of a graph G in a graphon W. The aforementioned result of Erdős, Lovász and Spencer [8] asserts that for every $k \in \mathbb{N}$, there exist $x_0 \in [0,1]^{\mathcal{G}_k^c}$ and $\varepsilon > 0$ such that for every $x \in B_\varepsilon(x_0) \subseteq [0,1]^{\mathcal{G}_k^c}$, there exists a graphon W such that $t(G, W)_{G\in\mathcal{G}_k^c} = x$. In addition, there exists a function $f : [0,1]^{\mathcal{G}_k^c} \to [0,1]^{\mathcal{G}_k}$, independent of W, and such that $f(t(G, W)_{G\in\mathcal{G}_k^c}) = t(G, W)_{G\in\mathcal{G}_k}$. In other words, the dimension of the feasible region of homomorphic densities of graphs with at most k vertices in graphons (large graphs) is equal to the number of connected graphs with at most k vertices.

We determine the dimension of the feasible region of densities of k-patterns in permutations; again we refer to Section 2 for definitions. Glebov et al. [10] showed that this dimension is at least the number of non-trivial indecomposable permutations of size at most k. Borga and the last author [2] observed utilizing a result of Vargas [20] that this dimension is at most the number of non-trivial Lyndon permutations of size at most k, and conjectured [2, Conjecture 1.3] that this bound is tight. Our main result asserts that this is indeed the case. Similarly to [10], our argument is based on perturbing a permuton comprised of blow-ups of indecomposable permutations. However, to be able to control the densities of the larger set of all Lyndon permutations, we choose a suitable order of the blow ups of indecomposable permutations and analyze the interplay between the blow-ups using unique decomposition properties into Lyndon words [19].

2 Combinatorial limits

We now introduce notation used throughout this extended abstract. In addition to the monograph by Lovász [16], which provides a comprehensive introduction to the theory of graph limits, we refer the reader to [3–5, 17, 18] for basic results concerning graph limits and to [1, 6, 9, 11–15] for results developing and concerning permutation limits.

2.1 Graph limits

If H and G are two graphs, the *homomorphism density* of H in G, denoted by $t(H, G)$, is the probability that a uniformly random function $f : V(H) \to V(G)$, is a *homomorphism* of H to G. A sequence $(G_n)_{n\in\mathbb{N}}$ of graphs is *convergent* if the number of vertices of G_n tends to infinity and the values of $t(H, G_n)$ converge for every H.

A *graphon* is a symmetric measurable function $W : [0,1]^2 \to [0,1]$, i.e., $W(x, y) = W(y, x)$ for $(x, y) \in [0,1]^2$. The *homomorphism density* of a graph H in a graphon W is defined by

$$t(H, W) = \int_{[0,1]^{V(H)}} \prod_{uv\in E(H)} W(x_u, x_v) \mathrm{d}x_{V(H)}.$$

A graphon W is a *limit* of a convergent sequence $(G_n)_{n\in\mathbb{N}}$ of graphs if $t(H, W)$ is the limit of $t(H, G_n)$ for every graph H. Every convergent sequence of graphs has a limit graphon and every graphon is a limit of a convergent sequence of graphs as shown by Lovász and Szegedy [17]; also see [7] for a relation to exchangeable arrays.

2.2 Permutations

A *permutation* of *size* n is a bijective function π from $[n]$ to $[n]$ (we use $[n]$ to denote the set of the first n positive integers). The permutation π is often viewed as a word $\pi(1)\pi(2)\cdots\pi(n)$ and its size is denoted by $|\pi|$. The *pattern* induced by elements $1 \leq k_1 < \cdots < k_m \leq n$ is the unique permutation $\sigma : [m] \to [m]$ such that $\sigma(i) < \sigma(i')$ if and only if $\pi(k_i) < \pi(k_{i'})$ for all $i, i' \in [m]$. The *density* of a permutation σ in a permutation π, denoted by $d(\sigma, \pi)$, is the probability that the pattern induced by $|\sigma|$ elements chosen uniformly at random is equal to σ. Similarly to the graph case, we say that a sequence $(\pi_n)_{n\in\mathbb{N}}$ of permutations is *convergent* if the sizes of π_n tend to infinity and the sequence of densities $d(\sigma, \pi_n)$ converges for every permutation σ.

We say that a permutation is *non-trivial* if its size is at least two. The *direct sum* of two permutations π_1 and π_2 is the permutation π of size $|\pi_1| + |\pi_2|$ such that $\pi(k) = \pi_1(k)$ for $k \in [|\pi_1|]$ and $\pi(|\pi_1| + k) = |\pi_1| + \pi_2(k)$ for $k \in [|\pi_2|]$; the permutation π is denoted by $\pi_1 \oplus \pi_2$. A permutation is *indecomposable* if it is not a direct sum of two permutations; note that every permutation is a (possibly iterated) direct sum of indecomposable permutations.

A word $w_1 \cdots w_n$ is *Lyndon* if no proper suffix of the word $w_1 \cdots w_n$ is smaller (in the lexicographic order) than the word $w_1 \cdots w_n$ itself. For example, the word *aab* is Lyndon but the word *aba* is not. We want to use indecomposable permutations as the alphabet to form Lyndon words. For this we introduce an order $\prec$ on the set of indecomposable permutations such that indecomposable permutations of smaller size precede those of larger size. Indecomposable permutations of the same size are ordered lexicographically. Hence, the first five letters are associated with the following five (indecomposable) permutations: $1 \prec 21 \prec 231 \prec 312 \prec 321$. As mentioned above every permutation can be uniquely decomposed into a direct sum of indecomposable permutations and therefore corresponds to a word over the alphabet consisting of indecomposable permutations. A permutation π is *Lyndon* if the word corresponding to the decomposition of π into indecomposable permutations is Lyndon. For example, the permutation $21 \oplus 231 = 21453$ is Lyndon but the permutations $21 \oplus 1 = 213$ and $21 \oplus 21 = 2143$ are not. Note that all indecomposable permutations are Lyndon.

2.3 Permutation limits

A *permuton* is a probability measure Π on the σ-algebra of Borel subsets from $[0,1]^2$ that has uniform marginals, i.e.,

$$\Pi([a,b] \times [0,1]) = \Pi([0,1] \times [a,b]) = b - a$$

for all $0 \leq a \leq b \leq 1$. A Π-*random permutation* of size n is the permutation σ obtained by sampling n points according to the measure Π, sorting them according to their x-coordinates, say $(x_1, y_1), \ldots, (x_n, y_n)$ for $x_1 < \cdots < x_n$ (note that the x-coordinates are pairwise distinct with probability 1), and defining σ so that $\sigma(i) < \sigma(j)$ if and only if $y_i < y_j$ for $i, j \in [n]$. Finally, the *density* of a permutation σ in a permuton Π, which is denoted by $d(\sigma, \Pi)$, is the probability that the Π-random permutation of size $|\sigma|$ is σ.

A permuton Π is a *limit* of a convergent sequence $(\pi_n)_{n\in\mathbb{N}}$ of permutations if, for every permutation σ, $d(\sigma, \Pi)$ is the limit of $d(\sigma, \pi_n)$. Every permuton is a limit of a convergent sequence of permutations and every convergent sequence of permutations has a limit permuton [11, 12].

3 Main result

Let $\mathcal{P}_k$ be the set of all permutations of size at most k, $\mathcal{P}_k^L$ the set of all non-trivial Lyndon permutations of size at most k. Our main result is the following.

Theorem 1. *For every $k \in \mathbb{N}$, there exists $x_0 \in [0,1]^{\mathcal{P}_k^L}$ and $\varepsilon > 0$ such that for every $x \in B_\varepsilon(x_0) \subseteq [0,1]^{\mathcal{P}_k^L}$ there exists a permuton Π such that*

$$d(\sigma, \Pi)_{\sigma\in\mathcal{P}_k^L} = x.$$

In addition, there exists a function $f : [0,1]^{\mathcal{P}_k^L} \to [0,1]^{\mathcal{P}_k}$ such that

$$f\left(d(\sigma, \Pi)_{\sigma\in\mathcal{P}_k^L}\right) = d(\sigma, \Pi)_{\sigma\in\mathcal{P}_k}$$

for every permuton Π.

We next sketch the proof of Theorem 1. We start with the existence of the function f; we remark that the existence of the function f follows from the results presented in the extended abstract [20], and we outline the argument here. Let π be a permutation and let $\pi = \pi_1 \oplus \cdots \oplus \pi_k$ be the (unique) direct sum formed by indecomposable permutations. Further, let $w_1 \cdots w_k$ be the word corresponding to $\pi_1 \oplus \cdots \oplus \pi_k$; it is well-known that the word $w_1 \cdots w_k$ can be *uniquely* expressed as a concatenation of Lyndon words in non-increasing lexicographic order, and let $\pi'_1, \ldots, \pi'_\ell$ be the permutations corresponding to these Lyndon words. For example, if $\pi = 1324576 = 1 \oplus 21 \oplus 1 \oplus 1 \oplus 21$, then π'_1 is $1 \oplus 21 = 132$ and π'_2 is $1 \oplus 1 \oplus 21 = 1243$ which are both Lyndon. It can be shown using [19, Theorem 3.1.1(a)] that the constituents of the product of $\pi'_1 \times \ldots \times \pi'_\ell$ (in the flag algebra sense) are only permutations that either are direct sums of fewer than k indecomposable permutations or are direct sums of k indecomposable permutations but are lexicographically at least as large as π. It follows that every permutation σ that is not Lyndon can be expressed as a polynomial of Lyndon permutations of size at most $|\sigma|$ (in the flag algebra sense), which implies the existence of the function f; in fact, the function f is polynomial.

We next sketch the proof of the main part of Theorem 1, which yields the (matching) lower bound on the dimension on the feasible region of pattern densities. For the lower bound, we use a different mapping of indecomposable permutations to letters; note that this

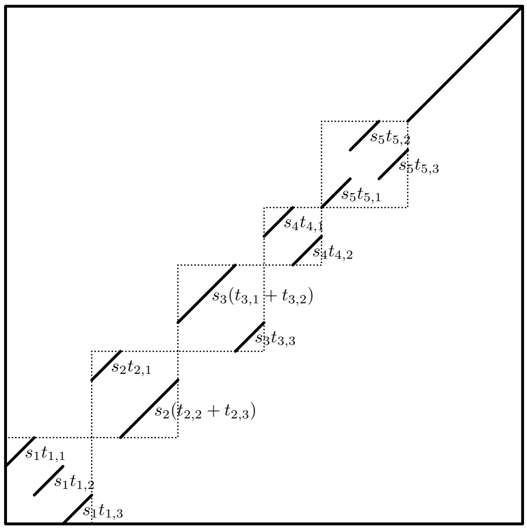


Figure 1: The permuton Π comprised of the "blow-up permutons" of the permutations 321, 312, 231, 21 and 132; the scaling factors s_i and $t_{i,j}$ are placed near their associated parts.

changes which permutations are Lyndon. The *compression* of a permutation π, which is denoted by $\widehat{\pi}$, is the permutation obtained by (iteratively) "merging" consecutive elements that increase by one; for example $\widehat{231} = 21$, $\widehat{3412} = 21$, $\widehat{2341} = 21$, and $\widehat{1342} = 132$. The new order $<$ on indecomposable permutations is defined using $\prec$ on their compressions, and if two different indecomposable permutations have the same compression, then $\prec$ is used directly. For example, $3412 < 321$, and so the letter associated with 3412 precedes the letter associated with 321. Note that while the permutation $321 \oplus 3412 = 3216745$ is Lyndon with respect to $\prec$ it is not with respect to $<$. However, it can be shown that the number of Lyndon permutations of size k is the same with respect to $\prec$ and to $<$.

Fix k and let $\pi_1, \ldots, \pi_N$ be all non-trivial Lyndon permutations of size at most k listed in the decreasing (lexicographic) order of the words corresponding to their indecomposable blocks; we emphasize that the modified order $<$ is used both to define which permutations are Lyndon and to order the Lyndon permutations. For $s_1, \ldots, s_N \in [0,1]$ and $t_{i,j} \in [0,1]$, $i \in [N]$ and $j \in [|\pi_i|]$ such that the sum of $t_{i,j}$'s is at most one, we define a permuton Π to be the permuton comprised of the "blow-up permutons" of the permutations $\pi_1, \ldots, \pi_N$. For each $i \in [N]$ the "blow-up permuton" uses a segment of horizontal length $t_{i,j}$ corresponding to the j'th point of the permutation π_i, $j \in [|\pi_i|]$. The "blow-up permutons" then get scaled by $s_1, \ldots, s_N$, respectively; see Figure 1 for illustration. We next consider the Jacobian matrix of the densities $d(\pi_1, \Pi), \ldots, d(\pi_N, \Pi)$ viewed as functions of $s_1, \ldots, s_N$ and observe that its determinant is a polynomial in the variables s_i and $t_{i,j}$ and the coefficient of the monomial formed by the product of all $t_{i,j}$ is non-zero; the latter is argued by making use of [19, Theorem 3.1.1(a)]. Hence, the Jacobian determinant is not identically zero and so there exists a choice of s_i and $t_{i,j}$ such that the determinant is non-zero, which implies the existence of the point $x_0 \in [0,1]^{\mathcal{P}_k^L}$ and the real $\varepsilon > 0$ from the statement of Theorem 1.

References

[1] J. Balogh, P. Hu, B. Lidický, O. Pikhurko, B. Udvari and J. Volec: *Minimum number of monotone subsequences of length 4 in permutations*, Combin. Probab. Comput. **24** (2015), 658–679.

[2] J. Borga and R. Penaguiao: *The feasible regions for consecutive patterns of pattern-avoiding permutations*, Discrete Math. **346** (2023), 113219.

[3] C. Borgs, J. Chayes, L. Lovász, V. T. Sós, B. Szegedy and K. Vesztergombi: *Graph limits and parameter testing*, Proc. 38th annual ACM Symposium on Theory of computing (STOC) (2006), 261–270.

[4] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós and K. Vesztergombi: *Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing*, Adv. Math. **219** (2008), 1801–1851.

[5] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós and K. Vesztergombi: *Convergent sequences of dense graphs II. Multiway cuts and statistical physics*, Ann. of Math. (2) (2012), 151–219.

[6] T. Chan, D. Král', J. A. Noel, Y. Pehova, M. Sharifzadeh and J. Volec: *Characterization of quasirandom permutations by a pattern sum*, Random Struct. Algor. **57** (2020), 920–939.

[7] P. Diaconis and S. Janson: *Graph limits and exchangeable random graphs*, Rend. Mat. Appl. **28** (2008), 33–61.

[8] P. Erdős, L. Lovász and J. Spencer: *Strong independence of graphcopy functions*, Graph theory and related topics (1979), 165–172.

[9] R. Glebov, A. Grzesik, T. Klimošová and D. Král': *Finitely forcible graphons and permutons*, J. Combin. Theory Ser. B **110** (2015), 112–135.

[10] R. Glebov, C. Hoppen, T. Klimošová, Y. Kohayakawa, D. Král' and H. Liu: *Densities in large permutations and parameter testing*, European J. Combin. **60** (2017), 89–99.

[11] C. Hoppen, Y. Kohayakawa, C. G. T. de A. Moreira, B. Ráth and R. M. Sampaio: *Limits of permutation sequences*, J. Combin. Theory Ser. B **103** (2013), 93–113.

[12] C. Hoppen, Y. Kohayakawa, C. G. T. de A. Moreira and R. M. Sampaio: *Testing permutation properties through subpermutations*, Theor. Comput. Sci. **412** (2011), 3555–3567.

[13] R. Kenyon, D. Král', C. Radin and P. Winkler: *Permutations with fixed pattern densities*, Random Struct. Algor. **56** (2020), 220–250.

[14] D. Kráľ and O. Pikhurko: *Quasirandom permutations are characterized by 4-point densities*, Geom. Funct. Anal. **23** (2013), 570–579.

[15] M. Kurečka: *Lower bound on the size of a quasirandom forcing set of permutations*, Combin. Probab. Comput. **31** (2022), 304–319.

[16] L. Lovász: Large Networks and Graph Limits, *Colloquium Publications*, volume 60, 2012.

[17] L. Lovász and B. Szegedy: *Limits of dense graph sequences*, J. Combin. Theory Ser. B **96** (2006), 933–957.

[18] L. Lovász and B. Szegedy: *Testing properties of graphs and functions*, Israel J. Math. **178** (2010), 113–156.

[19] D. A. Radford: *A natural ring basis for the shuffle algebra and an application to group schemes*, J. Algebra **58** (1979), 432–454.

[20] Y. Vargas: *Hopf algebra of permutation pattern functions*, DMTCS Proceedings vol. AT, 26th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2014) (2014), 839–850.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

On the structure and values of betweenness centrality in dense betweenness-uniform graphs

(Extended abstract)

Babak Ghanbari* David Hartman*† Vít Jelínek*
Aneta Pokorná*† Robert Šámal* Pavel Valtr‡

Abstract

Betweenness centrality is a network centrality measure based on the amount of shortest paths passing through a given vertex. A graph is betweenness-uniform (BUG) if all vertices have an equal value of betweenness centrality. In this contribution, we focus on betweenness-uniform graphs with betweenness centrality below one. We disprove a conjecture about the existence of a BUG with betweenness value α for any rational number α from the interval $(3/4, \infty)$ by showing that only very few betweenness centrality values below $6/7$ are attained for at least one BUG. Furthermore, among graphs with diameter at least three, there are no betweenness-uniform graphs with a betweenness centrality smaller than one. In graphs of smaller diameter, the same can be shown under a uniformity condition on the components of the complement.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-066`

*Computer Science Institute of Charles University, 118 00 Prague, Czech Republic. E-mail: `{babak,hartman,jelinek,pokorna,samal}@iuuk.mff.cuni.cz`. DH, AP and RS supported by ERC grant agreement No 810115. RS supported by the Czech Science Foundation Grant No.22-17398S.

†Computer Science Institute, Czech Academy of Sciences, 182 00 Prague, Czech Republic. E-mail: `{hartman, pokorna}@cs.cas.cz`. DH and AP supported by the Czech Science Foundation Grant No. 23-07074S.

‡Department of Applied Mathematics, Charles University, 118 00 Prague, Czech Republic. E-mail: `valtr@kam.mff.cuni.cz`

1 Introduction and preliminaries

In network science, it is useful to have the ability to evaluate the nodes according to their importance in the given network with respect to some criterion. Centrality measures, a tool for this evaluation, are based on various properties of nodes such as the sizes of their neighborhoods or their distances to other parts of the network. The choice of the most suitable centrality heavily depends on what is being represented by the corresponding network, resulting in the impossibility to directly compare these measures.

In this work, we study the properties of betweenness centrality, which is based on the fraction of shortest paths passing through the given vertex. More precisely, for a vertex $u \in V(G)$ of a graph $G = (V, E)$ the *betweenness centrality* of u (or shortly just betweenness) is

$$B(u) = \sum_{\{v,w\} \in V(G)} \frac{\sigma_{vw}(u)}{\sigma_{vw}},$$

where σ_{vw} denotes the number of shortest paths between v and w and $\sigma_{vw}(u)$ is the number of such paths passing through u [4]. The betweenness centrality of a graph G, denoted $B(G)$, is defined as the average betweenness centrality of its vertices. Although having a wide range of applications, for example in the assessment of electric-grid vulnerability [1], in the measuring of dependencies in object-oriented software systems [10], in arms transfer [2] or in citation networks [12], there has not been many theoretical results about properties of betweenness centrality.

One of the few theoretical aspects studied is the distribution of betweenness centrality values and specifically its extremal cases, be it either distinct betweenness for all vertices [9] or betweenness centrality being the same for all vertices [6, 7, 8]. In the latter case, such graph G is called *betweenness-uniform*, or shortly a BUG, and its complement $\overline{G}$ is called a coBUG. Our main result is in understanding of the structure of betweenness-uniform graphs with betweenness value below one and in disproving the following conjecture.

Conjecture 1 (Coroničová-Hurajová, Madaras (2013) [6]). *For any rational value α in the interval $(3/4, \infty)$, there exists a betweenness-uniform graph with betweenness value of α.*

Note that the authors of the above-mentioned conjecture use a definition of betweenness that counts contributions of ordered pairs of vertices, whereas our definition counts contributions of unoriented pairs, resulting in betweenness values smaller by a factor of one-half.

In the following text, we use standard graph-theoretic notation, namely the *diameter* $\operatorname{diam}(G)$ being the longest of the shortest paths inside G, and $\overline{G} = (V, \binom{n}{2} \setminus E)$ being the complement of the graph G. A graph is *k-regular* if each vertex has exactly k neighbors. The fact that graphs G and H are isomorphic is denoted by $G \approx H$. Also, K_n denotes the complete graph on n vertices and $K_{a,b}$ denotes the complete bipartite graph with parts of sizes a and b.

2 Known constructions of betweenness-uniform graphs and their values

It is an open problem to characterize all betweenness-uniform graphs and the spectrum of betweenness values that are attained on their vertices. We introduce some of the known classes of BUGs to give the reader some intuition.

The simplest BUGs are the *vertex-transitive* graphs, i.e., the graphs in which for each pair of vertices there is an automorphism of the graph mapping one vertex onto the other. [6, 11] This holds for example for cycles, complete graphs, or complete bipartite graphs with the same sizes of parts.

It has been shown that the average betweenness centrality of an arbitrary connected graph G is related to the average pairwise distances of its vertices: $B(G) = \frac{1}{2}(n-1)(l(G)-1)$, where $l(G)$ is the average distance between pairs of vertices G and n is the number of its vertices [3, 5]. Using this expression and the fact that average betweenness is equal to betweenness of any vertex in a BUG, it is not hard to see that any half-integer can be obtained as a value of betweenness in a vertex-transitive graph.

Moreover, it has been shown that for a given n, there are superpolynomially many BUGs that are not vertex-transitive [6]. A class containing many such graphs is the class of distance-regular graphs. A graph is *distance-regular* if for any two vertices x and y, the number of vertices in distance a from x and in distance b from y depends only on the triple $(a, b, \mathrm{dist}(x, y))$.Any distance-regular graph is a BUG [6].

There is a construction allowing the creation of a BUG G' from a smaller k-regular BUG G. Let H be a disjoint union of ℓ distinct cliques of multiplicity n_i and order r_i for each $i \in \{1, \dots, \ell\}$. Then replace each vertex of G by H and for H_x and H_y being the copies of H that replaced x and y, make a full-join of H_x and H_y whenever $xy \in E(G)$ [6]. The betweenness value in G' is given by

$$B(G') = \frac{1}{2}\Big(mB(G) + m - 1 - \sum_{i=1}^{\ell} \frac{n_i r_i(r_i - 1)}{m}\Big),$$

where $m = |V(H)| = \sum_{i=1}^{\ell} n_i r_i$. Take $G \approx K_n$ for $n \geq 2$ and consider graphs G', G'' obtained by doing the above-mentioned construction with disjoint unions of cliques $H' = \bigcup_{a=1}^{c} K_{r'_a}$ and $H'' = \bigcup_{a=1}^{c-1} K_{r''_a}$ where for $i \leq c-2$, $r'_i = r''_i$ and $r''_{c-1} = r'_{c-1} + r'_c$. Then with an increasing number of vertices in the disjoint unions of cliques, we can construct G', G'' with decreasing $|B(G') - B(G'')|$.

3 Betweenness-uniform graphs with betweenness value below one

Based on the fact that only non-adjacent pairs of vertices can contribute to the betweenness of other vertices, we show the following relation between the density of the complement of G, $\overline{G}$, and the resulting average betweenness centrality.

Lemma 2. *For a connected graph G, we have $B(G) \geq \frac{|E(\overline{G})|}{|V(\overline{G})|}$, with equality if and only if G has diameter at most 2.*

3.1 No BUGs with diameter ≥ 3 and betweenness below one

By Lemma 2, betweenness-uniform graphs with betweenness-centrality below one are quite dense. For dense graphs, it is often easier to analyze their structure and properties by studying their complements. We start by observing that complements of graphs with diameter at least three have a simple structure.

Theorem 3. *A connected graph G satisfies diam$(G) \geq 3$ if and only if $\overline{G}$ is connected and contains a spanning tree which is a double star.*

By combining Theorem 3 and Lemma 2 and observing that a complement of a double star is not a BUG leads to the following corollary.

Corollary 4. *There is no connected betweenness-uniform graph of diameter greater than two having betweenness centrality below one.*

3.2 Structure of BUGs with diameter 2 and betweenness below one

Due to Corollary 4, we can restrict the remaining analysis to betweenness-uniform graphs with diameter two. In these graphs, each shortest path contributes to exactly one vertex.

Consider $\overline{G}$ a complement of a BUG of diameter two with $|V(\overline{G})| = n$. We say that a vertex v is *close* to an edge e in $\overline{G}$ if v is adjacent to at least one endpoint of the given edge; in particular, the two endpoints of e are close to e. Let $C_{\overline{G}}(e)$ be the set of vertices close to the edge e, and let $C_{\overline{G}}(v)$ be the set of edges that are close to the vertex v.

Observe that for a vertex v and an edge $e = \{x, y\}$ in $\overline{G}$, v is close to e if and only if no shortest path from x to y in G passes through v, and thus x and y do not contribute to the betweenness of v in G. In particular, for an edge $e = \{x, y\}$ of $\overline{G}$, there are $n - |C_{\overline{G}}(e)|$ shortest paths from x to y in G, each passing through a different vertex of $V \setminus C_{\overline{G}}(e)$. We denote the contribution of the edge e to each vertex of $V(G) \setminus C_{\overline{G}}(e)$ as the *weight* of the edge e, $w(e) = 1/(n - |C_{\overline{G}}(e)|)$.

The *weight* of a vertex v, $w(v) = \sum_{e \in C_{\overline{G}}(v)} w(e)$, is closely related to the betweenness of v.

Proposition 5. *Let G be of diameter two. Then for all $x \in V(G)$,*

$$B(x) = \Big(\sum_{e \in E} w(e)\Big) - w(x).$$

Corollary 6. *G is a BUG if and only if all vertices of $\overline{G}$ have the same weight.*

Furthermore, as we are interested only in betweenness values below one, $\overline{G}$ must have more vertices than edges by Lemma 2.

Observation 7. *If G is a BUG with $B(G) < 1$, then $\overline{G}$ has some components which are trees.*

Considering x a leaf adjacent to y, $x, y \in V(G)$, by comparing $C_{\overline{G}}(x)$ and $C_{\overline{G}}(y)$, we obtain a restriction on the structure of the tree components of $\overline{G}$.

Lemma 8. *For G a BUG, any vertex of degree one appears only in a star component of $\overline{G}$.*

Conditioning on the vertices having the same weight, we even prohibit stars of different sizes.

Proposition 9. *All the tree components of a coBUG $\overline{G}$ are stars of the same size.*

Let us call a BUG G *exotic* if it has $B(G) < 1$ and $\overline{G}$ contains a component different from a star.

Conjecture 10. *There are no exotic betweenness-uniform graphs.*

A graph $\overline{G}$ is *(m,t)-uniform* if $|C_{\overline{G}}(v)| = m$ for every $v \in V(\overline{G})$ and $|C_{\overline{G}}(e)| = t$ for every $e \in E(\overline{G})$. Note that if $\overline{G}$ is disconnected and (m,t)-uniform, then G is always a BUG with betweenness value $^m/_t$. Indeed, in an (m,t)-uniform graph $\overline{G}$, every vertex has weight $\frac{m}{n-t}$. However, we can show that there can be no exotic BUGs whose complements have nontrivial (m,t)-uniform components.

Theorem 11. *There is no exotic BUG with a complement containing an (m,t)-uniform component other than a star.*

Note that there exist infinitely many BUGs of betweenness exactly one whose complements have both a star component and an (m,t)-uniform non-star component. There are also BUGs whose complements have a star component and a non-(m,t)-uniform component, but we have not found any such BUG with betweenness below one.

3.3 Only values $\frac{k}{k+1}$ on the interval $\langle 0, \frac{6}{7} \rangle$

Apart from showing that a non-star component of a coBUG with density less than one would have to be non-(m,t)-uniform, we prove that some small stars cannot occur as a component of a coBUG with any other types of components, by which we show that there are no exotic BUGs with betweenness below $\frac{6}{7}$.

By comparing the vertices and edges in the closeness relation and their weights, we can infer the following forbidden structures in components of coBUGs.

Lemma 12. *Let H be a component of a coBUG. Then*

- *H has no vertex of degree one unless $H \approx K_{1,\ell}$ for some $\ell \geq 1$.*
- *H has no vertex of degree two whose neighbours are adjacent unless $H \approx K_{1,1,\ell}$ for some $\ell \geq 1$.*
- *H has no two adjacent vertices a, b of degree two, whose neighbours c and d are adjacent, i.e. $ab, ac, bd, cd \in E(H)$, unless $H \approx C_4$.*
- *H has no adjacent vertices a, b of degree two, whose neighbours e and c have a common neighbour d, i.e. $ab, bc, cd, de, ea \in H$, unless $H \approx C_5$*
- *H has no vertices a, b, c of degree two such that $ab, bc \in E(H)$*

By using Lemma 12 and by counting of edges we can conclude the following lemma.

Lemma 13. *Let v be a vertex of a non-star component of a coBUG $\overline{G}$. If* $\deg(v) = d$, *then $|C_{\overline{G}}(v)| \geq 2d$.*

By considering the weight $\frac{\ell}{n-(\ell+1)}$ of vertices in a star $K_{1,\ell}$, by careful consideration of obtainable weights, by using of Lemmas 12 and 13 and by a case analysis we can prove the following theorem.

Theorem 14. *Let H be one of the following: K_1, K_2, $K_{1,2}$, $K_{1,3}$, $K_{1,4}$, $K_{1,5}$, $K_{1,6}$. If $\overline{G}$ is a coBUG of density less than 1 containing H as a connected component, then all the other components of $\overline{G}$ are isomorphic to H as well.*

This result allows us to both disprove the Conjecture 1 and to generalize previous result by Hurajová and Madaras [6] claiming that there are no betweenness-uniform graphs in the interval $(0, \frac{1}{2})$.

Corollary 15. *If G is a BUG with $B(G) \leq \frac{6}{7}$, then $B(G) \in \{0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{6}{7}\}$. Moreover, any such BUG is a complement of a disjoint union of stars of the same size.*

Acknowledgements

This research was started during workshop KAMAK 2022, we are grateful to its organizers.

References

[1] J. C. Birk, C. J. Bresloff, and R. Darbali-Zamora. Electric Grid Vulnerability Analysis to Identify Communities Prone to Wildfires. *IEEE Access*, 2023.

[2] C. C. Chou, C.-S. Teng, and H. H. Tung. How do alliances trade arms? Political alliance networks and global arms transfers. *PLOS ONE*, 18(3):1–14, 2023.

[3] F. Comellas and S. Gago. Spectral bounds for the betweenness of a graph. *Linear Algebra and its Applications*, 423(1):74 – 80, 2007. Special Issue devoted to papers presented at the Aveiro Workshop on Graph Spectra.

[4] L. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40:35–41, 1977.

[5] S. Gago. *Métodos espectrales y nuevas medidas, modelos y parámetros en grafos pequeño-mundo invariantes de escala.* PhD thesis, Universitat Politècnica de Catalunya, 2006.

[6] S. Gago, J. Hurajova-Coronicova, and T. Madaras. On betweenness-uniform graphs. *Czechoslovak Mathematical Journal*, 63(3):629–642, 2013.

[7] D. Hartman, A. Pokorná, and P. Valtr. On the connectivity and the diameter of betweenness-uniform graphs. In A. Mudgal and C. R. Subramanian, editors, *Algorithms and Discrete Applied Mathematics*, pages 317–330, Cham, 2021. Springer.

[8] J. Hurajova-Coronicova and T. Madaras. More on betweenness-uniform graphs. *Czechoslovak Mathematical Journal*, 68(2):293–306, 2018.

[9] R. Lopez, J. Worrell, H. Wickus, R. Florez, and D. A. Narayan. Towards a characterization of graphs with distinct betweenness centralities. *Australasian Journal Of Combinatorics*, 68(2):285–303, 2017.

[10] A. Ouellet and M. Badri. Combining object-oriented metrics and centrality measures to predict faults in object-oriented software: An empirical validation. *Journal of Software: Evolution and Process*, 2023.

[11] A. Pokorná. Characteristics of network centralities. Master's thesis, Charles University, Prague, Czech Republic, 2020.

[12] Y. Sarabi and M. Smith. Gender diversity and publication activity—an analysis of STEM in the UK. *Research Evaluation*, 2023.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

The Minimum Degree Removal Lemma Thresholds

(Extended abstract)

Lior Gishboliner* Zhihan Jin* Benny Sudakov*

Abstract

The graph removal lemma is a fundamental result in extremal graph theory which says that for every fixed graph H and $\varepsilon > 0$, if an n-vertex graph G contains εn^2 edge-disjoint copies of H then G contains $\delta n^{v(H)}$ copies of H for some $\delta = \delta(\varepsilon, H) > 0$. The current proofs of the removal lemma give only very weak bounds on $\delta(\varepsilon, H)$, and it is also known that $\delta(\varepsilon, H)$ is not polynomial in ε unless H is bipartite. Recently, Fox and Wigderson initiated the study of minimum degree conditions guaranteeing that $\delta(\varepsilon, H)$ depends polynomially or linearly on ε. We answer several questions of Fox and Wigderson on this topic.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-067`

1 Introduction

The graph removal lemma, first proved by Ruzsa and Szemerédi [22], is a fundamental result in extremal graph theory. It also has important applications to additive combinatorics and property testing. The lemma states that for every fixed graph H and $\varepsilon > 0$, if an n-vertex graph G contains εn^2 edge-disjoint copies of H then G it contains $\delta n^{v(H)}$ copies of H, where $\delta = \delta(\varepsilon, H) > 0$. Unfortunately, the current proofs of the graph removal lemma give only very weak bounds on $\delta = \delta(\varepsilon, H)$ and it is a very important problem to understand the dependence of δ on ε. The best known result, due to Fox [11], proves that $1/\delta$ is at most a tower of exponents of height logarithmic in $1/\varepsilon$. Ideally, one would like to have better

*Department of Mathematics, ETH, Zürich, Switzerland. Research supported in part by SNSF grant 200021_196965. Email: **{lior.gishboliner, zhihan.jin, benjamin.sudakov}@math.ethz.ch**.

bounds on $1/\delta$, where an optimal bound would be that δ is polynomial in ε. However, it is known [2] that $\delta(\varepsilon, H)$ is polynomial in ε only if H is bipartite. This situation led Fox and Wigderson [12] to initiate the study of minimum degree conditions which guarantee that $\delta(\varepsilon, H)$ depends polynomially or linearly on ε. Formally, let $\delta(\varepsilon, H; \gamma)$ be the maximum $\delta \in [0, 1]$ such that if G is an n-vertex graph with minimum degree at least γn and with εn^2 edge-disjoint copies of H, then G contains $\delta n^{v(H)}$ copies of H.

Definition 1.1. *Let H be a graph.*

1. *The* linear removal threshold *of H, denoted $\delta_{\textit{lin-rem}}(H)$, is the infimum γ such that $\delta(\varepsilon, H; \gamma)$ depends linearly on ε, i.e. $\delta(\varepsilon, H; \gamma) \geq \mu\varepsilon$ for some $\mu = \mu(\gamma) > 0$ and all $\varepsilon > 0$.*

2. *The* polynomial removal threshold *of H, denoted $\delta_{\textit{poly-rem}}(H)$, is the infimum γ such that $\delta(\varepsilon, H; \gamma)$ depends polynomially on ε, i.e. $\delta(\varepsilon, H; \gamma) \geq \mu\varepsilon^{1/\mu}$ for some $\mu = \mu(\gamma) > 0$ and all $\varepsilon > 0$.*

Trivially, $\delta_{\text{lin-rem}}(H) \geq \delta_{\text{poly-rem}}(H)$. Fox and Wigderson [12] initiated the study of $\delta_{\text{lin-rem}}(H)$ and $\delta_{\text{poly-rem}}(H)$, and proved that $\delta_{\text{lin-rem}}(K_r) = \delta_{\text{poly-rem}}(K_r) = \frac{2r-5}{2r-3}$ for every $r \geq 3$, where K_r is the clique on r vertices. They further asked to determine the removal lemma thresholds of odd cycles. Here we completely resolve this question. The following theorem handles the polynomial removal threshold.

Theorem 1.2. $\delta_{\textit{poly-rem}}(C_{2k+1}) = \frac{1}{2k+1}$.

Theorem 1.2 also answers another question of Fox and Wigderson [12], of whether $\delta_{\text{lin-rem}}(H)$ and $\delta_{\text{poly-rem}}(H)$ can only obtain finitely many values on r-chromatic graphs H for a given $r \geq 3$. Theorem 1.2 shows that $\delta_{\text{poly-rem}}(H)$ obtains infinitely many values for 3-chromatic graphs. In contrast, $\delta_{\text{lin-rem}}(H)$ obtains only three possible values for 3-chromatic graphs. Indeed, the following theorem determines $\delta_{\text{lin-rem}}(H)$ for every 3-chromatic H. An edge xy of H is called *critical* if $\chi(H - xy) < \chi(H)$.

Theorem 1.3. *For a graph H with $\chi(H) = 3$, it holds that*

$$\delta_{\textit{lin-rem}}(H) = \begin{cases} \frac{1}{2} & H \textit{ has no critical edge,} \\ \frac{1}{3} & H \textit{ has a critical edge and contains a triangle,} \\ \frac{1}{4} & H \textit{ has a critical edge and } \text{odd-girth}(H) \geq 5. \end{cases}$$

Theorems 1.2 and 1.3 show a separation between the polynomial and linear removal thresholds, giving a sequence of graphs (i.e. $C_5, C_7, \dots$) where the polynomial threshold tends to 0 while the linear threshold is constant $\frac{1}{4}$. The proof of Theorem 1.3 appears in the full version of this paper.

The parameters $\delta_{\text{poly-rem}}$ and $\delta_{\text{lin-rem}}$ are related to two other well-studied minimum degree thresholds: the chromatic threshold and the homomorphism threshold. The chromatic threshold of a graph H is the infimum γ such that every n-vertex H-free graph G with

$\delta(G) \geq \gamma n$ has bounded cromatic number, i.e., there exists $C = C(\gamma)$ such that $\chi(G) \leq C$. The study of the chromatic threshold originates in the work of Erdős and Simonovits [10] from the '70s. Following multiple works [4, 14, 15, 7, 5, 24, 25, 18, 6, 13, 19], the chromatic threshold of every graph was determined by Allen et al. [1].

Moving on to the homomorphism threshold, we define it more generally for families of graphs. The *homomorphism threshold* of a graph-family $\mathcal{H}$, denoted $\delta_{\text{hom}}(\mathcal{H})$, is the infimum γ for which there exists an $\mathcal{H}$-free graph $F = F(\gamma)$ such that every n-vertex $\mathcal{H}$-free graph G with $\delta(G) \geq \gamma n$ is homomorphic to F. When $\mathcal{H} = \{H\}$, we write $\delta_{\text{hom}}(H)$. This parameter was widely studied in recent years [17, 21, 16, 8, 23]. It turns out that δ_{hom} is closely related to $\delta_{\text{poly-rem}}(H)$, as the following theorem shows. For a graph H, let $\mathcal{I}_H$ denote the set of all minimal (with respect to inclusion) graphs H' such that H is homomorphic to H'.

Theorem 1.4. *For every graph H, $\delta_{\text{poly-rem}}(H) \leq \delta_{\text{hom}}(\mathcal{I}_H)$.*

Note that $\mathcal{I}_{C_{2k+1}} = \{C_3, C_5, \dots, C_{2k+1}\}$. Using this, the upper bound in Theorem 1.2 follows immediately by combining Theorem 1.4 with the result of Ebsen and Schacht [8] that $\delta_{\text{hom}}(\{C_3, C_5, \dots, C_{2k+1}\}) = \frac{1}{2k+1}$. The lower bound in Theorem 1.2 was established in [12].

2 Proof of Theorem 1.4

We say that an n-vertex graph G is ε*-far* from a graph property $\mathcal{P}$ (e.g. being H-free for a given graph H, or being homomorphic to a given graph F) if one must delete at least εn^2 edges to make G satisfy $\mathcal{P}$. Trivially, if G has εn^2 edge-disjoint copies of H, then it is ε-far from being H-free. The following result is from [20].

Theorem 2.1. *For every graph F on f vertices and for every $\varepsilon > 0$, there is $q = q_F(\varepsilon) =$ poly(f/ε), such that the following holds. If a graph G is ε-far from being homomorphic to F, then for a sample of q vertices $x_1, \dots, x_q \in V(G)$, taken uniformly with repetitions, it holds that $G[\{x_1, \dots, x_q\}]$ is not homomorphic to F with probability at least $\frac{2}{3}$.*

Theorem 2.1 is proved in Section 2 of [20]. In fact, [20] proves a more general result on property testing of the so-called 0/1-partition properties. Such a property is given by an integer f and a function $d : [f]^2 \to \{0, 1, \perp\}$, and a graph G satisfies the property if it has a partition $V(G) = V_1 \cup \dots \cup V_f$ such that for every $1 \leq i, j \leq f$ (possibly $i = j$), it holds that (V_i, V_j) is complete if $d(i, j) = 1$ and (V_i, V_j) is empty if $d(i, j) = 0$ (if $d(i, j) = \perp$ then there are no restrictions). One can express the property of having a homomorphism into F in this language, simply by setting $d(i, j) = 0$ for $i = j$ and $ij \notin E(F)$. In [20], the class of these partition properties is denoted $\mathcal{GPP}_{0,1}$, and every such property is shown to be testable by sampling poly(f/ε) vertices. This implies Theorem 2.1.

For a graph H on $[h]$ and integers $s_1, s_2, \dots, s_h > 0$, we denote by $H[s_1, \dots, s_h]$ the blow-up of H where each vertex $i \in V(H)$ is replaced by a set S_i of size s_i. The following lemma is standard, and follows from the hypergraph version of the Kővári-Sós-Turán theorem [9].

Lemma 2.2. *Let H be a fixed graph on vertex set $[h]$ and let $s_1, s_2, \dots, s_h \in \mathbb{N}$. There exists a constant $c = c(H, s_1, \dots, s_h) > 0$ such that the following holds. Let G be an n-vertex graph and $V_1, \dots, V_h \subseteq V(G)$. Suppose that G contains at least ρn^h copies of H mapping i to V_i for all $i \in [h]$. Then G contains at least $c\rho^{\frac{1}{c}} \cdot n^{s_1+\dots+s_h}$ copies of $H[s_1, \dots, s_h]$ mapping S_i to V_i for all $i \in [h]$.*

Proof of Theorem 1.4. Recall that $\mathcal{I}_H$ is the set of minimal graphs H' (with respect to inclusion) such that H is homomorphic to H'. For convenience, put $\delta := \delta_{\text{hom}}(\mathcal{I}_H)$. Our goal is to show that $\delta_{\text{poly-rem}}(H) \leq \delta + \alpha$ for every $\alpha > 0$. So fix $\alpha > 0$ and let G be a graph with minimum degree $\delta(G) \geq (\delta + \alpha)n$ and with εn^2 edge-disjoint copies of H. By the definition of the homomorphism threshold, there is an $\mathcal{I}_H$-free graph F (depending only on $\mathcal{I}_H$ and α) such that if a graph G_0 is $\mathcal{I}_H$-free and has minimum degree at least $(\delta + \frac{\alpha}{2}) \cdot |V(G_0)|$, then G_0 is homomorphic to F. Observe that if a graph G_0 is homomorphic to F then G_0 is H-free, because F is free of any homomorphic image of H. It follows that G is ε-far from being homomorphic to F, because G is ε-far from being H-free. Now we apply Theorem 2.1. Let $q = q_F(\varepsilon)$ be given by Theorem 2.1. We assume that $q \gg \frac{\log(1/\alpha)}{\alpha^2}$ and $n \gg q^2$ without loss of generality. Sample q vertices $x_1, \dots, x_q \in V(G)$ with repetition and let $X = \{x_1, \dots, x_q\}$. By Theorem 2.1, $G[X]$ is not homomorphic to F with probability at least $2/3$. As $n \gg q^2$, the vertices $x_1, \dots, x_q$ are pairwise-distinct with probability at least 0.99. Also, for every $i \in [q]$, the number of indices $j \in [q] \setminus \{i\}$ with $x_i x_j \in E(G)$ dominates a binomial distribution $\mathrm{B}(q-1, \frac{\delta(G)}{n})$. By the Chernoff bound (see e.g. [3, Appendix A]) and as $\delta(G) \geq (\delta + \alpha)n$, the number of such indices is at least $(\delta + \frac{\alpha}{2})q$ with probability $1 - e^{-\Omega(q\alpha^2)}$. Taking the union bound over $i \in [q]$, we get that $\delta(G[X]) \geq (\delta + \frac{\alpha}{2})|X|$ with probability at least $1 - qe^{-\Omega(q\alpha^2)} \geq 0.9$, as $q \gg \frac{\log(1/\alpha)}{\alpha^2}$. Hence, with probability at least $\frac{1}{2}$ it holds that $\delta(G[X]) \geq (\delta + \frac{\alpha}{2})|X|$ and $G[X]$ is not homomorphic to F. If this happens, then $G[X]$ is not $\mathcal{I}_H$-free (by the choice of F), hence $G[X]$ contains a copy of some $H' \in \mathcal{I}_H$. By averaging, there is $H' \in \mathcal{I}_H$ such that $G[X]$ contains a copy of H' with probability at least $\frac{1}{2|\mathcal{I}_H|}$. Put $k = |V(H')|$ and let M be the number of copies of H' in G. The probability that $G[X]$ contains a copy of H' is at most $M(\frac{q}{n})^k$. Using the fact that $q = \text{poly}_{H,\alpha}(\frac{1}{\varepsilon})$, we conclude that $M \geq \frac{1}{2|\mathcal{I}_H|} \cdot (\frac{n}{q})^k \geq \text{poly}_{H,\alpha}(\varepsilon)n^k$. As $H \to H'$, there exists H'', a blow-up of H', such that H'' have the same number of vertices as H, and that $H \subset H''$. By Lemma 2.2 for H' with $V_i = V(G)$ for all i, there exist $\text{poly}_{H,\alpha}(\varepsilon)n^{v(H'')}$ copies of H'' in G, and thus $\text{poly}_{H,\alpha}(\varepsilon)n^{v(H)}$ copies of H. This completes the proof. $\square$

3 Concluding remarks and open questions

It would be interesting to determine the possible values of $\delta_{\text{poly-rem}}(H)$ for 3-chromatic graphs H. So far we know that $\frac{1}{2k+1}$ is a value for each $k \geq 1$. Is there a graph H with $\frac{1}{5} < \delta_{\text{poly-rem}}(H) < \frac{1}{3}$? Also, is it true that $\delta_{\text{poly-rem}}(H) > \frac{1}{5}$ if H is not homomorphic to C_5?

Another question is whether the inequality in Theorem 1.4 is always tight, i.e. is it always true that $\delta_{\text{poly-rem}}(H) = \delta_{\text{hom}}(\mathcal{I}_H)$?

Finally, we wonder whether the parameters $\delta_{\text{poly-rem}}(H)$ and $\delta_{\text{lin-rem}}(H)$ are monotone, in the sense that they do not increase when passing to a subgraph of H. We are not aware of a way of proving this without finding $\delta_{\text{poly-rem}}(H), \delta_{\text{lin-rem}}(H)$.

References

[1] Peter Allen, Julia Böttcher, Simon Griffiths, Yoshiharu Kohayakawa, and Robert Morris. The chromatic thresholds of graphs. *Advances in Mathematics*, 235:261–295, 2013. 1

[2] Noga Alon. Testing subgraphs in large graphs. *Random Structures & Algorithms*, 21(3-4):359–370, 2002. 1

[3] Noga Alon and Joel H. Spencer. *The probabilistic method.* John Wiley & Sons, 2016. 2

[4] Béla Andrásfai, Paul Erdös, and Vera T Sós. On the connection between chromatic number, maximal clique and minimal degree of a graph. *Discrete Mathematics*, 8(3):205–218, 1974. 1

[5] Stephan Brandt. On the structure of dense triangle-free graphs. *Combinatorics, Probability and Computing*, 8(3):237–245, 1999. 1

[6] Stephan Brandt and Stéphan Thomassé. Dense triangle-free graphs are four-colorable: A solution to the Erdős-Simonovits problem. *preprint*, 2011. 1

[7] Chuan-Chong Chen, Guoping P Jin, and Khee Meng Koh. Triangle-free graphs with large degree. *Combinatorics, Probability and Computing*, 6(4):381–396, 1997. 1

[8] Oliver Ebsen and Mathias Schacht. Homomorphism thresholds for odd cycles. *Combinatorica*, 40(1):39–62, 2020. 1, 1

[9] Paul Erdős. On extremal problems of graphs and generalized graphs. *Israel Journal of Mathematics*, 2(3):183–190, 1964. 2

[10] Paul Erdős and Miklós Simonovits. On a valence problem in extremal graph theory. *Discrete Mathematics*, 5(4):323–334, 1973. 1

[11] Jacob Fox. A new proof of the graph removal lemma. *Annals of Mathematics*, pages 561–579, 2011. 1

[12] Jacob Fox and Yuval Wigderson. Minimum degree and the graph removal lemma. *Journal of Graph Theory*, 2021. 1, 1, 1, 1

[13] Wayne Goddard and Jeremy Lyle. Dense graphs with small clique number. *Journal of Graph Theory*, 66(4):319–331, 2011. 1

[14] Roland Häggkvist. Odd cycles of specified length in non-bipartite graphs. In *North-Holland Mathematics Studies*, volume 62, pages 89–99. Elsevier, 1982. 1

[15] Guoping Jin. Triangle-free four-chromatic graphs. *Discrete Mathematics*, 145(1-3):151–170, 1995. 1

[16] Shoham Letzter and Richard Snyder. The homomorphism threshold of $\{C_3, C_5\}$-free graphs. *Journal of Graph Theory*, 90(1):83–106, 2019. 1

[17] Tomasz Łuczak. On the structure of triangle-free graphs of large minimum degree. *Combinatorica*, 26(4):489–493, 2006. 1

[18] Tomasz Łuczak and Stéphan Thomassé. Coloring dense graphs via VC-dimension. *arXiv preprint arXiv:1007.1670*, 2010. 1

[19] Jeremy Lyle. On the chromatic number of H-free graphs of large minimum degree. *Graphs and Combinatorics*, 27(5):741–754, 2011. 1

[20] Yonatan Nakar and Dana Ron. On the testability of graph partition properties. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018. 2, 2

[21] Heiner Oberkampf and Mathias Schacht. On the structure of dense graphs with bounded clique number. *Combinatorics, Probability and Computing*, 29(5):641–649, 2020. 1

[22] Imre Z. Ruzsa and Endre Szemerédi. Triple systems with no six points carrying three triangles. In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II, Colloq. Math. Soc. János Bolyai,*, volume 18, pages 939–945. North-Holland, Amsterdam-New York, 1978. 1

[23] Maya Sankar. Homotopy and the homomorphism threshold of odd cycles. *arXiv preprint arXiv:2206.07525*, 2022. 1

[24] Carsten Thomassen. On the chromatic number of triangle-free graphs of large minimum degree. *Combinatorica*, 22(4):591–596, 2002. 1

[25] Carsten Thomassen. On the chromatic number of pentagon-free graphs of large minimum degree. *Combinatorica*, 27(2):241–243, 2007. 1

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Hamilton cycles in pseudorandom graphs

(Extended abstract)

Stefan Glock* David Munhá Correia† Benny Sudakov‡

Abstract

Finding general conditions which ensure that a graph is Hamiltonian is a central topic in graph theory. An old and well known conjecture in the area states that any d-regular n-vertex graph G whose second largest eigenvalue in absolute value $\lambda(G)$ is at most d/C, for some universal constant $C > 0$, has a Hamilton cycle. We obtain two main results which make substantial progress towards this problem. Firstly, we settle this conjecture in full when the degree d is at least a small power of n. Secondly, in the general case we show that $\lambda(G) \leq d/C(\log n)^{1/3}$ implies the existence of a Hamilton cycle, improving the 20-year old bound of $d/\log^{1-o(1)} n$ of Krivelevich and Sudakov. We use in a novel way a variety of methods, such as a robust Pósa rotation-extension technique, the Friedman-Pippenger tree embedding with rollbacks and the absorbing method, combined with additional tools and ideas.

Our results have several interesting applications, giving best bounds on the number of generators which guarantee the Hamiltonicity of random Cayley graphs, which is an important partial case of the well known Hamiltonicity conjecture of Lovász. They can also be used to improve a result of Alon and Bourgain on additive patterns in multiplicative subgroups.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-068`

*Fakultät für Informatik und Mathematik, Universität Passau, Germany. E-mail: `stefan.glock@uni-passau.de`.

†Department of Mathematics, ETH, Zürich, Switzerland. E-mail: `david.munhacanascorreia@math.ethz.ch`. Research supported in part by SNSF grant 200021_196965.

‡Department of Mathematics, ETH, Zürich, Switzerland. E-mail: `benjamin.sudakov@math.ethz.ch`.
Research supported in part by SNSF grant 200021_196965.

1 Introduction

A *Hamilton cycle* in a graph G is a cycle passing through all the vertices of G. If it exists, then G is called *Hamiltonian.* Being one of the most central notions in Graph Theory, it has been extensively studied by numerous researchers, see e.g., [1, 9, 13, 15, 16, 20, 23, 28, 31, 32, 34, 38], and the surveys [22, 33]. In particular, the problem of deciding Hamiltonicity of a graph is known to be NP-complete and thus, finding general conditions which ensure that G has a Hamilton cycle is one of the most popular topics in Graph Theory. For instance, two famous theorems of this nature are the celebrated result of Dirac [19], which states that if the minimum degree of an n-vertex graph G is at least $n/2$, then G contains a Hamilton cycle, and the criterion of Chvátal and Erdős [13] that a graph is Hamiltonian if its connectivity number is at least as large as its independence number.

In fact, most of the classical criteria for Hamiltonicity focus on rather dense graphs. A prime example of this is clearly Dirac's theorem stated above, but also the Chvátal-Erdős condition requires the graph to be relatively dense, of average degree $\Omega(\sqrt{n})$. In contrast, sufficient conditions that ensure Hamiltonicity of sparse graphs seem much more difficult to obtain. A natural starting point towards this topic is to consider sparse random graphs, to which a lot of research has been dedicated in the last 50 years. In a breakthrough paper in 1976, Pósa [38] proved that the binomial random graph model $G(n,p)$ with $p \geq C \log n/n$ for some large constant C almost surely contains a Hamilton cycle. In doing so, he invented the influential *rotation-extension technique* for finding long cycles and paths, which has found numerous further applications since then. Pósa's result was later refined by Korshunov [26] and in 1983, a more precise threshold for Hamiltonicity was obtained by Bollobás [8] and Komlós and Szemerédi [25], who independently showed that if $p = (\log n + \log\log n + \omega(1))/n$, then $G(n,p)$ is almost surely Hamiltonian. It is a standard exercise to note that this is essentially tight - indeed, if $p = (\log n + \log\log n - \omega(1))/n$, then $G(n,p)$ almost surely has a vertex with degree at most 1, and hence is not Hamiltonian. In parallel, significant attention has also been given to the Hamiltonicity of the random d-regular graph model $G_{n,d}$ - it is known that $G_{n,d}$ almost surely contains a Hamilton cycle for all values of $3 \leq d \leq n-1$. For this result, the reader is referred to Cooper, Frieze and Reed [14] and Krivelevich, Sudakov, Vu and Wormald [30] and their references.

Given the success of the study of Hamilton cycles in sparse random graphs, it became natural to then consider pseudorandom graphs, which are deterministic graphs that resemble random graphs in various important properties. A convenient way to express pseudorandomness is via spectral techniques and was introduced by Alon. An (n,d,λ)-*graph* is an n-vertex d-regular graph G whose second largest eigenvalue in absolute value, $\lambda(G)$, is such that $\lambda(G) \leq \lambda$. Roughly speaking, $\lambda(G)$ is a measure of how "smooth" the edge-distribution of G is, and the smaller its value, the closer to "random" G behaves. The reader is referred to [29] for a detailed survey concerning pseudorandom graphs.

In a rather influential paper, Krivelevich and Sudakov [27] employed Pósa's rotation-extension technique to prove the very general result that (n,d,λ)-graphs are Hamiltonian, provided λ is significantly smaller than d. Precisely, they showed that if n is sufficiently large, then

$$d/\lambda \geq \frac{1000 \log n(\log\log\log n)}{(\log\log n)^2} \tag{1}$$

guarantees that any (n, d, λ)-graph contains a Hamilton cycle. It is worth mentioning that Hefetz, Krivelevich and Szabó [23] provided a more general sufficient condition for Hamiltonicity in terms of expansion and some variant of high connectivity, yet for (n, d, λ)-graphs their condition essentially reduces to (1).

The above result on Hamiltonicity of (n, d, λ)-graphs has found numerous applications in the last 20 years towards some well-known problems, some of which we will discuss later. Given its significance and generality, it leads to the very natural and fundamental question of whether a smaller ratio of d/λ is already sufficient to imply Hamiltonicity. Krivelevich and Sudakov [27] conjectured that it should suffice that d/λ is only a large enough constant.

Conjecture 1.1. *There exists an absolute constant $C > 0$ such that any (n, d, λ)-graph with $d/\lambda \geq C$ contains a Hamilton cycle.*

2 Main results

Despite the plethora of incentives, there has been no improvement until now on the Krivelevich and Sudakov bound stated in (1). We make significant progress towards Conjecture 1.1 in two ways. First, we improve on the Krivelevich and Sudakov bound in general by showing that a spectral ratio of order $(\log n)^{1/3}$ already guarantees Hamiltonicity.

Theorem 2.1. *There exists a constant $C > 0$ such that any (n, d, λ)-graph with $d/\lambda \geq C(\log n)^{1/3}$ contains a Hamilton cycle.*

The proof of the above result will rely on the Pósa rotation-extension method with various new ideas. Namely, we will need to develop some techniques in order to use this method in a *robust* manner.

Secondly, we confirm Conjecture 1.1 in full when the degree is polynomial in the order of the graph.

Theorem 2.2. *For every constant $\alpha > 0$, there exists a constant $C > 0$ such that any (n, d, λ)-graph with $d \geq n^\alpha$ and $d/\lambda \geq C$ contains a Hamilton cycle.*

In fact, Theorem 2.2 is a corollary of a more general statement that we will prove which in particular states that (n, d, λ)-graphs with linearly many vertex-disjoint cycles are Hamiltonian.

3 Applications and related problems

Both Theorem 2.1 and Theorem 2.2 immediately yield improvements in several applications which made use of the result of Krivelevich and Sudakov. One application is an important special case of a famous open question of Lovász [35] from 1969 concerning the Hamiltonicity of a certain class of well-behaved graphs (see e.g., [17] and its references for more background on the problem).

Conjecture 3.1. *Every connected vertex-transitive graph contains a Hamilton path, and, except for five known examples, a Hamilton cycle.*

Since Cayley graphs are vertex-transitive and none of the five known exceptions in Lovász's conjecture is a Cayley graph, the conjecture in particular includes the following, which was asked much earlier in 1959 by Rapaport Strasser [39].

Conjecture 3.2. *Every connected Cayley graph is Hamiltonian.*

For these conjectures, a proof is currently out of sight. Indeed, notable progress towards them in their full generality are a result of Babai [5] that every vertex-transitive n-vertex graph contains a cycle of length $\Omega(\sqrt{n})$ (see [18] for a recent improvement) and a result of Christofides, Hladký and Máthé [12] that every vertex-transitive graph of linear minimum degree contains a Hamilton cycle.

Given this, it is natural to consider the "random" version of Conjecture 3.2. Indeed, Alon and Roichman [4] showed that in any group G, a random set S of $O(\log |G|)$ elements is such that the Cayley graph generated by them, $\Gamma(G, S)$, is almost surely connected. Therefore, a particular instance of Conjecture 3.2 is to show that $\Gamma(G, S)$ is almost surely Hamiltonian, which is itself a conjecture of Pak and Radoičić [37]. In fact, this relates directly to Conjecture 1.1 since it can be shown, generalizing the result of Alon and Roichman, that if $|S| \geq C \log |G|$ for some large constant C, then $\Gamma(G, S)$ is almost surely an (n, d, λ)-graph with $d/\lambda \geq K$ for some large constant K. Hence, Conjecture 1.1 would imply the Hamiltonicity of $\Gamma(G, S)$. Improving on several earlier results [11, 27, 36] we will show how Theorem 2.1 can be used to prove that if $|S|$ is of order $\log^{5/3} n$, then $\Gamma(G, S)$ is almost surely Hamiltonian. We will also give an improved bound on a related problem of Akbari, Etesami, Mahini, and Mahmoody [3] concerning Hamilton cycles in coloured complete graphs which use only few colours.

Another application of our results concerns one of the central themes in Additive Combinatorics, the interplay between the two operations sum and product. A well-known fact in this area is that any multiplicative subgroup A of the finite field $\mathbb{F}_q$ of size at least $q^{3/4}$ must contain two elements x, y such that $x + y$ also belongs to A. Motivated by this, Alon and Bourgain [3] studied more complex additive structures in multiplicative subgroups. In particular, they proved that when a subgroup has size $|A| \geq q^{3/4}(\log q)^{1/2-o(1)}$, then there is a cyclic ordering of the elements of A such that the sum of any two consecutive elements is also in A. Using Theorem 2.2, we can improve on Alon and Bourgain's result, showing that the additional polylog-factor can be avoided. This shows that when $|A|$ is of order $q^{3/4}$, not only does it contain $x, y, x + y \in A$ but also much more complex structures.

Finally, we give an application of our techniques to another problem related to Conjecture 3.2. Motivated by this conjecture, Pak and Radoičić [37] showed that every group G has a set of generators S of size at most $\log_2 |G|$ such that the Cayley graph $\Gamma(G, S)$

is Hamiltonian, which is optimal since there are groups that do not have generating sets of size smaller than $\log_2 |G|$. Since their proof relies on the classification of finite simple groups, they asked to find a classification-free proof of this result. Using the methods developed for the proof of Theorem 2.2 we give a classification-free proof that there is always such a set S with $|S| = O(\log n)$.

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi. First occurrence of Hamilton cycles in random graphs. *North-Holland Math. Stud.*, 115(C):173–178, 1985.

[2] S. Akbari, O. Etesami, H. Mahini, and M. Mahmoody. On rainbow cycles in edge colored complete graphs. *Australas. J. Combin.*, 37:33, 2007.

[3] N. Alon and J. Bourgain. Additive patterns in multiplicative subgroups. *Geom. Funct. Anal.*, 24:721–739, 2014.

[4] N. Alon and Y. Roichman. Random Cayley graphs and expanders. *Random Structures Algorithms*, 5(2):271–284, 1994.

[5] L. Babai. Long cycles in vertex-transitive graphs. *J. Graph Theory*, 3(3):301–304, 1979.

[6] I. Balla, A. Pokrovskiy, and B. Sudakov. A remark on Hamilton cycles with few colors. *Moscow Journal of Combinatorics and Number Theory*, 7(3):73–77, 2017.

[7] B. Bollobás. Modern graph theory. *Springer Science & Business Media*, 184, 1998.

[8] B. Bollobás. The evolution of sparse graphs. *Graph theory and combinatorics (Cambridge, 1983)*, 35–57, 1984.

[9] B. Bollobás, T. I. Fenner, and A. M. Frieze. An algorithm for finding Hamilton paths and cycles in random graphs. *Combinatorica*, 7:327–341, 1987.

[10] D. Christofides and K. Markström. Expansion properties of random Cayley graphs and vertex transitive graphs via matrix martingales. *Random Structures Algorithms*, 32(1):88–100, 2008.

[11] D. Christofides and K. Markstrom. Random Latin square graphs. *Random Structures Algorithms*, 41(1):47–65, 2012.

[12] D. Christofides, J. Hladký, and A. Máthé. Hamilton cycles in dense vertex-transitive graphs. *J. Combin. Theory Ser. B*, 109:34–72, 2014.

[13] V. Chvátal and P. Erdős. A note on Hamiltonian circuits. *Discrete Math.*, 2(2):111–113, 1972.

[14] C. Cooper, A. Frieze, and B. Reed. Random regular graphs of non-constant degree: connectivity and Hamiltonicity. *Combin. Probab. Comput.*, 11(3):249–261, 2002.

[15] B. Csaba, D. Kühn, A. Lo, D. Osthus, and A. Treglown. Proof of the 1-factorization and Hamilton decomposition conjectures. *Mem. Amer. Math. Soc.*, 244, monograph 1154, 164 pages, 2016.

[16] B. Cuckler and J. Kahn. Hamiltonian cycles in Dirac graphs. *Combinatorica*, 29:299–326, 2009.

[17] S. J. Curran and J. A. Gallian. Hamiltonian cycles and paths in Cayley graphs and digraphs—a survey. *Discrete Math.*, 156(1-3):1–18, 1996.

[18] M. DeVos. Longer cycles in vertex transitive graphs. *arXiv:2302.04255*, 2023.

[19] G. A. Dirac. Some theorems on abstract graphs. *Proc. Lond. Math. Soc.*, 3(1):69–81, 1952.

[20] A. Ferber, E. Long, and B. Sudakov. Counting Hamilton decompositions of oriented graphs. *Int. Math. Res. Not. IMRN*, (22):6908–6933, 2018.

[21] A. Frieze and M. Krivelevich. Hamilton cycles in random subgraphs of pseudo-random graphs. *Discrete Math.*, 256(1-2):137–150, 2002.

[22] R. J. Gould. Recent advances on the Hamiltonian problem: Survey III. *Graphs Combin.*, 30 :1–46, 2014.

[23] D. Hefetz, M. Krivelevich and T. Szabó. Hamilton cycles in highly connected and expanding graphs. *Combinatorica*, 29(5):547–568, 2009.

[24] F. Knox, D. Kühn, and D. Osthus. Approximate Hamilton decompositions of random graphs. *Random Structures Algorithms*, 40(2):133–149, 2012.

[25] J. Komlós and E. Szemerédi. Limit distribution for the existence of Hamiltonian cycles in a random graph. *Discrete Math.*, 43(1):55–63, 1983.

[26] A. Korshunov. Solution of a problem of Erdős and Rényi on Hamilton cycles in non-oriented graphs. *Soviet Math Dokl*, 17:760–764, 1976.

[27] M. Krivelevich and B. Sudakov. Sparse pseudo-random graphs are Hamiltonian. *J. Graph Theory*, 42(1):17–33, 2003.

[28] M. Krivelevich. The critical bias for the Hamiltonicity game is $(1 + o(1))$. *J. Amer. Math. Soc.*, 24(1):125–131, 2011.

[29] M. Krivelevich and B. Sudakov. Pseudo-random graphs. In *More sets, graphs and numbers*, pages 199–262. Springer, 2006.

[30] M. Krivelevich, B. Sudakov, V. H. Vu, and N. C. Wormald. Random regular graphs of high degree. *Random Structures Algorithms*, 18(4):346–363, 2001.

[31] M. Krivelevich, C. Lee, and B. Sudakov. Robust Hamiltonicity of Dirac graphs. *Trans. Amer. Math. Soc.*, 366(6):3095–3130, 2014.

[32] D. Kühn and D. Osthus. Hamilton decompositions of regular expanders: a proof of Kelly's conjecture for large tournaments. *Adv. Math.*, 237:62–146, 2013.

[33] D. Kühn and D. Osthus. Hamilton cycles in graphs and hypergraphs: an extremal perspective. In *Proceedings of the International Congress of Mathematicians2014, Seoul, Korea.*, 4: 381–406, 2014.

[34] D. Kühn, J. Lapinskas, D. Osthus, and V. Patel. Proof of a conjecture of Thomassen on Hamilton cycles in highly connected tournaments. *Proc. Lond. Math. Soc.*, 109(3):733–762, 2014.

[35] L. Lovász. Combinatorial structures and their applications. In *Proc. Calgary Internat. Conf., Calgary, Alberta*, 243–246, 1970.

[36] J. Meng and Q. Huang. Almost all Cayley graphs are Hamiltonian. *Acta Math. Sinica*, 12 (2):151–155, 1996.

[37] I. Pak and R. Radoičić. Hamiltonian paths in Cayley graphs. *Discrete Math.*, 309(17): 5501–5508, 2009.

[38] L. Pósa. Hamiltonian circuits in random graphs. *Discrete Math.*, 14(4):359–364, 1976.

[39] E. Rapaport Strasser. Cayley color groups and Hamilton lines. *Scripta Math.*, 24:51–58, 1959.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Random perfect matchings in regular graphs

(Extended abstract)

Bertille Granet* Felix Joos*

Abstract

We prove that in all regular robust expanders G, every edge is asymptotically equally likely contained in a uniformly chosen perfect matching M. We also show that given any fixed matching or spanning regular graph N in G, the random variable $|M \cap E(N)|$ is approximately Poisson distributed. This in particular confirms a conjecture and a question due to Spiro and Surya, and complements results due to Kahn and Kim who proved that in a regular graph every vertex is asymptotically equally likely contained in a uniformly chosen matching. Our proofs rely on the switching method and the fact that simple random walks mix rapidly in robust expanders.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-069`

1 Introduction

A remarkable result due to Kahn and Kim [5] says that in *any* d-regular graph G, the probability that a vertex is contained in a uniformly chosen matching in G is $1-(1+o_d(1))d^{-\frac{1}{2}}$. This shows that the structure of a d-regular graph has essentially no impact on the probability that a vertex is contained in a uniformly chosen matching.

In this paper we are interested in uniformly chosen *perfect* matchings. Then, surely, each vertex is contained in every perfect matching. Hence, as the statement for vertices is

*Institut für Informatik, Universität Heidelberg, 69120 Heidelberg, Germany. E-mail: `{granet,joos}@informatik.uni-heidelberg.de`. Supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 428212407 and by the DFG under Germany's Excellence Strategy EXC-2181/1 – 390900948 (the Heidelberg STRUCTURES Cluster of Excellence).

trivial, what about the probability that an edge is contained in a random perfect matching? Is each edge equally likely contained in random perfect matching? A moment of thought reveals that this is wrong in a very strong sense. In every odd-regular graph with exactly one bridge, the bridge is contained in every perfect matching, while the edges adjacent to the bridge are contained in none of the perfect matchings. Therefore, in order to avoid a trivial statement further conditions are needed.

Hall's condition for the existence of perfect matchings in bipartite graphs says that the neighbourhood of an (independent) set should be at least as large as the set itself, which is clearly also a necessary condition. Here, we assume that this property is present in a robust sense in order to avoid the trivial scenarios mentioned above. More precisely, let $\nu, \tau > 0$ and G be a graph on n vertices. Then, we define the *ν-robust neighbourhood* $RN_{\nu,G}(S)$ of a set $S \subseteq V(G)$ in G to be the set of vertices of G which have at least νn neighbours in S. We say that G is a *robust (ν, τ)-expander* if $RN_{\nu,G}(S) \geq |S| + \nu n$ for each $S \subseteq V(G)$ satisfying $\tau n \leq |S| \leq (1-\tau)n$. Robust expansion is a fairly mild assumption and consequently it proved to be useful in several situations, see for example [3, 6, 7].

We denote by $\mathcal{P}(G)$ the set of all perfect matchings in G and write $M \sim U(\mathcal{P}(G))$ to refer to a uniformly chosen matching from $\mathcal{P}(G)$. Our main result implies that such matchings M are extremely well-distributed in robust expanders.

Theorem 1. *For any $\delta > 0$, there exists $\tau > 0$ such that for all $\nu > 0$, there exists $n_0 \in \mathbb{N}$ for which the following holds. Let $n \geq n_0$ be even and $d \geq \delta n$. Then, for any d-regular robust (ν, τ)-expander G on n vertices, $M \sim U(\mathcal{P}(G))$, and $e \in E(G)$, we have*

$$\mathbb{P}[e \in M] = (1 + o_n(1))d^{-1}.$$

In fact much more is true. Fix any matching N in G, let $M \sim U(\mathcal{P}(G))$, and consider $X := |M \cap N|$. Then, linearity of expectation and Theorem 1 imply that $\mathbb{E}[X] = (1 + o_n(1))d^{-1}|N|$. Employing the heuristic that each edge is *independently* present in $M \sim U(\mathcal{P}(G))$ with probability d^{-1}, then we expect that X has a binomial distribution with parameters $|N|$ and d^{-1}. This is approximated by a Poisson distribution with parameter $d^{-1}|N|$, whenever $|N|$ grows with n. Our next result confirms this.

To this end, we define the *total variation distance* of two integer-valued random variables Y and Z as $d_{\mathrm{TV}}(Y, Z) := \frac{1}{2}\sum_{k\in\mathbb{Z}} |\mathbb{P}[Y = k] - \mathbb{P}[Z = k]|$, which measures how close two distributions are. Moreover, we write $Y \sim \mathrm{Po}(\lambda)$ if Y is a random variable which follows a Poisson distribution with parameter λ.

Theorem 2. *For any $\delta > 0$, there exists $\tau > 0$ such that for all $\nu > 0$, there exists $n_0 \in \mathbb{N}$ for which the following holds. Let $n \geq n_0$ be even and $d \geq \delta n$. Then, for any d-regular robust (ν, τ)-expander G on n vertices, $M \sim U(\mathcal{P}(G))$, any matching N in G, $X := |M \cap N|$, and $Y \sim \mathrm{Po}(d^{-1}|N|)$, we have $d_{\mathrm{TV}}(X, Y) = o_n(1)$.*

The fact that N is a matching is not crucial for our argument, however note for example that if N is a star, then X is a $\{0, 1\}$-valued random variable. Hence, X can only converge to a Poisson distribution if N is somewhat spread out. In particular, when N is a spanning

r-regular graph for some fixed r, we can derive an analogue of Theorem 2 (see Section 2), which answers a question of Spiro and Surya [8].

Theorem 2 has some interesting consequences. We define $\mathrm{pm}(G) := |\mathcal{P}(G)|$ and suppose G and M are as in Theorem 2. Let N be a perfect matching in G. Then, Theorem 2 implies that

$$\frac{\mathrm{pm}(G-N)}{\mathrm{pm}(G)} = \mathbb{P}[M \cap N = \emptyset] = (1+o_n(1))e^{-\frac{n}{2d}}.$$

For a graph G with a perfect matching, we denote by G° a subgraph of G where one perfect matching is removed. Various combinatorial problems can be expressed as determining $\frac{\mathrm{pm}(G^\circ)}{\mathrm{pm}(G)}$. For example, when $G = K_{\frac{n}{2},\frac{n}{2}}$, this ratio is equal to the probability that a random permutation of order $\frac{n}{2}$ is fixed-point-free, and it is well known that this probability equals $(1+o_n(1))e^{-1}$. The case when $G = K_n$ also has a combinatorial interpretation, see [4].

Let $K_{a\times b}$ denote the complete multipartite graph with a parts, each of size b. As an interpolation between the cases $K_{\frac{n}{2},\frac{n}{2}}$ and K_n, one may ask whether $\mathrm{pm}(K^\circ_{r\times\frac{n}{r}})(\mathrm{pm}(K_{r\times\frac{n}{r}}))^{-1}$ converges to a limit. Johnston, Kayll, and Palmer [4] formulated this as a conjecture (and conjectured the limit value). Recently this was resolved by Spiro and Surya [8]. As all these graphs are robust expanders (excluding $K_{\frac{n}{2},\frac{n}{2}}$; we discuss bipartite graphs in Section 2), Theorem 2 reproves the result due to Spiro and Surya [8].

In fact, Spiro and Surya [8] also speculate whether for any $\alpha > \frac{1}{2}$, all regular graphs G on an even number n of vertices with $\delta(G) \geq \alpha n$ satisfy $\frac{\mathrm{pm}(G^\circ)}{\mathrm{pm}(G)} \to e^{-\frac{1}{2\alpha}}$, but consider this statement far too strong to be true. As it is trivial to show that graphs on n vertices with $\delta(G) \geq (\frac{1}{2}+o_n(1))n$ are robust expanders, Theorem 2 shows that this statement is actually true.

Our proof strategy is as follows (see the full version of this article [2] for more details). Let G, M, N, and X be as in the statement of Theorem 2. We estimate the ratios of the form $\frac{\mathbb{P}[X=k]}{\mathbb{P}[X=k-1]}$ via the so-called switching method. Knowing all relevant fractions of this type already exhibits the distribution of X, which has the advantage that the probabilities $\mathbb{P}[X=k]$ do not need to be calculated directly.

The switching method is implemented as follows. Fix a positive integer k and denote by $\mathcal{M}_k$ and $\mathcal{M}_{k-1}$ the sets of perfect matchings in G which contain precisely k and $k-1$ edges of N, respectively. Then, construct an auxiliary bipartite graph H on vertex classes $\mathcal{M}_k$ and $\mathcal{M}_{k-1}$ by joining two perfect matchings $M \in \mathcal{M}_k$ and $M' \in \mathcal{M}_{k-1}$ if there is a cycle C of length 2ℓ in G which contains precisely one edge of N and alternates between edges of M and M'. (In other words, $M \in \mathcal{M}_k$ and $M' \in \mathcal{M}_{k-1}$ are adjacent in H if $N \cap M' \subseteq N \cap M$ and the extra edge in $(N \cap M) \setminus M'$ can be 'switched out' of M to obtain M' by exchanging ℓ edges of M for ℓ edges of M', where these 2ℓ edges altogether form a cycle.)

Note that if all perfect matchings in $\mathcal{M}_k$ have degree (roughly) d_k in H, while all perfect matchings in $\mathcal{M}_{k-1}$ have degree (roughly) d_{k-1}, then $d_k|\mathcal{M}_k| \approx e(H) \approx d_{k-1}|\mathcal{M}_{k-1}|$. Hence, $\frac{\mathbb{P}[X=k]}{\mathbb{P}[X=k-1]} = \frac{|\mathcal{M}_k|}{|\mathcal{M}_{k-1}|} \approx \frac{d_{k-1}}{d_k}$. Therefore, the crux of the proof consists in precisely estimating the number of such alternating cycles.

Counting the number of cycles of a certain length can be achieved using random walks as follows. Given a d-regular graph, note that the number of walks of length ℓ starting at u is precisely d^ℓ, and so the probability that a simple random walk that starts in u is in v after ℓ steps is equal to the number of walks from u to v of length ℓ divided by d^ℓ. Since simple random walks are rapidly mixing in robust expanders, one can precisely estimate such probabilities, and therefore the number of such walks. A simple counting argument can eliminate those walks which are not paths, and so we can accurately count the number of cycles of fixed length in a regular robust expander. In practice, we have to consider simple random walks that use in every second step an edge from a fixed perfect matching M. However, this additional technicality does not affect the mixing properties of such walks and so we can still precisely count them.

We remark that Spiro and Surya [8] also used the switching method, which is common for this type of problems. Our contribution is to use longer cycles and perform the analysis with Markov chains; although the intuition is that the estimations become less precise with larger cycles, we employ key properties of Markov chains to show that in fact the opposite is true. Besides the fact that our results are substantially more general, the analysis also becomes significantly shorter and cleaner.

2 Extensions

In the full version of this paper we showed that uniformly chosen perfect matchings in robust expanders contain each edge asymptotically equally likely. In fact, for a larger set of disjoint edges, these events are approximately independent. As robust expanders are a fairly large class of graphs, this in particular contains graphs G on n vertices with $\delta(G) \geq (\frac{1}{2} + o_n(1))n$, which confirms a question of Spiro and Surya [8] in a strong form.

2.1 Regular subgraphs

Spiro and Surya [8] also suggest to estimate the probability that a uniformly chosen perfect matching of Turán graphs intersects a fixed spanning r-regular subgraph.

Theorem 3. *For any $\delta > 0$, there exists $\tau > 0$ such that for all $\nu > 0$, there exists $n_0 \in \mathbb{N}$ for which the following holds. Let $n \geq n_0$ be even, $d \geq \delta n$, and let $r \leq n^{\frac{1}{50}}$ be a positive integer. Then, for any d-regular robust (ν, τ)-expander G on n vertices, $M \sim U(\mathcal{P}(G))$, any spanning r-regular subgraph N in G, $X \coloneqq |M \cap E(N)|$, and $Y \sim \mathrm{Po}(\frac{rn}{2d})$, we have $d_{\mathrm{TV}}(X, Y) = o_n(1)$.*

As a corollary, one can calculate the probability that r perfect matchings, each chosen independently and uniformly at random, are (edge-)disjoint. This relates to a problem of Ferber, Hänni, and Jain [1], which asks for the probability of selecting r edge-disjoint copies of a graph H in a host graph G. They answer this question for Hamilton cycles in the complete graph. The following corollary is an analogue for perfect matchings in the more general class of robust expanders. The proof follows immediately from Theorem 3 by induction on r.

Corollary 4. *For any $\delta > 0$, there exists $\tau > 0$ such that for all $\nu > 0$, there exists $n_0 \in \mathbb{N}$ for which the following holds. Let $n \geq n_0$ be even, $d \geq \delta n$, and $r \leq n^{\frac{1}{50}}$. Then, for any d-regular robust (ν, τ)-expander G on n vertices and independent $M_1, \ldots, M_r \sim U(\mathcal{P}(G))$, we have*

$$\mathbb{P}[M_1, \ldots, M_r \text{ are disjoint}] = (1 + o_n(1))e^{-\frac{n}{2d}\binom{r}{2}}.$$

2.2 Bipartite graphs

Of particular interest are perfect matchings in (balanced) bipartite graphs, but bipartite graphs are not robust expanders as the neighbourhood of one of the partition classes is only at most as large as the class itself. However, the notion of robust expanders can be adapted to bipartite graphs. Let G be a bipartite graph with vertex partition (A, B) and $|A| = |B| = n$. We say that G is a *bipartite robust (ν, τ)-expander* if $RN_{\nu,G}(S) \geq |S| + \nu n$ for each $S \subseteq A$ satisfying $\tau n \leq |S| \leq (1 - \tau)n$.

The following is an analogue of Theorems 1–3 for bipartite graphs. This then also includes an approximation for the number of derangements.

Theorem 5. *For any $\delta > 0$, there exists $\tau > 0$ such that for all $\nu > 0$, there exists $n_0 \in \mathbb{N}$ for which the following holds. Let $n \geq n_0$, $d \geq \delta n$, and $r \leq n^{\frac{1}{50}}$. Let G be a balanced bipartite d-regular robust (ν, τ)-expander on $2n$ vertices and suppose that N is a matching in G or a spanning r-regular subgraph of G. Let $M \sim U(\mathcal{P}(G))$, let $X := |M \cap E(N)|$, let $Y \sim \mathrm{Po}(d^{-1}e(N))$, and let $e \in E(G)$. Then, $\mathbb{P}[e \in M] = (1 + o_n(1))d^{-1}$ and $d_{\mathrm{TV}}(X, Y) = o_n(1)$.*

References

[1] A. Ferber, K. Hänni, and V. Jain. The probability of selecting k edge-disjoint Hamilton cycles in the complete graph. *arXiv:2001.01149*, 2020.

[2] B. Granet and F. Joos. Random perfect matchings in regular graphs. *arXiv:2301.10131*, 2023.

[3] V. Gruslys and S. Letzter. Cycle partitions of regular graphs. *Combin. Probab. Comput.*, 30:526–549, 2021.

[4] D. Johnston, P. M. Kayll, and C. Palmer. Deranged matchings: proofs and conjectures. *arXiv:2209.11319*, 2022.

[5] J. Kahn and J. H. Kim. Random matchings in regular graphs. *Combinatorica*, 18:201–226, 1998.

[6] D. Kühn, A. Lo, D. Osthus, and K. Staden. The robust component structure of dense regular graphs and applications. *Proc. Lond. Math. Soc.*, 110:19–56, 2015.

[7] D. Kühn and D. Osthus. Hamilton decompositions of regular expanders: a proof of Kelly's conjecture for large tournaments. *Adv. Math.*, 237:62–146, 2013.

[8] S. Spiro and E. Surya. Counting deranged matchings. *arXiv:2211.01872*, 2022.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Forcing Generalized Quasirandom Graphs Efficiently

(Extended abstract)

Andrzej Grzesik* Daniel Král'† Oleg Pikhurko‡

Abstract

We study generalized quasirandom graphs whose vertex set consists of q parts (of not necessarily the same sizes) with edges within each part and between each pair of parts distributed quasirandomly; such graphs correspond to the stochastic block model studied in statistics and network science. Lovász and Sós showed that the structure of such graphs is forced by homomorphism densities of graphs with at most $(10q)^q + q$ vertices; subsequently, Lovász refined the argument to show that graphs with $4(2q+3)^8$ vertices suffice. Our results imply that the structure of generalized quasirandom graphs with $q \geq 2$ parts is forced by homomorphism densities of graphs with at most $4q^2 - q$ vertices, and, if vertices in distinct parts have distinct degrees, then $2q+1$ vertices suffice. The latter improves the bound of $8q - 4$ due to Spencer.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-070`

1 Introduction

Quasirandom graphs play an important role in structural and extremal graph theory. The notion of quasirandom graphs can be traced to the works of Rödl [38], Thomason [42, 43]

*Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. E-mail: `Andrzej.Grzesik@uj.edu.pl`. Supported by the National Science Centre grant number 2021/42/E/ST1/00193.

†Faculty of Informatics, Masaryk University, Botanická 68A, 602 00 Brno, Czech Republic. E-mail: `kral@fi.muni.cz`. Supported by the MUNI Award in Science and Humanities (MUNI/I/1677/2018) of the Grant Agency of Masaryk University.

‡Mathematics Institute and DIMAP, University of Warwick, Coventry CV4 7AL, United Kingdom. E-mail: `o.pikhurko@warwick.ac.uk`. Supported by ERC Advanced Grant 101020255 and Leverhulme Research Project Grant RPG-2018-424.

and Chung, Graham and Wilson [9] in the 1980s, and is also deeply related to Szemerédi's Regularity Lemma [40]. Indeed, the Regularity Lemma asserts that each graph can be approximated by partitioning into a bounded number of quasirandom bipartite graphs. There is also a large body of literature concerning quasirandomness of various kinds of combinatorial structures such as groups [24], hypergraphs [5, 6, 22, 23, 28, 31, 37, 39], permutations [4, 10, 32, 33], Latin squares [11, 17, 20, 25], subsets of integers [8], tournaments [3, 7, 13, 14, 27, 26], etc. Many of these notions have been treated in a unified way in the recent paper by Coregliano and Razborov [15].

The starting point of our work is the following classical result [9] on quasirandom graphs: a sequence of graphs $(G_n)_{n\in\mathbb{N}}$ is quasirandom with density p if and only if the homomorphism densities of the single edge K_2 and the 4-cycle C_4 in $(G_n)_{n\in\mathbb{N}}$ converge to p and p^4, i.e., to their expected densities in the Erdős-Rényi random graph with density p. In particular, quasirandomness is forced by homomorphism densities of graphs with at most 4 vertices. We consider a generalization of quasirandom graphs, which corresponds to the stochastic block model in statistics. In this model, the edge density of a (large) graph is not homogeneous as in the Erdős-Rényi random graph model, however, the graph can be partitioned into q parts such that the edge density is homogeneous inside each part and between each pair of the parts. Lovász and Sós [35] established that the structure of such graphs is forced by homomorphism densities of graphs with at most $(10q)^q + q$ vertices. Lovász [34, Theorem 5.33] refined this result by showing that homomorphism densities of graphs with at most $4(2q + 3)^8$ vertices suffice. Our main result (Theorem 1) improves this bound: the structure of generalized quasirandom graphs with $q \geq 2$ parts is forced by homomorphism densities of graphs with at most $4q^2 - q$ vertices. Our line of arguments substantially differs from that in [35, 34], in particular, it is more explicit and so of a more constructive nature, which is of importance in relation to applications [2, 19, 29, 30].

Spencer [41] considered generalized quasirandom graphs with q parts with an additional assumption that vertices in distinct parts have distinct degrees, and established that the structure of such graphs is forced by homomorphism densities of graphs with at most $8q-4$ vertices. Addressing a question posed in [41], we show (Theorem 2) that graphs with at most $\max\{2q + 1, 4\}$ vertices suffice in this restricted setting.

We present our arguments using the language of the theory of graph limits, which is introduced in Section 2. We remark that similarly to arguments presented in [35, 34], although not explicitly stated there, our arguments also apply in a more general setting of kernels in addition to graphons (see Section 2 for the definitions of the two notions). In Section 3, we state our main results and sketch the main ideas of their proofs.

2 Notation

We now introduce the notions and tools from the theory of graph limits that we need to present our results; we refer the reader to the monograph by Lovász [34] for a more comprehensive introduction. We also rephrase results concerning quasirandom graphs and generalized quasirandom graphs with q parts presented in Section 1 in the language of the theory of graph limits.

If H and G are two graphs, the *homomorphism density* of H in G, denoted by $t(H,G)$, is the probability that a random mapping of the vertex set of H to the vertex set of G is a homomorphism of H to G. A sequence $(G_n)_{n\in\mathbb{N}}$ of graphs is *convergent* if the number of vertices of G_n tends to infinity and the values of $t(H,G_n)$ converge for every graph H as $n\to\infty$. A sequence $(G_n)_{n\in\mathbb{N}}$ of graphs is *quasirandom with density* p if it is convergent and the limit of $t(H,G_n)$ is equal to $p^{|E(H)|}$ for every graph H, where $E(H)$ denotes the edge set of H. If the particular value of p is irrelevant, we just say that a sequence of graphs is quasirandom instead of quasirandom with density p.

The theory of graph limits provides analytic ways of representing sequences of convergent graphs. A *kernel* is a bounded measurable function $U:[0,1]^2\to\mathbb{R}$ that is *symmetric*, i.e., $U(x,y)=U(y,x)$ for all $(x,y)\in[0,1]^2$. The points in the domain $[0,1]$ of a kernel are often referred to as *vertices*. A *graphon* is kernel whose values are restricted to $[0,1]$. The *homomorphism density* of a graph H in a kernel U is defined as follows:

$$t(H,U)=\int_{[0,1]^{V(H)}}\prod_{uv\in E(H)}U(x_u,x_v)\mathrm{d}x_{V(H)};$$

we often just briefly say the *density* of a graph H in a kernel U rather than the homomorphism density of H in U. A graphon W is a *limit* of a convergent sequence $(G_n)_{n\in\mathbb{N}}$ of graphs if $t(H,W)$ is the limit of $t(H,G_n)$ for every graph H. Every convergent sequence of graphs has a limit graphon and every graphon is a limit of a convergent sequence of graphs [36]; also see [16] for relation to exchangeable arrays. Two kernels (or graphons) U_1 and U_2 are *weakly isomorphic* if $t(H,U_1)=t(H,U_2)$ for every graph H. Note that any two limits of a convergent sequence of graphs are weakly isomorphic, and we refer particularly to [1] for results on the structure of weakly isomorphic graphons and more generally kernels.

We now revisit the notion of quasirandom graphs using the language of the theory of graph limits. Observe that a sequence of graphs is quasirandom with density p if and only if the sequence is convergent and its limit is the graphon equal to p everywhere. Hence, the following holds for every graphon W: W is weakly isomorphic to the graphon equal to p everywhere if and only if $t(K_2,W)=p$ and $t(C_4,W)=p^4$. More strongly, we say that a kernel (or graphon) U is *forced* by graphs contained in a set $\mathcal{H}$ if every kernel U' such that $t(H,U')=t(H,U)$ for every graph $H\in\mathcal{H}$ is weakly isomorphic to U. In particular, the constant graphon is forced by the graphs K_2 and C_4.

A *q-step kernel* U is a kernel such that $[0,1]$ can be partitioned to q non-null measurable sets $A_1,\dots,A_q$ such that U is constant on $A_i\times A_j$ for all $1\le i,j\le q$ but no such partition into $q-1$ parts exists; a *q-step graphon* is a q-step kernel that is also a graphon. If the value of q is not important, we just briefly say a *step kernel* or a *step graphon*. Observe that step graphons correspond to stochastic block models and so to generalized quasirandom graphs discussed in Section 1. In particular, the result of Lovász and Sós [35] mentioned in Section 1 asserts that every q-step graphon is forced by graphs with at most $(10q)^q+q$ vertices, and the result of Lovász [34, Theorem 5.33] that every q-step graphon is forced by graphs with at most $4(2q+3)^8$ vertices.

3 Results

We now state our two main results and sketch the ideas behind their proofs.

Theorem 1. *The following holds for every $q \geq 2$ and every q-step kernel U: if the density of each graph with at most $4q^2 - q$ vertices in a kernel U' is the same as in U, then the kernels U and U' are weakly isomorphic.*

To sketch the proof of Theorem 1, we need to recall the notion of a quantum graph: a *quantum graph* is a finite linear combination of graphs (called *constituents*) and the *density* of a quantum graph G in a kernel U is the linear combination of densities of graphs forming G in U with the coefficients as in G. Fix now a q-step kernel U, and let U' be another kernel such that the density of each graph with at most $4q^2-q$ vertices in U' is the same as in U. Lovász [34, Proposition 14.44] established the existence of a quantum graph Q_k with constituents having $k(k+1)$ vertices such that $t(Q_k, U'') = 0$ if and only if U'' is weakly isomorphic to a step kernel with at most $k-1$ parts. It follows that $t(Q_q, U) \neq 0$ and $t(Q_{q+1}, U) = 0$ and so $t(Q_q, U') \neq 0$ and $t(Q_{q+1}, U') = 0$, which yields that U' is a q-step kernel.

The main step of our argument is a construction of a quantum graph $P_{s_1,\dots,s_q}$ with $s_1 + \cdots + s_q$ roots, which are split into q groups of $s_1, \dots, s_q$ roots, with the following property: when each root of $P_{s_1,\dots,s_q}$ is assigned a vertex of a q-step kernel, i.e., a point of $[0,1]$, the rooted quantum graph $P_{s_1,\dots,s_q}$ evaluates to zero unless the roots in each of the q groups are chosen from the same part of the step kernel. We show that there exists a quantum rooted graph $P_{s_1,\dots,s_q}$ for each choice of parameters $s_1, \dots, s_q$ between $q+2$ and $2q+2$ such that

- each constituent of $P_{s_1,\dots,s_q}$ has at most $s_1 + \cdots + s_q + 2q(q-1)$ vertices, and
- if the roots in the same group are chosen from the same part but roots from different groups are from different parts, then the value of $P_{s_1,\dots,s_q}$ is non-zero and does not depend on the parameters $s_1, \dots, s_q$.

By introducing edges between some of the roots of $P_{s_1,\dots,s_q}$, it is possible to extract the values of the densities of U' within the q parts and between the pairs of the parts, and so these values need to be the same as the corresponding values in U. If we consider different choices of the parameters $s_1, \dots, s_q$ in addition to introducing edges between the roots, it is also possible to extract a system of q equations that determines the sizes of the parts of U' uniquely, which yields that the kernels U and U' are weakly isomorphic. Finally, the analysis of the range of parameters $s_1, \dots, s_q$ needed in the argument yields the bound given in Theorem 1 on the number of vertices of graphs that need to be considered.

To state our second result, recall that if U is a kernel and $x \in [0,1]$ is a vertex of U, then the *degree* of x is

$$\int_{[0,1]} U(x,y)\mathrm{d}y.$$

Theorem 2. *The following holds for every $q \geq 2$ and every q-step kernel U such that the degrees of vertices in different parts are different: if the density of each graph with at most $2q+1$ vertices in a kernel U' is the same as in U, then the kernels U and U' are weakly isomorphic.*

We now sketch the proof of Theorem 2. Fix $q \geq 2$ and a q-step kernel U with properties given in the statement of Theorem 2 and let U' be another kernel such that the density of each graph with at most $2q+1$ vertices in U' is the same as in U. To prove Theorem 2, we construct for every choice of reals $d_1, \ldots, d_q \in \mathbb{R}$ a quantum graph $G_{d_1,\ldots,d_q}$ with $2q+1$ vertices such that the density of $G_{d_1,\ldots,d_q}$ in a kernel is zero if and only if the degree of almost every vertex of the kernel is equal to one of the values $d_1, \ldots, d_q$. The assumption of Theorem 2 now yields that the sets of the degrees of the vertices of the kernels U and U' are the same. We next construct a quantum graph with q vertices, one of them being a root, which forces the root to be from a part of a step kernel with a specific degree. These rooted quantum graphs are then used to force the sizes of the parts, the densities within the parts and between all pairs of the parts. Finally, we use the fact that a step kernel (see [12, Lemma 11], also see [34, Proposition 14.14]) is the minimizer of the density of C_4 among all partitioned kernels with same sizes of the parts, densities within the parts and between the pairs of the parts, to conclude that the kernels U and U' are weakly isomorphic.

We conclude by stating as an open problem whether it suffices in Theorem 1 to consider homomorphism densities of graphs with $o(q^2)$ vertices. To supplement the open problem, we show that the order of graphs needs to be at least linear in q. Our argument is similar to that used in analogous scenarios, e.g., in [18, 21]. For reals $a_1, \ldots, a_q > 0$ such that $a_1 + \cdots + a_q < 1$, let $U_{a_1,\ldots,a_q}$ be the $(q+1)$-step graphon with parts whose sizes are $a_1, \ldots, a_q$ and $1 - a_1 - \cdots - a_q$, and that is equal to one within each of the first q parts and to zero elsewhere. Observe that if H is a graph that, after removing isolated vertices, consists of k components with respectively $n_1, \ldots, n_k$ vertices then

$$t\left(H, U_{a_1,\ldots,a_q}\right) = \prod_{i=1}^{k} \sum_{j=1}^{q} a_j^{n_i}.$$

It follows that if

$$t\left(K_{\ell+1}, U_{a_1,\ldots,a_q}\right) = t\left(K_{\ell+1}, U_{a'_1,\ldots,a'_q}\right) \text{ for every } \ell = 1, \ldots, q-1, \tag{1}$$

then the homomorphism density of every graph with at most q vertices is the same in $U_{a_1,\ldots,a_q}$ and in $U_{a'_1,\ldots,a'_q}$. View $(t(K_{\ell+1}, U_{a_1,\ldots,a_q}))_{\ell=1}^{q-1} \in \mathbb{R}^{q-1}$ as a function of $a_1, \ldots, a_{q-1}$. If its arguments $a_1, \ldots, a_{q-1}$ are distinct, then the Jacobian matrix can be shown to be invertible and the Implicit Function Theorem gives, for every a'_q sufficiently close to a_q, a vector $(a'_1, \ldots, a'_{q-1})$ close to $(a_1, \ldots, a_{q-1})$ such that (1) holds. It follows that there are two non-weakly-isomorphic $(q+1)$-step graphons that have the same homomorphism density of every graph with at most q vertices.

References

[1] C. Borgs, J. Chayes and L. Lovász: *Moments of two-variable functions and the uniqueness of graph limits*, Geom. Funct. Anal. **19** (2010), 1597–1619.

[2] C. Borgs, J. T. Chayes, H. Cohn and S. Ganguly: *Consistent nonparametric estimation for heavy-tailed sparse graphs*, Ann. Statist. **49** (2021), 1904–1930.

[3] M. Bucić, E. Long, A. Shapira and B. Sudakov: *Tournament quasirandomness from local counting*, Combinatorica **41** (2021), 175–208.

[4] T. Chan, D. Král', J. A. Noel, Y. Pehova, M. Sharifzadeh and J. Volec: *Characterization of quasirandom permutations by a pattern sum*, Random Structures & Algorithms **57** (2020), 920–939.

[5] F. R. K. Chung and R. L. Graham: *Quasi-random hypergraphs*, Random Structures & Algorithms **1** (1990), 105–124.

[6] F. R. K. Chung and R. L. Graham: *Quasi-random set systems*, Journal of the American Mathematical Society **4** (1991), 151–196.

[7] F. R. K. Chung and R. L. Graham: *Quasi-random tournaments*, Journal of Graph Theory **15** (1991), 173–198.

[8] F. R. K. Chung and R. L. Graham: *Quasi-random subsets of Z_n*, Journal of Combinatorial Theory, Series A **61** (1992), 64–86.

[9] F. R. K. Chung, R. L. Graham and R. M. Wilson: *Quasi-random graphs*, Combinatorica **9** (1989), 345–362.

[10] J. N. Cooper: *Quasirandom permutations*, Journal of Combinatorial Theory, Series A **106** (2004), 123–143.

[11] J. W. Cooper, D. Král', A. Lamaison and S. Mohr: *Quasirandom Latin squares*, Random Structures & Algorithms **61** (2022), 298–308.

[12] J. W. Cooper, D. Král' and T. Martins: *Finitely forcible graph limits are universal*, Adv. Math. **340** (2018), 819–854

[13] L. N. Coregliano, R. F. Parente and C. M. Sato: *On the maximum density of fixed strongly connected subtournaments*, The Electronic Journal of Combinatorics **26** (2019), P1.44.

[14] L. N. Coregliano and A. A. Razborov: *On the density of transitive tournaments*, Journal of Graph Theory **85** (2017), 12–21.

[15] L. N. Coregliano and A. A. Razborov: *Natural quasirandomness properties*, preprint arXiv:2012.11773 (2020).

[16] P. Diaconis and S. Janson: *Graph limits and exchangeable random graphs*, Rend. Mat. Appl. **28** (2008), 33–61.

[17] S. Eberhard, F. Manners and R. Mrazović: *Transversals in quasirandom Latin squares*, preprint arXiv:2209.02180 (2022).

[18] P. Erdős, L. Lovász and J. Spencer: *Strong independence of graphcopy functions*, Graph theory and related topics (1979), 165–172.

[19] C. Gao, Y. Lu and H. H. Zhou: *Rate-optimal graphon estimation*, Ann. Statist. **43** (2015), 2624–2652.

[20] F. Garbe, R. Hancock, J. Hladký and M. Sharifzadeh: *Limits of Latin squares*, preprint arXiv:2010.07854 (2020).

[21] R. Glebov, A. Grzesik, T. Klimošová and D. Kráľ: *Finitely forcible graphons and permutons*, J. Combin. Theory Ser. B **110** (2015), 112–135.

[22] W. T. Gowers: *Quasirandomness, counting and regularity for 3-uniform hypergraphs*, Combinatorics, Probability and Computing **15** (2006), 143–184.

[23] W. T. Gowers: *Hypergraph regularity and the multidimensional Szemerédi theorem*, Annals of Mathematics, Second Series **166** (2007), 897–946.

[24] W. T. Gowers: *Quasirandom groups*, Combinatorics, Probability and Computing **17** (2008), 363–387.

[25] W. T. Gowers and J. Long: *Partial associativity and rough approximate groups*, Geometric and Functional Analysis **30** (2020), 1–65.

[26] A. Grzesik, D. Il'kovič, B. Kielak and D. Kráľ: *Quasirandom forcing orientations of cycles*, preprint arXiv:2212.09343 (2022).

[27] R. Hancock, A. Kabela, D. Kráľ, T. Martins, R. Parente, F. Skerman and J. Volec: *No additional tournaments are quasirandom-forcing*, European Journal of Combinatorics **108** (2023), 103632.

[28] J. Haviland and A. Thomason: *Pseudo-random hypergraphs*, Discrete Math. **75** (1989), 255–278.

[29] O. Klopp, A. B. Tsybakov and N. Verzelen: *Oracle inequalities for network models and sparse graphon estimation*, Ann. Statist. **45** (2017), 316–354.

[30] O. Klopp and N. Verzelen: *Optimal graphon estimation in cut distance*, Probab. Theory Related Fields **174** (2019), 1033–1090.

[31] Y. Kohayakawa, V. Rödl and J. Skokan: *Hypergraphs, quasi-randomness, and conditions for regularity*, Journal of Combinatorial Theory, Series A **97** (2002), 307–352.

[32] D. Král' and O. Pikhurko: *Quasirandom permutations are characterized by 4-point densities*, Geometric and Functional Analysis **23** (2013), 570–579.

[33] M. Kurečka: *Lower bound on the size of a quasirandom forcing set of permutations*, Combinatorics, Probability and Computing **31** (2022), 304–319.

[34] L. Lovász: Large Networks and Graph Limits, *Colloquium Publications*, volume 60, 2012.

[35] L. Lovász and V. Sós: *Generalized quasirandom graphs*, J. Combin. Theory Ser. B **98** (2008), 146–163.

[36] L. Lovász and B. Szegedy: *Limits of dense graph sequences*, J. Combin. Theory Ser. B **96** (2006), 933–957.

[37] B. Nagle, V. Rödl and M. Schaht: *The counting lemma for regular k-uniform hypergraphs*, Random Structures & Algorithms **28** (2006), 113–179.

[38] V. Rödl: *On universality of graphs with uniformly distributed edges*, Discrete Mathematics **59** (1986), 125–134.

[39] V. Rödl and J. Skokan: *Regularity lemma fork-uniform hypergraphs*, Random Structures & Algorithms **25** (2004), 1–42.

[40] M. Simonovits and V. T. Sós: *Szemerédi's partition and quasirandomness*, Random Structures & Algorithms **2** (1991), 1–10.

[41] J. Spencer: *Quasirandom multitype graphs*, in: An irregular mind, *Bolyai Soc. Math. Stud.*, volume 21 (2010), 607–617.

[42] A. Thomason: *Pseudo-random graphs*, Annals of Discrete Mathematics **144** (1987), 307–331.

[43] A. Thomason: *Random graphs, strongly regular graphs and pseudo-random graphs*, in: Surveys in Combinatorics, *London Mathematical Society Lecture Note Series*, volume 123 (1987), 173–196.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Refined list version of Hadwiger's Conjecture

(Extended abstract)

Yangyan Gu[*] Yiting Jiang[†] David R. Wood[‡] Xuding Zhu[§]

Abstract

Assume $\lambda = \{k_1, k_2, \ldots, k_q\}$ is a partition of $k_\lambda = \sum_{i=1}^{q} k_i$. A λ-list assignment of G is a k_λ-list assignment L of G such that the colour set $\bigcup_{v\in V(G)} L(v)$ can be partitioned into $|\lambda| = q$ sets $C_1, C_2, \ldots, C_q$ such that for each i and each vertex v of G, $|L(v) \cap C_i| \geq k_i$. We say G is *λ-choosable* if G is L-colourable for any λ-list assignment L of G. The concept of λ-choosability is a refinement of choosability that puts k-choosability and k-colourability in the same framework. If $|\lambda|$ is close to k_λ, then λ-choosability is close to k_λ-colourability; if $|\lambda|$ is close to 1, then λ-choosability is close to k_λ-choosability. This paper studies Hadwiger's Conjecture in the context of λ-choosability. Hadwiger's Conjecture is equivalent to saying that every K_t-minor-free graph is $\{1 \star (t-1)\}$-choosable for any positive integer t. We prove that for $t \geq 5$, for any partition λ of $t-1$ other than $\{1 \star (t-1)\}$, there is a K_t-minor-free graph G that is not λ-choosable. We then construct several types of K_t-minor-free graphs that are not λ-choosable, where $k_\lambda - (t-1)$ gets larger as $k_\lambda - |\lambda|$ gets larger.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-071`

1 Introduction

Given graphs H and G, we say H is *a minor* of G (or G has *an H-minor*) if a graph isomorphic to H can be obtained from a subgraph of G by contracting edges. Let K_t

[*]School of Mathematical Sciences, Zhejiang Normal University, China. E-mail: `yangyan@zjnu.edu.cn`.
[†]School of Mathematical Sciences, Nanjing Normal University, China. E-mail: `ytjiang@njnu.edu.cn`.
[‡]School of Mathematics, Monash University, Melbourne, Australia. E-mail: `david.wood@monash.edu`.
Research supported by the Australian Research Council.
[§]School of Mathematical Sciences, Zhejiang Normal University, China. E-mail: `xdzhu@zjnu.edu.cn`.
Supported by National Natural Science Foundation of China grant NSFC 11971438 and U20A2068.

be the t-vertex complete graph. A graph G is *K_t-minor-free* if G has no K_t-minor. In 1943, Hadwiger [8] conjectured the following upper bound on the chromatic number of K_t-minor-free graphs:

Conjecture 1 (Hadwiger's Conjecture)**.** *For every integer $t \geq 1$, every K_t-minor-free graph is $(t-1)$-colourable.*

This conjecture is a deep generalization of the Four Colour Theorem, and has motivated many developments in graph colouring and graph minor theory. Hadwiger [8] and Dirac [6] independently showed that Hadwiger's Conjecture holds for $t \leq 4$. Wagner [27] proved that for $t = 5$ the conjecture is equivalent to the Four Colour Theorem, which was subsequently proved by Appel, Haken and Koch [2,3] and Robertson, Sanders, Seymour and Thomas [20], both using extensive computer assistance. Robertson, Seymour and Thomas [21] went one step further and proved Hadwiger's Conjecture for $t = 6$, also by reducing it to the Four Colour Theorem. The conjecture for $t \geq 7$ is open and seems to be extremely challenging. For more on Hadwiger's Conjecture, see the survey of Seymour [23].

The evident difficulty of Hadwiger's Conjecture has inspired many researchers to study the following natural weakening (cf. [9, 10, 19]):

Conjecture 2 (Linear Hadwiger's Conjecture)**.** *There exists a constant $C > 0$ such that for every integer $t \geq 1$, every K_t-minor-free graph is Ct-colourable.*

For many decades, the best general upper bound on the chromatic number of K_t-minor-free graphs was $O(t\sqrt{\log t})$, which was proved independently by Kostochka [12, 13] and Thomason [24] in the 1980s. In 2019, Norine, Postle and Song [15] broke this barrier, and proved that the maximum chromatic number of K_t-minor-free graphs is in $O(t(\log t)^{1/4+o(1)})$. Following a series of improvements, [14, 16–18] the best known bound is $O(t \log\log t)$ due to Delcourt and Postle [5].

A list assignment of a graph G is a mapping L that assigns to each vertex v of G a set $L(v)$ of permissible colours. An *L-colouring* of G is a proper colouring f of G such that for each vertex v of G, $f(v) \in L(v)$. We say G is *L-colourable* if G has an L-colouring. A *k-list assignment* of G is a list assignment L with $|L(v)| \geq k$ for each vertex v. We say G is *k-choosable* if G is L-colourable for any k-list assignment L of G. The *choice-number* of G is the minimum integer k such that G is k-choosable.

Hadwiger's Conjecture is also widely considered in the setting of list colourings. Voigt [26] constructed planar graphs (hence K_5-minor-free) with choice-number 5. Hence the list version of Hadwiger's Conjecture is false. Nevertheless, the list version of Linear Hadwiger's Conjecture, proposed by Kawarabayashi and Mohar [10] in 2007, remains open.

Conjecture 3 (List Hadwiger's Conjecture)**.** *There exists a constant $C > 0$ such that for every integer $t \geq 1$, every K_t-minor-free graph is Ct-choosable.*

The current state-of-the-art upper bound on the choice-number of K_t-minor-free graphs is $O(t(\log\log t)^6)$ [18]. If Conjecture 3 is true, then a natural problem is to determine the minimum value of C. Barát, Joret and Wood [4] constructed K_t-minor-free graphs that are not $4(t-3)/3$-choosable, implying $C \geq \frac{4}{3}$ in Conjecture 3. Improving upon this result, Steiner [22] recently proved that the maximum choice-number of K_t-minor-free graphs is at least $2t - o(t)$, and hence $C \geq 2$ in Conjecture 3.

This paper considers Hadwiger's Conjecture in the context of λ-choosability, which was introduced by Zhu [28]. In general, k-colourability and k-choosability behave very differently. Indeed, bipartite graphs can have arbitrary large choice-number. λ-choosability is a refinement of the concept of choosability that puts k-choosability and k-colourability in the same framework and considers a more complex hierarchy of colouring parameters.

Definition 1. *Let $\lambda = \{k_1, k_2, \ldots, k_q\}$ be a multiset of positive integers. Let $k_\lambda = \sum_{i=1}^{q} k_i$ and $|\lambda| = q$. A* λ-list assignment *of G is a list assignment L such that the colour set $\bigcup_{v \in V(G)} L(v)$ can be partitioned into q sets $C_1, C_2, \ldots, C_q$ such that for each i and each vertex v of G, $|L(v) \cap C_i| \geq k_i$. We say G is* λ-choosable *if G is L-colourable for any λ-list assignment L of G.*

Note that for each vertex v, $|L(v)| \geq \sum_{i=1}^{q} k_i = k_\lambda$. So a λ-list assignment L is a k_λ-list assignment with some restrictions on the set of possible lists.

For a positive integer a, let $m_\lambda(a)$ be the multiplicity o f a in λ. If $m_\lambda(a) = m$, then instead of writing m times the integer a, we write $a \star m$. For example, $\lambda = \{1 \star k_1, 2 \star k_2, 3\}$ means that λ is the multiset consisting of k_1 copies of 1, k_2 copies of 2 and one copy of 3. If $\lambda = \{k\}$, then λ-choosability is the same as k-choosability; if $\lambda = \{1 \star k\}$, then λ-choosability is equivalent to k-colourability. So the concept of λ-choosability puts k-choosability and k-colourability in the same framework.

For $\lambda = \{k_1, k_2, \ldots, k_q\}$ and $\lambda' = \{k'_1, k'_2, \ldots, k'_p\}$, we say λ' is a *refinement* of λ if $p \geq q$ and there is a partition $I_1, I_2, \ldots, I_q$ of $\{1, 2, \ldots, p\}$ such that $\sum_{j \in I_t} k'_j = k_t$ for $t = 1, 2, \ldots, q$. We say λ' is obtained from λ by *increasing* some parts if $p = q$ and $k_t \leq k'_t$ for $t = 1, 2, \ldots, q$. We write $\lambda \leq \lambda'$ if λ' is obtained from a refinement of λ by increasing some parts. It follows from the definitions that if $\lambda \leq \lambda'$, then every λ-choosable graph is λ'-choosable. Conversely, Zhu [28] proved that if $\lambda \nleq \lambda'$, then there is a λ-choosable graph that is not λ'-choosable. In particular, λ-choosability implies k_λ-colourability, and if $\lambda \neq \{1 \star k_\lambda\}$, then there are k_λ-colourable graphs that are not λ-choosable.

All the partitions λ of a positive integer k are sandwiched between $\{k\}$ and $\{1 \star k\}$ in the above order. As observed above, $\{k\}$-choosability is the same as k-choosability, and $\{1 \star k\}$-choosability is equivalent to k-colourability. By considering each partition λ of k, λ-choosability provides a complex hierarchy of colouring parameters that interpolate between k-colourability and k-choosability.

The framework of λ-choosability provides room to explore strengthenings of colourability and choosability results. For example, Kermnitz and Voigt [11] proved that there are planar graphs that are not $\{1, 1, 2\}$-choosable. This result strengthens Voigt's result that there are non-4-choosable planar graphs, and shows that the Four Colour Theorem is sharp in the sense that for any partition λ of 4 other than $\{1 \star 4\}$, there is a planar graph that is not λ-choosable. This paper considers Hadwiger's Conjecture in the context of λ-choosability.

2 Results

This paper constructs several examples of K_t-minor-free graphs that are not λ-choosable where $k_\lambda \geq t-1$ and q is close to k_λ. In particular, if the multiplcity of 1 in λ is large enough, then the number of parts of λ will be close to k_λ.

First we strengthen the above-mentioned result of Kermnitz and Voigt to K_t-minor-free graphs for $t \geq 5$ as follows:

Theorem 1. *For every integer $t \geq 5$, there exists a K_t-minor-free graph that is not $\{1 \star (t-3), 2\}$-choosable.*

If λ is a partition of $t-1$ other than $\{1 \star (t-1)\}$, then $\{1 \star (t-3), 2\}$ is a refinement of λ. Hence we have the following corollary.

Corollary 2. *If λ is a partition of $t-1$ other than $\{1 \star (t-1)\}$, then there is a K_t-minor-free graph that is not λ-choosable.*

For a multiset λ of positive integers, let $h(\lambda)$ be the maximum t such that every K_t-minor-free graph is λ-choosable. Since $K_{k_\lambda+1}$ is not k_λ-colourable and hence not λ-choosable, we know that $h(\lambda) \leq k_\lambda + 1$.

For a multiset λ of positive integers, $k_\lambda - |\lambda|$ measures the "distance" of λ-choosability from k_λ-colourability. Hadwiger's Conjecture says that if $k_\lambda - |\lambda| = 0$, then $h(\lambda) = k_\lambda + 1$. By Theorem 1, if $k_\lambda - |\lambda| \geq 1$, then $h(\lambda) \leq k_\lambda$, provided that $k_\lambda \geq 5$. It seems natural that if $k_\lambda - |\lambda|$ gets bigger, then $k_\lambda - h(\lambda)$ also gets bigger, provided that k_λ is sufficiently large. The next result shows this is true for various λ.

Theorem 3. *For each integer $a \geq 0$, there exists an integer $t_1 = t_1(a)$ such that for every integer $t \geq t_1$, there exists a K_t-minor-free graph that is not $\{1 \star (t-2a-6), 3a+6\}$-choosable.*

For the λ in Theorem 3, $k_\lambda = t+a$, $h(\lambda) \leq t-1$ and $|\lambda| = t-(2a+5)$. As $k_\lambda - |\lambda| = 3a+5$ tends to infinity, the difference $k_\lambda - h(\lambda) \geq a+1$ also tends to infinity, provided that $k_\lambda \geq \phi(k_\lambda - |\lambda|)$, where ϕ is a certain given function. It remains open whether such a conclusion holds for all λ. We conjecture a positive answer.

Conjecture 4. *There are functions $\phi, \psi : \mathbb{N} \to \mathbb{N}$ for which the following hold:*

- $\lim_{n\to\infty} \psi(n) = \infty$.
- *For any multiset λ of positive integers, if $k_\lambda \geq \phi(k_\lambda - |\lambda|)$, then $k_\lambda - h(\lambda) \geq \psi(k_\lambda - |\lambda|)$.*

It is easy to see that if $k_\lambda - |\lambda| = b$, then $\{1 \star (k_\lambda - 2b'), 2 \star b'\}$ is a refinement of λ, where $b \geq b' \geq b/2$. Thus to prove Conjecture 4, it suffices to prove it for λ of the form $\{1 \star k_1, 2 \star k_2\}$.

Theorem 4 below shows that Conjecture 4 holds for any λ of the form $\{1 \star k_1, 3 \star k_2\}$.

Theorem 4. *For each integer $a \geq 0$, there exists an integer $t_2 = t_2(a)$ such that for every integer $t \geq t_2$, there exists a K_t-minor-free graph that is not $\{1 \star (t-5a-9), 3 \star (2a+3)\}$-choosable.*

As $|\lambda|$ becomes very small compared to k_λ, say $|\lambda|$ is constant and k_λ tends to infinity, then λ-choosability becomes very close to k_λ-choosability. The following result, which generalizes the main result of Steiner [22], deals with such λ.

Theorem 5. *For every $\varepsilon \in (0,1)$ and $q \in \mathbb{N}$, there exists an integer $t_3 = t_3(q,\varepsilon)$ such that for every integer $t \geq t_3$ and $k_1, k_2, \ldots, k_q \in \mathbb{N}$ satisfying*

$$\sum_{j=1}^{q} k_j \leq (2-\varepsilon)t,$$

there exists a K_t-minor-free graph G that is not $\{k_1, k_2, \ldots, k_q\}$-choosable.

The $q = 1$ case of above theorem was proved by Steiner [22].

Acknowledgements

This research was partially completed at the Structural Graph Theory Downunder (https://www.matrix-inst.org.au/events/structural-graph-theory-downunder/) workshop at the Mathematical Research Institute MATRIX (November 2019). Xuding Zhu is partially supported by NSFC Grants 11971438, U20A2068, and ZJNSFC grant LD19A010001.

References

[1] N. Alon, Zs. Tuza, M. Voigt. *Choosability and fractional chromatic numbers*, Discrete Math. 165/166, 31–38, 1997.

[2] K. Appel, W. Haken. *Every planar map is four colourable. I. Discharging.* Illinois J. Math., 21(3):429–490, 1977.

[3] K. Appel, W. Haken, J. Koch. *Every planar map is four colourable. II. Reducibility.* Illinois J. Math., 21(3):491–567, 1977.

[4] J. Barát, G. Joret, D.R. Wood. *Disproof of the list Hadwiger conjecture.* Electronic J. Combin., 18(1), P232, 2011.

[5] M. Delcourt, L. Postle. *Reducing linear Hadwiger's conjecture to coloring small graphs.* arXiv:2108.01633.

[6] G. A. Dirac. *A property of 4-chromatic graphs and some remarks on critical graphs.* J. London Math. Soc., 27:85–92, 1952.

[7] P. Erdős, A.L. Rubin, H. Taylor, *Choosability in graphs*, in: Proc. West Coast Conf. on Combinatorics, Graph Theory and Computing, Congressus Numerantium XXVI, 1979, pp. 125–157.

[8] H. Hadwiger. *Über eine Klassifikation der Streckenkomplexe*. Vierteljschr. Naturforsch. Ges. Zürich, 88:133–142, 1943.

[9] K. Kawarabayashi. *On the connectivity of minimum and minimal counterexamples to Hadwiger's Conjecture*. J. Combin. Theory Ser. B, 97(1):144–150, 2007.

[10] K. Kawarabayashi, B. Mohar. *Some recent progress and applications in graph minor theory*. Graphs Combin., 23(1):1–46, 2007.

[11] A. Kemnitz, M. Voigt. *A note on non-4-list colourable planar graphs*, Electron. J. Combin. 25(2), #P2.46, 2018.

[12] A. V. Kostochka. *The minimum Hadwiger number for graphs with a given mean degree of vertices*. Metody Diskret. Analiz., (38):37–58, 1982.

[13] A. V. Kostochka. *Lower bound of the Hadwiger number of graphs by their average degree*. Combinatorica, 4(4):307–316, 1984.

[14] S. Norin, L. Postle. *Connectivity and choosability of graphs with no K_t minor*. J. Combin. Theory, Ser. B 158(1):283–300, 2021.

[15] S. Norin, L. Postle, and Z. Song. *Breaking the degeneracy barrier for colouring graphs with no K_t minor*. arXiv:1910.09378v2, 2019.

[16] L. Postle. *Further progress towards Hadwiger's conjecture*. arXiv:2006.11798, 2020.

[17] L. Postle. *An even better density increment theorem and its application to Hadwiger's conjecture*. arXiv:2006.14945, 2020.

[18] L. Postle. *Further progress towards the list and odd versions of Hadwiger's conjecture*. arXiv:2010.05999, 2020.

[19] B. Reed, P. Seymour. *Fractional colouring and Hadwiger's conjecture*. J. Combin. Theory Ser. B, 74(2):147–152, 1998.

[20] N. Robertson, D.P. Sanders, P. Seymour, R. Thomas. *The four-colour theorem*. J. Combin. Theory Ser. B, 70(1):2–44, 1997.

[21] N. Robertson, P. Seymour, R. Thomas. *Hadwiger's conjecture for K_6-free graphs*. Combinatorica, 13(3):279–361, 1993.

[22] R. Steiner. *Improved lower bound for the list chromatic number of graphs with no K_t minor*. Combin. Probab. Comput. 32(2): 326–333, 2023.

[23] P. Seymour, *Hadwiger's conjecture*. In: Open Problems in Mathematics, Eds. J. F. Nash Jr. and M. Th. Rassias, 417–437, Springer, 2016.

[24] A. Thomason. *An extremal function for contractions of graphs*. Math. Proc. Cambridge Philos. Soc., 95(2):261–265, 1984.

[25] V.G. Vizing, *Coloring the vertices of a graph in prescribed colours*, Diskret. Analiz 29 Metody Diskret. Anal. v Teorii Kodov i Shem (1976) 3–10, p. 101, (in Russian).

[26] M. Voigt. *List colourings of planar graphs*. Discrete Math., 120 (1–3), 215–219, 1993.

[27] K. Wagner. *Über eine Eigenschaft der ebenen Komplexe*. Mathematische Annalen, 114:570–590, 1937.

[28] X. Zhu, *A refinement of choosability of graphs*. J. Combin. Theory, Ser. B 141 (2020) 143–164.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

A GENERAL APPROACH TO TRANSVERSAL VERSIONS OF DIRAC-TYPE THEOREMS

(EXTENDED ABSTRACT)

Pranshu Gupta* Fabian Hamann† Alp Müyesser‡ Olaf Parczyk§
Amedeo Sgueglia‡

Abstract

Given a collection of hypergraphs $\mathbf{H} = (H_1, \ldots, H_m)$ with the same vertex set, an m-edge graph $F \subset \cup_{i\in[m]} H_i$ is a transversal if there is a bijection $\phi : E(F) \to [m]$ such that $e \in E(H_{\phi(e)})$ for each $e \in E(F)$. How large does the minimum degree of each H_i need to be so that $\mathbf{H}$ necessarily contains a copy of F that is a transversal? Each H_i in the collection could be the same hypergraph, hence the minimum degree of each H_i needs to be large enough to ensure that $F \subseteq H_i$. Since its general introduction by Joos and Kim [Bull. Lond. Math. Soc., 2020, 52(3): 498–504], a growing body of work has shown that in many cases this lower bound is tight. In this paper, we give a unified approach to this problem by providing a widely applicable sufficient condition for this lower bound to be asymptotically tight. This is general enough to recover many previous results in the area and obtain novel transversal variants of several classical Dirac-type results for (powers of) Hamilton cycles. For example, we derive that any collection of rn graphs on an n-vertex set, each with minimum degree at least $(r/(r+1)+o(1))n$, contains a transversal copy of the r-th power of a Hamilton

*University of Passau, Faculty of Computer Science and Mathematics, Passau, Germany. *E-mail:* pranshu.gupta@uni-passau.de. This research was conducted while PG was affiliated with the Hamburg University of Technology.

†Hamburg University of Technology, Institute of Mathematics, Am Schwarzenberg-Campus 3, 21073 Hamburg, Germany. *E-mail:* fabian.hamann@tuhh.de.

‡Department of Mathematics, University College London, London WC1E 6BT, UK. *E-mail:* {alp.muyesser.21|a.sgueglia}@ucl.ac.uk. This research was conducted while AS was a PhD student at the London School of Economics.

§Department of Mathematics and Computer Science, Freie Universität Berlin, Arnimallee 3, 14195 Berlin, Germany. *E-mail:* parczyk@mi.fu-berlin.de.

cycle. This can be viewed as a rainbow version of the Pósa-Seymour conjecture.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-072`

1 Introduction

Given an integer $m \geq 1$, we say that $\mathbf{H} = (H_1, \dots, H_m)$ is a *hypergraph collection on vertex set* V if, for each $i \in [m]$, the hypergraph H_i has vertex set V. We call the collection a *graph collection* if each hypergraph in the collection has uniformity two. Given an m-edge hypergraph F on V, we say that $\mathbf{H}$ has a *transversal* copy of F if there is a bijection $\phi\colon E(F) \to [m]$ such that $e \in H_{\phi(e)}$ for each $e \in E(F)$. We will also use the adjective *rainbow* for a transversal copy of F. Indeed, we can think of the edges of hypergraph H_i to be coloured with colour i and, in this framework, a transversal copy of F is a copy of F in $\bigcup_{i\in[m]} H_i$ with edges of pairwise distinct colours. We are interested in the following general question formulated originally by Joos and Kim [6].

Question 1. Let F be an m-edge hypergraph with vertex set V, $\mathcal{H}$ be a family of hypergraphs, and $\mathbf{H} = (H_1, \dots, H_m)$ be a hypergraph collection on vertex set V with $H_i \in \mathcal{H}$ for each $i \in [m]$. Which conditions on $\mathcal{H}$ guarantee a transversal copy of F in $\mathbf{H}$?

By taking $H_1 = H_2 = \cdots = H_m$, it is clear that such a property needs to guarantee that each hypergraph in $\mathcal{H}$ contains F as a subhypergraph. However, this alone is not always sufficient, not even asymptotically. For example, Aharoni, DeVos, de la Maza, Montejano and Šámal [1] showed that if $\mathbf{G} = (G_1, G_2, G_3)$ is a graph collection on $[n]$ with $e(G_i) > (\frac{26-2\sqrt{7}}{81})n^2$ for each $i \in [3]$, then $\mathbf{G}$ contains a transversal which is a triangle. As shown in [1], the constant $\frac{26-2\sqrt{7}}{81} > 1/4$ is optimal. On the other hand, Mantel's theorem states that any graph with at least $n^2/4$ edges must contain a triangle.

Instead of a lower bound on the total number of edges, it is also natural to investigate what can be guaranteed with a lower bound on the minimum degree. It turns out that even in this more restrictive setting, there can be a discrepancy between the uncoloured and the rainbow versions of the problem. To make this more precise, we give the following two definitions, where, for a k-uniform hypergraph H and $1 \leq d < k$, we let $\delta_d(H)$ be the minimum number of edges of H that any set of d vertices of $V(H)$ is contained in. Moreover, for a hypergraph collection $\mathbf{H} = (H_1, \dots, H_m)$, we denote $|\mathbf{H}| = m$ and $\delta_d(\mathbf{H}) = \min_{i\in[m]} \delta_d(H_i)$.

Definition 1.1 (Uncoloured minimum degree threshold)**.** Let $\mathcal{F}$ be an infinite family of k-uniform hypergraphs. By $\delta_{\mathcal{F},d}$ we denote, if it exists, the smallest real number δ such that for all $\alpha > 0$ and for all but finitely many $F \in \mathcal{F}$ the following holds. Let $n = |V(F)|$ and H be any n-vertex k-uniform hypergraph with $\delta_d(H) \geq (\delta + \alpha)n^{k-d}$. Then H contains a copy of F.

For example, if $\mathcal{F}$ is the family of graphs consisting of a cycle on n vertices for each $n \in \mathbb{N}$, then we have $\delta_{\mathcal{F},1} = 1/2$. Indeed, this follows from Dirac's theorem which states that any graph with minimum degree at least $n/2$ has a Hamilton cycle.

Definition 1.2 (Rainbow minimum degree threshold)**.** Let $\mathcal{F}$ be an infinite family of k-uniform hypergraphs. By $\delta^{\mathrm{rb}}_{\mathcal{F},d}$ we denote, if it exists, the smallest real number δ such that for all $\alpha > 0$ and for all but finitely many $F \in \mathcal{F}$ the following holds. Let $n = |V(F)|$ and $\mathbf{H}$ be any k-uniform hypergraph collection on n vertices with $|\mathbf{H}| = |E(F)|$ and $\delta_d(\mathbf{H}) \geq (\delta + \alpha)n^{k-d}$. Then $\mathbf{H}$ contains a transversal copy of F.

If the two values are well-defined, it must be that $\delta^{\mathrm{rb}}_{\mathcal{F},d} \geq \delta_{\mathcal{F},d}$. Indeed, if H contains no copy of F, the collection $\mathbf{H}$ consisting of $|E(F)|$ copies of H does not contain a transversal copy of H either. However, Montgomery, Müyesser, and Pehova [11] made the following observation which shows that $\delta^{\mathrm{rb}}_{\mathcal{F},d}$ can be much larger than $\delta_{\mathcal{F},d}$. Set $\mathcal{F} = \{k \times (K_{2,3} \cup C_4) \colon k \in \mathbb{N}\}$ where $k \times G$ denotes the graph obtained by taking k vertex-disjoint copies of G. It follows from a result of Kühn and Osthus [7] that $\delta_{\mathcal{F},1} = 4/9$. Consider the graph collection $\mathbf{H} = (H_1, \ldots, H_m)$ on V obtained in the following way. Partition V into two almost equal vertex subsets, say A and B, and suppose that $H_1 = H_2 = \cdots = H_{m-1}$ are all disjoint unions of a clique on A and a clique on B. Suppose that H_m is a complete bipartite graph between A and B. Observe that each H_i in this resulting graph collection has minimum degree $\lfloor |V|/2 \rfloor$. Further observe that if $\mathbf{H}$ contains a transversal copy of some $F \in \mathcal{F}$, the edge of $K_{2,3}$ or C_4 that gets copied to an edge of H_m would be a bridge (an edge whose removal disconnects the graph) of F. However, neither $K_{2,3}$ nor C_4 contains a bridge. Hence, $\delta^{\mathrm{rb}}_{\mathcal{F},d} \geq 1/2$.

On the other hand, there are many natural instances where $\delta^{\mathrm{rb}}_{\mathcal{F},d} = \delta_{\mathcal{F},d}$. When this equality holds, we say that the corresponding family $\mathcal{F}$ is *d-colour-blind*, or just *colour-blind* in the case $\mathcal{F}$ is a family of graphs (and $d = 1$). For example, Joos and Kim [6], improving a result of Cheng, Wang, and Zhao [4] and confirming a conjecture of Aharoni [1], showed that, if $n \geq 3$, then any n-vertex graph collection $\mathbf{G} = (G_1, \ldots, G_n)$ with $\delta(G_i) \geq n/2$ for each $i \in [n]$ has a transversal copy of a Hamilton cycle. This generalises Dirac's classical theorem and implies that the family $\mathcal{F}$ of n-cycles is colour-blind[1]. There are many more families of colour-blind (hyper)graphs. In particular, matchings [2, 8, 9, 10], Hamilton ℓ-cycles [3], factors [2, 11], and spanning trees [11] have been extensively studied. We recall that for $1 \leq \ell < k$, a k-uniform hypergraph is called an ℓ-cycle if its vertices can be ordered cyclically such that each of its edges consists of k consecutive vertices and every two consecutive edges (in the natural order of the edges) share exactly ℓ vertices.

Building on techniques introduced by Montgomery, Müyesser, and Pehova [11], we give a widely applicable sufficient condition for a family of hypergraphs $\mathcal{F}$ to be colour-blind. Our condition gives a unified proof of several of the aforementioned results, as well as many new rainbow Dirac-type results. The following theorem lists some applications, though we believe that our setting can capture even more families of hypergraphs.

[1] In fact, in this particular case, the corresponding thresholds are *exactly* the same, and there is no need for an error term.

Theorem 1.3. *The following families of hypergraphs are all d-colour-blind.*

(A) The family of the r-th powers of Hamilton cycles for fixed $r \geq 2$ (and $d = 1$).

(B) The family of k-uniform Hamilton ℓ-cycles for the following ranges of k, ℓ, and d.

(B1) $1 < \ell < k/2$ and $d = k - 2$;

(B2) $1 \leq \ell < k/2$ or $\ell = k - 1$, and $d = k - 1$;

(B3) $\ell = k/2$ and $k/2 < d \leq k - 1$ with k even.

Remark 1. Theorem 1.3 (B2) when $\ell = k-1$ was originally proven by Cheng, Han, Wang, Wang, and Yang [3], who raised the problem of obtaining the rainbow minimum degree threshold for a wider range of $\ell \in [k-2]$. Moreover, the case of Hamilton cycles in graphs (i.e. $k = 2$ and $d = \ell = 1$) was previously proven by Cheng, Wang, and Zhao [4] (and their result was sharpened by Joos and Kim [6]).

2 Towards the statement of the main theorem

The precise statement of our main theorem is quite technical, therefore we provide some intuition here and refer the interested reader to the arXiv preprint [5]. Firstly, we look at hypergraph families $\mathcal{F}$ with a 'cyclic' structure. That is, we assume there exists a k-uniform hypergraph $\mathcal{A}$ such that all $F \in \mathcal{F}$ can be obtained by gluing several copies of $\mathcal{A}$ in a Hamilton cycle fashion, and in this case we say that F is a Hamilton $\mathcal{A}$-*cycle*. Similarly, an $\mathcal{A}$-*chain* is a graph obtained by gluing several copies of $\mathcal{A}$ in a path-like fashion. Moreover, the first (resp. last) copy of $\mathcal{A}$ in that chain is called the *start* (resp. the *end*) of the chain. For example, for k-uniform Hamilton cycles, $\mathcal{A}$ would be a single k-uniform edge (see Figure 1), whereas for the r-th power of a Hamilton cycle, $\mathcal{A}$ would be a a clique on r vertices (see Figure 2).

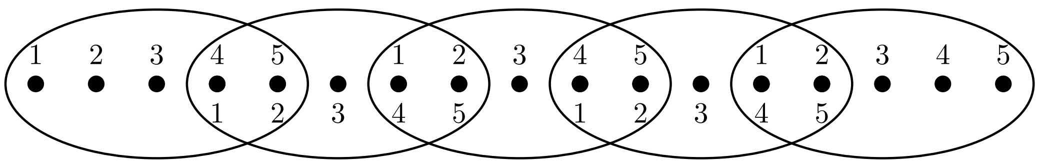


Figure 1: A 5-uniform 2-path is an $\mathcal{A}$-chain, with $\mathcal{A}$ being (any ordering of) a single 5-uniform edge. The numbering of the vertices in each edge denotes the (ordered) isomorphism between that edge and $\mathcal{A}$.

In the uncoloured setting, most of the well-studied problems fit into this framework, including everything listed in Theorem 1.3. A common framework for embedding such hypergraphs with cyclic structure is the absorption method. Our main result essentially states that if there is an absorption-based proof that δ is the uncoloured minimum d-degree threshold for some $\mathcal{F}$ with cyclic structure, then the rainbow minimum d-degree threshold of $\mathcal{F}$ is equal to δ. While some partial progress towards such an abstract statement was

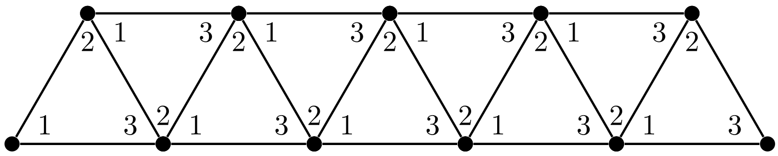


Figure 2: The square of a path is an $\mathcal{A}$-chain, with $\mathcal{A}$ being (any ordering of) a triangle.

already made in [11], our approach does not require the need to make ad-hoc strengthenings to the uncoloured version of the result, allowing for a very short proof of Theorem 1.3. To achieve this, we codify what it means for there to be streamlined absorption proof for the uncoloured result, and we use the existence of such a proof as a black-box. We do so through two properties: **Ab** and **Con**. Property **Ab** states that every k-uniform hypergraph with minimum d-degree at least $(\delta + \alpha)n^{k-d}$ contains an absorber for $\mathcal{A}$, i.e. a set of vertices which can absorb any small set of vertices into an $\mathcal{A}$-chain. Property **Con** states that in every k-uniform hypergraph with minimum d-degree at least $(\delta + \alpha)n^{k-d}$, any two copies $\mathcal{S}$ and $\mathcal{T}$ of $\mathcal{A}$ can be connected into an $\mathcal{A}$-chain of bounded length with start $\mathcal{S}$ and end $\mathcal{T}$.

In addition to properties **Ab** and **Con** which guarantee we can rely on a streamlined absorption proof for the uncoloured result, our main theorem assumes another property, which we call property **Fac**. One reason why transversal versions of Dirac-type results are more difficult is that every single hypergraph in the collection as well as every single vertex of the host graph needs to be utilised in the target spanning structure (the transversal). This is crucial as demonstrated by the construction given after Definition 1.2. In this construction, the possibility of finding a transversal copy of $\mathcal{F}$ is ruled out by showing that a particular graph in the collection (namely the hypergraph H_m) cannot be used in a transversal copy of a $K_{2,3}$ or C_4. Therefore, in addition to some properties which are related to the uncoloured case and where colours do not play any role, we require a property concerning the coloured case which we call **Fac**. This roughly states that vertex-disjoint copies of $\mathcal{A}$ (the building block of the hypergraph we are trying to find) can be found in a rainbow fashion using a fixed, adversarially specified set of hypergraphs from the collection. This ensures that we never get stuck while trying to use up every single colour/hypergraph that we start with.

Our main theorem claims that if properties **Ab**, **Con** and **Fac** hold for $\mathcal{A}$, then the family of Hamilton $\mathcal{A}$-cycles is d-colour-blind. In all our applications (see Theorem 1.3), in order to ensure that properties **Ab** and **Con** hold, we rely on existing lemmas in the literature without having to do any extra work. Moreover, when $\mathcal{A}$ is a single edge (as it is the case for Theorem 1.3 (B)), the property **Fac** is trivial to check. For powers of Hamilton cycles, however, this property is more delicate and, in order to verify it, we rely on a non-trivial coloured property from [11].

3 Proof overview

We will now attempt to give a self-contained account of the main ideas of our proof strategy. For the purposes of the proof sketch, it will be conceptually (and notationally) simpler to imagine that we are trying to prove that the family of (2-uniform) Hamilton cycles is colour-blind. Observe that a Hamilton cycle is an $\mathcal{A}$-cycle with $\mathcal{A}$ being an edge.

Proposition 3.1 (Theorem 2 in [4])**.** *For any $\alpha > 0$, there exists $n_0 \in \mathbb{N}$ such that the following holds. Let $\mathbf{G}$ be a graph collection on vertex set $[n]$ with $|\mathbf{G}| = n$ and $\delta(\mathbf{G}) \geq (1/2+\alpha)n$. Then $\mathbf{G}$ contains a transversal copy of a Hamilton cycle.*

The basic premise of our approach, which is shared with [11], is that Proposition 3.1 becomes significantly easier to prove if we assume that $|\mathbf{G}| = (1+o(1))n$, that is, if we have a bit more colours than we need to find a rainbow Hamilton cycle on n vertices. Thus, relying on the hypergraph analogue of Lemma 3.4 from [11], it is enough to show the following.

Proposition 3.2. *Let $1/n \ll \zeta \ll \kappa, \alpha$. Let $\mathbf{G}$ be a graph collection on $[n]$ with $|\mathbf{G}| = (1+\kappa-\zeta)n$ and $\delta(\mathbf{G}) \geq (1/2+\alpha)n$. Let $a, b \in [n]$ be distinct vertices. Then, $\mathbf{G}$ contains a rainbow Hamilton path with a and b as its endpoints, using every colour G_i with $i \in [(1-\zeta)n]$.*

Unfortunately, due to the technicalities present in the statement, Proposition 3.2 is far from trivial to show. Most of the novelty in the proof of our main theorem is the way we approach Proposition 3.2 for arbitrary $\mathcal{A}$-chains satisfying **Ab**, **Con**, and **Fac**. We now proceed to explain briefly how we achieve this, and how the three properties come in handy.

Firstly, in the setting of Proposition 3.2, it is quite easy to find a few rainbow paths using most of the colours from the set $[(1-\zeta)n]$. Below is a formal statement of a version of this for arbitrary $\mathcal{A}$-chains, where we write $s(\mathcal{A})\cdot n$ for the number of edges of an $\mathcal{A}$-cycle spanning n vertices.

Lemma 3.3. *Let $1/n \ll 1/T \ll \omega, \alpha$. Let $\mathcal{A}$ be k-uniform graph and $d \in [k-1]$. Let δ be the minimum d-degree threshold for the containment of a Hamilton $\mathcal{A}$-cycle. Let $\mathbf{H}$ be a k-uniform hypergraph collection on $[n]$ with $\delta_d(\mathbf{H}) \geq (\delta+\alpha)n^{k-d}$ and suppose that $|\mathbf{H}| \geq s(\mathcal{A})\cdot n$. Then $\mathbf{H}$ contains a rainbow collection of T-many pairwise vertex-disjoint $\mathcal{A}$-chains covering all but at most ωn vertices of $\mathbf{H}$.*

Although it is easy to use most of the colours coming from a colour set using the above result, a challenge in Proposition 3.2 is that we need to use all of the colours coming from the set $[(1-\zeta)n]$. As we are currently concerned with the case when $\mathcal{A}$ consists of a single edge, this will not be a major issue. Indeed, using the minimum degree condition on each of the colours, we can greedily find rainbow matchings using small colour subsets of $[(1-\zeta)n]$. For arbitrary $\mathcal{A}$, we would like to proceed in the same way; however, say when $\mathcal{A}$ is a triangle, the situation becomes considerably more complicated. This is why the property **Fac** is built into the assumptions of the main theorem.

Our ultimate goal is to build a single $\mathcal{A}$-chain connecting specific ends, not just a collection of $\mathcal{A}$-chains. Hence, we rely on the property **Con** to connect the ends of the paths we obtained via Lemma 3.3 (as well as the greedy matching we found for the purpose of exhausting a specific colour set). An issue is that **Con** is an uncoloured property, whereas we would like to connect these ends in a rainbow manner. Here we rely on the following trick: in hypergraph collections where each hypergraph has good minimum d-degree conditions, we can pass down to the following auxiliary hypergraph $\mathcal{K}$ which also has good minimum d-degree conditions. An edge appears in $\mathcal{K}$ if and only if that edge belongs to $\Omega(n)$ many colours in the original hypergraph collection. We can use the property **Con** on $\mathcal{K}$ to connect ends via short uncoloured paths, and later assign greedily one of the many available colours to the edges on this path.

As is the case with many absorption-based arguments, the short connecting paths we find will be contained in a pre-selected random set. After all the connections are made, there will remain many unused vertices inside this random set. To include these vertices inside a path, we use the property **Ab**. Similarly to **Con**, property **Ab** is an uncoloured property, but we can use again the trick of passing down to an appropriately chosen auxiliary hypergraph.

References

[1] Ron Aharoni, Matt DeVos, Sebastián González Hermosillo de la Maza, Amanda Montejano, and Robert Šámal. A rainbow version of Mantel's theorem. *Adv. Comb.*, pages Paper No. 2, 12, 2020.

[2] Yangyang Cheng, Jie Han, Bin Wang, and Guanghui Wang. Rainbow spanning structures in graph and hypergraph systems. arXiv:2105.10219, 2021.

[3] Yangyang Cheng, Jie Han, Bin Wang, Guanghui Wang, and Donglei Yang. Rainbow Hamilton cycle in hypergraph systems. arXiv:2111.07079, 2021.

[4] Yangyang Cheng, Guanghui Wang, and Yi Zhao. Rainbow pancyclicity in graph systems. *Electron. J. Combin.*, 28(3):Paper No. 3.24, 9, 2021.

[5] Pranshu Gupta, Fabian Hamann, Alp Müyesser, Olaf Parczyk, and Amedeo Sgueglia. A general approach to transversal versions of Dirac-type theorems. arXiv:2209.09289, 2022.

[6] Felix Joos and Jaehoon Kim. On a rainbow version of Dirac's theorem. *Bull. Lond. Math. Soc.*, 52(3):498–504, 2020.

[7] Daniela Kühn and Deryk Osthus. The minimum degree threshold for perfect graph packings. *Combinatorica*, 29(1):65–107, 2009.

[8] Hongliang Lu, Yan Wang, and Xingxing Yu. Co-degree threshold for rainbow perfect matchings in uniform hypergraphs. arXiv:2111.00372, 2021.

[9] Hongliang Lu, Yan Wang, and Xingxing Yu. Rainbow perfect matchings for 4-uniform hypergraphs. *SIAM Journal on Discrete Mathematics*, 36(3):1645–1662, 2022.

[10] Hongliang Lu, Xingxing Yu, and Xiaofan Yuan. Rainbow matchings for 3-uniform hypergraphs. *J. Combin. Theory Ser. A*, 183:105489, 2021.

[11] Richard Montgomery, Alp Müyesser, and Yani Pehova. Transversal factors and spanning trees. *Adv. Comb.*, 2022.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

The maximum number of copies of an even cycle in a planar graphs

(Extended abstract)

Zequn Lv* Ervin Győri† Zhen He‡ Nika Salia§ Casey Tompkins¶
Xiutao Zhu‖

Abstract

We resolve a conjecture of Cox and Martin by determining asymptotically for every $k \geq 2$ the maximum number of copies of C_{2k} in an n-vertex planar graph.

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-073

1 Introduction

A fundamental problem in extremal combinatorics is maximizing the number of occurrences of subgraphs of a certain type among all graphs from a given class. In the case of n-vertex planar graphs, Hakimi and Schmeichel [8] determined the maximum possible number of cycles of length 3 and 4 exactly and showed that for any $k \geq 3$, the maximum number of

*Department of Mathematical Sciences, Tsinghua University, Beijing, China. Email: lvzq19@mails.tsinghua.edu.cn.

†Alféd Rényi Institute, Budapest, Hungary. Email: gyori.ervin@renyi.hu. Supported by the National Research, Development and Inno- vation Office NKFIH, grants K132696 and SNN-135643.

‡Department of Mathematical Sciences, Tsinghua University, Beijing, China. Email: hz18@mails.tsinghua.edu.cn.

§Extremal Combinatorics and Probability Group, Institute for Basic Science, Daejeon, South Korea. Email: nikasalia@yahoo.com. Supported by the National Research, Development and Inno- vation Office NKFIH, grants K132696 and SNN-135643.

¶Alféd Rényi Institute, Budapest, Hungary. Email: ctompkins496@gmail.com. Supported by NKFIH grant K135800.

‖Department of Mathematics, Nanjing University, Nanjing, China. Email: zhuxt@smail.nju.edu.cn.

k-cycles is $\Theta(n^{\lfloor k/2 \rfloor})$. Moreover, they proposed a conjecture for the maximum number of 5-cycles in an n-vertex planar graph which was verified much later by Győri *et al.* in [6]. The maximum number of 6-cycles and 8-cycles was settled asymptotically by Cox and Martin in [3], and later the same authors [4] also determined the maximum number of 10-cycles and 12-cycles asymptotically.

Following the work of Hakimi and Schmeichel [8], Alon and Caro [1] considered the general problem of maximizing copies of a given graph H among n-vertex planar graphs. Wormald [11] and later independently Eppstein [5] showed that for 3-connected H, the maximum number of copies of H is $\Theta(n)$. The order of magnitude in the case when H is a tree was determined in [7], and the order of magnitude for an arbitrary graph was settled by Huynh, Joret and Wood [9]. Note that by Kuratowski's theorem [10] such problems can be thought of as generalized Turán problems where we maximize the number of copies of the graph H while forbidding all subdivisions of K_5 and $K_{3,3}$.

Given that the order of magnitude of the maximum number of copies of any graph H in an n-vertex planar graph is determined, it is natural to look for sharp asymptotic results. While in recent times a number of results have been obtained about the asymptotic number of H-copies in several specific cases, less is known for general classes of graphs. Cox and Martin [3] introduced some general tools for studying such problems and conjectured that in the case of an even cycle C_{2k} with $k \geq 3$, the maximum number of copies is asymptotically n^k/k^k. We confirm their conjecture.

Theorem 1. *For every $k \geq 3$, the maximum number of copies of C_{2k} in an n-vertex planar graph is*

$$\frac{n^k}{k^k} + o(n^k).$$

A construction containing this number of copies of C_{2k} is obtained by taking a C_{2k} and replacing every second vertex by an independent set of approximately n/k vertices, each with the same neighborhood as the original vertex. Cox and Martin [3] proved that a weaker upper bound of $\frac{n^k}{k!} + o(n^k)$ holds for the number of copies of C_{2k} and introduced a general method for (asymptotically) maximizing the number of copies of a large variety of graphs in a planar graph. We will discuss this method in Section 2 and present another conjecture of Cox and Martin which implies Theorem 1. In Section 3, we prove this stronger conjecture (Theorem 2). We have learned that Asaf Cohen Antonir and Asaf Shapira have independently obtained a bound within a factor of e of the optimal bound attained in Theorem 2.

2 Reduction lemma of Cox and Martin

For a positive integer n we will consider functions $w : E(K_n) \to \mathbb{R}$ satisfying the conditions:

1. For all $e \in E(K_n)$, $w(e) \geq 0$,
2. $\sum_{e\in E(K_n)} w(e) = 1$.

For a subgraph H' of K_n and a function w satisfying Conditions 1 and 2, let

$$p_w(H') := \prod_{e \in E(H')} w(e).$$

Also for a fixed graph H and w satisfying Conditions 1 and 2 let

$$\beta(w, H) := \sum_{H \cong H' \subseteq K_n} p_w(H').$$

For simplicity of notation, we will often omit statements about isomorphism in the sums. Cox and Martin proved several reduction lemmas for pairs of graphs H and K, in which an optimization problem involving $\beta(w, K)$ implies a corresponding upper bound on the maximum number of copies of the graph H among n-vertex planar graphs. We state the reduction lemma which Cox and Martin proved for cycles. For an integer $k \geq 3$, let

$$\beta(k) = \sup_w \beta(w, C_k),$$

where w is allowed to vary across all n and all weight functions satisfying Conditions 1 and 2.

Lemma 1 (Cox and Martin [3]). *For all $k \geq 3$, the number of $2k$-cycles in a planar graph is at most*

$$\beta(k)n^k + o(n^k).$$

Cox and Martin conjectured that $\beta(k) \leq \frac{1}{k^k}$. By Lemma 1 such a bound immediately implies Theorem 1. In Section 3, we prove that this bound indeed holds.

Theorem 2. *For all $k \geq 3$,*

$$\beta(k) \leq \frac{1}{k^k}.$$

Equality is attained only for weight functions satisfying $w(e) = \frac{1}{k}$ for $e \in E(C)$ and $w(e) = 0$ otherwise, where C is a fixed cycle of length k of K_n.

3 Proof of Theorem 2

Proof. Let us fix an integer n, a complete graph K_n and a function w satisfying Conditions 1 and 2. Let us assume w maximizes $\sum_{C_k \subseteq K_n} p_w(C_k)$. Let P_j be a path with j vertices. A $(j+2)$-vertex path with terminal vertices u and v is denoted by uP_jv. For vertices u and v, a subgraph H of K_n and an integer j such that $2 \leq j \leq n$, we define

$$f_H(j, u, v) = \sum_{uP_{j-2}v \subseteq H} p_w(uP_{j-2}v),$$

and

$$f_H(j,u) = \sum_{v \in V(H)\setminus\{u\}} f(j,u,v).$$

In the case when H is the complete graph K_n we simply write $f(j,u,v)$ and $f(j,u)$. The following lemma will be essential in the proof of Theorem 2. Related lemmas were also deduced in the original paper of Cox and Martin [3] (see Lemmas 4.5 and 4.6 in their paper), and both approaches can be used to deduce an upper bound of $1/k$ on the weight of every edge.

Lemma 2. *Let $k \geq 2$, and let $e_1 = u_1v_1$ and $e_2 = u_2v_2$ be distinct edges of K_n such that $w(e_1) > 0$ and $w(e_2) > 0$. Then we have $f(k,u_1,v_1) = f(k,u_2,v_2)$.*

Proof of Lemma 2. Omitted for space. Full proofs can be found in the Arxiv preprint of the same title. □

From Lemma 2, for an edge uv with non-zero weight $w(uv) > 0$ we may assume $f(j,u,v) = \mu$ for some fixed constant μ. Hence we have

$$\sum_{C_k \subseteq K_n} p_w(C_k) = \frac{1}{k} \sum_{uv \in E(K_n)} w(uv) f(j,u,v) = \frac{\mu}{k} \sum_{uv \in E(K_n)} w(uv) = \frac{\mu}{k}. \tag{1}$$

Furthermore $w(e) \leq 1/k$ for every edge $e \in E(K_n)$. Indeed,

$$w(e)\mu = \sum_{e \in C_k} p_w(C_k) \leq \sum_{C_k \subseteq K_n} p_w(C_k) = \frac{\mu}{k}.$$

For any subgraph G of K_n and any vertex $v \in V(K_n)$ we denote $\sum_{u \in V(G)} w(uv)$ by $d_G(v)$. Furthermore, for a vertex set $S \subseteq V(G)$, we denote the graph $G[V(G) \setminus S]$ by $G \setminus S$. Also for an edge $e \in E(G)$, the graph with vertex set $V(G)$ and edge set $E(G) \setminus \{e\}$ is denoted by $G \setminus e$.

Lemma 3. *For a fixed integer r such that $3 \leq r \leq n$ and distinct vertices v_1 and u there exists a sequence $v_2, v_3, \dots, v_{r-1}$ of distinct vertices such that for every integer t satisfying $1 \leq t \leq r-1$, where $G_1 = K_n \setminus v_1u$ and $G_i = K_n \setminus \{v_1, v_2, \dots, v_{i-1}\}$ for every $i = 2, 3, \dots, r-1$, we have*

$$f_{G_1}(r, v_1, u) \leq d_{G_1}(v_1) d_{G_2}(v_2) \cdots d_{G_{t-1}}(v_{t-1}) f_{G_t}(r-t+1, v_t, u).$$

Proof. Omitted for space. □

Lemma 4. *For any non-negative weight function $w : E(K_n) \to \mathbb{R}$ and for every vertex v and integer r with $2 \leq r \leq n$, we have*

$$f(r,v) \leq \left(\frac{\sum_{e \in E(K_n)} w(e)}{r-1} \right)^{r-1}.$$

Remark 1. *In Lemma 4, we do not require that* $\sum_{e\in E(K_n)} w(e) = 1$*, only that the weights are non-negative.*

Proof. Omitted for space. □

In order to finish the proof of Theorem 2 it is sufficient to show that $\mu \leq \frac{1}{k^{k-1}}$ by (1).

Choose an edge v_0v_1 with the maximum weight $w(v_0v_1)$. Let us denote the graph $K_n \setminus v_0v_1$ by G_1. By Lemma 3 we have a sequence of vertices $v_2, v_3, \dots, v_{k-1} \in V(K_n)$ satisfying the following inequality for every t where $1 \leq t \leq k-1$ and $G_i = K_n \setminus \{v_1, v_2, \dots, v_{i-1}\}$ for all $i \in \{2, 3, \dots, r-1\}$:

$$f_{G_1}(k, v_1, v_0) \leq d_{G_1}(v_1)d_{G_2}(v_2)\cdots d_{G_{t-1}}(v_{t-1})f_{G_t}(k-t+1, v_t, v_0). \tag{2}$$

Here we distinguish the following two cases.

Case 1: Suppose that $d_{G_1}(v_1)+d_{G_2}(v_2)+\cdots+d_{G_{k-2}}(v_{k-2}) \leq \frac{k-2}{k}$. Then by the inequality of the arithmetic and geometric means we have

$$\prod_{i=1}^{k-2} d_{G_i}(v_i) \leq \left(\frac{\sum_{i=1}^{k-2} d_{G_i}(v_i)}{k-2}\right)^{k-2} \leq \frac{1}{k^{k-2}}.$$

From (2) we obtain the desired inequality

$$\mu = f_{G_1}(k, v_1, v_0) \leq \left(\prod_{i=1}^{k-2} d_{G_i}(v_i)\right) \cdot f_{G_{k-1}}(2, v_{k-1}, v_0) \leq \frac{1}{k^{k-2}}\frac{1}{k} \leq \frac{1}{k^{k-1}}.$$

Even more the inequality holds with equality if and only if $w(v_0v_1) = w(v_1v_2) = \cdots = w(v_{k-2}v_{k-1}) = w(v_{k-1}v_0) = 1/k$ (here we use that for all e, $w(e) \leq 1/k$). Therefore equality is attained in Theorem 2 only for weight functions satisfying $w(e) = \frac{1}{k}$ for $e \in E(C)$ and $w(e) = 0$ otherwise, where C is a fixed cycle of length k of K_n.

Case 2: Suppose that $d_{G_1}(v_1) + d_{G_2}(v_2) + \cdots + d_{G_{k-2}}(v_{k-2}) > \frac{k-2}{k}$. Proof of this case is omitted for space. □

References

[1] Noga Alon and Yair Caro. On the number of subgraphs of prescribed type of planar graphs with a given number of vertices. In *North-Holland Mathematics Studies*, volume 87, pages 25–36. Elsevier, 1984.

[2] Asaf Cohen Antonir and Asaf Shapira. Bounding the number of odd paths in planar graphs via convex optimization. *arXiv preprint arXiv:2208.02097*, 2022.

[3] Christopher Cox and Ryan R Martin. Counting paths, cycles, and blow-ups in planar graphs. *Journal of Graph Theory*, 101(3):521–558, 2022.

[4] Christopher Cox and Ryan R Martin. The maximum number of 10-and 12-cycles in a planar graph. *Discrete Mathematics*, 346(2):113245, 2023.

[5] David Eppstein. Connectivity, graph minors, and subgraph multiplicity. *Journal of Graph Theory*, 17(3):409–416, 1993.

[6] Ervin Győri, Addisu Paulos, Nika Salia, Casey Tompkins, and Oscar Zamora. The maximum number of pentagons in a planar graph. *arXiv preprint arXiv:1909.13532*, 2019.

[7] Ervin Győri, Addisu Paulos, Nika Salia, Casey Tompkins, and Oscar Zamora. Generalized planar tur\'an numbers. *arXiv preprint arXiv:2002.04579*, 2020.

[8] Seifollah Louis Hakimi and Edward F Schmeichel. On the number of cycles of length k in a maximal planar graph. *Journal of Graph Theory*, 3(1):69–86, 1979.

[9] Tony Huynh, Gwenaël Joret, and David R Wood. Subgraph densities in a surface. *Combinatorics, Probability and Computing*, 31(5):812–839, 2022.

[10] Casimir Kuratowski. Sur le probleme des courbes gauches en topologie. *Fundamenta mathematicae*, 15(1):271–283, 1930.

[11] Nicholas C Wormald. On the frequency of 3-connected subgraphs of planar graphs. *Bulletin of the Australian Mathematical Society*, 34(2):309–317, 1986.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

3-uniform linear hypergraphs without a long Berge path

(Extended abstract)

Ervin Győri* Nika Salia†

Abstract

Extensions of the Erdős-Gallai theorem for general hypergraphs are well studied. In this work, we prove the extension of the Erdős-Gallai theorem for linear hypergraphs. In particular, we show that the number of hyperedges in an n-vertex 3-uniform linear hypergraph, without a Berge path of length k as a subgraph is at most $\frac{(k-1)}{6}n$ for $k \geq 4$. This is an extended abstract for EUROCOMB23 of the manuscript arXiv:2211.16184.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-074`

1 Introduction

Finding the maximum number of edges in a graph with fixed order not containing another graph as a subgraph is a central problem in extremal combinatorics. This work considers problems where a path of fixed length is forbidden. This problem is well understood for graphs and r-uniform hypergraphs. The Erdős–Gallai theorem states that a graph of order n containing no path of length k as a subgraph contains at most $\frac{k-1}{2}n$ edges. This bound is sharp for infinitely many n. In particular, equality holds if and only if n is a multiple of k and the graph is isomorphic to the union of $\frac{n}{k}$ cliques of size k. This theorem was extended to r-uniform hypergraphs by Győri, Katona and Lemons [11]. In order to state their result, we will introduce the necessary definitions.

*Alfréd Rényi Institute of Mathematics. E-mail: `gyori.ervin@renyi.hu`. Supported by the National Research, Development and Innovation Office – NKFIH, grant K116769, K132696 and SNN117879.

†E-mail: `salianika@gmail.com`. Supported by the National Research, Development, and Innovation Office – NKFIH, grant K132696.

For an integer r, a hypergraph $\mathcal{H}$ is r-uniform if it is a family of r-element sets of finite family $V(\mathcal{H})$. We will use the following extension of this definition. For a set of integers R, a hypergraph $\mathcal{H}$ is R-uniform if it is a family of sets of the finite family $V(\mathcal{H})$, such that the sizes of the sets are elements of R. Paths in hypergraphs can be defined in a number of ways. In this paper, we follow the definition of Berge [2]. A Berge path of length k in a hypergraph $\mathcal{H}$ is an alternating sequence $v_1, h_1, v_2, \dots, h_k, v_{k+1}$ of distinct vertices and hyperedges such that $\{v_i, v_{i+1}\} \subseteq h_i$ for all $i \in [k]$. A Berge cycle of length k is also defined similarly. The vertices v_i, $i \in [k+1]$, are defining vertices of the Berge path and the hyperedges h_i, $i \in [k]$, are defining hyperedges of the Berge path.

Theorem (Győri, Katona and Lemons [11])**.** *Let $\mathcal{H}$ be an n-vertex r-uniform hypergraph containing no Berge path of length k as a subgraph. Then if $r \geq k > 2$ then the number of hyperedges of $\mathcal{H}$ is at most $\frac{k-1}{r+1}n$. If $k > r+1 > 2$ then the number of hyperedges of $\mathcal{H}$ is at most $\frac{\binom{k}{r}}{k}n$.*

The remaining case $k = r+1$ was settled later in [3], the bound matches with the bound in Theorem 1 for $k > r+1$ case. Forbidden path problems for connected graphs and hypergraphs including their stability versions are well studied, we refer interested readers to [16, 1, 13, 6, 15, 8, 7, 9]. Uniform hypergraphs with bounded circumference was studied in [5, 12] and references therein.

Here we introduce some necessary technical definitions. For a hypergraph $\mathcal{H}$ let $E(\mathcal{H})$ be the hyperedge set and $V(\mathcal{H})$ be the vertex set, we denote their sizes by $e(\mathcal{H})$ and $v(\mathcal{H})$ accordingly. The hypergraph $\mathcal{H}$ is linear if for any two distinct hyperedges h_1, h_2 we have $|h_1 \cap h_2| \leq 1$. For a vertex set V, $V \subseteq V(\mathcal{H})$, we define another hypergraph $\mathcal{H}_V$. Where $V(\mathcal{H}_V) = V$ and $E(\mathcal{H}_V) = \{h \setminus V : h \in E(\mathcal{H}), |h \setminus V| \geq 2\}$. Note that if $\mathcal{H}$ is $\{2,3\}$-uniform linear hypergraph then $\mathcal{H}_V$ is $\{2,3\}$-uniform linear hypergraph also. The induced hypergraph on the vertex set V is denoted by $\mathcal{H}[V]$. For a hypergraph $\mathcal{H}$ we denote two-shadow of $\mathcal{H}$ by $\partial\mathcal{H}$. It is a graph on the same vertex set as $\mathcal{H}$ and the set of edges is $\{\{u,v\} : \{u,v\} \subseteq h \in E(\mathcal{H})\}$. The degree of a vertex v in a hypergraph $\mathcal{H}$ is the number of hyperedges incident to the vertex v and is denoted by $d_{\mathcal{H}}(v)$. The minimum degree of a vertex in a hypergraph $\mathcal{H}$ is denoted by $\delta_{\mathcal{H}}(v)$. The circumference of $\mathcal{H}$ is the length of the longest Berge cycle in a hypergraph $\mathcal{H}$ and is denoted by $c(\mathcal{H})$. The neighborhood of a vertex v in a hypergraph $\mathcal{H}$ is denoted by $N_{\mathcal{H}}(v)$. For a hypergraph $\mathcal{H}$ and sub-hypergraph $\mathcal{H}'$ we denote the hypergraph on the same vertex set as $\mathcal{H}$ and hyperedge set $E(\mathcal{H}) \setminus E(\mathcal{H}')$ by $\mathcal{H} \setminus \mathcal{H}'$.

2 Main results

Recently Gyárfás, Ruszinkó, and Sárközy [10] initiated the study of three uniform linear hypergraphs not containing a linear path, a matching, and a small tree. In particular, they proved that the maximum number of hyperedges in an n vertex three uniform linear hypergraph not containing a linear path of k edges is $1.5nk$. In this paper, we prove the extension of Erdős-Gallai theorem for linear 3-uniform hypergraphs but instead of forbidding linear paths, we forbid Berge paths.

Theorem 1. *Let $\mathcal{H}$ be an n vertex 3-uniform linear hypergraph, containing no Berge path of length $k \geq 4$. Then the number of hyperedges in $\mathcal{H}$ is at most $\frac{k-1}{6}n$.*

Note that the upper bound is sharp for infinitely many k and n. In particular for all k for which there exists a Steiner Triple System (a 3-uniform hypergraph that every pair of vertices is covered by precisely one hyperedge) and n multiple of k, there exists an n-vertex 3-uniform linear hypergraph $\mathcal{H}$, containing no Berge path of length k with $\frac{k-1}{6}n$ hyperedges. Where $\mathcal{H}$ is the disjoint union of $\frac{n}{k}$ copies of k-vertex Steiner Triple Systems.

In order to prove Theorem 1 with induction for k, we need a stronger and more general statement of the theorem.

Theorem 2. *Let $\mathcal{H}$ be an n vertex $\{2,3\}$-uniform linear hypergraph, containing no Berge path of length $k \geq 4$. Then the number of edges in $\partial\mathcal{H}$ is at most $\frac{k-1}{2}n$.*

Note that Theorem 1 is a direct corollary of Theorem 2. The following remark shows that the condition $k \leq 4$ in Theorem 2 is necessary since for $k < 4$ we have different bounds.

Remark. *Let $\mathcal{H}$ be an n vertex linear $\{2,3\}$-uniform hypergraph, containing no Berge path of length k.*

- *If $k = 1$ then $e(\partial\mathcal{H}) = 0$;*
- *If $k = 2$ then $e(\partial\mathcal{H}) \leq v(\mathcal{H})$; The upper-bound is sharp and the equality is achieved if and only if is $v(\mathcal{H})$ multiple of 3 and $\mathcal{H}$ is $\frac{v(\mathcal{H})}{3}$ independent hyperedges of size three.*
- *If $k = 3$ then $e(\partial\mathcal{H}) \leq 3\frac{v(\mathcal{H})-1}{2}$. The upper-bound is sharp and the equality is achieved if and only if $v(\mathcal{H})$ is odd and $\mathcal{H}$ is $\frac{v(\mathcal{H})-1}{2}$ hyperedges of size three sharing the same vertex for every $n \geq 3$.*

We find it challenging to obtain the precise bound for the problem initiated by Gyárfása, Ruszinkó, and Sárközy [10]. Consequently, we would like to put forth a natural conjecture.

Conjecture 3. *Let $\mathcal{H}$ be an n vertex 3-uniform linear hypergraph, containing no linear path of length $k \geq 5$. Then the number of hyperedges in $\mathcal{H}$ is at most $\frac{2k-1}{6}n$.*

Note that, this bound is sharp for infinitely many pairs of n and k. In particular for every k such that there exists a Steiner Triple System on $2k$ vertices and for every n multiple of $2k$. The hypergraph containing $\frac{n}{2k}$ copies of a Steiner Triple System on $2k$ vertices achieves the desired bound.

3 Proof of Theorem 2

For the full proof see manuscript [14].

We prove Theorem 2 by induction on k. At first, we consider the base case when $k = 4$. We may assume $\mathcal{H}$ is a connected hypergraph since the upper bound is linear for n and the additive constant is 0. If $\mathcal{H}$ is Berge cycle free then $e(\partial\mathcal{H}) \leq \frac{3(n-1)}{2}$ (the upper-bound is attained by hyperedges of size three sharing a fixed vertex). If $\mathcal{H}$ contains a Berge cycle it must be a Berge cycle of length 3 or 4 since it is a linear hypergraph. If $\mathcal{H}$ contains Berge cycle of length 4 then by connectivity $v(\mathcal{H}) \leq 4$, hence $e(\mathcal{H}) \leq \binom{4}{2} = \frac{3n}{2}$. If $\mathcal{H}$ contains a cycle of length 3, we denote it by C_3. Cycle C_3 is a linear cycle since $\mathcal{H}$ is a linear hypergraph. If all of the hyperedges of C_3 are size three then by the connectivity of $\mathcal{H}$ we have $\mathcal{H} = C_3$ and $e(\partial\mathcal{H}) = 9 = \frac{3n}{2}$. If two of the hyperedges are size three then by the connectivity of $\mathcal{H}$ we have $\mathcal{H} = C_3$ and $e(\partial\mathcal{H}) = 7 < \frac{3n}{2}$. If at most one hyperedge is size three then we have $e(\partial\mathcal{H}) \leq \frac{3n}{2}$. So the base case $k = 4$ is done.

Let $\mathcal{H}$ be an n-vertex linear $\{2,3\}$-uniform hypergraph containing no Berge path of length k for some integer $k > 4$. Suppose by way of contradiction that $e(\partial\mathcal{H}) > \frac{n(k-1)}{2}$. Without loss of generality, we may assume n is minimal, in particular, we assume all linear $\{2,3\}$-uniform hypergraphs containing no Berge path of length k with n' vertices, $n' < n$, contain at most $\frac{n(k-1)}{2}$ edges in the shadow. Note that from the minimality of n we have the hypergraph $\mathcal{H}$ is connected. Even more, for each vertex v, $\mathcal{H}_{V(\mathcal{H})\setminus\{v\}}$ contains no Berge path of length k, thus from the minimality of n we have $d_{\partial\mathcal{H}}(v) > \frac{k-1}{2}$. Hence we have $\delta_{\partial\mathcal{H}}(v) \geq \lceil \frac{k}{2} \rceil$. Note that since $e(\partial\mathcal{H}) > \frac{n(k-1)}{2}$ the longest path of $\mathcal{H}$ is length $k-1$ by the induction hypothesis.

We omit the proof of the following Claims.

Claim 4. $c(\mathcal{H}) \geq \lceil \frac{k+1}{2} \rceil$.

Let $\mathcal{C}_\ell := v_1, h_1, v_2, h_2, \ldots h_{\ell-1}, v_\ell, h_\ell, v_1$ be a longest Berge cycle of $\mathcal{H}$. Some $\mathcal{C}_\ell$ defining hyperedges h_i are size three, let us denote the third vertex by x_i, that is $h_i = \{v_i, v_{i+1}, x_i\}$ for hyperedges of size three. From Claim 4 we have $\ell \geq \lceil \frac{k+1}{2} \rceil$. Let us denote the hypergraph $\mathcal{H}_{V(\mathcal{H})\setminus\{v_i : i \in [\ell]\}}$ by $\mathcal{H}'$.

Claim 5. *The hypergraph $\mathcal{H}'$ is $\mathcal{BP}_{k-\ell}$-free.*

If $k - \ell \geq 4$ then by Claim 5 and induction hypothesis for hypergraph $\mathcal{H}'$ we have

$$e(\partial\mathcal{H}') \leq \frac{(n-\ell)(k-\ell-1)}{2}. \tag{1}$$

For a vertex $u \in V(\mathcal{H}')$ we define the set $S(u) := N_{\mathcal{H}\setminus C_\ell}(u) \cap V(C_\ell)$, $L(u) := \{v_i : u = x_i\}$ and $R(u) := \{v_{i+1} : u = x_i\}$. For a vertex set S such that $S \subseteq V(C_\ell)$ let S^+ be a set S shifted right, in particular $S^+ := \{v_i : v_{i-1} \in S\}$, the indices are taken module ℓ. Similarly we definite S^-, in particular S^- is a set for which $S = (S^-)^+$. Naturally we denote the set $(S^-)^-$ with S^{--} and the set $(S^+)^+$ with S^{++}. Note that $L(u)^+ = R(u)$, thus the size of $L(u)$ and $R(u)$ are the same.

In what follows we are going to estimate the number of edges in $\partial\mathcal{H}$, in the following way

$$e(\partial\mathcal{H}) = e(\partial\mathcal{H}_{V(C_\ell)}) + e_{\partial\mathcal{H}}(V(C_\ell), V(\mathcal{H}')) + e(\partial\mathcal{H}'). \tag{2}$$

Noting that $e_G(A, B)$ denotes the number of edges between vertex set A and B in G. In most cases, we will use a naive upper bound for $e(\partial\mathcal{H}_{V(C_\ell)}) \le \binom{\ell}{2}$. For $k - \ell \ge 4$, we estimate $e(\partial\mathcal{H}')$ by the induction hypotheses as in Equation 1. We estimate the number of edges from $V(\mathcal{H}')$ to $V(C_\ell)$, for each vertex $u \in V(\mathcal{H}')$ in $\partial\mathcal{H}$. In particular the number of adjacent vertices to u is $|L(u)| + |R(u)| + |S(u)|$. Since each defining hyperedge of C_ℓ provides at most two edges crossing between the vertices $V(\mathcal{H}')$ and $V(C_\ell)$ we have a naive upper bound for $e_{\partial\mathcal{H}}(V(C_\ell), V(\mathcal{H}'))$ which is enough for most of the cases.

$$e_{\partial\mathcal{H}}(V(C_\ell), V(\mathcal{H}')) \le 2\ell + \sum_{u\in V(\mathcal{H}')} |S(u)|\,. \tag{3}$$

Since C_ℓ is a longest Berge cycle of $\mathcal{H}$ we are able to get an upper bound for $|S(u)|$ from the following claim.

Claim 6. *For a vertex $u \in V(\mathcal{H}')$ we have $(S(u) \cup L(u)) \cap S(u)^- = \emptyset$.*

Note that if a vertex $v_i \in S(u)$ then $v_{i+1} \notin S(u)$ from Claim 6. Thus we have $|S(u)| \le \frac{\ell}{2}$ for each vertex u of $\mathcal{H}'$. Therefore $e_{\partial\mathcal{H}}(V(C_\ell), V(\mathcal{H}')) \le 2\ell + \frac{\ell}{2}(n - \ell)$ from Equation 3. If $k - \ell \ge 4$ then by Equation 2 and 1 we have a contradiction

$$e(\partial\mathcal{H}) \le \binom{\ell}{2} + 2\ell + \frac{\ell(n-\ell)}{2} + \frac{(n-\ell)(k-\ell-1)}{2} = \frac{n(k-1)}{2} + \frac{\ell}{2}(\ell+4-k) \le \frac{n(k-1)}{2}.$$

We study the rest of the possible values of ℓ separately, $\ell \in \{k-3, k-2, k-1, k\}$. Let x be the number of defining hyperedges of C_ℓ incident to a vertex of $\mathcal{H}'$. Note that $0 \le x \le \ell$.

If $\ell = k$ then $\mathcal{C}_\ell = \mathcal{H}$ otherwise we have a Berge path of length k in $\mathcal{H}$ by the connectivity of $\mathcal{H}$. Thus we have $n = k = \ell$ and

$$e(\partial\mathcal{H}) \le \binom{\ell}{2} = \frac{n(k-1)}{2}.$$

If $\ell = k - 1$ then $\mathcal{H}'$ contains no hyperedge by Claim 5. Since $\mathcal{H}$ does not contain a Berge path of length k, if a hyperedge h_i adjacent to a vertex from $V(\mathcal{H}')$, then neither v_i nor v_{i+1} is a vertex of $S(u)$, for all $u \in V(\mathcal{H})$. In particular for $u, u' \in V(\mathcal{H}')$ we have $L(u) \cap (S(u))^- = \emptyset$. By this observation and Claim 6 every vertex of $V(\mathcal{H}')$ is adjacent to at most $\frac{k-1-x}{2}$ vertices of C_ℓ with a non-defining hyperedge, that is $|S(u)| \le \frac{k-1-x}{2}$. Thus by Equation 2 we have

$$e(\partial\mathcal{H}) \le \binom{k-1}{2} + 2x + \frac{k-1-x}{2}(n-(k-1)).$$

Hence if $n \ge k+2$ or $n = k+1$ and $x \le \frac{k-1}{2}$ then we have $e(\partial\mathcal{H}) \le \frac{n(k-1)}{2}$, since $x \le k-1$. As $e(\partial\mathcal{H}) > \frac{n(k-1)}{2}$ we have $n \ge k+1$.

If $n = k+1$ and $x > \frac{k-1}{2}$ then there are two C_ℓ non-defining hyperedges h_i and h_{i+1} such that $\{x_i, x_{i+1}\} = V(\mathcal{H}')$. Since $\mathcal{H}$ does not contain a Berge path of length k, if a defining vertex of C_ℓ is incident to both vertices of $\mathcal{H}'$, either both incidences are from a defining hyperedge or both incidences are from a non-defining hyperedge. If v_j is incident to both vertices of $\mathcal{H}'$ with h_{j-1} and h_j such that $j \neq i-1, i, i+1$ then v_j is not incident to v_{i+1}. Otherwise, if there is a hyperedge f' incident to v_j and v_{j+1}, then it is a non-defining hyperedge and the following is a Berge path or a Berge cycle of length k,

$$x_{i+1}, h_{i+1}, v_{i+2}, \ldots, v_j, f', v_{i+1}, h_i, v_i, \ldots, v_{j+1}, h_j, x_j.$$

If a vertex v_j is adjacent to x_1 or x_2 with a non-defining hyperedge then v_{j+1} is not adjacent to a vertex from $\{x_1, x_2\}$. Thus for each vertex $v_j \in V(C_\ell)$, $j \notin \{i-1, i, i+1\}$, either there is at most one vertex from $V(\mathcal{H}')$ adjacent to it, or if there are two then $v_j v_{i+1}$ is not an edge of $\partial\mathcal{H}$ or v_{j+1} is not adjacent to any vertex of $V(\mathcal{H}')$. Note that if there is a defining hyperedge of C_ℓ not incident to a vertex of $\mathcal{H}'$ then we may choose i such that $i-1$ has exactly one neighbor in $V(\mathcal{H}')$. If all defining hyperedges of C_ℓ are incident to a vertex of $\mathcal{H}'$ then we may choose any i from $[k-1]$. Thus we have a contradiction from Equation 2

$$e(\partial\mathcal{H}) \leq \binom{k-1}{2} + k - 1 + 2 \leq \frac{n(k-1)}{2}.$$

The proof of remaining cases $\ell = k-2$ and $\ell = k-3$ involves more structural study and can be seen in the original manuscript.

References

[1] Paul N Balister, E Győri, Jenő Lehel, and Richard H Schelp. Connected graphs without long paths. *Discrete Mathematics*, 308(19):4487–4494, 2008.

[2] Claude Berge. *Graphs and hypergraphs.* North-Holland, 1973.

[3] Akbar Davoodi, Ervin Győri, Abhishek Methuku, and Casey Tompkins. An Erdős-Gallai type theorem for uniform hypergraphs. *European Journal of Combinatorics*, 69:159–162, 2018.

[4] Zoltán Füredi, Alexandr Kostochka, and Ruth Luo. Avoiding long Berge cycles ii, exact bounds for all n. *arXiv preprint arXiv:1807.06119*, 2018.

[5] Zoltán Füredi, Alexandr Kostochka, and Ruth Luo. Avoiding long Berge cycles. *Journal of Combinatorial Theory, Series B*, 137:55–64, 2019.

[6] Zoltán Füredi, Alexandr Kostochka, and Ruth Luo. On 2-connected hypergraphs with no long cycles. *arXiv preprint arXiv:1901.11159*, 2019.

[7] Zoltán Füredi, Alexandr Kostochka, Ruth Luo, and Jacques Verstraëte. Stability in the Erdős–Gallai theorem on cycles and paths, II. *Discrete Mathematics*, 341(5):1253–1263, 2018.

[8] Zoltán Füredi, Alexandr Kostochka, and Jacques Verstraëte. Stability in the Erdős–Gallai theorems on cycles and paths. *Journal of Combinatorial Theory, Series B*, 121:197–228, 2016.

[9] Dániel Gerbner, Dániel Nagy, Balázs Patkós, Nika Salia, and Máté Vizer. Stability of extremal connected hypergraphs avoiding Berge-paths. *arXiv preprint arXiv:2008.02780*, 2020.

[10] András Gyárfás, Miklós Ruszinkó, and Gábor N Sárközy. Linear Turán numbers of acyclic triple systems. *European Journal of Combinatorics*, 99:103435, 2022.

[11] Ervin Győri, Gyula Y Katona, and Nathan Lemons. Hypergraph extensions of the Erdős-Gallai theorem. *European Journal of Combinatorics*, 58:238–246, 2016.

[12] Ervin Győri, Nathan Lemons, Nika Salia, and Oscar Zamora. The structure of hypergraphs without long Berge cycles. *Journal of Combinatorial Theory, Series B*, 2020.

[13] Ervin Győri, Abhishek Methuku, Nika Salia, Casey Tompkins, and Máté Vizer. On the maximum size of connected hypergraphs without a path of given length. *Discrete Mathematics*, 341(9):2602–2605, 2018.

[14] Ervin Győri and Nika Salia. Linear three-uniform hypergraphs with no berge path of given length. *arXiv preprint arXiv:2211.16184*, 2022.

[15] Ervin Győri, Nika Salia, and Oscar Zamora. Connected hypergraphs without long berge-paths. *European Journal of Combinatorics*, 96:103353, 2021.

[16] GN Kopylov. On maximal paths and cycles in a graph. In *Doklady Akademii Nauk*, volume 234, pages 19–21. Russian Academy of Sciences, 1977.

[17] Alexandr Kostochka and Ruth Luo. On r-uniform hypergraphs with circumference less than r. *Discrete Applied Mathematics*, 276:69–91, 2020.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Rooting algebraic vertices of convergent sequences*

(Extended abstract)

David Hartman[†] [‡] Tomáš Hons[†] Jaroslav Nešetřil[†]

Abstract

Structural convergence is a framework for convergence of graphs by Nešetřil and Ossona de Mendez that unifies the dense (left) graph convergence and Benjamini-Schramm convergence. They posed a problem asking whether for a given sequence of graphs (G_n) converging to a limit L and a vertex r of L it is possible to find a sequence of vertices (r_n) such that L rooted at r is the limit of the graphs G_n rooted at r_n. A counterexample was found by Christofides and Král', but they showed that the statement holds for almost all vertices r of L. We offer another perspective to the original problem by considering the size of definable sets to which the root r belongs. We prove that if r is an algebraic vertex (i.e. belongs to a finite definable set), the sequence of roots (r_n) always exists.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-075`

1 Introduction

The field of graph convergence studies asymptotic properties of large graphs. The goal is to define a well-behaved notion of a limit structure that describes the limit behavior of a convergent sequence of graphs. Several different approaches are studied. The two most prominent types of convergence are defined for sequences of dense [2][7][6] and sparse graphs

*Supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC Synergy Grant DYNASNET, grant agreement No 810115).

†Computer Science Institute, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

‡Institute of Computer Science of the Czech Academy of Sciences, Prague, Czech Republic

[1][4]. The recently introduced notion of structural convergence by Nešetřil and Ossona de Mendez offers a generalizing framework for these cases using ideas from analysis, model theory and probability [8][9].

Structural convergence is a framework of convergence for general relational structures; however, we follow the usual approach that we restrict to the of language graphs and rooted graphs without loss of generality. Our arguments remain valid in the general case (e.g. as in [3]). The *Stone pairing* of a formula ϕ in the language of graphs and a finite graph G, denoted by $\langle \phi, G \rangle$, is the probability that ϕ is satisfied by a tuple of vertices of G selected uniformly at random (for ϕ sentence, we set $\langle \phi, G \rangle = 1$ if $G \models \phi$, and $\langle \phi, G \rangle = 0$ otherwise). A sequence of finite graphs (G_n) is said to be FO-*convergent* if the sequence $(\langle \phi, G_n \rangle)$ converges for each ϕ. The limit structure L, called *modeling*, is a graph with measure ν on a standard Borel space satisfying that all the first-order definable are sets measurable. The value $\langle \phi, L \rangle$ is defined as the measure of the set $\phi(L)$, the set of solutions of ϕ in L, using the appropriate power of the measure ν. A modeling L is a limit of an FO-convergent sequence (G_n) if $\lim \langle \phi, G_n \rangle = \langle \phi, L \rangle$ for each ϕ. A modeling limit does not exist for each convergent sequence. It is known to exist for all sequences of graphs from a class $\mathcal{C}$ if and only if $\mathcal{C}$ is a nowhere dense class [10].

The authors of this framework asked in [8] the following question: given a sequence (G_n) converging to a modeling L and a vertex of r of L, is there a sequence of vertices (r_n) such that the graphs G_n rooted at r_n converge to L rooted at r? Christofides and Král' provided an example that the answer is negative in general. However, they also proved that it is possible to find such a sequence (r_n) for almost all choices of the vertex r. That is, if the root of L is chosen at random (according to the measure ν), the vertices (r_n) exist with probability 1 [3].

In this paper, we refine the original problem by considering the root r to be an *algebraic* vertex of L. That is, r belongs to a finite definable set of L. We prove that the sequence of roots (r_n) always exists under such condition. Our main result reads as follows:

Theorem 1. *Let (G_n) be an* FO-*convergent sequence of graphs with a modeling limit L and r be an algebraic vertex of L. Then there is a sequence (r_n), $r_n \in V(G_n)$, such that (G_n, r_n)* FO-*converges to (L, r).*

Note that Theorem 1 deals with full FO-convergence and not just convergence with respect to sentences (called *elementary convergence*), for which it is a trivial statement (see the case of $p = 0$ in Lemma 3).

2 Notation

All graphs are finite except modelings, which are of size continuum. The vertex set of a graph G is denoted by $V(G)$. We use $\mathbb{N} = \{1, 2, \dots\}, \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and $[n] = \{1, 2, \dots, n\}, [n]_0 = [n] \cup \{0\}$. The set of formulas in p free variables in the language of graphs is denoted by FO_p and $\mathrm{FO} = \bigcup_{p \in \mathbb{N}_0} \mathrm{FO}_p$ is the set of all formulas. Tuples of vertices, free variables, etc. are denoted by boldface letters, e.g. $\boldsymbol{x} = (x_1, \dots, x_p)$. Multiset is a set that allows multiplicities of its elements. The power set of a set X is denoted by 2^X.

Let G be an arbitrary graph and r one of its vertices. By (G, r) we denote the graph G rooted at r. Formally, considering G as a structure in the language of graphs, we add a new constant "Root" to the vocabulary and interpret it as r. We refer to the extended language as the language of rooted graphs. The set of formulas in the extended language is denoted by FO^+. Note that $\mathrm{FO}_p \subseteq \mathrm{FO}_p^+$.

Let L be a modeling. A formula $\phi \in \mathrm{FO}_p$ is algebraic in L if $\phi(L)$ is finite, where $\phi(L) = \{\boldsymbol{v} \in V(L)^p : L \models \phi(\boldsymbol{v})\}$ is the set of solutions of ϕ in L. A vertex of L is algebraic if it satisfies an algebraic formula.

3 Rooting in algebraic sets

We prove the following statement, which is equivalent to Theorem 1.

Theorem 2. *Let (G_n) be an* FO*-convergent sequence of graphs with a modeling limit L and $\xi(x)$ be an algebraic formula in L. Then there is a sequence (r_n), $r_n \in V(G_n)$, and a vertex $r \in \xi(L)$ such that (G_n, r_n)* FO*-converges to (L, r).*

Obviously, Theorem 2 is implied by Theorem 1. The converse follows from fact that ξ has only finitely many solutions in L and we can iteratively root them one by one until we reach r.

Fix (G_n), L and ξ for the rest of the paper. Without loss of generality, assume that $|\xi(G_n)| = |\xi(L)|$ for each n and $\xi(L)$ is an inclusion-minimal definable set in L. We prove Theorem 2 in three steps. First, we consider a single formula ϕ in the language of rooted graphs and show that we can find the roots (r_n) and r such that $\lim\langle\phi, (G_n, r_n)\rangle = \langle\phi, (L, r)\rangle$. Then we consider an arbitrary finite collection of formulas $\phi_1, \ldots, \phi_k$ and construct a single formula ψ with the property that convergence of $\langle\psi, (G_n, r_n)\rangle$ to $\langle\psi, (L, r)\rangle$ implies convergence of each $\langle\phi_i, (G_n, r_n)\rangle$ to $\langle\phi_i, (L, r)\rangle$. Finally, a routine use of compactness extends the previous to all formulas, which proves the theorem.

3.1 Single formula

For a formula $\phi(\boldsymbol{x}) \in \mathrm{FO}_p^+$, let $\phi^-(\boldsymbol{x}, y) \in \mathrm{FO}_{p+1}$ be the formula created from ϕ by replacing each occurrence of the term "Root" by "y" (we assume that y does not appear in ϕ).

Lemma 3. *For a given $\phi \in \mathrm{FO}_p^+$ there is a sequence (r_n), $r_n \in \xi(G_n)$, and a vertex $r \in \xi(L)$ such that $\lim\langle\phi, (G_n, r_n)\rangle = \langle\phi, (L, r)\rangle$.*

Proof. If $p = 0$, then either the sentence $(\forall y)(\xi(y) \to \phi^-(y))$ or $(\forall y)(\xi(y) \to \neg\phi^-(y))$ is satisfied in L (using the assumption that $\xi(L)$ is an inclusion-minimal definable set); hence, it holds in each G_n from a certain index on. Therefore, an arbitrary choice of $r_n \in \xi(G_n)$ and $r \in \xi(L)$ meets the conclusion.

Let ν be the measure associated to the modeling L. Define $f_L : V(L)^p \to 2^{\xi(L)}$ to be the function that sends $\boldsymbol{v}$ to the set $\{u \in \xi(L) : L \models \phi^-(\boldsymbol{v}, u)\}$. Consider the pushforward measure μ_L on $2^{\xi(L)}$ of the p-th power of ν by f_L. Viewing $2^{\xi(L)}$ as a lattice, we are mostly interested in the measure of the filter generated by atoms of $2^{\xi(L)}$. Let $X^{\uparrow}$ denote the filters generated by $X \in 2^{\xi(L)}$. Observe that for $u \in \xi(L)$ we have $\mu_L(\{u\}^{\uparrow}) = \langle \phi, (L, u)\rangle$. Suppose that $|\xi(L)| = t$ and define an ordering $R_L = (u_1, u_2, \dots, u_t)$ such that $\mu_L(R_L) = (\mu_L(\{u_i\}^{\uparrow}))_{i\in[t]}$ satisfies $\mu_L(\{u_1\}^{\uparrow}) \geq \mu_L(\{u_2\}^{\uparrow}) \geq \cdots \geq \mu_L(\{u_t\}^{\uparrow})$. Define similarly for each n the function $f_n : V(G_n)^p \to 2^{\xi(G_n)}$, measure μ_n (as the pushforward of the uniform measure) and the vector R_n.

We claim that the sequence $\big(\mu_n(R_n)\big) \subset ([0,1]^t, \|\cdot\|_\infty)$ converges to $\mu_L(R_L)$. Then an arbitrary choice of an index $i \in [t]$ yields the sequence (r_n) and vertex r as the i-th elements of the vectors R_n, resp. R_L.

The claim follows from the fact that the vectors $\mu_n(R_n)$ continuously depend on the values $\langle \psi_{k,\ell}, G_n\rangle$, where $\psi_{k,\ell}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_k) \in \mathrm{FO}_{k\cdot p}$ is

$$(\exists y_1, \dots, y_\ell)\left(\bigwedge_{i=1}^{l} \xi(y_i) \wedge \bigwedge_{1\leq i<j\leq \ell} y_i \neq y_j \wedge \bigwedge_{i=1}^{k}\bigwedge_{j=1}^{\ell} \phi^-(\boldsymbol{x}_i, y_j)\right)$$

for $\ell \in [m]_0, k \in \left[\binom{m}{\ell}\right]$ and that $\langle \psi_{k,\ell}, G_n\rangle \to \langle \psi_{k,\ell}, L\rangle$. This continuous dependency can be proved by inclusion-exclusion with a help of classical results from combinatorics and complex analysis: Girard-Newton formulas [11] and the continuous dependency of the roots of a polynomial on its coefficients [12]. $\square$

3.2 Finite collection of formulas

In this part, we use Lemma 3 to prove an analogous statement for a finite collection of formulas.

Lemma 4. *For given formulas $\phi_1, \dots, \phi_k$ there is a sequence (r_n), $r_n \in \xi(G_n)$, and a vertex $r \in \xi(L)$ such that* $\lim\langle \phi_i, (G_n, r_n)\rangle = \langle \phi_i, (L, r)\rangle$ *for each ϕ_i.*

Proof. Since for sentences any choice of (r_n) and r works, we assume that neither of $\phi_1, \dots, \phi_k$ is a sentence.

Consider an inclusion-maximal set $I \subseteq [k]$ for which there is $v \in \xi(L)$ such that every $i \in I$ satisfies $\langle \phi_i, (L, v)\rangle > 0$, denote $|I|$ by k'. If $I = \emptyset$, we can choose (r_n) and r arbitrarily; hence, assume otherwise. For $i \in I$ set $A_i = \{\langle \phi_i, (L, u)\rangle : u \in \xi(L)\} \cap (0, 1]$. Take a vector $\boldsymbol{e} \in \mathbb{N}^{k'}$ of exponents with the property that for each distinct $\boldsymbol{a}, \boldsymbol{b} \in \bigtimes_{i\in I} A_i$ we have $\prod_{i\in I} a_i^{e_i} \neq \prod_{i\in I} b_i^{e_i}$. Such a vector exists as each A_i is finite and contains only positive values. The set of *bad* choices of rational exponents that make the values for particular $\boldsymbol{a}, \boldsymbol{b}$ coincide form a $(k'-1)$-dimensional hyperplane in $\mathbb{Q}^{k'}$. We can surely avoid finitely many of such hyperplanes (one for each choice of $\boldsymbol{a}$ and $\boldsymbol{b}$) to find a *good* vector of positive rational exponents and scale them to integers.

Use Lemma 3 for the formula ψ of the form

$$\bigwedge_{i\in I}\bigwedge_{j=1}^{e_i}\phi_i(\boldsymbol{x}_{i,j}),$$

where all the tuples $\boldsymbol{x}_{i,j}$ are pairwise disjoint, to obtain roots (r_n) and r. In particular, we can take the vertex r such that $\langle\psi,(L,r)\rangle > 0$ (due to our choice of I).

We have $\lim\langle\phi_i,(G_n,r_n)\rangle = \langle\phi_i,(L,r)\rangle > 0$ for each $i \in I$ as

$$\langle\psi,(L,r)\rangle = \prod_{i\in I}\langle\phi_i,(L,r)\rangle^{e_i},$$

using our selection of exponents $\boldsymbol{e}$.

Also, it holds that $\lim\langle\phi_j,(G_n,r_n)\rangle = \langle\phi_j,(L,r)\rangle = 0$ for each $j \notin I$: for the formula $\chi = \bigwedge_{i\in I\cup\{j\}}\phi_i(\boldsymbol{x}_i)$, we have $\lim\langle\chi,(G_n,r_n)\rangle = \langle\chi,(L,r)\rangle = 0$ due to the maximality of I (this is for *any* choice of (r_n) and r). We have

$$\langle\chi,(G_n,r_n)\rangle = \prod_{i\in I\cup\{j\}}\langle\phi_i,(G_n,r_n)\rangle$$

and as for some $\varepsilon > 0$ there is n_0 such that $\langle\phi_i,(G_n,r_n)\rangle > \varepsilon$ for each $i \in I$ and $n \geq n_0$, the factor $\langle\phi_j,(G_n,r_n)\rangle$ must tend to 0. □

We remark that the rationalization of the fact that the sequence $\big(\langle\phi_j,(G_n,r_n)\rangle\big)$ for $j \notin I$ even converge is the reason why we are proving Theorem 2 instead of Theorem 1 directly. We are using the fact that we can choose the set I (and the root r for the formula ψ) such that any rooting (r_n) makes the sequence $\langle\chi,(G_n,r_n)\rangle$ converge to 0.

4 Concluding remarks

An iterative use of Theorem 1 or 2 allows us to gain complete control over the algebraic elements as we can consider each of them separately.

We note that it is possible to root solutions of algebraic formulas with multiple free variables as the projection to each coordinate yields an algebraic set. Moreover, the natural modification of Theorem 2 remains valid for FO-convergent sequences (G_n) without a modeling limit. The proofs are analogous except that the set I in Lemma 4 is defined as an inclusion-maximal set for which there are roots (r_n) such that $\lim\langle\bigwedge_{i\in I}\phi_i(\boldsymbol{x}_i),(G_n,r_n)\rangle > 0$.

Besides the original problem in [8], our motivation was the study of structural convergence of sequences created via gadget construction, see [5]. Using the result of this paper, we conclude that FO-convergence is preserved if the gadgets replace only finitely many edges (under natural additional assumptions).

In the typical case, the modeling L is of size continuum and the set of algebraic vertices (which is at most countable) has measure 0. Hence, our results reveal only a negligible portion of vertices of L for which the roots (r_n) exist, which shows that there is still room for further research.

References

[1] Itai Benjamini and Oded Schramm. "Recurrence of Distributional Limits of Finite Planar Graphs". In: *Electronic Journal of Probability* 6.none (Jan. 2001). DOI: `10.1214/ejp.v6-96`.

[2] Christian Borgs et al. "Counting Graph Homomorphisms". In: *Topics in Discrete Mathematics*. Springer Berlin Heidelberg, 2006, pp. 315–371. ISBN: 978-3-540-33700-3. DOI: `https://doi.org/10.1007/3-540-33700-8_18`.

[3] Demetres Christofides and Daniel Kráľ. "First-Order Convergence and Roots". In: *Combinatorics, Probability and Computing* 25.2 (Feb. 2015), pp. 213–221. DOI: `10.1017/s0963548315000048`.

[4] Gábor Elek. "Note on limits of finite graphs". In: *Combinatorica* 27.4 (July 2007), pp. 503–507. DOI: `10.1007/s00493-007-2214-8`.

[5] David Hartman, Tomáš Hons, and Jaroslav Nešetřil. *Gadget construction and structural convergence*. 2022. DOI: `10.48550/ARXIV.2212.10985`.

[6] László Lovász. *Large networks and graph limits*. American Mathematical Society, 2012. ISBN: 9780821890851.

[7] László Lovász and Balázs Szegedy. "Limits of dense graph sequences". In: *Journal of Combinatorial Theory, Series B* 96.6 (Nov. 2006), pp. 933–957. DOI: `10.1016/j.jctb.2006.05.002`.

[8] Jaroslav Nešetřil and Patrice Ossona de Mendez. "A Unified Approach to Structural Limits and Limits of Graphs with Bounded Tree-Depth". In: *Memoirs of the American Mathematical Society* 263.1272 (Jan. 2020). DOI: `10.1090/memo/1272`.

[9] Jaroslav Nešetřil and Patrice Ossona de Mendez. "Cluster analysis of local convergent sequences of structures". In: *Random Structures & Algorithms* 51.4 (Apr. 2017), pp. 674–728. DOI: `10.1002/rsa.20719`.

[10] Jaroslav Nešetřil and Patrice Ossona de Mendez. "Existence of Modeling Limits for Sequences of Sparse Structures". In: *The Journal of Symbolic Logic* 84.02 (Mar. 2019), pp. 452–472. DOI: `10.1017/jsl.2018.32`.

[11] Richard P. Stanley and Sergey Fomin. *Enumerative Combinatorics*. Cambridge University Press, Jan. 1999. DOI: `10.1017/cbo9780511609589`.

[12] Hassler Whitney. *Complex Analytic Varieties*. Addsion Weslsey Longman Publishing Group, 1972, p. 399. ISBN: 9780201086539.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Colouring complete multipartite and Kneser-type digraphs

(Extended abstract)

Ararat Harutyunyan* Gil Puig i Surroca†

Abstract

The dichromatic number of a digraph D is the smallest k such that D can be partitioned into k acyclic subdigraphs, and the dichromatic number of an undirected graph is the maximum dichromatic number over all its orientations. We present bounds for the dichromatic number of Kneser graphs and Borsuk graphs, and for the list dichromatic number of certain classes of Kneser graphs and complete multipartite graphs. The bounds presented are sharp up to a constant factor. Additionally, we give a directed analogue of Sabidussi's theorem on the chromatic number of graph products.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-076`

We consider graphs and digraphs without loops or multiple edges/arcs. A *proper k-colouring* of a graph $G = (V, E)$ is a mapping $f : V \to [k] = \{1, ..., k\}$ such that $f^{-1}(i)$ is an independent set for any $i \in [k]$. The *chromatic number* of G, denoted by $\chi(G)$, is the minimum k for which G has a proper k-colouring. A *proper k-colouring* of a digraph $D = (V, A)$ is a mapping $f : V \to [k]$ such that $f^{-1}(i)$ is acyclic for any $i \in [k]$, and the *dichromatic number* of D, denoted by $\vec{\chi}(D)$, is the minimum k for which D has a proper k-colouring. Note that this definition generalizes the usual colouring, in the sense that the chromatic number of a graph is equal to the dichromatic number of its corresponding bidirected digraph. The notion was introduced by Neumann-Lara in 1982 [17] and it

*LAMSADE, Université Paris Dauphine - PSL, 75775 Paris Cedex 16, France. E-mail: `ararat.harutyunyan@lamsade.dauphine.fr`.

†LAMSADE, Université Paris Dauphine - PSL, 75775 Paris Cedex 16, France. E-mail: `gil.puig-i-surroca@dauphine.eu`.

was later rediscovered by Mohar [14]. Since then it has been shown that many classical results hold also in this setting [3, 8, 9, 10]. However, some fundamental questions remain unanswered. The *dichromatic number* of an undirected graph G, denoted by $\vec{\chi}(G)$, is the maximum dichromatic number over all its orientations. Erdős and Neumann-Lara conjectured the following.

Conjecture 1. [5] *For every integer k there exists an integer $r(k)$ such that $\vec{\chi}(G) \geq k$ for any undirected graph G satisfying $\chi(G) \geq r(k)$.*

For instance, $r(1) = 1$ and $r(2) = 3$. But it is already unknown whether $r(3)$ exists. Mohar and Wu [15] managed to prove the fractional analogue of Conjecture 1.

The *Kneser graph* with parameters n, k, denoted by $KG(n, k)$, is the graph with vertex set $\binom{[n]}{k}$ (i.e. the set of subsets of $[n]$ of size k) where two vertices u, v are adjacent if and only if $u \cap v = \emptyset$. It is well-known [7, 12, 13] that $\chi(KG(n, k)) = n - 2k + 2$ for $1 \leq k \leq \frac{n}{2}$, as Kneser conjectured [11, 20]. Providing further evidence for Conjecture 1, Mohar and Wu showed that Kneser graphs with large chromatic number have large dichromatic number.

Theorem 2. [15] *For any positive integers n, k with $1 \leq k \leq \frac{n}{2}$ we have that $\vec{\chi}(KG(n, k)) \geq \left\lfloor \frac{n-2k+2}{8 \log_2 \frac{n}{k}} \right\rfloor$.*

Note that, since $\chi(KG(n, k)) \geq \vec{\chi}(KG(n, k))$, this estimate is sharp up to a constant factor when k is a constant fraction of n. Improving Theorem 2 asymptotically, we show that the dichromatic number of Kneser graphs is of the order of their chromatic number in general.

Theorem 3. *There exists a positive integer n_0 such that, for all $n \geq n_0$ and $2 \leq k \leq \frac{n}{2}$, we have that $\vec{\chi}(KG(n, k)) \geq \left\lfloor \frac{1}{16} \chi(KG(n, k)) \right\rfloor$.*

We did not try to optimize the constant $\frac{1}{16}$. The proof of Theorem 3 is based on Greene's proof of Kneser's conjecture, but it also relies on Theorem 2 for solving the case of large k. Note that the bound cannot be extended to $k = 1$ (see Theorem 11).

Kneser's conjecture was an open problem for more than two decades [11, 20]. The famous resolution by Lóvasz [12] was inspired by the analogy between Kneser graphs and Borsuk graphs. Let n be a natural number and $a < 2$ a positive real number. The *Borsuk graph* with parameters $n + 1$ and a, denoted by $BG(n + 1, a)$, is the undirected graph with vertex set $\mathbb{S}^n = \{x \in \mathbb{R}^{n+1} \mid \|x\| = 1\}$ where two vertices x, y are adjacent if and only if $\mathrm{dist}_{\mathbb{R}^{n+1}}(x, y) \geq a$. It is known that $\chi(BG(n + 1, a)) \geq n + 2$, which in fact is equivalent to the Borsuk–Ulam theorem [13]. On the other hand, if a is not too small, an $(n + 2)$-colouring of $BG(n + 1, a)$ can be obtained by projecting the faces of an inscribed $(n + 1)$-dimensional simplex. Regarding the dichromatic number of Borsuk graphs, we can show the following.

Theorem 4. $\vec{\chi}(BG(n + 1, a)) \geq n + 2$ *for any $n \geq 1$.*

Next, we look at list colourings. They were introduced by Erdős, Rubin and Taylor [6], and, independently, by Vizing [19]. A *k-list assignment* to a graph $G = (V, E)$ (or to a digraph $D = (V, A)$) is a mapping $L : V \to \binom{\mathbb{Z}^+}{k}$. A colouring (a mapping) $f : V \to \mathbb{Z}^+$ is said to be *accepted* by L if $f(v) \in L(v)$ for every $v \in V$. G (resp. D) is *k-list colourable* if every k-list assignment accepts a proper colouring. The *list chromatic number* of G (resp. the *list dichromatic number* of D), also called its *choice number*, is the minimum k such that G (resp. D) is k-list colourable, and it is denoted by $\chi_\ell(G)$ (resp. $\vec{\chi}_\ell(D)$). Similarly, the *list dichromatic number* of G, denoted by $\vec{\chi}_\ell(G)$, is the maximum list dichromatic number over all orientations of G. Bensmail, Harutyunyan and Le [2] gave a sample of instances where the list dichromatic number of digraphs behaves as its undirected counterpart.

Recently, Bulankina and Kupavskii [4] studied the list chromatic number of Kneser graphs. They proved the following two results.

Theorem 5. [4] *For any positive integers n, k with $1 \le k \le \frac{n}{2}$ we have that $\chi_\ell(KG(n,k)) \le n \ln \frac{n}{k} + n$.*

Theorem 6. [4] *Let $s \ge 3$ be an integer. If n is sufficiently large and $3 \le k \le n^{1/2-1/s}$, then $\chi_\ell(KG(n,k)) \ge \frac{1}{2s^2} n \ln n$. For $k = 2$, we have that $\chi_\ell(KG(n,k)) \ge \frac{1}{32} n \ln n$ for sufficiently large n.*

However, good bounds for larger k are still unknown. Using the arguments of Bulankina and Kupavskii, as well as ideas from [15], we can prove the directed version of Theorem 6.

Theorem 7. *For every $\varepsilon \in \mathbb{R}^+$ there exists a constant $c_\varepsilon \in \mathbb{R}^+$ such that $\vec{\chi}_\ell(KG(n,k)) \ge c_\varepsilon n \ln n$ for all $n \ge 2k$ with $2 \le k \le n^{1/2-\varepsilon}$.*

Dense Kneser graphs have some similarities with complete multipartite graphs. Denote by K_{m*r} the complete r-partite graph with m vertices on each part. Alon determined, up to a constant factor, the list chromatic number of K_{m*r}, answering a question of Erdős, Rubin and Taylor [6].

Theorem 8. [1] *There exist two positive constants c_1 and c_2 such that for every $m \ge 2$ and for every $r \ge 2$*

$$c_1 r \ln m \le \chi_\ell(K_{m*r}) \le c_2 r \ln m.$$

His proof can be adapted to find an analogous bound for the list dichromatic number of K_{m*r} when m is not too small.

Theorem 9. *For every $\rho > 3$, there exist constants $C_1, C_2 \in \mathbb{R}^+$ such that if $r \ge 2$ and $m \ge \ln^\rho r$ then*

$$C_1 r \ln m \le \vec{\chi}_\ell(K_{m*r}) \le C_2 r \ln m.$$

In what follows we present a proof of Theorem 9. The following probabilistic result will be required.

Theorem 10. (Simple Concentration Bound, [16]) *Let X be a random variable determined by n independent trials, and satisfying the property that changing the outcome of any single trial can affect X by at most c. Then*

$$\mathbb{P}(|X - \mathbb{E}X| > t) \leq 2e^{-\frac{t^2}{2c^2 n}}.$$

Proof of Theorem 9. The upper bound is implied by Theorem 8. We may assume that m is large enough. Let $V_1, ..., V_r$ be the parts of K_{m*r}.

Claim. There is a constant c and an orientation D of K_{m*r} such that, if $\ell \geq c\ln(rm)$, then

(i) each subgraph of K_{m*r} isomorphic to K_ℓ has a directed cycle in D;

(ii) for each $U_i \subseteq V_i$ and $U_j \subseteq V_j$ with $|U_i| = |U_j| = \ell$ and $i \neq j$, $D[U_i \cup U_j]$ has a directed cycle.

Proof. Orient the edges of K_{m*r} at random, independently and with probability $\frac{1}{2}$. Let E, E' be the events that (i), (ii) hold, respectively. Put $\ell = \lceil c\ln(rm)\rceil$. There are $\binom{r}{\ell}m^\ell$ copies of K_ℓ in K_{m*r}, and $\binom{r}{2}\binom{m}{\ell}^2$ subgraphs of the form $K_{m*r}[U_i \cup U_j]$. Furthermore, K_ℓ (resp. $K_{m*r}[U_i \cup U_j]$) has $2^{\frac{\ell(\ell-1)}{2}}$ orientations (resp. 2^{ℓ^2}), among which $\ell!$ (resp. at most $(2\ell)!$) are acyclic. Therefore,

$$\mathbb{P}(E^{\mathrm{c}}) \leq \binom{r}{\ell} m^\ell \ell!\, 2^{-\frac{\ell(\ell-1)}{2}} \leq \left(rm2^{-\frac{\ell-1}{2}}\right)^\ell \leq \left(e^{\frac{\ell}{c}}2^{-\frac{\ell-1}{2}}\right)^\ell < \frac{1}{2} \text{ and}$$

$$\mathbb{P}(E'^{\mathrm{c}}) \leq \binom{r}{2}\binom{m}{\ell}^2 (2\ell)!\, 2^{-\ell^2} \leq \left(2er^2m^22^{-\ell}\right)^\ell \leq \left(e^{\frac{2\ell}{c}+1}2^{-\ell+1}\right)^\ell < \frac{1}{2}$$

if c is large enough. Hence $\mathbb{P}(E \cap E') > 0$ for some c. □

Let $k = \lfloor Cr\ln m\rfloor$, where $0 < C \leq 1$ is a constant for now unspecified. We start by showing that there exists an assignment of k-lists from a palette $\mathscr{C}$ of $\lfloor r\ln m\rfloor$ colours such that, for any given set $A \subseteq \mathscr{C}$ of at most $\frac{4}{3}\ln m$ colours, each part has at least $\frac{1}{2}m^{1-\delta}$ vertices that avoid the colours from A on their lists, where $\delta = 2C\ln 5$.

Assign to each vertex v of D a random k-list $L(v)$ chosen independently and uniformly among the $\binom{|\mathscr{C}|}{k}$ possible k-lists. Given $i \in [r]$ and $A \subseteq \mathscr{C}$, consider the random variable $X_{i,A} = |\{v \in V_i \mid L(v) \cap A = \emptyset\}|$. Note that there are exactly $\binom{|\mathscr{C}|-|A|}{k}$ k-lists avoiding the colours in A. Devoting ourselves to the case $|A| = \left\lfloor \frac{4}{3}\ln m\right\rfloor$, we have that

$$\mathbb{E}X_{i,A} = m\frac{\binom{|\mathscr{C}|-|A|}{k}}{\binom{|\mathscr{C}|}{k}} \geq m\left(\frac{|\mathscr{C}|-|A|-k}{|\mathscr{C}|-k}\right)^k = m\left(1-\frac{|A|}{|\mathscr{C}|-k}\right)^k$$

$$\geq m\left(1-\frac{\frac{4}{3}\ln m}{(1-C)r\ln m - 1}\right)^{Cr\ln m} \geq m\left(1-\frac{4}{5}\right)^{2C\ln m} = m^{1-\delta}$$

if m is large enough and C is not too large. By the Simple Concentration Bound (Theorem 10),

$$\mathbb{P}(X_{i,A} < \frac{1}{2}m^{1-\delta}) \leq \mathbb{P}(|X_{i,A} - \mathbb{E}X_{i,A}| > \frac{1}{2}m^{1-\delta}) \leq 2e^{-\frac{1}{8}m^{1-2\delta}}.$$

Let E be the event that $X_{i,A} < \frac{1}{2}m^{1-\delta}$ for some $i \in [r]$ and $A \subseteq \mathscr{C}$ with $|A| \leq \frac{4}{3}\ln m$. We have that

$$\mathbb{P}(E) \leq r\binom{|\mathscr{C}|}{\lfloor \frac{4}{3}\ln m \rfloor} 2e^{-\frac{1}{8}m^{1-2\delta}} \leq r\left(e\frac{\lfloor r\ln m \rfloor}{\lfloor \frac{4}{3}\ln m \rfloor}\right)^{\lfloor \frac{4}{3}\ln m \rfloor} 2e^{-\frac{1}{8}m^{1-2\delta}}$$

$$\leq r(er)^{\frac{4}{3}\ln m} 2e^{-\frac{1}{8}m^{1-2\delta}} \leq 2e^{5\ln r\ln m - \frac{1}{8}m^{1-2\delta}} \leq 2e^{5m^{\frac{1}{\rho}}\ln m - \frac{1}{8}m^{1-2\delta}}$$

if m is large enough. Consequently, if $\delta < \frac{1}{2}(1-\frac{1}{\rho})$ and m is large enough, there exists a list assignment L' satisfying the desired property. This is the assignment that we are going to use.

Now let f be a proper colouring of D. We claim that there exists a set of indices $I \subseteq [r]$ of size at least $\frac{3r}{4}$ such that $|f(V_i)| \leq 4c\ln^2(rm)$ for each $i \in I$. Indeed, if more than $\frac{r}{4}$ parts are coloured with more than $4c\ln^2(rm)$ colours each, then one of the colours appears on more than $\frac{cr\ln^2(rm)}{|\mathscr{C}|} \geq c\frac{\ln^2(rm)}{\ln m} \geq c\ln(rm)$ parts. By the choice of D, f is not proper, a contradiction.

For each $i \in [r]$ define the set $A_i = \{\gamma \in \mathscr{C} \mid |V_i \cap f^{-1}(\gamma)| \geq c\ln(rm)\}$. We claim that if f is acceptable then $|A_i| > \frac{4}{3}\ln m$ for every $i \in I$. Indeed, otherwise, by the choice of the lists, at least $\frac{1}{2}m^{1-\delta}$ vertices of V_i have been coloured with colours not from A_i. Thus one of these colours is used at least

$$\frac{\frac{1}{2}m^{1-\delta}}{4c\ln^2(rm)} \leq c\ln(rm)$$

times on V_i. If m is large enough, this implies that

$$m^{1-\delta} \leq 8c^2\ln^3(rm) \leq 8c^2(m^{\frac{1}{\rho}} + \ln m)^3 \leq 9c^2 m^{\frac{3}{\rho}}.$$

If we further assume that $\delta < 1 - \frac{3}{\rho}$, we get a contradiction when m is large. Therefore $|A_i| > \frac{4}{3}\ln m$ for every $i \in I$.

Now, by the choice of D, the sets $A_1, ..., A_r$ are mutually disjoint. But then

$$|\mathscr{C}| \geq \sum_{i=1}^{r}|A_i| \geq \sum_{i\in I}|A_i| > \frac{4}{3}|I|\ln m \geq r\ln m \geq |\mathscr{C}|.$$

This contradiction shows that there is no acceptable proper colouring for the k-list assignment L'. □

We do not know what happens with other values of m, r. What is clear is that the bound of Theorem 9 is not valid in general. Indeed, if $m \leq \ln r$ then Theorem 11 implies that $\vec{\chi}(K_{m*r}) \leq \vec{\chi}(K_{mr}) \leq cr$ for some constant c.

Theorem 11. [2] *Let T be a tournament of order n. Then $\vec{\chi}_\ell(T) \leq \frac{n}{\log_2 n}(1 + o(1))$.*

Some of our proofs rely on graph products. Let G, H be graphs (resp. digraphs). The *Cartesian product* of G and H is the graph (resp. digraph) $G \square H$ with vertex set $V(G) \times V(H)$ where there is an edge between (u, x) and (v, y) (resp. an arc from (u, x) to (v, y)) if and only if either $u = v$ and $\{x, y\} \in E(H)$ (resp. and $(x, y) \in A(H)$), or $x = y$ and $\{u, v\} \in E(G)$ (resp. and $(u, v) \in A(G)$). A well-known theorem of Sabidussi [18] states that for any two graphs G and H the chromatic number of its product is $\chi(G \square H) = \max\{\chi(G), \chi(H)\}$. His proof can be adapted to show an analogous result for digraphs.

Theorem 12. *Let G and H be digraphs. Then $\vec{\chi}(G \square H) = \max\{\vec{\chi}(G), \vec{\chi}(H)\}$.*

Acknowledgements. We thank Stéphane Bessy and Stéphan Thomassé for many previous conversations around Conjecture 1, which motivated us to start this work. We also wish to thank Denis Cornaz for exciting and encouraging discussions. The authors are supported by ANR grant 21-CE48-0012 DAGDigDec (DAGs and Digraph Decompositions). GPS was also partially supported by the Spanish Ministerio de Ciencia e Innovación through grant PID2019-194844GB-I00.

References

[1] N. Alon. Choice Numbers of Graphs: a Probabilistic Approach. *Combinatorics, Probability and Computing*, 1(2):107–114, 1992.

[2] J. Bensmail, A. Harutyunyan, and N. K. Le. List coloring digraphs. *Journal of Graph Theory*, 87(4):492–508, 2018.

[3] D. Bokal, G. Fijavz, M. Juvan, P. Kayll, and B. Mohar. The circular chromatic number of a digraph. *Journal of Graph Theory*, 46(3):227–240, 2004.

[4] V. Bulankina and A. Kupavskii. Choice number of Kneser graphs. *Discrete Mathematics*, 345(11), 2022.

[5] P. Erdős. Problems and results in number theory and graph theory. In *Proceedings of the 9th Manitoba Conference on Numerical Mathematics and Computing*, pages 3–21, 1979.

[6] P. Erdős, A. L. Rubin, and H. Taylor. Choosability in graphs. In *Proceedings of the West Coast Conference on Combinatorics, Graph Theory and Computing, Congressus Numerantium 26*, pages 125–157, 1980.

[7] J. E. Greene. A New Short Proof of Kneser's Conjecture. *The American Mathematical Monthly*, 109(10):918–920, 2002.

[8] A. Harutyunyan and B. Mohar. Gallai's Theorem for List Coloring of Digraphs. *SIAM Journal on Discrete Mathematics*, 25(1):170–180, 2011.

[9] A. Harutyunyan and B. Mohar. Strengthened Brooks' Theorem for Digraphs of Girth at least Three. *Electr. J. Comb.*, 18, 10 2011.

[10] A. Harutyunyan and B. Mohar. Two results on the digraph chromatic number. *Discrete Mathematics*, 312(10):1823–1826, 2012.

[11] M. Kneser. Aufgabe 360. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 58(27):107–116, 1955.

[12] L. Lovász. Kneser's Conjecture, Chromatic Number, and Homotopy. *Journal of Combinatorial Theory, Series A*, 25(3):319–324, 1978.

[13] J. Matoušek, A. Björner, and G. M. Ziegler. *Using the Borsuk-Ulam Theorem*. Berlin: Springer, 2003.

[14] B. Mohar. Circular colorings of edge-weighted graphs. *Journal of Graph Theory*, 43:107–116, 2003.

[15] B. Mohar and H. Wu. Dichromatic number and fractional chromatic number. *Forum of Mathematics, Sigma*, 4, 2016.

[16] M. Molloy and B. Reed. *Graph Colouring and the Probabilistic Method*. Springer, 2002.

[17] V. Neumann-Lara. The Dichromatic Number of a Digraph. *Journal of Combinatorial Theory*, 33:265–270, 1982.

[18] G. Sabidussi. Graphs with Given Group and Given Graph-Theoretical Properties. *Canadian Journal of Mathematics*, 9:515–525, 1957.

[19] V. G. Vizing. Vertex colorings with given colors (in Russian). *Metody Diskret. Analiz.*, 29:3–10, 1976.

[20] G. M. Ziegler. *"Aufgabe 360: The Kneser Conjecture"*, pages 743–749. Springer International Publishing, 2021.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

The hitting time of clique factors

(Extended abstract)

Annika Heckel* Marc Kaufmann† Noela Müller‡ Matija Pasch§

Abstract

In [13], Kahn gave the strongest possible, affirmative, answer to Shamir's problem, which had been open since the late 1970s: Let $r \geqslant 3$ and let n be divisible by r. Then, in the random r-uniform hypergraph process on n vertices, as soon as the last isolated vertex disappears, a perfect matching emerges. In the present work, we prove the analogue of this result for clique factors in the random graph process: At the time that the last vertex joins a copy of the complete graph K_r, the random graph process contains a K_r-factor. Our proof draws on a novel sequence of couplings which embeds the random hypergraph process into the cliques of the random graph process. An analogous result is proved for clique factors in the s-uniform hypergraph process ($s \geqslant 3$).

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-077

*Matematiska institutionen, Uppsala universitet, Box 480, 751 06 Uppsala, Sweden. Email: annika.heckel@math.uu.se. The research leading to these results has received funding from the European Research Council, ERC Grant Agreement 772606–PTRCSP, and from the Swedish Research Council, reg. nr. 2022-02829.

†Institut für Theoretische Informatik, ETH Zürich, Zürich, Switzerland. Email: marc.kaufmann@inf.ethz.ch. The author gratefully acknowledges support by the Swiss National Science Foundation [grant number 200021_192079].

‡Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands. Email: n.s.muller@tue.nl. Research supported by NWO Gravitation project NETWORKS under grant no. 024.002.003.

§Mathematisches Institut, Ludwig-Maximilians-Universität München, Theresienstr. 39, 80333 München, Germany. Email: pasch@math.lmu.de. The research leading to these results has received funding from the European Research Council, ERC Grant Agreement 772606–PTRCSP.

1 Introduction

When can we cover the vertices of a graph with disjoint isomorphic copies of a small subgraph? The study of this question goes back at least to 1891, when Julius Petersen, in his *Theorie der regulären graphs* [16], provided sufficient conditions for a graph to contain a perfect matching, that is, a cover of the vertices with pairwise disjoint edges. Let $H_r(n,\pi)$ be the random r-uniform hypergraph on the vertex set $V=[n]$ where each of the $N_r=\binom{n}{r}$ possible hyperedges of size r is present independently with probability π. The binomial random graph in this notation is then $G(n,p)=H_2(n,p)$. In 1979, Shamir asked the following natural question, as reported by Erdős [6]:

Question 1.1. How large does $\pi=\pi(n)$ need to be for $H_r(n,\pi)$ to contain a perfect matching whp[1], that is, a collection of n/r vertex-disjoint hyperedges[2]?

A closely related question, posed by Ruciński [18] and Alon and Yuster [1], is:

Question 1.2. For which $p=p(n)$ does the random graph $G(n,p)$ contain a K_r-factor whp?

That is, for which $p(n)$ does $G(n,p)$ contain a collection of n/r vertex-disjoint copies of K_r? In the following, we will also call a copy of K_r an r-clique. For $r=2$, the two questions are the same — and thanks to Erdős and Rényi [5], we have known since 1966 that there is a sharp threshold[3] for the existence of a perfect matching at $p_0=\frac{\log n}{n}$. The lower bound for this is immediate: At $p=(1-\varepsilon)p_0$, some vertices in the graph are still isolated, so there cannot be a perfect matching. The upper bound relies on Tutte's Theorem, for which there is no known hypergraph analogue.

On the other hand, for the case $r\geqslant 3$, these questions remained some of the most prominent open problems in random (hyper-)graph theory. Initial results on perfect matchings in random r-uniform hypergraphs were obtained by Schmidt and Shamir [19] - guaranteeing a perfect matching for hypergraphs with expected degree $\omega(\sqrt{n})$, with improvements by Frieze and Janson [7] to $\omega(n^{\frac{1}{3}})$ and further to $\omega\big(n^{1/(5+2/(r-1))}\big)$ by Kim [14]. For clique factors, even determining the special case of triangle factors proved hard, despite partial results by Alon and Yuster [1], Ruciński [18] and Krivelevich [15]. Finally, both questions were jointly resolved up to constant factors by Johansson, Kahn and Vu in their seminal paper [11]. It had long been assumed that, as in the case $r=2$, the main obstacle in finding a perfect matching in $H_r(n,\pi)$ were *isolated vertices*, that is, vertices not contained in any hyperedge. In the clique factor setting, the obstacle corresponding to isolated vertices are vertices not contained in any r-clique. Let

$$\pi_0=\pi_0(r)=\frac{\log n}{\binom{n-1}{r-1}}\quad\text{and}\quad p_0=p_0(r)=\pi_0^{1/\binom{r}{2}};$$

[1] We say that a sequence of events $(E_n)_{n\geqslant 1}$ holds with high probability (whp) if $P(E_n)\to 1$ as $n\to\infty$.

[2] Here and in the following, we implicitly assume $n\in r\mathbb{Z}_+$ whenever necessary.

[3] Recall that a sequence $p^*=p^*(n)$ is called a *sharp threshold* for a graph property $\mathcal{P}$, if for all fixed $\epsilon>0$ we have $G(n,p)\notin\mathcal{P}$ whp if $p(n)<(1-\epsilon)p^*(n)$, and $G(n,p)\in\mathcal{P}$ whp if $p(n)>(1+\epsilon)p^*(n)$. For a (weak) threshold, the conditions become $p=o(p^*)$ and $p^*=o(p)$, respectively.

then π_0 and p_0 are known to be sharp thresholds for the properties 'minimum degree at least 1' in $H_r(n,\pi)$ and 'every vertex is covered by an r-clique' in $G(n,p)$, respectively [4, 10]. Johansson, Kahn and Vu [11] showed that π_0 and p_0 are indeed (weak) thresholds for the existence of a perfect matching in $H_r(n,\pi)$ and for the existence of an r-clique factor in $G(n,p)$, respectively.

Recently, Kahn [12] proved that π_0 is in fact a *sharp* threshold for the existence of a perfect matching in $H_r(n,\pi)$. Indeed, he was able to confirm the conjecture that isolated vertices are essentially the only obstacle, and thereby answer Shamir's question, in the strongest possible sense:

Let $\mathbf{h}_1,\dots,\mathbf{h}_{N_r}$ be a uniformly random order of the hyperedges in $\binom{V}{r}$, then the *random r-uniform hypergraph process* $(H_t^r)_{t=0}^{N_r}$ is given by $H_t^r = \{\mathbf{h}_1,\dots,\mathbf{h}_t\}$. Let

$$T_H = \min\{t : H_t^r \text{ has no isolated vertices}\}$$

be the hyperedge cover hitting time, i.e., the time t where the last isolated vertex 'disappears' by being included in a hyperedge. In the graph case $r=2$, Bollobás and Thomason [2] proved in 1985 that this hitting time whp coincides with the hitting time for a perfect matching. Kahn [13] showed that this is indeed also the case when $r \geqslant 3$:

Theorem 1.3 ([13])**.** *Let $r \geqslant 3$ and $n \in r\mathbb{Z}_+$, then whp $H_{T_H}^r$ has a perfect matching.*

Can we get a similarly strong answer to the clique factor question? For $r=3$, the question whether a triangle factor exists in the random graph process as soon as every vertex is covered by a triangle was attributed to Erdős and Spencer in [3, §5.4]. This question seems much harder than its Shamir counterpart because, unlike hyperedges in the random hypergraph, cliques do not appear independently of each other. However, for sharp thresholds it has indeed been possible to *reduce* the clique factor problem to the perfect matching problem, using the following coupling result of Riordan (for $r \geqslant 4$) and the first author (for $r=3$):

Theorem 1.4 ([8, 17])**.** *Let $r \geqslant 3$. There are constants $\varepsilon(r), \delta(r) > 0$ such that, for any $p = p(n) \leqslant n^{-2/r+\varepsilon}$, letting $\pi = p^{\binom{r}{2}}(1-n^{-\delta})$, we may couple the random graph $G = G(n,p)$ with the random r-uniform hypergraph $H = H_r(n,\pi)$ so that, whp, for every hyperedge in H there is a copy of K_r in G on the same vertex set.*[4]

Together with Kahn's sharp threshold result [12], the following corollary is immediate.

Corollary 1.5. *There is a sharp threshold for the existence of a K_r-factor at p_0.*

In the same spirit, we wish to transfer Kahn's hitting time theorem, Theorem 1.3, directly to the random graph process setting, showing its clique factor analogue. Such a derivation of the factor result from its Shamir counterpart was believed to be out of reach — Kahn remarks in [12] that 'there seems little chance of anything analogous' for

[4]In [8, 17], Theorem 1.4 was given with an unspecified $o(1)$-term in place of $n^{-\delta}$; the formulation above is Remark 4 in [17] and in the case $r=3$, an unnumbered remark near the end of [8].

Theorem 1.3, and in [13] that the connection between the factor version and the Shamir version of the result 'seems unlikely to extend to' Theorem 1.3. One important reason for this is that the original coupling provides merely a one-way bound. While it guarantees a copy of K_r in $G = G(n,p)$ on the same vertex set for every h in $H = H_r(n,\pi)$, we cannot, as observed by Riordan [17], expect to find a corresponding hyperedge of H for every K_r in G, since there we will find roughly $n^{2r-2}p^{2\binom{r}{2}-1}$ pairs of K_r sharing two vertices, which is much larger than the expected number $n^{2r-2}\pi^2$ of pairs of hyperedges of H sharing two vertices. A second obstacle is that whenever we do have such a pair of overlapping hyperedges in H, the corresponding cliques in G will not appear independently of each other in the associated random graph process — for example the shared edge could be the last to appear, and then those cliques emerge simultaneously in the random graph process. And indeed, extra cliques and pairs of overlapping cliques do pose a challenge, but they will not appear 'near' those candidate vertices which may be among the last vertices to be covered by cliques.

Now, let $(G_t)_{t=0}^{N_2}$ be the random graph process, which is the random r-uniform hypergraph process for $r = 2$. Denote the hitting time of an r-clique cover by

$$T_G = \min\{t : \text{ every vertex in } G_t \text{ is contained in at least one } r\text{-clique}\}.$$

Then, to apply Kahn's hitting time result to the clique factor setting, we need to find a copy of $H^r_{T_H}$ within the cliques of G_{T_G}. That this can be achieved is our main result:

Theorem 1.6. *Let $r \geqslant 3$. We may couple the random graph process $(G_t)_{t=0}^{N_2}$ with the random r-uniform hypergraph process $(H^r_t)_{t=0}^{N_r}$ so that, whp, for every hyperedge in $H^r_{T_H}$ there is a clique in G_{T_G} on the same vertex set. In particular, whp G_{T_G} contains a K_r-factor.*

What is more, a simplification of the proof of Theorem 1.6 yields a corresponding result for $K_r^{(s)}$-factors. For this, let $r > s \geqslant 3$ and $K_r^{(s)}$ denote the complete s-uniform hypergraph on r vertices. Let $(G_t)_{t=1}^{N_s} = (H^s_t)_{t=1}^{N_s}$ and denote the hitting time of a $K_r^{(s)}$-cover by T_G. Then:

Theorem 1.7. *Let $r > s \geqslant 3$. We may couple the stopped random r-uniform hypergraph process H_{T_H} and the stopped random s-uniform hypergraph process G_{T_G} so that, whp, for every hyperedge in H_{T_H} there is copy of $K_r^{(s)}$ in G_{T_G} on the same vertex set. In particular, whp G_{T_G} has a $K_r^{(s)}$-factor.*

2 Preliminaries

In the remainder, we fix $r \geqslant 3$ and suppress the dependence on r writing H_t instead of H^r_t, etc. Let $M = N_r = \binom{n}{r}$ and $N = N_2 = \binom{n}{2}$. By an *r-uniform hypergraph H* on the vertex set $V = [n]$, we mean a subset of $\binom{V}{r}$, the set of all r-subsets of V. That is, we will use H as a set (of sets of vertices of size r) for convenient notation. For a hypergraph H on the vertex set $V = [n]$ and $v \in [n]$, we use $d(v)$ to denote the degree of v in H. In a graph G, an r-*clique* is a clique on r vertices. We denote by $\mathrm{cl}(G)$ the set of vertex sets from $\binom{V}{r}$ which span r-cliques in G (so $\mathrm{cl}(G)$ is an r-uniform hypergraph in the aforementioned sense). Throughout the paper, we fix an arbitrary function $g(n)$ satisfying

$$g(n) = o\big(\log n/\log\log n\big) \quad \text{and} \quad g(n) = \omega(1). \tag{1}$$

2.1 The standard coupling and the critical window

It will be useful to work with the following standard device which gives a convenient coupling of the random hypergraphs $H(n, \pi)$ for all $\pi \in [0, 1]$ and the random hypergraph process.

Definition 2.1 (Standard coupling)**.** For every $h \in \binom{V}{r}$, let U_h be an independent random variable, uniform from $[0, 1]$. Let

$$H_\pi = (V, \{h : U_h \leqslant \pi\}).$$

Then $H_\pi \sim H(n, \pi)$. Almost surely all values U_h, $h \in \binom{V}{r}$, are distinct, yielding an instance of the random hypergraph process $(H_t)_{t=0}^M$, as we can add the hyperedges in ascending order of U_h.

We will operate within the following *critical window*: Define π_- and π_+ by setting

$$\pi_\pm = \frac{\log n \pm g(n)}{\binom{n-1}{r-1}}, \tag{2}$$

where $g(n)$ is the function which was fixed globally in (1), and, using δ, ε from Theorem 1.4, let

$$p_\pm = \big(\pi_\pm/(1 - n^{-\delta})\big)^{1/\binom{r}{2}}. \tag{3}$$

For n large enough we have $p_+ \leqslant n^{-2/r+\epsilon}$, so Theorem 1.4 applies with $p = p_+$ and $\pi = \pi_+$.

It is well-known that (π_-, π_+) is the 'critical window' for the disappearance of the last isolated vertex in a random r-uniform hypergraph (see [4, Lemma 5.1(a)]), and so by Theorem 1.3 for the appearance of a perfect matching. So if we couple as in Definition 2.1, then whp we have

$$G_{p_-} \subset G_{T_G} \subset G_{p_+} \quad \text{and} \quad H_{\pi_-} \subset H_{T_H} \subset H_{\pi_+}. \tag{4}$$

2.2 Proof overview

Define p_+, π_+ as in equations (2) and (3). Our starting point is the coupling of $G \sim G(n, p_+)$ and $H \sim H(n, \pi_+)$ given by Theorem 1.4. We review this coupling in §3. In §4, the heart of the proof, we take the coupled $G \sim G(n, p_+)$ and $H \sim H(n, \pi_+)$ and proceed by carefully coupling uniform orders of the edges of G and hyperedges of H. Since p_+ and π_+ are the upper ends of the respective critical windows (see §2.1), whp this couples (copies of) the stopped graph process G_{T_G} and the stopped hypergraph process H_{T_H}. This coupling *almost* does what we want: for all hyperedges $h \in H_{T_H}$, except those in a small exceptional set $\mathcal{E}$, there is an r-clique in G_{T_G} on the same vertex set. Moreover, we show that whp all $h \in \mathcal{E}$ have a partner hyperedge which appears between time T_H and time $T_H + \lfloor g(n)n \rfloor$. To prove Theorem 1.6, we are left to show that we can get rid of the hyperedges in $\mathcal{E}$ and still have an instance of the stopped random hypergraph process. To this end, $\mathcal{E}$ can be whp embedded into a binomial random subset $\mathcal{R} \subset H_{T_H}$ where each hyperedge $h \in H_{T_H}$ is included independently with a small probability. We proceed to show that if we remove the hyperedges in $\mathcal{R}$ from the hypergraph process up to time T_H, whp this essentially does not change the hitting time T_H, and in particular whp $H_{T_H} \setminus \mathcal{R}$ is still an instance of the stopped random hypergraph process. Chaining the couplings together then proves Theorem 1.6. The necessary modifications in the proof of Theorem 1.7 are detailed in [9].

3 Coupling of $G(n, p_+)$ and $H(n, \pi_+)$

In §3.1 we briefly review Riordan's coupling from Theorem 1.4 for $r \geqslant 4$.[5] We let $\pi = \pi_+$ from equation (2) and $p = p_+$ from equation (3).

3.1 The coupling algorithm for $r \geqslant 4$

Order the $M = \binom{n}{r}$ potential hyperedges in some arbitrary way as $h_1, \dots, h_M$, and for $1 \leqslant j \leqslant M$, let A_j be the event that there is an r-clique in $G \sim G(n,p)$ on the vertex set of h_j. We construct the coupling of $G \sim G(n,p)$ and $H \sim H(n,\pi)$ step by step; in step j revealing whether or not $h_j \in H$, as well as some information about A_j.

Coupling algorithm: For each j from 1 to M:

- Calculate π_j, the conditional probability of A_j given all the information revealed so far.
- If $\pi_j \geqslant \pi$, toss a coin which lands heads with probability π/π_j, independently of everything else. If the coin lands heads, then test whether A_j holds (which it does with probability exactly π_j). Include the hyperedge h_j in H if and only if the coin lands heads and A_j holds. (Note that the probability of including h_j is exactly $\pi/\pi_j \cdot \pi_j = \pi$.)
- If $\pi_j < \pi$, then toss a coin which lands heads with probability π (independently of everything else), and declare h_j present in H if and only if the coin lands heads. If this happens for any j, we say that the coupling has failed.

After steps $j = 1, \dots, M$, we have decided all hyperedges of H, and revealed information on the events $A_1, \dots, A_M$ of G. Now choose G conditional on the revealed information on the events A_j.

[5] For the modifications in the case $r = 3$ we refer the reader to [8].

4 Process coupling

Building upon Theorem 1.4, we couple the random graph process with the random hypergraph process. Roughly speaking, we may couple the random graph process and the random hypergraph process so that there is *almost* a copy of H_{T_H} within the r-cliques of G_{T_G}: for all hyperedges in H_{T_H} except those in a set $\mathcal{E}$ (the *exceptional* hyperedges), there is an r-clique in G_{T_G} on the same vertex set. Moreover, the hyperedges in $\mathcal{E}$ all gain a partner hyperedge shortly after time T_H.

Proposition 4.1. *We may couple the random graph process $(G_t)_{t=0}^N$ and the random hypergraph process $(H_t)_{t=0}^M$ so that whp the following holds. There is a set of hyperedges $\mathcal{E} \subset H_{T_H}$ so that*

a) $H_{T_H} \setminus \mathcal{E} \subset \mathrm{cl}(G_{T_G})$*, and*

b) for every $h_1 \in \mathcal{E}$ *there is a* $h_2 \in H_{T_H+\lfloor g(n)n \rfloor} \setminus H_{T_H}$ *so that* $|h_1 \cap h_2| = 2$.

Now whp, we can embed the set $\mathcal{E}$ of 'exceptional' hyperedges from Proposition 4.1 into a random set $\mathcal{R}$ which includes every $h \in H_{T_H}$ independently with a small probability.

Proposition 4.2. *We may couple the random r-uniform hypergraph process $(H_t)_{t=0}^M$ and a set $\mathcal{R} \subset \binom{[n]}{r}$ of hyperedges so that both of the following properties hold.*

a) We have $\mathcal{R} \subseteq H_{T_H}$, and (given only H_{T_H}) each hyperedge $h \in H_{T_H}$ is included in $\mathcal{R}$ independently with probability $\pi_{\mathcal{R}} = \frac{10r^4 g(n)}{n}$.

b) Let $\mathcal{F} \subset H_{T_H}$ be the set of hyperedges in H_{T_H} with a partner hyperedge in $H_{T_H+\lfloor g(n)n \rfloor} \setminus H_{T_H}$. Then, whp, $\mathcal{F} \subset \mathcal{R}$.

As the final puzzle piece, we find that after removing every hyperedge from H_{T_H} independently with a small probability, whp we still have an instance of the stopped random hypergraph process.

Proposition 4.3. *Let H_{T_H} be the stopped random hypergraph process, and let $\mathcal{R} \subset H_{T_H}$ be a subset of hyperedges where we include every $h \in H_{T_H}$ independently with probability $\pi_{\mathcal{R}} = \frac{10r^4 g(n)}{n}$. We may couple H_{T_H} and $\mathcal{R}$ with another instance $H'_{T'_H}$ of the stopped random hypergraph process so that, whp, $H_{T_H} \setminus \mathcal{R} = H'_{T'_H}$.*

Combining Propositions 4.1, 4.2 and 4.3 yields a chain of couplings that whp embeds the stopped hypergraph process into the cliques of the stopped graph process, completing the proof:

$$H'_{T'_H} \quad \overset{\text{whp}}{=} \quad H_{T_H} \setminus \mathcal{R} \quad \overset{\text{whp}}{\subseteq} \quad H_{T_H} \setminus \mathcal{F} \quad \overset{\text{whp}}{\subseteq} \quad \mathrm{cl}(G_{T_H}).$$

Acknowledgments This project was initiated during the research workshop of Angelika Steger's group in Buchboden, August 2021. We are grateful to Oliver Riordan for a helpful discussion.

References

[1] Noga Alon and Raphael Yuster. Threshold functions for H-factors. *Combin. Probab. Comput.*, 2(2):137–144, 1993.

[2] Béla Bollobás and Andrew Thomason. Random graphs of small order. In *Random graphs '83 (Poznań, 1983)*, volume 118 of *North-Holland Math. Stud.*, pages 47–97. North-Holland, Amsterdam, 1985.

[3] Fan Chung and Ron Graham. *Erdős on graphs: His legacy of unsolved problems.* AK Peters/CRC Press, 1998.

[4] Pat Devlin and Jeff Kahn. Perfect fractional matchings in k-out hypergraphs. *Electron. J. Combin.*, 24(3):Paper No. 3.60, 12, 2017.

[5] Pál Erdős and Alfréd Rényi. On the existence of a factor of degree one of a connected random graph. *Acta Math. Acad. Sci. Hungar.*, 17:359–368, 1966.

[6] Pál Erdős. On the combinatorial problems which I would most like to see solved. *Combinatorica*, 1(1):25–42, 1981.

[7] Alan Frieze and Svante Janson. Perfect matchings in random s-uniform hypergraphs. *Random Structures & Algorithms*, 7(1):41–57, 1995.

[8] Annika Heckel. Random triangles in random graphs. *Random Structures & Algorithms*, 59(4):616–621, 2021.

[9] Annika Heckel, Marc Kaufmann, Noela Müller, and Matija Pasch. The hitting time of clique factors. *arXiv:2302.08340*, 2023.

[10] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[11] Anders Johansson, Jeff Kahn, and Van Vu. Factors in random graphs. *Random Structures Algorithms*, 33(1):1–28, 2008.

[12] Jeff Kahn. Asymptotics for Shamir's problem. *arXiv:1909.06834*, 2019.

[13] Jeff Kahn. Hitting times for Shamir's problem. *Trans. Amer. Math. Soc.*, 375(1):627–668, 2022.

[14] Jeong Han Kim. Perfect matchings in random uniform hypergraphs. *Random Struct. Algorithms*, 23(2):111–132, 2003.

[15] Michael Krivelevich. Triangle factors in random graphs. *Combinatorics, Probability and Computing*, 6:337 – 347, 1997.

[16] Julius Petersen. Die theorie der regulären graphs. *Acta Math.*, 15:193–220, 1891.

[17] Oliver Riordan. Random cliques in random graphs and sharp thresholds for F-factors. *Random Structures & Algorithms*, 61(4):619–637, 2022.

[18] Andrzej Ruciński. Matching and covering the vertices of a random graph by copies of a given graph. *Discrete Math.*, 105(1-3):185–197, 1992.

[19] Jeanette Schmidt and Eli Shamir. A threshold for perfect matchings in random d-pure hypergraphs. *Discrete Mathematics*, 45(2):287–295, 1983.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Roudneff's Conjecture in Dimension 4

(Extended abstract)

Rangel Hernández-Ortiz* Kolja Knauer† Luis Pedro Montejano‡
Manfred Scheucher§

Abstract

J.-P. Roudneff conjectured in 1991 that every arrangement of $n \geq 2d+1 \geq 5$ pseudohyperplanes in the real projective space $\mathbb{P}^d$ has at most $\sum_{i=0}^{d-2}\binom{n-1}{i}$ *complete cells* (i.e., cells bounded by each hyperplane). The conjecture is true for $d = 2, 3$ and for arrangements arising from Lawrence oriented matroids. The main result of this manuscript is to show the validity of Roudneff's conjecture for $d = 4$. Moreover, based on computational data we conjecture that the maximum number of complete cells is only obtained by cyclic arrangements.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-078`

1 Introduction

A *projective* arrangement of n pseudohyperplanes $H(d, n)$ in the real projective space $\mathbb{P}^d$ is a finite collection of mildly deformed linear hyperplanes with several combinatorial properties, see Section 2.1 for the definition in terms of oriented matroids. In particular, no point belongs to every pseudohyperplane of $H(d, n)$. Any arrangement $H(d, n)$ decomposes $\mathbb{P}^d$

*Universitat Rovira i Virgili, Departament d'Enginyeria Informàtica i Matemàtiques, Tarragona, Spain. E-mail: `rangel.hernandez@urv.cat`.

†Departament de Matemàtiques i Informàtica, Universitat de Barcelona, Spain. E-mail: `kolja.knauer@ub.edu`. Supported by the Spanish *Ministerio de Economía, Industria y Competitividad* through grants RYC-2017-22701 and ALCOIN: PID2019-104844GB-I00.

‡Serra Húnter Fellow, Universitat Rovira i Virgili, Departament d'Enginyeria Informàtica i Matemàtiques, Tarragona, Spain. E-mail: `luispedro.montejano@urv.cat`. Supported by SGR Grant 2021 00115.

§Institut für Mathematik, Technische Universität Berlin, Germany. E-mail: `lastname@math.tu-berlin.de`. Supported by DFG Grant SCHE 2214/1-1.

into a d-dimensional cell complex and any d-cell c of $H(d,n)$ has at most n *facets* (that is, $(d-1)$-cells). We say that a d-cell c is a *complete cell* of $H(d,n)$ if c has exactly n facets, i.e., c is bounded by each pseudohyperplane of $H(d,n)$.

The *cyclic polytope* of dimension d with n vertices, discovered by Carathéodory [3], is the convex hull in $\mathbb{R}^d$ of $n \geq d+1 \geq 3$ different points $x(t_1), \dots, x(t_n)$ on the moment curve $x : \mathbb{R} \to \mathbb{R}^d,\ t \mapsto (t, t^2, \dots, t^d)$. Cyclic polytopes play an important role in combinatorial convex geometry due to their connection with certain extremal problems. See for example, the upper bound theorem due to McMullen [10]. *Cyclic arrangements* are defined as the dual of the cyclic polytopes. As for cyclic polytopes, cyclic arrangements also have extremal properties, see Section 2.1 for the definition in terms of oriented matroids. For instance, Shannon [14] introduced cyclic arrangements as examples of projective arrangements in dimension d which minimize the number of cells with $(d+1)$ facets.

Denote by $C_d(n)$ the number of complete cells of the cyclic arrangement of dimension d with n hyperplanes. Roudneff [13] proved that $C_d(n) \geq \sum_{i=0}^{d-2} \binom{n-1}{i}$ holds for $d \geq 2$ and that this bound is tight for all $n \geq 2d+1$. Moreover, he conjectured that in that case, cyclic arrangements maximize the number of complete cells.

Conjecture 1.1 ([13, Conjecture 2.2])**.** Every arrangement of $n \geq 2d+1 \geq 5$ pseudohyperplanes in $\mathbb{P}^d$ has at most $\sum_{i=0}^{d-2} \binom{n-1}{i}$ complete cells.

The conjecture is true for $d = 2$ (that is, any arrangement of n pseudolines in $\mathbb{P}^2$ contains at most one complete cell), Ramírez Alfonsín [12] proved the case $d = 3$, and in [11] the authors proved it for arrangements corresponding to Lawrence oriented matroids.

In [8] the exact number of complete cells of cyclic arrangements was calculated for any positive integers d and n with $n \geq d+1$, namely,

$$C_d(n) = \binom{d}{n-d} + \binom{d-1}{n-d-1} + \sum_{i=0}^{d-2} \binom{n-1}{i}.$$

Thus, in view of Roudneff's conjecture, the following question was asked in [11].

Question 1.2. *Is it true that every arrangement of $n \geq d+1 \geq 3$ pseudohyperplanes in $\mathbb{P}^d$ has at most $C_d(n)$ complete cells?*

Notice that there is a unique arrangement of 3 (resp. 4) lines in $\mathbb{P}^2$ with $C_2(3) = 4$ (resp. $C_2(4) = 3$) complete cells. Since Conjecture 1.1 is true for $d = 2$ and $n \geq 5$, Question 1.2 is answered affirmatively for $d = 2$.

As the main result of this paper, we give an affirmative answer to Question 1.2 for $d = 4$ and therefore prove Roudneff's conjecture for dimension 4, further supporting the general conjecture. In addition, with a few simple observations, we answer Question 1.2 for $d = 3$ and further strengthen Roudneff's conjecture.

2 Oriented matroids

Let us give some basic notions and definitions in oriented matroid theory. We assume some knowledge and standard notation of the theory of oriented matroids, for further reference the reader can consult the textbook [2]. A *signed set* or *signed vector* X on ground set E is a set $\underline{X} \subseteq E$ together with a partition (X^+, X^-) of $\underline{X}$ into two distinguished subsets: X^+, the set of *positive elements* of X, and X^-, its set of *negative elements.* The set $\underline{X} = X^+ \cup X^-$ is the *support* of X. We denote by $-X$ the sign-vector such that $-X^+ = X^-$ and $-X^- = X^+$. An oriented matroid $\mathcal{M} = (E, \mathcal{C})$ is a pair of a finite ground set E and a collection of signed sets on E called *circuits*, satisfying the following axioms:

- $\emptyset \notin \mathcal{C}$,
- if $X \in \mathcal{C}$ then $-X \in \mathcal{C}$,
- if $X, Y \in \mathcal{C}$ and $\underline{X} \subseteq \underline{Y}$ then $X = \pm Y$,
- if $X, Y \in \mathcal{C}$, $X \neq -Y$, and $e \in X^+ \cap Y^-$, then there is $Z \in \mathcal{C}$, with $e \notin \underline{Z}$ and $Z^+ \subseteq X^+ \cap Y^+$ and $Z^- \subseteq X^- \cap Y^-$.

We say that $X \in \mathcal{C}$ is a *positive circuit* if $X^- = \emptyset$. We call the set of all reorientations of $\mathcal{M}$ its *reorientation class.* We say that $\mathcal{M}$ is *acyclic* if it does not contain positive circuits (otherwise, $\mathcal{M}$ is called *cyclic*). A *reorientation* of $\mathcal{M}$ on $R \subseteq E$ is performed by changing the signs of the elements in R in all the circuits of $\mathcal{M}$. It is easy to check that the new set of signed circuits is also the set of circuits of an oriented matroid, usually denoted by $\mathcal{M}_R$. A reorientation is *acyclic* if $\mathcal{M}_R$ is acyclic. Recall that oriented matroid on n elements is *uniform of rank* r if the set of supports of its circuits consists of all $(r+1)$-element subsets of E. Given a uniform oriented matroid $\mathcal{M}$ of rank r on $n = |E|$ elements, we denote its dual by $\mathcal{M}^*$, which is another uniform oriented matroid of rank $n - r$ on n elements.

A characterization of oriented matroids in terms of basis orientations (that we will not make explicit here) was given by Lawrence [9]. Let $r \geq 1$ be an integer and $E = \{1, \ldots, n\}$ be a set. A mapping $\chi : E^r \to \{-1, 0, 1\}$ (where we will abbreviate it by $\{-, 0, +\}$) is a basis orientation of an oriented matroid of rank r on E if and only if χ is a *chirotope*, that is, a special alternating mapping not identically zero. It is known that $\chi : E^r \to \{-, +\}$ is a chirotope if and only if χ is a basis orientation of a rank r uniform oriented matroid on E. Moreover, if $\chi(B) = +$ for any ordered basis $B = (b_1, \ldots, b_r)$ of $\mathcal{M}$ with $b_1 < \ldots < b_r$, then the uniform matroid $\mathcal{M}$ is known to be the *alternating oriented matroid* of rank r on n elements. In that case, the signs of each circuit alternate along the ordering of E.

Given two sign-vectors $X, Y \in \{+, -, 0\}^E$, their *separation* is the set $S(X, Y) = \{e \in E \mid X_e \cdot Y_e = -\}$, where X_e and Y_e are the signs of the element e in $\underline{X}$ and $\underline{Y}$, respectively. We denote by $X \perp Y$ and say that X and Y are *orthogonal* if the sets $S(X, Y)$ and $S(X, -Y)$ are either both empty or both non-empty. Maximal covectors of an oriented matroid $\mathcal{M}$ are usually called *topes.* It is known that a sign-vector $T \in \{+, -\}^E$ is a tope of $\mathcal{M}$ if and only if $T \perp X$ for all circuit $X \in \mathcal{C}$ (see [2, Section 1.2, page 14]). Moreover, T is a tope of $\mathcal{M}$ if and only if $S(T, X)$ and $S(T, -X)$ are both non-empty, for every circuit $X \in \mathcal{C}$.

2.1 Topological Representation Theorem

The combinatorial properties of arrangements of pseudohyperplanes can be studied in the language of oriented matroids. The Folkman–Lawrence topological representation theorem [7] states that the reorientation classes of oriented matroids on n elements and rank r (without loops or parallel elements) are in one-to-one correspondence with the classes of isomorphism of arrangements of n pseudospheres in S^{r-1} (see [2, Theorem 1.4.1]). There is a natural identification between pseudospheres and pseudohyperplanes as follows. Recall that $\mathbb{P}^{r-1}$ is the topological space obtained from S^d by identifying all pairs of antipodal points. The double covering map $\pi : S^{r-1} \to \mathbb{P}^{r-1}$, given by $\pi(x) = \{x, -x\}$, gives an identification of centrally symmetric subsets of S^{r-1} and general subsets of $\mathbb{P}^{r-1}$. This way centrally symmetric pseudospheres in S^{r-1} correspond to pseudohyperplanes in $\mathbb{P}^{r-1}$. Hence, the topological representation theorem can also be stated in terms of pseudohyperplanes in $\mathbb{P}^{r-1}$, i.e., the reorientation classes of oriented matroids on n elements and rank r (without loops or parallel elements) are in one-to-one correspondence with the classes of isomorphism of arrangements of n pseudohyperplanes in $\mathbb{P}^{r-1}$ (see [2, Exercise 5.8]).

An arrangement $H(d, n)$ is called *simple* if every intersection of d pseudohyperplanes is a unique distinct point. Simple arrangements correspond to uniform oriented matroids. The d-cells of any arrangement $H(d, n)$ are usually called topes since they are in one-to-one correspondence with the topes of each of the oriented matroids $\mathcal{M}$ of rank $r = d + 1$ on n elements of its corresponding reorientation class. It is known that a tope of $\mathcal{M}$ (i.e, a d-cell of its corresponding arrangement) corresponds to an acyclic reorientation of $\mathcal{M}$ having as interior elements precisely those pseudohyperplanes not bordering the tope. Moreover, a tope T of $\mathcal{M}$ is a complete cell if reorienting any single element of T, the resulting sign-vector is also a tope of $\mathcal{M}$. Cyclic arrangements of n hyperplanes in $\mathbb{P}^d$ are equivalent to alternating oriented matroids of rank $r = d + 1$ on n elements, which hence have exactly $2C_{r-1}(n)$ complete cells. Summarizing, Question 1.2 (and hence Roudneff's conjecture) can be stated in the following form:

> *Every rank r oriented matroid $\mathcal{M}$ on $n \geq r + 1$ elements has at most $2C_{r-1}(n)$ complete cells.*

We summarize for later usage: Given a rank r oriented matroid $\mathcal{M} = (E, \mathcal{C})$, the following three conditions hold.

(a) A tope of $\mathcal{M}$ is a sign-vector $T \in \{+, -\}^E$ such that $T \perp X$ for all circuit $X \in \mathcal{C}$.

(b) A tope T of $\mathcal{M}$ is a complete cell if reorienting any single element of T, the resulting sign-vector is also a tope of $\mathcal{M}$.

(c) If the corresponding arrangement of n pseudohyperplanes in $\mathbb{P}^{r-1}$ of $\mathcal{M}$ is simple, then $\mathcal{M}$ is uniform.

3 Previous results

We will use the following result due to Roudneff.

Proposition 3.1 ([13]). *To prove Conjecture 1.1 for dimension d, it suffices to verify it for all simple arrangements of $n = 2d + 1$ pseudohyperplanes in $\mathbb{P}^d$.*

From the proof of the above proposition, it can be seen that even for any arrangement H with $n \leq 2d$ pseudohyperplanes in $\mathbb{P}^d$, we may also perturb each hyperplane of H a bit in order to obtain a simple arrangement H' with at least the same number of complete cells as H (see Proposition 2.3 of [13]). This shows that also for Question 1.2, we can restrict ourselves to simple arrangements.

Remark 3.2. *To answer Question 1.2 for dimension d in the affirmative, it suffices to verify it for simple arrangements of pseudohyperplanes in $\mathbb{P}^d$.*

Thus, by condition (c) and by Remark 3.2, it is sufficient to prove Question 1.2 for uniform oriented matroids. The following observation will be useful in this work.

Remark 3.3. *There is only one reorientation class of uniform rank r oriented matroids on $n \leq r + 2$ elements.*

Proof. The number of reorientation classes of a uniform oriented matroid $\mathcal{M}$ of rank r on n elements is equal to the number of reorientation classes of its dual $\mathcal{M}^*$. Now, if $\mathcal{M}$ has rank r and $n \leq r+2$ elements, then $\mathcal{M}^*$ has rank at most 2. Hence, $\mathcal{M}^*$ and therefore $\mathcal{M}$ has only one reorientation class. □

Thus, every acyclic uniform oriented matroid on at most $r + 2$ elements is in the reorientation class of the alternating oriented matroid and hence, they all have the same number of complete cells. As a consequence of Remarks 3.2 and 3.3, we can answer affirmatively Question 1.2 for $n \leq r+2$. In particular, as for $r = 4$ (dimension $d = 3$) Conjecture 1.1 is true for $n \geq 7$, we obtain the following.

Corollary 3.4. *Every arrangement of $n \geq 4$ pseudohyperplanes in $\mathbb{P}^3$ has at most $C_3(n)$ complete cells.*

4 Main result

Given a uniform rank r oriented matroid $\mathcal{M} = (E, \mathcal{C})$ on $n = |E|$ elements, we explain the procedure to obtain the set of all complete cells of its corresponding arrangement of n pseudohyperplanes in $\mathbb{P}^d$ via the signed bases of $\mathcal{M}$. We start with the signature of all the bases of $\mathcal{M}$ and then, we obtain all its signed circuits. After that, we get the set of topes of $\mathcal{M}$ and finally, we obtain the set of all complete cells of $\mathcal{M}$ as follows:

Bases $\rightarrow$ Circuits: From the chirotope, we may obtain that $\chi(B) = -X_{b_i} \cdot X_{b_{i+1}} \cdot \chi(B')$, where $\underline{X} = \{b_1, ..., b_{r+1}\}$ is the support of an ordered circuit of $\mathcal{M}$ and $B = \underline{X} - b_i$ and $B' = \underline{X} - b_{i+1}$ are two bases of $\mathcal{M}$ (see [2, Section 3.5]). Hence, given $\chi(B)$, for any basis B of $\mathcal{M}$, we obtain the signed circuit X and since $\mathcal{M}$ is uniform, we can proceed to obtain all the signed circuits of $\mathcal{M}$.

Circuits → Topes: For any sign-vector $T \in \{+,-\}^n$, we verify condition (a) to confirm that T is a tope of $\mathcal{M}$, i.e., we check for all circuit $X \in \mathcal{C}$ of $\mathcal{M}$ if $S(T,X)$ and $S(T,-X)$ are both non-empty (see [2, Section 1.2, page 14]).

Topes → Complete cells: For any tope T, we verify condition (b) to confirm that T is a complete cell of $\mathcal{M}$. That is, we reorient any single element of T, check if the resulting sign-vector is a tope of $\mathcal{M}$ and verify this for each of the n entries of T.

Finschi and Fukuda [5, 6] enumerated the signed bases of all the reorientation classes of uniform rank 5 oriented matroids on 8 and 9 elements. While the data for 8 elements is available on the website [4], the data for 9 elements and also their source code for the enumeration is available upon request from Lukas Finschi. We follow the procedure explained above with a computer program (available at [1]) which gives us the number of complete cells of each acyclic reorientation class. After about 26 CPU days of computing time (i.e., few days with parallelization), we obtain the following.

Theorem 4.1. *Each of the* 135 *reorientation classes of uniform rank* 5 *oriented matroids on* 8 *elements has at most* $2C_4(8)$ *complete cells. Moreover, the class of the alternating oriented matroid is the only one with exactly* $2C_4(8)$ *complete cells.*

Theorem 4.2. *Each of the* 9276595 *reorientation classes of uniform rank* 5 *oriented matroids on* 9 *elements has at most* $2C_4(9)$ *complete cells. Moreover, the class of the alternating oriented matroid is the only one with exactly* $2C_4(9)$ *complete cells.*

We can now prove our main result:

Theorem 4.3. *Every arrangement of* $n \geq 5$ *pseudohyperplanes in* $\mathbb{P}^4$ *has at most* $C_4(n)$ *complete cells.*

Proof. By Proposition 3.2, it is sufficient to prove the theorem for simple arrangements, that is, for uniform oriented matroids (see condition (c)). Thus, by Remark 3.3 and Theorem 4.1, the result holds for $n = 5, 6, 7$ and 8. Finally, by Proposition 3.1 it suffices to verify it for $n = 9$. Therefore, the result holds by Theorem 4.2. □

Finally, we have used our computer program to verify that the cyclic arrangement is the unique example which maximizes the number of complete cells for $d = 2$ and $n \leq 10$, for $d = 3$ and $n \leq 7$, and for $d = 4$ and $n \leq 9$. Based on our computational evidence, we conclude this article with the following strengthening of Roudneff's conjecture and Question 1.2:

Conjecture 4.4. Every arrangement of $n \geq d+1 \geq 3$ pseudohyperplanes in $\mathbb{P}^d$ has at most $C_d(n)$ complete cells. Moreover, among all arrangements of n pseudohyperplanes in $\mathbb{P}^d$ the cyclic arrangement is (up to isomorphism) the only one with $C_d(n)$ complete cells.

Last but not least, as the proof of Proposition 3.1, it suffices to verify Conjecture 4.4 for simple arrangements of pseudohyperplanes in $\mathbb{P}^d$. However, we do not know whether the setting can also be restricted to $n \leq 2d + 1$ without loss of generality.

References

[1] Supplemental source code and data. `https://github.com/manfredscheucher/supplemental-roudneff4`.

[2] A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G. M. Ziegler. *Oriented Matroids, 2nd edition*, volume 46 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 2 edition, 1999.

[3] C. Carathéodory. Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen. *Mathematische Annalen*, 64:95–115, 1907.

[4] L. Finschi. Webpage: Homepage of oriented matroids. `https://finschi.com/math/om/?p=catom&filter=nondeg`.

[5] L. Finschi. *A graph theoretical approach for reconstruction and generation of oriented matroids*. PhD thesis, ETH Zürich, Switzerland, 2001.

[6] L. Finschi and K. Fukuda. Generation of oriented matroids–a graph theoretical approach. *Discrete & Computational Geometry*, 27(1):117–136, 2002.

[7] J. Folkman and J. Lawrence. Oriented matroids. *Journal of Combinatorial Theory, Series B*, 25(2):199–236, 1978.

[8] D. Forge and J. L. Ramírez Alfonsín. On counting the k-face cells of cyclic arrangements. *European Journal of Combinatorics*, 22(3):307–312, 2001.

[9] J. Lawrence. Oriented matroids and multiply ordered sets. *Linear Algebra and its Applications*, 48:1–12, 1982.

[10] P. McMullen. The maximum numbers of faces of a convex polytope. *Mathematika*, 17(2):179–184, 1970.

[11] L. P. Montejano and J. L. Ramírez Alfonsín. Roudneff's Conjecture for Lawrence Oriented Matroids. *Electronic Journal of Combinatorics*, 22(2):Paper #P2.3, 4 pages, 2015.

[12] J. L. Ramírez Alfonsín. Cyclic arrangements and Roudneff's conjecture in the space. *Information Processing Letters*, 71(5):179–182, 1999.

[13] J.-P. Roudneff. Cells with many facets in arrangements of hyperplanes. *Discrete Mathematics*, 98(3):185–191, 1991.

[14] R. W. Shannon. Simplicial cells in arrangements of hyperplanes. *Geometriae Dedicata*, 8(2):179–187, 1980.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Locality in Sumsets

(Extended abstract)

Peter van Hintum* Peter Keevash †

Abstract

Motivated by the Polynomial Freiman-Ruzsa (PFR) Conjecture, we develop a theory of locality in sumsets, with several applications to John-type approximation and stability of sets with small doubling. One highlight shows that if $A \subset \mathbb{Z}$ with $|A+A| \leq (1-\epsilon)2^d|A|$ is non-degenerate then A is covered by $O(2^d)$ translates of a d-dimensional generalised arithmetic progression (d-GAP) P with $|P| \leq O_{d,\epsilon}(|A|)$; thus we obtain one of the polynomial bounds required by PFR, under the non-degeneracy assumption that A is not efficiently covered by $O_{d,\epsilon}(1)$ translates of a $(d-1)$-GAP.

We also prove a stability result showing for any $\epsilon, \alpha > 0$ that if $A \subset \mathbb{Z}$ with $|A+A| \leq (2-\epsilon)2^d|A|$ is non-degenerate then some $A' \subset A$ with $|A'| > (1-\alpha)|A|$ is efficiently covered by either a $(d+1)$-GAP or $O_\alpha(1)$ translates of a d-GAP. This 'dimension-free' bound for approximate covering makes for a surprising contrast with exact covering, where the required number of translates not only grows with d, but does so exponentially. Another highlight shows that if $A \subset \mathbb{Z}$ is non-degenerate with $|A+A| \leq (2^d+\ell)|A|$ and $\ell \leq 0.1 \cdot 2^d$ then A is covered by $\ell+1$ translates of a d-GAP P with $|P| \leq O_d(|A|)$; this is tight, in that $\ell+1$ cannot be replaced by any smaller number.

The above results also hold for $A \subset \mathbb{R}^d$, replacing GAPs by a suitable common generalisation of GAPs and convex bodies, which we call generalised convex progressions. In this setting the non-degeneracy condition holds automatically, so we obtain essentially optimal bounds with no additional assumption on A. Here we show that if $A \subset \mathbb{R}^k$ satisfies $|\frac{A+A}{2}| \leq (1+\delta)|A|$ with $\delta \in (0,1)$, then $\exists A' \subset A$ with $|A'| \geq (1-\delta)|A|$ so that $|\operatorname{co}(A')| \leq O_{k,1-\delta}(|A|)$. This is a dimensionally independent sharp stability result for the Brunn-Minkowski inequality for equal sets, which hints towards a possible analogue for the Prékopa-Leindler inequality.

*New College, University of Oxford, UK. E-mail: peter.vanhintum@new.ox.ac.uk
†Mathematical Institute, University of Oxford, UK. Supported by ERC Advanced Grant 883810.

These results are all deduced from a unifying theory, in which we introduce a new intrinsic structural approximation of any set, which we call the 'additive hull', and develop its theory via a refinement of Freiman's theorem with additional separation properties. A further application that will be published separately is a proof of Ruzsa's Discrete Brunn-Minkowski Conjecture [vHKT23].

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-079`

1 Introduction

A foundational result in Additive Combinatorics is *Freiman's Theorem* [Fre59] that any subset A of integers with bounded doubling is a dense subset of a generalised arithmetic progression (GAP) P of bounded dimension (see the book of Tao and Vu [TV06] for definitions and background). This gives a satisfactory qualitative description of A: it can be approximated by some P belonging to a simple class of sets with bounded doubling. However, the doubling of P may be much larger than that of A, so the quest for a more quantitative version of Freiman's Theorem has been a major driving force in the development of Additive Combinatorics. A little thought reveals that one must allow P to come from a broader class of sets than just GAPs. One natural attempt is to approximate A by sets $X + P$ with $|X|$ bounded, i.e. the union of a bounded number of translates of P. One might hope to find a such an approximation with polynomial bounds, i.e. if $|A + A| \leq e^{O(d)}|A|$ then we could find such X, P with $\dim(P) = O(d)$, $|X| \leq e^{O(d)}$ and $|X + P| \leq e^{O(d)}|A|$. However, an example of Lovett and Regev [LR17] shows that this is not always possible. Improving the bounds in Freiman's theorem has been the subject of a rich body of research [Ruz94, Bil99, Cha02, GT06, San08, Sch11]. The best known bounds, due to Sanders [San12], have $\dim(P)$ and $\log(|X + P|/|A|)$ about $O(d^6)$.

The *Polynomial Freiman-Ruzsa (PFR) Conjecture* (see [Gre07]) attempts to approximate by translates of a *convex progression*, i.e. a set P of the form $\phi(C \cap \mathbb{Z}^k)$ for some convex set $C \subset \mathbb{R}^k$ and linear map $\phi : \mathbb{Z}^k \to \mathbb{Z}$, for which polynomial bounds may be true. The conjecture states that if $|A + A| \leq e^{O(d)}|A|$ then one can find such P with $k = O(d)$ and $|P| \leq e^{O(d)}|A|$ such that $A \subset X + P$ for some X with $|X| \leq e^{O(d)}$. Below we will describe three perspectives on the PFR Conjecture that provide a thematic overview of our results; these are (1) John-type approximation, (2) Stability, (3) Locality. The third theme of locality is our primary focus, i.e. most of the technical work goes into developing the theory of locality, which is then used to deduce the results discussed within the first two themes. Our results hold both for the discrete setting $A \subset \mathbb{Z}$ considered in PFR and the continuous setting $A \subset \mathbb{R}^k$. For now we will continue to focus on the discrete setting (in some sense the hardest case; we achieve better bounds in the continuous setting).

Our first perspective interprets PFR as a *John-type approximation.* In general terms, a John-type theorem says that any object in some class is approximated efficiently (i.e. up to some constant factor) by some object from some simpler class. Some examples are John's Theorem approximating convex bodies by ellipsoids, Freiman's Theorem approximating sets of small doubling by GAPs, and a theorem of Tao and Vu [TV08] approximating convex progressions by GAPs.

Theme 1: John-type approximation. One question that we address is the existence of John-type approximations P for sets of bounded doubling A. E.g. if $A \subset \mathbb{Z}$ is non-degenerate with $|A+A| \le (2^d+\ell)|A|$ and $\ell \le 0.1 \cdot 2^d$ we show that A is covered by $\ell+1$ translates of a d-GAP P with $|P| \le O_d(|A|)$. Here $\ell+1$ translates is optimal (as shown by adding ℓ scattered points to a d-GAP), so in the sense of John-type approximation we have a precise characterisation of such sets A. We also show (see Theorem 1.5) that if $A \subset \mathbb{Z}$ with $|A+A| \le (1-\epsilon)2^d|A|$ is non-degenerate then A is contained in $O(2^d)$ translates of a d-dimensional convex progression P with $|P| \le O_{d,\epsilon}(|A|)$; thus we obtain one of the polynomial bounds required by PFR.

Our second perspective sees PFR as a *stability* statement. In general terms, if an object in some class is close to maximising some function on the class, then it must be structurally close to some extremal example. The possible meanings of 'structurally close' are nicely expressed by terminology of Tao: we speak of 1%, 99% or 100% stability according to whether we approximate some constant fraction (1%), all bar some constant fraction (99%), or everything (100%). For sets of small doubling, stability results are only known when the doubling is quite close to the minimum possible, such as the celebrated Freiman $3k-4$ Theorem (see [Fre59]) and various results (described below) for 'non-degenerate' A in $\mathbb{R}^k$ or $\mathbb{Z}^k$ with $|A+A| \le (2^k+\delta)|A|$ for small δ. Ruzsa's Covering Lemma (see e.g. [Ruz99]) converts any 1% stability theorem into a 100% stability theorem. However, this argument is quantitatively weak, so we require an alternative approach for optimal bounds.

Theme 2: Stability. We approach 100% stability via 99% stability, i.e. we first seek a structural description for almost all of A, and then use it to deduce the remaining structure. This approach is well-known in Extremal Combinatorics (the 'stability method'), but we are not aware of applications to Freiman's Theorem. Our 99% stability result (see Theorem 1.6) shows for any $\epsilon, \alpha > 0$ that if $A \subset \mathbb{Z}$ with $|A+A| \le (2-\epsilon)2^d|A|$ is non-degenerate then some $A' \subset A$ with $|A'| > (1-\alpha)|A|$ is efficiently covered by either a $(d+1)$-GAP or $O_{\epsilon,\alpha}(1)$ translates of a d-GAP. This 'dimension-free' bound for approximate covering makes for a surprising contrast with exact covering, where the required number of translates not only grows with d, but does so exponentially.

Our third perspective on PFR sees it as describing the *locality* of A. We think of $|X|$ as the number of locations for A, taking the view that elements of the same convex progression are close additively, even though they need not be close metrically. This perspective is particularly clarifying for the continuous setting of $A \subset \mathbb{R}^k$. Here the classical Brunn-Minkowski inequality shows that[1] $|A+A| \ge 2^k|A|$, with equality if and only if A is convex up to a null set. There is a substantial literature on $A \subset \mathbb{R}^k$ with $|A+A| \le (2^k+\delta)|A|$ for small $\delta > 0$. For such A, Christ [Chr12a] showed that the convex hull $\mathrm{co}(A)$ satisfies $|\mathrm{co}(A)| \le (1+\epsilon(\delta))|A|$, where $\epsilon(\delta) \to 0$ as $\delta \to 0$. Improvements were obtained by Figalli and Jerison [FJ15, FJ21], recently culminating in a sharp stability result with optimal

[1]For now we use $|\cdot|$ notation for both (discrete) cardinality and (continuous) measure, but for clarity later we use $|\cdot|$ for measure and $\#(\cdot)$ for cardinality.

parameters by van Hintum, Spink and Tiba [vHST22]. Thus for small δ the locality of A is simply its convex hull, but for larger δ the picture becomes more complicated.

Theme 3: Locality. A natural starting point is to consider $A \subset \mathbb{R}^k$ with $|A + A| \leq (2^k + \delta)|A|$ and ask how large $\delta > 0$ may be for us to still efficiently cover A by a convex set. This is clearly impossible for $\delta \geq 1$: consider a convex set and add one faraway point. Our general theory shows that the threshold is exactly at 1, i.e. if $\delta < 1$ then $|\operatorname{co}(A)| < O_k(|A|)$. We have a simple proof for this case, which will be published separately in [vHK23b]. As δ increases we can efficiently cover A by translates of a convex set; moreover, while $\delta < 0.1 \cdot 2^k$ we can do so with $\delta + 1$ translates, which is optimal (similarly to the result in $\mathbb{Z}$ mentioned above; both are subsumed in a more general picture). However, when δ reaches 2^k, any fixed number of translates of a convex set will not suffice, as A may be of the form $X + P$ where P is convex and X is an AP.

Generalised convex progressions. Hence, we introduce a common generalisation of convex progressions and GAPs: a *convex (k, d)-progression* is a set $P = \phi(C \cap (\mathbb{R}^k \times \mathbb{Z}^d))$, where $C \subset \mathbb{R}^{k+d}$ is convex and $\phi : \mathbb{R}^{k+d} \to \mathbb{R}^k$ is a linear map. In Theorem 1.4 we consider non-degenerate $A \subset \mathbb{R}^k$ with $|A + A| \leq (2^{k+d} + \ell)|A|$ for $d, \ell \geq 0$ and find an efficient covering of A by $X + P$, where P is a convex (k, d)-progression and $|X|$ is tightly controlled in terms of ℓ; in particular, for $\ell \leq 0.1 \cdot 2^{k+d}$ we obtain the optimal bound $|X| \leq \ell + 1$. Setting $d = 0$ recovers our results discussed above for $A \subset \mathbb{R}^k$. We will also see that the results in $\mathbb{Z}$ follow from those in $\mathbb{R}$. Thus in our general setting we can still think of $|X|$ as measuring locality, provided that we think of elements of the same generalised convex progression as being close additively, if not metrically.

In the next subsection we introduce notation that is used to formally develop the above concepts and state our precise results during the remainder of this introduction. These are organised by subsection according to our main theme of locality, covering the results discussed above, and the following further results.

- For non-degenerate $A \subset \mathbb{R}^k$ with $|A + A| \leq (2^{k+d} + \delta)|A|$ where $\delta \in (0, 1)$ we find a convex (k, d)-progression P with $|P \setminus A| < O_{k,d}(\delta|A|)$ (see Theorem 1.9). Setting $d = 0$ recovers the previously mentioned sharp stability result of [vHST22] for $A \subset \mathbb{R}^k$, whereas setting $k = 0$ recovers a sharp stability result for non-degenerate $A \subset \mathbb{Z}^d$ by the same authors [vHST23a].

- We obtain a very precise structural description of sets $A \subset \mathbb{R}$ in the line with doubling less than 4 (see Theorem 1.8).

1.1 Overview and notation

The following generalised notion of convex hull will play a crucial role throughout the paper. For $A \subset \mathbb{R}^k$, we write $\operatorname{co}_t^{\mathbb{R}^k,d}(A) = X + P$, where P is a proper convex (k, d)-progression and $\#X \leq t$, choosing X and P so that $A \subset X + P$ and $|X + P|$ is minimal; we fix an arbitrary choice if $X + P$ is not unique. In some cases we will omit k, d if $d = 0$ and t if $t = 1$, e.g. $\operatorname{co}(X)$ should be understood as $\operatorname{co}_1^{\mathbb{R}^k,0}(X)$, where k is the dimension of the ambient space for X, so that it coincides with the common notion of the convex hull.

We also require the closely related notion $\operatorname{gap}_t^{\mathbb{R}^k,d}(A)$, defined as a minimum volume set $X+P+Q$ containing A such that $\#X \le t$, P is a proper d-GAP and Q is a parallelotope. This is roughly equivalent to the variant $\operatorname{sco}_t^{\mathbb{R}^k,d}(A)$ defined exactly as $\operatorname{co}_t^{\mathbb{R}^k,d}(A)$ but imposing the symmetry requirement $P=-P$. Indeed, $|\operatorname{co}_t^{\mathbb{R}^k,d}(A)| \le |\operatorname{sco}_t^{\mathbb{R}^k,d}(A)| \le |\operatorname{gap}_t^{\mathbb{R}^k,d}(A)|$ is clear, and we will show that $|\operatorname{gap}_t^{\mathbb{R}^k,d}(A)| = O_{d,k}(|\operatorname{sco}_t^{\mathbb{R}^k,d}(A)|)$. Most results in this paper will be stated using $\operatorname{gap}_t^{\mathbb{R}^k,d}$, are equivalent to the corresponding statement using $\operatorname{sco}_t^{\mathbb{R}^k,d}$, and imply the corresponding statement using $\operatorname{co}_t^{\mathbb{R}^k,d}$. However, for some very precise statements we require $\operatorname{co}_t^{\mathbb{R}^k,d}$.

To state our results in $\mathbb{Z}$, let $\operatorname{co}_t^{\mathbb{Z},d}, \operatorname{sco}_t^{\mathbb{Z},d}$, and $\operatorname{gap}_t^{\mathbb{Z},d}$ be the corresponding functions for subsets of $\mathbb{Z}$, replacing 'convex (k,d)-progression' by 'convex d-progression'. To be precise, $\operatorname{co}_t^{\mathbb{Z},d}(A) = X+P$ where P is a convex d-progression and $\#X \le t$, choosing X and P so that $A \subset X+P$ and $\#(X+P)$ minimal. Analogously define $\operatorname{sco}_t^{\mathbb{Z},d}(A)$ and $\operatorname{gap}_t^{\mathbb{Z},d}(A)$ with the additional requirement that P is origin symmetric and a generalised arithmetic progression, respectively. Intuitively, throughout the paper we think of k as 'continuous dimension' and d as 'discrete dimension'. To stress this connection we write $\operatorname{co}_t^{k,d}$ for $\operatorname{co}_t^{\mathbb{R}^k,d}$ and $\operatorname{co}_t^{0,d}$ for $\operatorname{co}_t^{\mathbb{Z},d}$.

As further illustrations of this notation we can restate Freiman's Theorem and PFR as follows.

$$\begin{aligned} \text{Freiman's Theorem.} \qquad & \text{If } A \subset \mathbb{Z} \text{ with } \#(A+A) \le K\#A, \\ & \qquad \text{then } \#\left(\operatorname{co}_{O_K(1)}^{0,O_K(1)}(A)\right) \le O_K(\#A). \\ \text{PFR.} \qquad & \text{If } A \subset \mathbb{Z} \text{ with } \#(A+A) \le e^{O(d)}\#A, \\ & \qquad \text{then } \#\left(\operatorname{co}_{e^{O(d)}}^{0,O(d)}(A)\right) \le e^{O(d)}\#A. \end{aligned}$$

Many of our theorems include a non-degeneracy condition which should be interpreted as follows

$$\#\left(\operatorname{gap}_{O_{d,\xi}(1)}^{0,d-1}(A)\right) \ge \Omega_{d,\xi}(\#A) \iff \forall d,\xi, \exists C,c : \#\left(\operatorname{gap}_c^{0,d-1}(A)\right) \ge C\#A, \tag{1}$$

i.e. there is no collection of few (c depending only on d, and ξ) translates of a $d-1$ dimensional generalized arithmetic progression covering A efficiently (exceeding the size of A by at most a factor C depending only on d and ξ).

We will state our results in the following subsections according to the theme of locality, i.e. with respect to the bounds on the parameter t in $\operatorname{co}_t^{k,d}(A)$. A rough summary of their contents is as follows:

- A big part of the set is in one place.
- The entire set is in few places.
- Almost all of the set is in a constant number of places.

- Sets in the line with doubling less than 4 are almost convex.
- A sharp doubling condition for almost convexity.

The theory of sumsets has been developed in several groups. Particular attention has been given to the continuous setting of $\mathbb{R}^k$ (e.g. [FMP09, FMP10, Chr12a, FJ15, FJ17, FJ21, vHST23c, vHST22]) and to the discrete setting of $\mathbb{Z}$ (e.g. [Fre59, Ruz94, Bil99, Cha02, GT06, San08, Sch11]). In the context of this paper, the setting does not make much difference to our results, so although we state some results below for $\mathbb{Z}$, for the proofs we will generally prefer to work in $\mathbb{R}^k$. This is justified as (a) the proofs are the same modulo the theory of the additive hull, and (b) the results in $\mathbb{Z}$ follow from the results in $\mathbb{R}$ via the following proposition.

Proposition 1.1. *For any $A \subset \mathbb{Z}$ and $d, t \in \mathbb{N}$ there is $\epsilon = \epsilon(A, d, t) > 0$ so that $B := A + [-\epsilon, \epsilon] \subset \mathbb{R}$ has $\left|\mathrm{gap}_t^{1,d}(B)\right| = \Theta_{d,t}\left(\epsilon \# \left(\mathrm{gap}_t^{0,d}(A)\right)\right)$.*

1.2 A big part of the set is in one place

We start with the continuous setting of $\mathbb{R}^k$, where we obtain a clean unified formulation of the 1% stability phenomenon to be discussed in this subsection (a big part in one place) and as $\delta \to 0$ the 99% stability phenomenon (for which we will describe sharper results below). Combining 1% stability with Ruzsa covering one obtains 100% stability (for which we will also describe sharper results below). Previous stability results for the Brunn-Minkowski inequality only applied for much smaller δ; in particular, we are not aware of any previous results of this kind where δ does not decrease with the dimension.

Theorem 1.2. *Let $A \subset \mathbb{R}^k$ with $\left|\frac{A+A}{2}\right| \leq (1+\delta)|A|$, where $\delta \in (0,1)$. Then there exists $A' \subset A$ with $|A'| \geq (1 - \min\{\delta, \delta^2(1+O(\delta))\})|A|$ and $|\mathrm{co}(A')| \leq O_{1-\delta,k}(|A|)$.*

Since the size of A' is also independent of the dimension, this result is closely related to the stability question of the Prékopa-Leindler inequality for equal functions, which can be seen as a dimensionally independent version of Brunn-Minkowski. We expand on this connection and formulate a conjectured extension of Theorem 1.2 in [vHK23a, Section 11.1].

Now we consider the integer setting, where the Freiman-Bilu theorem [Bil99] shows that for sets with small doubling a large part of the set is contained in a small GAP of low dimension. A quantitative version of Green and Tao [GT06] shows that for $A \subset \mathbb{Z}$ with $\#(A+A) \leq 2^d(2-\epsilon)\#A$ there exists some $A' \subset A$ with $\#\left(\mathrm{gap}^{0,d}(A')\right) \leq \#A$ and

$$\#A' \geq \exp\left(-O(8^d d^3)\right) \epsilon^{O(2^d)} \#A.$$

With the following result we establish a bound on $\#A'$ that is optimal up to lower order terms, at the cost of a non-degeneracy assumption and relaxing the bound on $\#\,\mathrm{gap}^{0,d}(A')$ in the spirit of our John-type theme. We emphasise that our bound on $\#A'$ does not depend on the dimension.

Theorem 1.3. *Fix $\beta > 0$. Suppose $A \subset \mathbb{Z}$ with $\#(A+A) \le 2^d(1+\delta)\#A$, where $\delta \in (0,1)$. If $\#\left(\mathrm{gap}^{0,d-1}_{O_{d,\delta,\beta}(1)}(A)\right) \ge \Omega_{d,\delta,\beta}(\#A)$ then there exists $A' \subset A$ with $\#\left(\mathrm{gap}^{0,d}(A')\right) \le O_{\delta,d}(\#A)$ and*

$$\frac{\#\left(A \setminus A'\right)}{\#A} \le \min\left\{(1+\beta)\delta, \delta^2 + 60\delta^3\right\}.$$

The above results are sharp up to lower order terms both as $\delta \to 0$ for large k and as $\delta \to 1$. For $\delta \to 0$, consider the union of two homothetic convex sets of volumes δ^2 and $1-\delta^2$; for $\delta \to 1$ consider an arithmetic progression of $\frac{1}{1-\delta}$ equal convex sets. The second example suggests a $k+1$ dimensional convex structure, which we indeed establish in Section 1.4.

1.3 The entire set is in few places

Now we consider the 100% stability problem: what is the maximum 'locality' for given doubling? Our fundamental example is a GAP together with some scattered points, i.e. $A = P \cup S \subset \mathbb{Z}$, where P is a proper d-dimensional GAP and $\#S = \ell$. Then $\#(A+A) \le (2^d + \ell)\#A$ and A has locality $\ell + 1$.

The following result shows for non-degenerate A that this example is exactly sharp for a large range of ℓ and asymptotically sharp when $2^d - \ell \gg 2^{d/2}$. We remark that even the much weaker bound of $O(2^d)$ for the locality is already sufficient to cover A by $X + P$ with doubling $O(2^{2d})$, i.e. our John-type approximation only loses a square in the doubling, whereas the PFR setting allows any polynomial loss.

Theorem 1.4. *Let $\ell \in (0, 2^d)$ and $A \subset \mathbb{Z}$ with $\#(A+A) \le (2^d+\ell)\#A$ and $\#\left(\mathrm{gap}^{d-1}_{O_d(1)}(A)\right) \ge \Omega_d(\#A)$. Then $\#\left(\mathrm{gap}^{0,d}_{\ell'}(A)\right) \le O_{d,\ell}(\#A)$, where*

$$\ell' \le \begin{cases} \ell + 1 & \text{for } \ell \in \mathbb{N} \text{ if } \ell \le 0.1 \cdot 2^d, \text{ or if } \ell \le 0.315 \cdot 2^d \text{ and } d \ge 13, \\ \ell\left(1 + O\left(\sqrt[3]{\frac{2^d}{(2^d-\ell)^2}}\right)\right) & \text{if } 0.1 \cdot 2^d \le \ell \le \left(1 - \frac{1}{\sqrt{2^d}}\right)2^d, \\ (1+o(1))\frac{d+1}{2\epsilon} & \text{where } \epsilon = \frac{2^d-\ell}{2^d} \text{ and } o(1) \to 0 \text{ as } \epsilon \to 0. \end{cases}$$

The final bound in Theorem 1.4 gives an asymptotically sharp result for the limiting case $\epsilon = \frac{2^d-\ell}{2^d} \to 0$, i.e. as the doubling approaches 2^{d+1}. Here the above fundamental example breaks down and a new example takes over, which can be thought of as a cone over a GAP; intuitively, this describes the 'most d-dimensional' $(d+1)$-dimensional construction.

The above results are very sharp for non-degenerate sets, but to make further progress towards PFR we need to weaken the non-degeneracy condition. Our next result takes a step in this direction, but its applicability is limited by the double-exponential dependence on d'.

Theorem 1.5. *Let $A \subset \mathbb{Z}$ with $\#(A+A) < 2^d\#A$ and $d' < d$. If $\#\left(\mathrm{gap}^{0,d-d'}_{O_d(1)}(A)\right) \ge \Omega_d(\#A)$ then*

$$\#\left(\mathrm{gap}^{0,d}_{2^d \exp\exp(O(d'))}(A)\right) = O_d(\#A).$$

1.4 Almost all of the set is in a constant number of places

Now we consider 99% stability. Here we find that the number of locations can be bounded by a constant independently of the doubling. This is a remarkable contrast with the 100% stability problem, for which we needed an exponential number of locations to cover the set unless it has close to the minimum possible doubling. Our fundamental example had many scattered points but essentially all of the mass of the set in one location, which hints that one should be able to do much better if one can discard a small part of the set. Furthermore, the second example that takes over as the doubling approaches 2^{d+1} is highly structured so that $\mathrm{co}^{0,d+1}(A)$ is small. The following shows that any non-degenerate set is approximately described by one of these two configurations: it is concentrated in a single $(d+1)$-progression or few d-progressions, where 'few' depends only on the approximation accuracy, not on d.

Theorem 1.6. *For any $\alpha, \epsilon > 0$ and $A \subset \mathbb{Z}$ with $\#(A+A) \leq 2^d(2-\epsilon)\#A$ there is $A' \subset A$ with $\#A' \geq (1-\alpha)\#A$ and*

$$\min\left\{\#\left(\mathrm{gap}^{0,d-1}_{O_d(1)}(A)\right), \#\left(\mathrm{gap}^{0,d}_{O(\alpha^{-2})}(A')\right), \#\left(\mathrm{gap}^{0,d+1}_{1}(A')\right)\right\} \leq O_{d,\epsilon,\alpha}(\#A).$$

Hence, we can always find arithmetic structure in an absolute constant fraction of the set.

Corollary 1.7. *Suppose $A \subset \mathbb{Z}$ with $\#(A+A) \leq 2^d(2-\epsilon)\#A$ and $\#\,\mathrm{gap}^{0,d-1}_{O_d(1)}(A) > \Omega_{d,\epsilon}(\#A)$. Then there is $A' \subset A$ with $\#A' \geq \frac{1}{50000}\#A$ and $\#\left(\mathrm{gap}^{0,d+1}(A')\right) \leq O_{d,\epsilon}(\#A')$.*

1.5 Linear stability results

We additionally prove the following results. For more details and background refer to [vHK23a].

We prove the following extension of Freiman's $3k-4$ theorem, characterizing sets $A \subset \mathbb{R}$ with $|A+A| < 4|A|$. Let $\mathrm{ap}_t(A)$ be a minimum size set containing A that is an AP of t intervals whose lengths are in arithmetic progression.

Theorem 1.8. *There is an absolute constant $C > 0$ such that the following holds. Suppose $A \subset \mathbb{R}$ with $|A+A| < 4|A|$ and $|\mathrm{co}(A)| \geq C|A|$. Let t be minimal so that $|\mathrm{co}_t(A)| < 2|A|$, and let $\delta := \frac{|A+A|}{|A|} - (4-2/t)$. Then $|\mathrm{co}^{1,1}(A) \setminus A| \leq \max\{150, 4t^2\}\delta|A|$. Moreover, if $\delta \leq (2t)^{-2}$ then $|\mathrm{ap}_t(A) \setminus A| \leq 100t\delta|A|$.*

We extend the main results from [vHST22, vHST23a] to establish the optimal bound on the doubling for which a set needs to be approximately convex. An independent proof of the corollary is published separately in [vHK23b].

Theorem 1.9. *For any $d \in \mathbb{N}$, $\gamma, \epsilon > 0$, $\delta \in (0, 1-\epsilon)$, if $A \subset \mathbb{Z}$ with $\#(A+A) \leq (2^d+\delta)\#A$ and $\#\left(\mathrm{gap}^{0,d-1}_{O_{d,\gamma,\epsilon}(1)}(A)\right) = \Omega_d(\#A)$ then $\#(\mathrm{co}^{0,d}(A) \setminus A)/\#(A) \leq O_d(\gamma+\delta)$.*

Corollary 1.10. *If $A \subset \mathbb{R}^k$ satisfies $|A+A| = (2^k+\delta)|A|$ with $\delta < 1$ then $|\mathrm{co}(A) \setminus A| \leq O_k(\delta)|A|$.*

References

[BEKR23] Ben Barber, Joshua Erde, Peter Keevash, and Alexander Roberts. Isoperimetric stability in lattices. *Forthcoming in the Proceedings of the American Mathematical Society*, 2023+.

[Bil98] Yuri Bilu. The $(\alpha + 2\beta)$-inequality on a torus. *Journal of the London Mathematical Society*, 57(3):513–528, 1998.

[Bil99] Yuri Bilu. Structure of sets with small sumset. *Astérisque*, 258(xi):77–108, 1999.

[CDR19] Pablo Candela and Anne De Roton. On sets with small sumset in the circle. *The Quarterly Journal of Mathematics*, 70(1):49–69, 2019.

[CGSG19] Pablo Candela, Diego González-Sánchez, and David Joseph Grynkiewicz. A step towards the $3k-4$ conjecture in $\mathbb{Z}/p\mathbb{Z}$ and an application to m-sum-free sets. *Acta Mathematica Universitatis Comenianae*, 88(3):521–525, 2019.

[Cha02] Mei-Chu Chang. A polynomial bound in Freiman's theorem. *Duke mathematical journal*, 113(3):399–419, 2002.

[Chr12a] Michael Christ. Near equality in the Brunn–Minkowski inequality. *arXiv preprint arXiv:1207.5062*, 2012.

[Chr12b] Michael Christ. Near equality in the two-dimensional Brunn–Minkowski inequality. *arXiv preprint arXiv:1206.1965*, 2012.

[CSS20] Pablo Candela, Oriol Serra, and Christoph Spiegel. A step beyond Freiman's theorem for set addition modulo a prime. *Journal de Théorie des Nombres de Bordeaux*, 32(1):275–289, 2020.

[EGM14] Sean Eberhard, Ben Green, and Freddie Manners. Sets of integers with no large sum-free subset. *Annals of Mathematics*, pages 621–652, 2014.

[FJ15] Alessio Figalli and David Jerison. Quantitative stability for sumsets in $\mathbb{R}^n$. *Journal of the European Mathematical Society*, 17(5):1079–1106, 2015.

[FJ17] Alessio Figalli and David Jerison. Quantitative stability for the Brunn–Minkowski inequality. *Advances in Mathematics*, 314:1–47, 2017.

[FJ21] Alessio Figalli and David Jerison. A sharp Freiman type estimate for semisums in two and three dimensional euclidean spaces. *Annales Scientifiques de l'École Normale Supérieure*, 54(4):235–257, 2021.

[FMP09] Alessio Figalli, Francesco Maggi, and Aldo Pratelli. A refined Brunn–Minkowski inequality for convex sets. In *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, volume 26, pages 2511–2519. Elsevier, 2009.

[FMP10] Alessio Figalli, Francesco Maggi, and Aldo Pratelli. A mass transportation approach to quantitative isoperimetric inequalities. *Inventiones mathematicae*, 182(1):167–211, 2010.

[Fre59] Gregory Freiman. The addition of finite sets. i. *Izvestiya Vysshikh Uchebnykh Zavedenii. Matematika*, (6):202–213, 1959.

[GR06] Ben Green and Imre Ruzsa. Sets with small sumset and rectification. *Bulletin of the London Mathematical Society*, 38(1):43–52, 2006.

[Gre07] Ben Green. The polynomial Freıman–Ruzsa conjecture. *Terence Tao's blog*, 2007.

[GT06] Ben Green and Terence Tao. Compressions, convex geometry and the Freiman–Bilu theorem. *Quarterly Journal of Mathematics*, 57(4):495–504, 2006.

[vHK23a] Peter van Hintum and Peter Keevash. Locality in sumsets. *arXiv preprint arXiv:2304.01189*, 2023.

[vHK23b] Peter van Hintum and Peter Keevash. The sharp doubling threshold for approximate convexity. *arXiv preprint arXiv:2304.01176*, 2023.

[vHKT23] Peter van Hintum, Peter Keevash, and Marius Tiba. On Ruzsa's discrete Brunn-Minkowski conjecture. *Manuscript*, 2023.

[vHST22] Peter van Hintum, Hunter Spink, and Marius Tiba. Sharp stability of Brunn–Minkowski for homothetic regions. *Journal of the European Mathematical Society*, 24(12):4207–4223, 2022.

[vHST23a] Peter van Hintum, Hunter Spink, and Marius Tiba. Sets in $\mathbb{Z}^k$ with doubling $2^k + \delta$ are near convex progressions. *Advances in Mathematics*, 413:108830, 2023.

[vHST23b] Peter van Hintum, Hunter Spink, and Marius Tiba. Sharp l^1 inequalities for sup-convolution. *Forthcoming in Discrete Analysis*, 2023+.

[vHST23c] Peter van Hintum, Hunter Spink, and Marius Tiba. Sharp quantitative stability of the planar Brunn–Minkowski inequality. *Forthcoming in Journal of the European Mathematical Society*, 2023+.

[Jin07] Renling Jin. Freiman's inverse problem with small doubling property. *Advances in Mathematics*, 216(2):711–752, 2007.

[Kee11] Peter Keevash. Hypergraph turan problems. *Surveys in combinatorics*, 392:83–140, 2011.

[LR17] Shachar Lovett and Oded Regev. A counterexample to a strong variant of the polynomial Freiman-Ruzsa conjecture in euclidean space. *Discrete Analysis*, 8:379–388, 2017.

[LS23] Vsevolod F Lev and Oriol Serra. Towards $3n - 4$ in groups of prime order. *arXiv preprint arXiv:2302.08465*, 2023.

[Mag08] Francesco Maggi. Some methods for studying stability in isoperimetric type problems. *Bulletin of the American Mathematical Society*, 45(3):367–408, 2008.

[MFY74] DA Moskvin, Gregory Freiman, and Aleksandr Aleksandrovich Yudin. Structure theory of set addition and local limit theorems for independent lattice random variables. *Theory of Probability & Its Applications*, 19(1):53–64, 1974.

[Ruz94] Imre Ruzsa. Generalized arithmetical progressions and sumsets. *Acta Mathematica Hungarica*, 8(4):1–6, 1994.

[Ruz99] Imre Ruzsa. An analog of Freiman's theorem in groups. *Astérisque*, 258(199):323–326, 1999.

[Ruz09] Imre Ruzsa. Sumsets and structure. *Combinatorial number theory and additive group theory*, pages 87–210, 2009.

[San08] Tom Sanders. Appendix to 'Roth's theorem on progressions revisited' by J. Bourgain. *Journal d'Analyse Mathématique*, 1(104):193–206, 2008.

[San12] Tom Sanders. On the Bogolyubov–Ruzsa lemma. *Analysis & PDE*, 5(3):627–655, 2012.

[Sch11] Tomasz Schoen. Near optimal bounds in Freiman's theorem. *Duke Mathematical Journal*, 158(1):1–12, 2011.

[Tao18] Terence Tao. An inverse theorem for an inequality of Kneser. *Proceedings of the Steklov Institute of Mathematics*, 303(1):193–219, 2018.

[TV06] Terence Tao and Van Vu. *Additive combinatorics*, volume 105. Cambridge University Press, 2006.

[TV08] Terence Tao and Van Vu. John-type theorems for generalized arithmetic progressions and iterated sumsets. *Advances in Mathematics*, 219(2):428–449, 2008.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Fractionally isomorphic graphs and graphons

(Extended abstract)

Jan Hladký* Eng Keat Hng†

Abstract

Fractional isomorphism is a well-studied relaxation of graph isomorphism with a very rich theory. Grebík and Rocha [Combinatorica 42, pp 365–404 (2022)] developed a concept of fractional isomorphism for graphons and proved that it enjoys an analogous theory. In particular, they proved that if $G_1, G_2, \ldots$ converge to a graphon U, $H_1, H_2, \ldots$ converge to a graphon W and each G_i is fractionally isomorphic to H_i, then U is fractionally isomorphic to W. Answering the main question from *ibid*, we prove the converse of the statement above: If U and W are fractionally isomorphic graphons, then there exist sequences of graphs $G_1, G_2, \ldots$ and $H_1, H_2, \ldots$ which converge to U and W respectively and for which each G_i is fractionally isomorphic to H_i. As an easy but convenient corollary of our methods, we get that every regular graphon can be approximated by regular graphs.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-080`

This is an extended abstract to a preprint [arXiv:2210.14097] which is currently being submitted to a journal for publication.

*Institute of Computer Science of the Czech Academy of Sciences, Pod Vodárenskou věží 2, 182 00, Praha 8, Czech Republic. E-mail: `hladky@cs.cas.cz`. Research supported by Czech Science Foundation Project GX21-21762X and with institutional support RVO:67985807.

†Institute of Computer Science of the Czech Academy of Sciences, Pod Vodárenskou věží 2, 182 00, Praha 8, Czech Republic. E-mail: `hng@cs.cas.cz`. Research supported by Czech Science Foundation Project GX21-21762X and with institutional support RVO:67985807.

1 Introduction and the statement of the main result

This work connects the notions of fractional isomorphism for graphs and for graphons. The former was introduced by Tinhofer in 1986 [7], and subsequently several important equivalent characterizations were added by Ramana, Scheinerman, and Ullman [6], by Dvořák [2] and by Dell, Grohe, and Rattan [1]. We recall these characterizations. Of these, (FGI-2) and (FGI-3) play an important role in our contribution to the corresponding theory for graphons. The remaining two are included only to illustrate the mathematical beauty of the theory, which has important applications in designing fast algorithms for fractional isomorphism testing (which is often used as a proxy to isomorphism testing).

Suppose that G and H are two graphs on the same vertex set V.

(FGI-1) *Characterization via bistochastic matrices.* G and H are fractionally isomorphic if and only if there is a bistochastic matrix S such that for the adjacency matrices A_G and A_H of the respective graphs we have $SA_G = A_H S$.

(FGI-2) *Characterization via counting trees.* For two graphs F and J, let $\hom(F, J)$ be the number of homomorphisms of F in J. The graphs G and H are fractionally isomorphic if and only if $\hom(T, G) = \hom(T, H)$ for every tree T.

(FGI-3) *Characterization via equitable partitions.* Let $V(F) = Y_1 \sqcup \ldots \sqcup Y_\ell$ be a partition of the vertex set of a graph F into nonempty sets. We say that $\mathcal{E} = (Y_1, \ldots, Y_\ell)$ is an *equitable partition* if there are numbers $(d_{i,j})_{i,j\in[\ell]}$ such that for every $i, j \in [\ell]$ and every $v \in Y_i$ we have
$$d_{i,j} = \deg_F(v, Y_j) \ . \tag{1}$$
We call the pair $(|Y_i|)_{i\in[\ell]}$ and $(d_{i,j})_{i,j\in[\ell]}$ *the parameters of* $\mathcal{E}$. The graphs G and H are fractionally isomorphic if and only if there are equitable partitions $\mathcal{E}_G$ of G and $\mathcal{E}_H$ of H that have the same parameters.

(FGI-4) *Characterization via iterated degree sequences.* For every vertex $v \in V$ first define $s_{1,G}(v) := \deg_G(v)$ and then inductively define multisets $s_{\ell+1,G}(v) := \{s_{\ell,G}(u) : u \in N_G(v)\}$. The *$\ell$-th iteration of the degree sequence of G* is the (multiset) collection $\mathcal{S}_{\ell,G} = \{s_{\ell,G}(v) : v \in V\}$. We can make analogous definitions for H. The graphs G and H are fractionally isomorphic if and only if $\mathcal{S}_{\ell,G} = \mathcal{S}_{\ell,H}$ for every $\ell \in \mathbb{N}$.

Let us now move to graphons. The basic theory of graphons and their role as limits of sequences of dense graphs is by now well-understood. We refer to [5] for the basics and borrow notation from there. Unless stated otherwise, the ground space for graphons is the square of a standard Borel space $(\Omega, \mathcal{B})$ equipped with a Borel probability measure π. Grebík and Rocha [4] developed a theory of fractional isomorphism for graphons. In particular, they showed that all the above characterizations of fractional graph isomorphism have graphon counterparts and are indeed equivalent. To formulate some of these counterparts, one needs to develop nontrivial analytic machinery. Here, we do that only for (FGI-2) and (FGI-3), which we require.

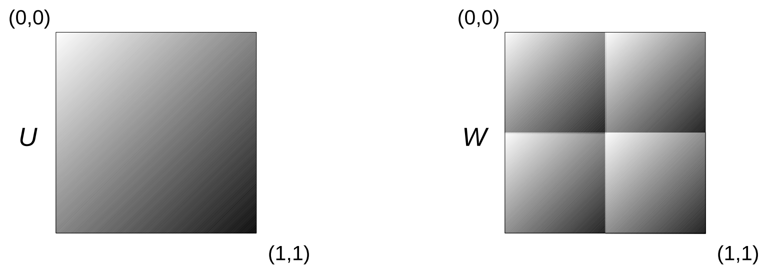


Figure 1: Graphons U and W from (FGI'-3).

(FGI'-2) Two graphons U and W are fractionally isomorphic if and only if $t(T,U) = t(T,W)$ for every tree T. Here, $t(\cdot,\cdot)$ is the usual homomorphism density function.

(FGI'-3) A naive counterpart to (FGI-3) would involve partitions $\Omega = Y_1 \sqcup \ldots \sqcup Y_\ell$ into sets of positive measure. This approach, however, does not work. Consider $\Omega = [0,1)$ with the Lebesgue measure π and two graphons U and W defined by $U(x,y) = (x+y)/2$ and $W(x,y) = 2((x+y) \mod 0.5)$; see Figure 1. It is easy to check that U and W are fractionally isomorphic in the sense of (FGI'-2). (As a matter of fact, U and W satisfy an even stronger condition called weak isomorphism.) But the requirement from (FGI-3) that all the vertices within one cell have the same degree dictates that the minimum (and only) equitable partition for U be $\{x\}_{x\in[0,1)}$ with uncountably many singleton cells. In W, we can pair up x and $x+\frac{1}{2}$ to get the minimum equitable partition $\left\{x, x+\frac{1}{2}\right\}_{x\in[0,\frac{1}{2})}$. This shows that one has to work with *sigma-algebras* instead of finite partitions into sets of positive measure. Here, we briefly recall the construction, referring to [3, 4] for details. We say that a sigma-algebra[*] $\mathcal{C} \subset \mathcal{B}$ is *U-invariant* if for every $f \in L^2(\Omega,\mathcal{C})$ and for the function $T_U f$ defined by the kernel operator T_U as $(T_U f)(x) = \int_y U(x,y)f(y)$ we have $T_U f \in L^2(\Omega,\mathcal{C})$. With this definition it can be shown that there is a unique minimum U-invariant sigma-algebra, denoted by $\mathcal{C}(U)$. Let us then consider the quotient space $\Omega/\mathcal{C}(U)$, the Borel probability measure $\pi/\mathcal{C}(U)$ on $\Omega/\mathcal{C}(U)$ and a measurable surjection $q_U\colon \Omega \to \Omega/\mathcal{C}(U)$ such that $\pi/\mathcal{C}(U)$ is the pushforward of π via q_U. With these notions, we can naturally transfer the conditional expectation $\mathbb{E}\left(U \,|\, \mathcal{C}(U) \times \mathcal{C}(U)\right)$ to the domain $(\Omega/\mathcal{C}(U))^2$ by requiring that for the resulting graphon $U/\mathcal{C} : (\Omega/\mathcal{C}(U))^2 \to [0,1]$ we have $\mathbb{E}\left(U \,|\, \mathcal{C}(U) \times \mathcal{C}(U)\right)(x,y) = U/\mathcal{C}(q_U(x), q_U(y))$ for all $x, y \in \Omega$. We can repeat the same construction for another graphon W. The graphons U and W are then fractionally isomorphic if there is a measure preserving bijection $b : \Omega/\mathcal{C}(U) \to \Omega/\mathcal{C}(W)$ so that $U/\mathcal{C}(U)(x,y) = (W/\mathcal{C}(W))(b(x), b(y))$ for every $x, y \in \Omega/\mathcal{C}(U)$. To summarize in nontechnical terms, the initial naive approach where the notion of grouping comes from a partition $\Omega = Y_1 \sqcup \ldots \sqcup Y_\ell$ has to be replaced by "grouping" according to sigma-algebra $\mathcal{C}(U)$ and conditional expectation $\mathbb{E}\left(U \,|\, \mathcal{C}(U) \times \mathcal{C}(U)\right)$ serves as the refined version of the numbers $d_{i,j}$ from (FGI-3).

[*]an additional technical condition is needed in the actual definition

We can start connecting the notions of fractional isomorphism for graphs and for graphons. The direction "finite graphs$\Rightarrow$graphons" was already observed in [4].

Proposition 1. *Suppose that $G_1, G_2, \ldots$ and $H_1, H_2, \ldots$ are sequences of graphs which converge in cut distance to graphons U and W respectively and for which G_i and H_i are fractionally isomorphic for each $i \in \mathbb{N}$. Then U and W are fractionally isomorphic.*

Since the proof in [4] is not very explicit ('follows from the fact that fractional isomorphism of graphons is an equivalence relation closed in the cut distance'), we give further details here.

Proof. Let T be an arbitrary tree on, say, k vertices. By (FGI-2) we have $\hom(T, G_i) = \hom(T, H_i)$ for each $i \in \mathbb{N}$. Since G_i and H_i are of the same order, we have equality of homomorphism densities, that is, $\frac{\hom(T,G_i)}{v(G_i)^k} = \frac{\hom(T,H_i)}{v(H_i)^k}$. Convergence in the cut distance implies convergence of all homomorphism densities, so in particular we have

$$t(T, U) = \lim_{i\to\infty} \frac{\hom(T, G_i)}{v(G_i)^k} = \lim_{i\to\infty} \frac{\hom(T, H_i)}{v(H_i)^k} = t(T, W) \ .$$

The fact that U and W are fractionally isomorphic now follows from (FGI'-2). $\square$

Let us look at the reverse direction "graphons$\Rightarrow$finite graphs". It is not true in general that if U and W are fractionally isomorphic graphons and $G_1, G_2, \ldots$ and $H_1, H_2, \ldots$ are sequences of graphs converging in cut distance to U and W respectively, then G_i and H_i are fractionally isomorphic for each $i \in \mathbb{N}$. Indeed, G_i and H_i might have different orders, which would automatically make them not fractionally isomorphic. Even if G_i and H_i had the same order, a single-edge edit of one of them would preserve convergence in cut distance but make them not fractionally isomorphic. Hence, a sensible question in this direction needs to have an existential quantification for the sequences instead of a universal one. Indeed, this was the main open question of [4].

Question 2 (Question 3.2 in [4])**.** *Suppose that U and W are fractionally isomorphic graphons. Do there exist sequences $\{G_i\}_{i\in\mathbb{N}}$ and $\{H_i\}_{n\in\mathbb{N}}$ of graphs which converge in cut distance to U and W respectively and for which G_i and H_i are fractionally isomorphic for each $i \in \mathbb{N}$?*

The main result of our work is a positive answer to Question 2. In fact, we prove a slightly stronger statement in which we simultaneously approximate an arbitrary (even infinite) family of mutually fractionally isomorphic graphons.

Theorem 3. *Suppose that $\mathcal{U}$ is a family of mutually fractionally isomorphic graphons. Then for each $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that for each $n \geq n_0$ there exists a family $\{H_U\}_{U\in\mathcal{U}}$ of mutually fractionally isomorphic graphs on vertex set $[n]$ with the property that for each $U \in \mathcal{U}$ the cut distance between U and H_U is at most ε.*

It is easy to check that for each $d \in [0,1]$ the family $\mathcal{U}_d$ of all d-regular graphons[†] is a family of mutually fractionally isomorphic graphons. Our proof of Theorem 3 gives the following corollary for d-regular graphons.

Theorem 4. *Suppose that $d \in [0,1]$ and $\mathcal{U}_d$ is the family of all d-regular graphons. Then for every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that for each $n \geq n_0$ there exists $D \in \mathbb{N}$ and a family $\{H_U\}_{U \in \mathcal{U}_d}$ of D-regular graphs on vertex set $[n]$ such that for each $U \in \mathcal{U}_d$ the cut distance between U and H_U is at most ε.*

2 Sketch of proof of Theorem 3

While we explained in (FGI'-3) that an infinitesimal approach using sigma-algebras is needed, we begin the following proof overview under the assumption that for a graphon $U \in \mathcal{U}$ the sigma-algebra $\mathcal{C}(U)$ is generated by a finite partition $\mathcal{Q} = \{Q_i\}_{i \in [M]}$ of Ω sets of positive measure; we shall return to the general case at the end of this exposition.

2.1 Desired graph profile from the graphon profile

The U-invariance of $\mathcal{C}(U)$ implies that for all $i, j \in [M]$ and $x \in Q_i$ we have a counterpart to (1), namely

$$\int_{z \in Q_j} U(x,z) = d_{i,j}\pi(Q_j) \tag{2}$$

where $d_{i,j} := \frac{1}{\pi(Q_i)\pi(Q_j)} \cdot \int_{y \in Q_i} \int_{z \in Q_j} U(z,y)$. The fact that each graphon $U' \in \mathcal{U}$ is fractionally isomorphic to U then means that there is a partition $\mathcal{Q}' = \{Q'_i\}_{i \in [M]}$ of Ω with $\pi(Q'_i) = \pi(Q_i)$ such that quantities $d'_{i,j}$ defined in analogy with $d_{i,j}$ satisfy $d'_{i,j} = d_{i,j}$. So, to prove the theorem in this simplified setting, it is enough to approximate (for a given $\varepsilon > 0$ and sufficiently large n) U in cut distance by an n-vertex graph H_U with an equitable partition (in the sense of (FGI-3)) whose parameters depend solely on the vector $\mathbf{r} = (\pi(Q_i))_{i \in [M]}$ and the matrix $\mathcal{D} = (d_{i,j})_{i,j \in [M]}$. Indeed, repeating the construction detailed below for any other $U' \in \mathcal{U}$ yields a graph $H_{U'}$ that has an equitable partition with the same parameters. Hence, it follows by (FGI-3) that $H_{U'}$ and H_U are fractionally isomorphic.

Our construction of H_U has two main steps, as detailed in the following two subsections.

2.2 Approximating by $\mathbb{G}(n,U)$

First, we use the inhomogeneous random graph model $\mathbb{G}(n,U)$ (see [5, Section 10.1]) to generate H^*_U. Let the points $x_1, \dots, x_n \in \Omega$ sampled in the procedure represent the respective vertices $1, \dots, n$ of $V(H^*_U)$. For $i \in [M]$ define $X^*_i := \{\ell \in V(H^*_U) : x_\ell \in Q_i\}$. By the 'Second sampling lemma' (see [5, Lemma 10.15]) H^*_U is close to U in cut distance

[†]A graphon W is d-regular, if for each $x \in \Omega$ we have $\int_y W(x,y) = d$.

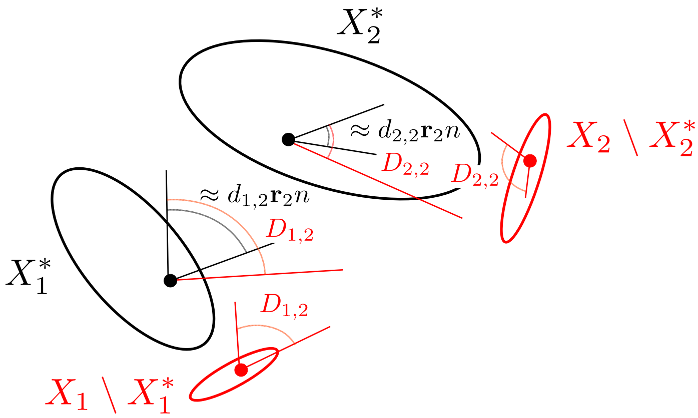


Figure 2: An example of the fine-tuning of the degree sequence. The figure focuses on the degrees within X_2 and from X_1 to X_2. The original graph H_U^* is shown in black. The parts added on the way to constructing H_U are shown in red. Theorems of Erdős–Gallai and of Gale–Ryser are used for the red parts.

with high probability. Further, basic concentration results tell us that with high probability, for each $i, j \in [M]$ and each $\ell \in X_i^*$ we have

$$|X_i^*| \approx \mathbf{r}_i n \text{ and} \tag{3}$$
$$\deg_{H_U^*}(\ell, X_j^*) \approx d_{i,j}\mathbf{r}_j n. \tag{4}$$

2.3 Fine-tuning the degree sequence

We shall modify H_U^* in several steps by adding $o(n)$ vertices to each set X_i and $o(n^2)$ edges inside each set X_i and inside each bipartite pair (X_i, X_j), with the aim to achieve after these modifications that

$$|X_i| = N_i \approx \mathbf{r}_i n \text{ and} \tag{5}$$
$$\deg_{H_U}(\ell, X_j) = D_{i,j} \approx d_{i,j}\mathbf{r}_j n \tag{6}$$

for some numbers $(N_i)_{i\in[M]}$ and $(D_{i,j})_{i,j\in[M]}$ which depend only on $\mathbf{r}$, $\mathcal{D}$ and n. Hence, we fulfil the task described in Section 2.1. We apply the classical theorem of Erdős and Gallai on graphic sequences and its bipartite counterpart due to Gale and Ryser. These theorems allow us to construct graphs within the sets $X_i \setminus X_i^*$ and in the pairs $(X_i \setminus X_i^*, X_j)$ (including the case $i = j$) with precisely controlled degree sequences to achieve a state in which each graph $H_U[X_i]$ is regular and each graph $H_U[X_i, X_j]$ is biregular as required by (6). An illustration is given in Figure 2.

2.4 From finite partitions to sigma-algebras

As explained in (FGI'-3), $U/\mathcal{C}(U)$ is defined in terms of a suitable sigma-algebra and does not usually correspond to a finite partition. Here, Szemerédi's regularity lemma comes to the rescue. Indeed, it is well-known that if $\{\tilde{Q}_i\}_{i\in[M]}$ is a δ-regular Szemerédi partition for a graphon Γ, then in particular we have an approximate version of (2) for most vertices $x \in \tilde{Q}_i$.[‡] We shall take $\delta \ll \varepsilon$.

So, given a graphon U, generate $H_U^* \sim \mathbb{G}(n, U)$ as in Section 2.2. Since we apply Szemerédi's regularity lemma merely to handle degrees, and since we want its application to work in the same way for all graphons in the class $\mathcal{U}$, we shall apply it to the graphon $\Gamma := U/\mathcal{C}(U)$. Let $\Omega/\mathcal{C}(U) = \tilde{Q}_1 \sqcup \ldots \sqcup \tilde{Q}_M$ be a δ-regular Szemerédi partition for Γ. Define $Q_1, \ldots, Q_M$ by $Q_i := q_U^{-1}(\tilde{Q}_i)$. Now, $\Omega = Q_1 \sqcup \ldots \sqcup Q_M$ is in general not a regular Szemerédi partition.[§] However, it can still be proved that an approximate version of (2) holds for most vertices $x \in Q_i$ with respect to the graphon U. This allows us to define again $X_i^* := \{\ell \in V(H_U^*) : x_\ell \in Q_i\}$ and then fine-tune the sequence in a spirit similar to that described in Section 2.3.

2.5 Proving Theorem 4

If U is d-regular then $\Omega/\mathcal{C}(U) = \{a\}$ consists of a single atom and $U/\mathcal{C}(U)(a, a) = d$. So, $M = 1$ and $Q_1 = \Omega$. In particular, the construction above guarantees that the graph $H_U = H_U[X_1]$ is regular, as needed.

[‡]In this overview, we neglect the issue that a Szemerédi partition may involve some irregular pairs. Also, we neglect that the usual regularity lemma does not control behaviour inside clusters whereas we shall need a counterpart to (2) even for $i = j$.

[§]This is perhaps best illustrated with an example. Let $\Omega = A_1 \sqcup A_2$, α_1, α_2 be two distinct numbers and U be such that for $i = 1, 2$ and for every $x \in A_i$ we have

$$\int_{y\in A_i} U(x, y) = \alpha_i \quad \text{and} \quad \int_{y\in A_{3-i}} U(x, y) = 0\,. \tag{7}$$

Indeed, in this example $\Omega/\mathcal{C}(U) = \{a_1, a_2\}$ consists of two atoms and we have $U/\mathcal{C}(U)(a_i, a_i) = \alpha_i$ and $U/\mathcal{C}(U)(a_i, a_{3-i}) = 0$. Obviously, the only possible Szemerédi regularization for $U/\mathcal{C}(U)$ has $M = 2$, $\tilde{Q}_1 = \{a_1\}$ and $\tilde{Q}_2 = \{a_2\}$. But the pullbacks $\{Q_i := q_U^{-1}(\tilde{Q}_i) = A_i\}_{i\in[2]}$ clearly need not form a Szemerédi regularization for U, since the restriction (7) leaves a lot of space for wildly structured graphons.

References

[1] H. Dell, M. Grohe, and G. Rattan. Lovász meets Weisfeiler and Leman. In *45th International Colloquium on Automata, Languages, and Programming*, volume 107 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 40, 14. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2018.

[2] Z. Dvořák. On recognizing graphs by numbers of homomorphisms. *J. Graph Theory*, 64(4):330–342, 2010.

[3] J. Grebík and I. Rocha. A graphon perspective for fractional isomorphism. *Acta Math. Univ. Comenian. (N.S.)*, 88(3):759–765, 2019.

[4] J. Grebík and I. Rocha. Fractional isomorphism of graphons. *Combinatorica*, 42(3):365–404, 2022.

[5] L. Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.

[6] M. V. Ramana, E. R. Scheinerman, and D. Ullman. Fractional isomorphism of graphs. *Discrete Math.*, 132(1-3):247–265, 1994.

[7] G. Tinhofer. Graph isomorphism and theorems of Birkhoff type. *Computing*, 36(4):285–300, 1986.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Permutation flip processes

(Extended abstract)

Jan Hladký* Hanka Řada†

Abstract

We introduce a broad class of stochastic processes on permutations which we call flip processes. A single step in these processes is given by a local change on a randomly chosen fixed-sized tuple of the domain. We use the theory of permutons to describe the typical evolution of any such flip process $\pi_0, \pi_1, \pi_2, \ldots$ started from any initial permutation $\pi_0 \in \mathrm{Sym}(n)$. More specifically, we construct trajectories $\Phi : \mathfrak{P} \times [0, \infty) \to \mathfrak{P}$ in the space of permutons with the property that if π_0 is close to a permuton γ then for any $T > 0$ with high probability π_{Tn} is close to $\Phi^T \gamma$. This view allows to study various questions inspired by dynamical systems.

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-081

1 Introduction

The theory of permutations offers many exciting structural, extremal and enumerative questions. For example, research centered around the Stanley–Wilf conjecture asks for the number of permutations of a given order avoiding a fixed pattern. What all these problems have in common is that they study static permutations. Following the success of the theory of dense graph limits, Hoppen, Kohayakawa, Moreira, Ráth, and Sampaio [4] developed a theory of permutation limits. The corresponding limit objects are called *permutons*. The theory of permutation limits allowed new results or streamlined proofs in the above areas

*Institute of Computer Science of the Czech Academy of Sciences. E-mail: hladky@cs.cas.cz. Supported by Czech Science Foundation Project 21-21762X.

†Institute of Computer Science of the Czech Academy of Sciences, and Department of Applied Mathematics, FIT, Czech Technical University in Prague. E-mail: hanka.rada@fit.cvut.cz. Supported by Czech Science Foundation Project 21-21762X.

(see e.g. [5, 3]), as well systematic treatment of many properties permutations coming from various random models.

There is another — *dynamic* — line of research of permutations. Much of this line of research is motivated by data structures, and by analysis of sorting algorithms in particular. In the dynamic setting, one studies evolution of sequences $\pi_0, \pi_1, \pi_2, \ldots$ of permutations (typically of the same order).

The main contribution of our work is a framework for capturing typical evolutions for a natural class of randomized local algorithms, which we call *flip processes*, using the theory of permutons.

2 Main concepts and results

2.1 Necessary notation

In order to state our results, we need to recall basics of the theory of permutons. All measure below on $[0,1]^2$ are tacitly assumed to be Borel. We write λ and λ^2 for the Lebesgue measure on $\mathbb{R}$ and on $\mathbb{R}^2$, respectively.

A *permuton* γ is a measure on $[0,1]^2$ with uniform marginals, that is, for each Borel set $Z \subset [0,1]$ we have that $\gamma([0,1] \times Z) = \gamma(Z \times [0,1])$ is the Lebesgue measure of Z. Permutons are an extension of permutations through the concept of permutation representation. Suppose that $\pi \in \mathrm{Sym}(n)$ is a permutation of order n. The *permuton representation* Γ_π of π is a measure defined

$$\Gamma_\pi(X) := n \cdot \lambda^2 \left(X \cap \bigcup_{i=1}^{n} [\tfrac{i-1}{n}, \tfrac{i}{n}] \times [\tfrac{\pi(i)-1}{n}, \tfrac{\pi(i)}{n}] \right)$$

for each $X \subset [0,1]^2$. Then Γ_π is indeed a permuton as the fact that each $i \in [n]$ appears exactly once in the domain and exactly once in the range corresponds the uniform marginals on the x-axis and the y-axis, respectively.

We write $\mathfrak{P}$ for the set of all permutons. Given two permutons α and β we define their *rectangular distance* by

$$\mathrm{d}_\square(\alpha,\beta) := \sup_{0 \le x_1 \le x_2 \le 1, 0 \le y_1 \le y_2 \le 1} \{|\alpha([x_1,x_2] \times [y_1,y_2]) - \beta([x_1,x_2] \times [y_1,y_2])|\} . \tag{1}$$

Three permutons which will appear in the text below are the two-dimensional Lebesgue measure λ^2, the *diagonal permuton* D, defined by $D(Z) = \lambda\{x \in [0,1] : (x,x) \in Z\}$ for $Z \subset [0,1]^2$ Borel, and the *antidiagonal permuton* A, defined by $A(Z) = \lambda\{x \in [0,1] : (x, 1-x) \in Z\}$.

2.2 Flip processes and permuton trajectories

To motivate our fairly broad class of flip processes, we start with a particular example, a specific ordering procedure. Suppose that π_0 is a permutation of order $n \ge 3$. Then in

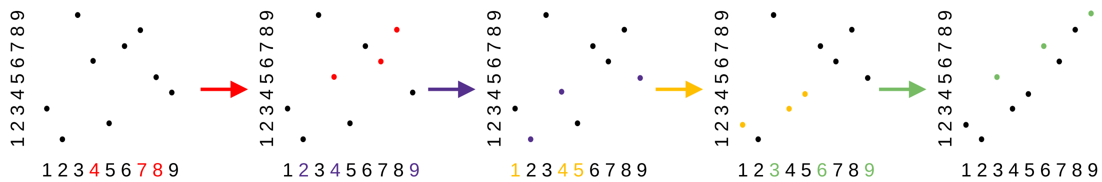


Figure 2.1: Five steps of the ordering process of order 3. The different colours indicate firstly the random choice of the elements and in the subsequent figure the corresponding reordering.

steps $\ell = 1, 2, \dots$ we take a uniform triple of distinct elements of $[n]$, say $i_1 < i_2 < i_3$. We take π_ℓ to be $\pi_{\ell-1}$, except at positions i_1, i_2, and i_3 and then we shuffle the values $\pi_{\ell-1}(i_1)$, $\pi_{\ell-1}(i_2)$, and $\pi_{\ell-1}(i_3)$ so that they are in the increasing order. We call this process the *ordering process of order* 3. An example is given in Figure 2.1.

Let us now proceed with a general definition. If $\pi \in \mathrm{Sym}(n)$ and $A \in \binom{[n]}{k}$ then *subpermutation of* π *restricted by* A $\pi_{\restriction A}$ is a permutation on $[k]$ such that for each $i, j \in [k]$ we have that $\pi_{\restriction A}(i) < \pi_{\restriction A}(j)$ if and only if for the i-th smallest element i_A of A and for the j-th smallest element j_A of A we have $\pi(i_A) < \pi(j_B)$. Next, we introduce a notion of *transplanting a subpermutation* $\psi \in \mathrm{Sym}(k)$ *into* π *on* A. This is a permutation $\widetilde{\pi} \in \mathrm{Sym}(n)$ such for each $i \in [n] \setminus A$ we have $\widetilde{\pi}(i) = \pi(i)$ and $\pi_{\restriction A} = \psi$.

So, the above ordering process can be defined as repeated transplantations of the identity permutation id_3 on randomly selected triples. General flip processes allow the transplanted subpermutation to depend on the sampled restricted subpermutation (and this choice can be a randomized one).

Let $k \in \mathbb{N}$. A *rule* is a stochastic matrix $\mathcal{R} \in [0,1]^{\mathrm{Sym}(k) \times \mathrm{Sym}(k)}$. Given an initial permutation $\pi_0 \in \mathrm{Sym}(n)$ (where $n \geq k$) the *flip process with rule* $\mathcal{R}$ works as follows. In each step $\ell = 1, 2, \dots$, pick a uniformly random k-tuple $A \in \binom{[n]}{k}$. Then pick a permutation $\psi \in \mathrm{Sym}(k)$ according to the probability distribution given by $\mathcal{R}$ on row $\pi_{\ell-1 \restriction A}$ and transplant it into $\pi_{\ell-1}$ on A. The resulting permutation is π_ℓ. To summarize, a flip process with a given rule is a discrete time-homogeneous Markov process $\pi_0, \pi_1, \pi_2, \dots$.

Suppose that we fix a flip process with a rule $\mathcal{R}$. The main result of our project states that there is a notion of 'trajectories', which are given as a two-variable function $\Phi : \mathfrak{P} \times [0, \infty) \to \mathfrak{P}$ (in which we write the second coordinate in the superscript, $\Phi^\alpha t$ for $\alpha \in \mathfrak{P}$ and $t \in [0, \infty)$) which predicts typical behaviour of the flip process $\mathcal{R}$ started from any permutation after linearly many steps (with respect to its order), up to a small error in the rectangular distance.

Theorem 1. *For every $k \in \mathbb{N}$ and for every permutation flip process $\mathcal{R}$ of order k, there exists a function $\Phi : \mathfrak{P} \times [0, \infty) \to \mathfrak{P}$ with the following property. For every $T > 0$, every $n \in \mathbb{N}$ and every $\pi_0 \in \mathrm{Sym}(n)$ we have with probability $1 - o_n(1)$ for the flip process $\pi_0, \pi_1, \dots$ with rule $\mathcal{R}$ that* $\max\left\{ \mathrm{d}_\square(\Gamma_{\pi_i}, \Phi^{\frac{i}{n}} \Gamma_{\pi_0}) \,:\, i \in (0, Tn] \cap \mathbb{N} \right\} = o_n(1)$.

That is, Theorem 1 establishes a correspondence between an analytic deterministic object and a discrete stochastic evolution.

Further, we can prove that the trajectories satisfy the following metric conditions for any $\alpha, \beta \in \mathfrak{P}$ and $s, t \in [0, \infty)$:

$$\exp\left(-\Theta(t)\right) \mathrm{d}_\square\left(\alpha, \beta\right) \leq \mathrm{d}_\square\left(\Phi^t\alpha, \Phi^t\beta\right) \leq \exp\left(\Theta(t)\right) \mathrm{d}_\square\left(\alpha, \beta\right) \quad \text{and ,} \tag{2}$$

$$\mathrm{d}_\square\left(\Phi^s\alpha, \Phi^t\alpha\right) \leq O(|s-t|) \; . \tag{3}$$

The upper bound in (2) implies that the evolution of the trajectory depends in a continuous fashion on the initial condition, and (3) says that it also depends in a continuous fashion on time. The lower bound in (2) in particular says that two different trajectories do not ever form a confluence (unless one is a subtrajectory of the other).

2.3 Flip processes as dynamical systems on $\mathfrak{P}$

Theorem 1 gives a potential of a comprehensive theory of permutation flip processes, richness of which reflects both the combinatorial and the dynamical systems facet of the area. We present several notions that we study. Due to space constraints we state only briefly and informally some of our results accompanying these notions as well as several open questions.

The first is a concept of destination. Suppose that $\mathcal{R}$ is a flip process and $\Phi : \mathfrak{P} \times [0, \infty) \to \mathfrak{P}$ are its trajectories. If γ is a permuton for which the limit (in the rectangular distance) $\lim_{t\to\infty} \Phi^t\gamma$ exists, then we call it the *destination of* γ, and write $\mathrm{dest}\,(\gamma)$. For example, it can be shown that the destination of any permuton in the above ordering process of order 3 is the diagonal permuton D. For most natural flip processes it appears that each permuton has a destination but we are also able to construct a flip process and an initial permuton γ for which $\lim_{t\to\infty} \Phi^t\gamma$ does not exist. More specifically, we are able to construct a flip process and argue that it contains a periodic trajectory, that is, a permuton γ and time T_0 so that $\Phi^t\gamma = \gamma$ for and only for times t that are multiples of T_0. It would be interesting to find other wild types of trajectories. For example, does there exist a flip process and an initial permuton α whose trajectory oscillates between the diagonal and the antidiagonal permuton, that is $\liminf_{t\to\infty} \mathrm{d}_\square(\Phi^t\alpha, D) = \liminf_{t\to\infty} \mathrm{d}_\square(\Phi^t\alpha, A) = 0$?

Destinations are connected with the notion of fixed points. For a flip process whose trajectories are $\Phi : \mathfrak{P} \times [0, \infty) \to \mathfrak{P}$, we call a permuton $\gamma \in \mathfrak{P}$ a *fixed point* if $\Phi^t\gamma = \gamma$ for all $t \geq 0$. It can be shown that if γ is a destination then it is also a fixed point (and obviously, if it is a fixed point then it is also its own destination). Is it true that every flip process has at least one fixed point?

Next, we explain the concept of origins. Suppose that $t > 0$, and α and β are permutons such that $\beta = \Phi^t\alpha$. In that case we write $\alpha = \Phi^{-t}\beta$. Let $\mathrm{age}(\beta)$ be the supremum of times $t \geq 0$ for which $\Phi^{-t}\beta$ exists as a permuton. It can be shown that if $\mathrm{age}(\beta) < \infty$ then there exists a permuton, denoted by $\mathrm{orig}\,(\beta)$, for which $\Phi^{\mathrm{age}(\beta)}(\mathrm{orig}\,(\beta)) = \beta$. We call $\mathrm{orig}\,(\beta)$ the *origin of* β. Another feature of interest is characterizing graphons with positive age. Indeed, the age of some permutons can be 0 as the example of the antidiagonal permuton A in the ordering process of any order $k \geq 2$ shows. On the positive side, we can show that if a permuton is absolutely continuous with respect to the Lebesgue measure and its Radon–Nikodym derivative is bounded then it is of positive age.

Of course, all the results and questions above would become more tractable if we could, for a any given flip process and any given initial graphon α and any time $t \geq 0$, explicitly compute the permuton $\Phi^t\alpha$. This task however involves solving a difficult system of differential equations and we were able to carry it out only for flip processes of order 2. Last, let us mention the question of the uniqueness of the rule. That is, suppose that we have rules $\mathcal{R}$ and $\mathcal{Q}$ whose respective trajectories are Φ and Ψ. If $\Phi = \Psi$, does it follow that $\mathcal{R} = \mathcal{Q}$?

Related work: Graph flip processes

Our work is similar to, and was in fact inspired by, the theory of flip processes for graphs, recently developed in [2]. Let us summarize that project. We write $\mathcal{H}_k$ for the family of graphs on vertex set $[k]$. A *rule $\mathcal{R}$ of a flip process of order k* is a stochastic matrix $\mathcal{R} \in [0,1]^{\mathcal{H}_k \times \mathcal{H}_k}$. For an *initial graph G_0* of order $n \geq k$, the *flip process with rule $\mathcal{R}$* is a discrete time process $(G_\ell)_{\ell=0}^{\infty}$ of graphs on the vertex set $[n]$ defined as follows. We get graph $G_{\ell+1}$ by sampling an ordered tuple $\mathbf{v} = (v_1, \cdots, v_k)$ of distinct vertices and sample a graph J from distribution $\mathcal{R}_{G_\ell[\mathbf{v}],*}$. We replace $G_\ell[\mathbf{v}]$ by J. The main result of [2] is that there exists trajectories $\Phi^{\cdot} W : [0, +\infty) \to \mathcal{W}_0$ with properties analogous to (2) and (3) such that if $T > 0$ is a constant and G_0 is an initial graph of large order n then with high probability for each $i \in \mathbb{N}$ with $i \leq Tn^2$ the graphon representation W_i of G_i is close to the trajectory started at the graphon representation W_0 of G_0 at time $\frac{i}{n^2}$ in the cut norm, that is, $\max\left\{\left\|W_i - \Phi^{\frac{i}{n^2}} W_0\right\|_{\square} : i \in (0, Tn^2] \cap \mathbb{N}\right\} = o(1)$.

There are substantial similarities between the proofs of the current permutation project and [2] in the overall strategy. In particular, the crucial construction of the trajectories is also based on an idea of a velocity operator (see Section 4). However, there are differences in technical execution of this overall strategy, which are mostly given by combinatorial differences between the cut norm distance for graphons and the rectangular distance for permutons, and of the underlying Banach spaces.[1] Also, families of natural and interesting flip processes seem to be quite different in both cases.[2]

3 Specific classes of flip processes

We give examples of several classes of flip processes. The purpose of this list is to show richness of scenarios that can be captured by flip processes and to hint to features that can be studied in the future. Many of these processes are counterparts to graph flip processes studied in [1].

[1]In the graphon case the Banach space of two-variable L^∞-functions, and in the permuton case the Banach space of signed measures as we describe in Section 4.

[2]Interesting graph flip processes are studied in [1].

The *ordering flip process of order k* is given by a rule in which $\mathcal{R}_{\psi,\mathrm{id}_k} = 1$ and $\mathcal{R}_{\psi,\rho} = 0$ for $\rho \neq \mathrm{id}_k$. All trajectories of this flip process converge to the identity permuton. While this is not the only process with this property, intuitively the speed of convergence to the identity permuton (which can be defined analogously to [2, Section 5.12]) is the fastest among all order-k flip processes. The *ignorant flip process* is a process in which the output distribution $\mathcal{R}_{\psi,*}$ does not depend on the input permutation ψ. An example of an ignorant process is the ordering process of any order. In [1, Section 4] 'ignorant graph processes' which have an analogous definition are studied. In the graph case, the trajectories can be explicitly described (see [1, Proposition 4.2]), however in the permutation setting this seems to be much more complicated. The *diagonal reversing flip process of order k* is a process designed to swap the order of the permutation in a particular, slow way. It outputs anti-diagonal for input being diagonal and in other cases does not change the permutation. For $k = 2$, it is a fatalistic flip process and the trajectories converge to the anti-diagonal permuton. For $k > 2$ the behaviour is more interesting. In the *complementing flip process* of order k, each input permutation is replaced by its reversal, that is $\mathcal{R}_{\psi,\overline{\psi}} = 1$ where for each $\psi \in \mathrm{Sym}(k)$, $\overline{\psi}$ is defined by $\overline{\psi}(i) := k + 1 - \psi(i)$. All the trajectories converge to the Lebesgue measure λ^2.

4 Proof of Theorem 1

We sketch the proof of Theorem 1, deferring other results announced in Section 2.3 for the full version of the paper. That is, for a flip process $\mathcal{R}$ of order k, we first need to construct the trajectories $\Phi : \mathfrak{P} \times [0, \infty) \to \mathfrak{P}$, and then we need to prove that a flip process started with π_0 stays with high probability within a thin sausage around $(\Phi^t \Gamma_{\pi_0})_{t \geq 0}$. Not suprisingly, our construction is tailored with respect to the latter property. More precisely, we work in the Banach space $\mathfrak{M}$ of finite signed Borel measures on $[0, 1]^2$ whose marginals are arbitrary multiples of the 1-dimensional Lebesgue measure, equipped with the distance $\mathrm{d}_\square$. We come up with a *velocity operator* $\nabla : \mathfrak{M} \to \mathfrak{M}$ whose defining formula (5) is explained below. We then require that for $\alpha \in \mathfrak{P}$ and $t \geq 0$ we have the following Banach-space valued equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi^t\alpha = \nabla\, \Phi^t\alpha \text{ (differential form), or equivalently } \Phi^t\alpha = \alpha + \int_0^t \nabla\, \Phi^\tau\alpha \,\mathrm{d}\tau \text{ (integral form).} \tag{4}$$

Using certain favorable properties of (4) it can be shown using the theory of Banach-space valued differential equations that it has a unique solution on the entire interval $[0, \infty)$.

We now turn to cooking up the defining formula for ∇. Recall that in Wormald's method of differential equation [6], one cooks up real-valued functions whose derivatives are idealizations of expected changes of tracked combinatorial parameters. Our idea is the same, except our derivatives are $\mathfrak{M}$-valued. That is, for $\alpha \in \mathfrak{M}$ we set

$$\nabla(\alpha)(Z) = \sum_{\omega \in \mathrm{Sym}(k)} \sum_{i \in [k]} \left(-t(\omega, Z, i; \alpha) + \sum_{\widetilde{\omega} \in \mathrm{Sym}(k)} \mathcal{R}_{\omega, \widetilde{\omega}} \cdot t(\omega \rightsquigarrow \widetilde{\omega}, Z, i; \alpha) \right), \tag{5}$$

for each Borel $Z \subset [0,1]^2$. Let us explain the motivation behind the quantities $t(\omega, Z, i; \alpha)$ and $t(\omega \rightsquigarrow \widetilde{\omega}, Z, i; \alpha)$, which we do under the assumption that α is a permuton.[3] The number $t(\omega, Z, i; \alpha)$ is the probability that when sampling[4] k points from α, we get a permutation ω and further the i-th leftmost point falls in Z. Likewise, $t(\omega \rightsquigarrow \widetilde{\omega}, Z, i; \alpha)$ is the probability that when sampling k points from α, we get a permutation ω and further, after swapping the y-coordinates of the sampled points from ω to $\tilde{\omega}$, the i-th leftmost point falls in Z. The corresponding formula (valid again for general $\alpha \in \mathfrak{M}$) for $t(\omega, Z, i; \alpha)$ is (writing $\alpha^{\otimes k}$ for the k-th power of α)

$$\begin{aligned} t(\omega, Z, i; \alpha) = & k! \cdot \alpha^{\otimes k} \Big(\Big\{ (x_1, y_1, \ldots, x_k, y_k) \in [0,1]^{2k} \ : \ x_1 < \ldots < x_k \text{ and } (x_i, y_i) \in Z \text{ and} \\ & \text{for all } \ell, j \in [k] \text{ we have } y_\ell < y_j \text{ if and only if } \omega(\ell) < \omega(j) \Big\} \Big), \end{aligned}$$

and a similar but more complicated formula can be written for $t(\omega \rightsquigarrow \widetilde{\omega}, Z, i; \alpha)$.

Having defined the trajectories Φ, we need to prove that a flip process started with a permutation π_0 stays within a thin sausage around $(\Phi^t \Gamma_{\pi_0})_{t \geq 0}$. Actually, we will only prove this for t small. Indeed, if we can prove that with high probability π_{tn} is close to $\Phi^t \Gamma_{\pi_0}$, then we can repeat this argument also starting with permutation $\widehat{\pi}_0 := \pi_{tn}$ and get that $\widehat{\pi}_{tn} = \pi_{2tn}$ is close to $\Phi^t \Gamma_{\widehat{\pi}_0} \approx \Phi^{2t} \Gamma_{\pi_0}$, and more generally, that for any constant $\ell \in \mathbb{N}$, with high probability $\pi_{\ell tn}$ is close to $\Phi^{\ell t} \Gamma_{\pi_0}$, as is needed. (Times between $(\ell - 1)tn$ and ℓtn can be dealt with easily as well.)

So, for $t > 0$ small we use Taylor series approximation of order 1, that is $\Phi^t \Gamma_{\pi_0} \approx \Gamma_{\pi_0} + t \cdot \nabla \Gamma_{\pi_0}$. Recalling that our distance is given by (1), we hence need to prove that with high probability for each $0 \leq x_1 \leq x_2 \leq 1, 0 \leq y_1 \leq y_2 \leq 1$ the quantity

$$|\{i \in [n] \cap [x_1 n, x_2 n] \ : \ \pi_{tn}(i) \in [y_1 n, y_2 n]\}| - |\{i \in [n] \cap [x_1 n, x_2 n] \ : \ \pi_0(i) \in [y_1 n, y_2 n]\}|$$

is approximately equal to $n \cdot t \cdot \nabla \Gamma_{\pi_0}([x_1, x_2] \times [x_2, y_2])$. This can be proved using concentration inequalities, and making use of the fact that $t(\omega, Z, i; \alpha)$ and $t(\omega \rightsquigarrow \widetilde{\omega}, Z, i; \alpha)$ were devised exactly to capture rates of deletions or insertions of points from a permutation in a single step from particular locations.

References

[1] P. Araújo, J. Hladký, E. K. Hng, and M. Šileikis. Prominent examples of flip processes. arXiv:2206.03884.

[3]Note that we ultimately, we need to solve (4) in the space of permutons only. However, as is usual in differential equations, to this end we need to work on an open domain. That is, at least some further elements of $\mathfrak{M}$ need to be dealt with.

[4]using the usual procedure of sampling points for a permuton, see Definition 3.2 in [4]

[2] F. Garbe, J. Hladký, M. Šileikis, and F. Skerman. From flip processes to dynamical systems on graphons. arxiv.org:2201.12272.

[3] F. Garbe, J. Hladký, G. Kun, and K. Pekárková. On pattern-avoiding permutons. arXiv.org:2208.12712.

[4] C. Hoppen, Y. Kohayakawa, C. G. T. de A. Moreira, B. Ráth, and R. M. Sampaio. Limits of permutation sequences. *J. Comb. Theory, Ser. B*, 103(1):93–113, 2013.

[5] D. Král' and O. Pikhurko. Quasirandom permutations are characterized by 4-point densities. *Geometric and Functional Analysis*, 23, 2013.

[6] N. C. Wormald. Differential equations for random processes and random graphs. *Ann. Appl. Probab.*, 5(4):1217–1235, 1995.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Twin-width of Planar Graphs; a Short Proof

(Extended abstract)

Petr Hliněný*

Abstract

The fascinating question of the maximum value of twin-width on planar graphs is nowadays not far from a final resolution; there is a lower bound of 7 coming from a construction by Král' and Lamaison [arXiv, September 2022], and an upper bound of 8 by Hliněný and Jedelský [arXiv, October 2022]. The upper bound (currently best) of 8, however, is rather complicated and involved. We give a short and simple self-contained proof that the twin-width of planar graphs is at most 11.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-082`

1 Introduction

The structural parameter twin-width was introduced in 2020 by Bonnet, Kim, Thomassé and Watrigant [2]. We consider it only for simple graphs (instead of general binary relational structures).

A *trigraph* is a simple graph G in which some edges are marked as *red*, and with respect to the red edges only, we naturally speak about *red neighbours* and *red degree* in G. However, when speaking about edges, neighbours and/or subgraphs without further specification, we count both ordinary and red edges together as one edge set. The edges of G which are not red are sometimes called (and depicted) black for distinction. For a pair of (possibly not adjacent) vertices $x_1, x_2 \in V(G)$, we define a *contraction* of the pair x_1, x_2 as the operation creating a trigraph G' which is the same as G except that x_1, x_2 are replaced with a new vertex x_0 (said to *stem from* x_1, x_2) such that:

*Faculty of Informatics, Masaryk University, Brno, Czech republic. E-mail: `hlineny@fi.muni.cz`

- the (full) neighbourhood of x_0 in G' (i.e., including the red neighbours), denoted by $N_{G'}(x_0)$, equals the union of the neighbourhoods $N_G(x_1)$ of x_1 and $N_G(x_2)$ of x_2 in G except x_1, x_2 themselves, that is, $N_{G'}(x_0) = (N_G(x_1) \cup N_G(x_2)) \setminus \{x_1, x_2\}$, and
- the red neighbours of x_0, denoted here by $N^r_{G'}(x_0)$, inherit all red neighbours of x_1 and of x_2 and add those in $N_G(x_1) \Delta N_G(x_2)$, that is, $N^r_{G'}(x_0) = \big(N^r_G(x_1) \cup N^r_G(x_2) \cup (N_G(x_1) \Delta N_G(x_2))\big) \setminus \{x_1, x_2\}$, where Δ denotes the symmetric set difference.

A *contraction sequence* of a trigraph G is a sequence of successive contractions turning G into a single vertex, and its *width* d is the maximum red degree of any vertex in any trigraph of the sequence. We also then say that it is a d-contraction sequence of G. The *twin-width* of a trigraph G is the minimum width over all possible contraction sequences of G. In other words, a graph has twin-width at most d, iff it admits a d-contraction sequence.

After the first implicit (and astronomical) upper bounds on the twin-width of planar graphs, e.g. [2], we have seen a stream of improving explicit bounds [1, 3, 4, 6], culminating with the current best upper bound of 8 by Hliněný and Jedelský [5]. This is complemented with a nearly matching lower bound of 7 by Král' and Lamaison [7], but the right maximum value (7 or 8?) is still an open question.

It comes without surprise that the gradually improving upper bounds have required stronger and more involved arguments, and the best ones are not easy to read for non-experts. In this paper, we take the opposite route; we give a slightly worse bound with a self-contained proof which is as short and simple as possible with the current knowledge:

Theorem 1. *The twin-width of any simple planar graph is at most* 11*.*

Due to page limits, some full proofs are left for the preprint version arXiv:2302.08938.

2 Layered Skeletal Trigraphs

We start with the key concept of our proof – of a "splendid layered skeletal trigraph".

We use standard terminology of graph theory, and assume every graph to be simple (without loops and multiple edges). A *BFS tree* of a graph G is a spanning tree defined by a run of the breadth-first-search algorithm on G.

For a (tri)graph G, an ordered partition $\mathcal{L} = (L_0, L_1, \ldots)$ of $V(G)$ is called a *layering of* G if, for every edge $\{v, w\}$ of G with $v \in L_i$ and $w \in L_j$, we have $|i - j| \le 1$. For example, every BFS tree $T \subseteq G$ with the root r naturally defines a layering; $L_0 = \{r\}$, and L_i for $i > 0$ consisting of all vertices of G at graph distance i from r.

If $T \subseteq G$ is a rooted tree (e.g., a BFS tree), a path $P \subseteq G$ is called T*-vertical* if $P \subseteq T$ is a subpath of some leaf-to-root path of T.

Definition 2 (Skeletal trigraph)**.** Let H be a trigraph and $S \subseteq H$ a 2-connected planar subgraph such that all edges of H induced by $V(S)$ are black (note; including the edges not in $E(S)$). Fix a plane embedding of S, and call S a *plane skeleton of* H. Further, consider a *face assignment* of H in S in which every connected component H_0 of $H - V(S)$

is assigned to some face ϕ of S, such that all neighbours of H_0 in $V(S)$ belong to ϕ. Denote by U_ϕ the union of the vertex sets of all components assigned to ϕ in this assignment.

If H and S satisfy the previous conditions for some face assignment, we call (H, S) a *skeletal trigraph*, and if $\mathcal{L}$ is a layering of H, then $(H, S, \mathcal{L})$ is a *layered skeletal trigraph.*

Definition 3 (Splendid layered skeletal trigraph)**.** Consider a layered skeletal trigraph $(H, S, \mathcal{L})$ as in Definition 2, and a face ϕ of S. We say that ϕ is *empty* if $U_\phi = \emptyset$ (i.e., if no connected component of $H - V(S)$ is assigned to ϕ), that ϕ is *reduced* if $|U_\phi \cap L_i| \le 1$ holds for every layer $L_i \in \mathcal{L}$, and that ϕ is *rich* if $|U_\phi \cap L_i| \le 3$ holds for every $L_i \in \mathcal{L}$.

A layered skeletal trigraph $(H, S, \mathcal{L})$ is *splendid* if either $S = \emptyset$ and $|V(H) \cap L_i| \le 4$ holds for all $L_i \in \mathcal{L}$, or $S \ne \emptyset$ and the following conditions are satisfied:

a) At most one face of the plane skeleton S is rich, and all other faces of S are empty or reduced. Every empty face of S is a triangle.
b) There exists a BFS tree $T \subseteq S$ of the skeleton S such that:
   - The layering defined by T in S is equal to the restriction of $\mathcal{L}$ to $V(S)$.
   - For every non-empty face ϕ of S, bounded by a cycle $C \subseteq S$, there exists an edge $e \in E(C)$ such that $C - e$ is the union of two T-vertical paths intersecting in one vertex $u \in V(C)$. Note that such u must be unique, and we call u the *sink of* ϕ.
c) For every non-empty face ϕ of S, and u the sink of ϕ; if $u \in L_i \in \mathcal{L}$, then all vertices of $U_\phi \cup V(C - u)$ belong to $L_{i+1} \cup L_{i+2} \cup \ldots$, and there is a black edge in H (but *no* red edge) from u to each vertex of $U_\phi \cap L_{i+1}$.
d) Assume ϕ is a rich face of S bounded by C. For every i such that $L_i \in \mathcal{L}$, every vertex v in $X := (U_\phi \cup V(C)) \cap L_i$ has in H at most 3 red edges into other vertices of X and at most 4 red edges into $U_\phi \cap (L_{i-1} \cup L_{i+1})$ (note; no $V(C)$ in the latter expression). Moreover, if $|U_\phi \cap L_{i+1}| > 1$, then $v \in X$ has at most 2 red edges into $U_\phi \cap L_{i-1}$.

Definition 3 is illustrated, with comments, in Figure 1.

The core of the paper is in the following two claims which follow directly from Definition 3. While the first one is easy and its proof is skipped here, a proof of the second one is sketched in the next section.

Lemma 4.* *Every splendid layered skeletal trigraph has maximum red degree at most* 11*.*

Lemma 5. *Every splendid layered skeletal trigraph admits an* 11*-contraction sequence.*

We now show how the claim implies our main result.

Proof of Theorem 1. Given a planar graph G, we fix any plane embedding of G. We construct a plane triangulation G^+ from G by adding new vertices to every face of G and connecting them to vertices of this face. Then G^+ is 2-connected. Choosing an arbitrary BFS tree of G^+, we take the layering $\mathcal{L} = (L_0, L_1, \ldots)$ of G^+ naturally defined by T. Then, trivially, $(G^+, G^+, \mathcal{L})$ is a splendid layered skeletal trigraph, and hence G^+ admits an 11-contraction sequence by Lemma 5. Restricting this sequence only to the contractions of pairs from $V(G)$ we, again trivially, obtain an 11-contraction sequence of G. □

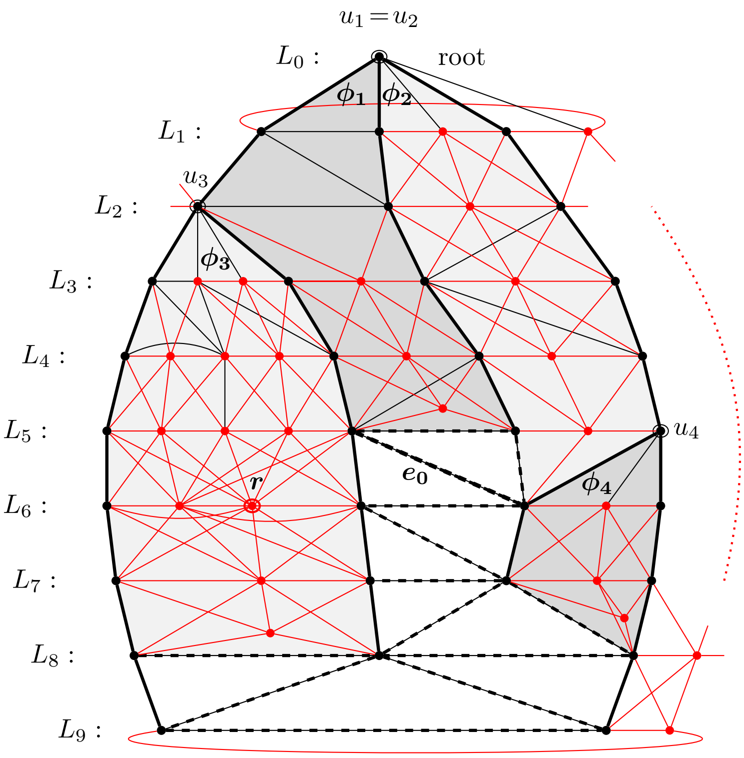


Figure 1: A picture of a splendid layered skeletal trigraph $(H, S, \mathcal{L})$, in which the skeleton S is depicted with black vertices and thick black edges such that the associated BFS tree $T \subseteq S$ is drawn with thick solid edges and the edges of $E(S) \setminus E(T)$ are thick dashed. T has its root at the top and its (ten) BFS layers are organized horizontally in the picture. There are four bounded non-empty faces in S, denoted by $\phi_1, \phi_2, \phi_3, \phi_4$ (with corresponding sinks u_1, u_2, u_3, u_4), and emphasized with gray shade. The unbounded face of S is also non-empty, but it is only sketched in the picture. There is one rich face in (H, S), namely ϕ_3, and it contains a red vertex r (emphasized with a circle around) that achieves the maximum red degree 11 allowed by Definition 3.

3 Proof of Lemma 5, a Sketch

Our proof starts with an auxiliary claim whose straightforward proof is skipped here.

Lemma 6.* *Let G be a 2-connected plane graph, and $T \subseteq G$ a BFS tree of G. Assume T that has at least 3 leaves, and that for every facial cycle C of G, we have $|E(C) \setminus E(T)| = 1$ or C is a triangle. Then there exists an edge $e \in E(G) \setminus E(T)$ such that, for the unique cycle $D_e \subseteq T + e$, one of the two faces of D_e contains (in its strict interior) precisely one leaf of T and not the root of T.*

For a proof of Lemma 5, consider a splendid layered skeletal trigraph $(H, S, \mathcal{L})$. For start, the maximum red degree of H is at most 11 by Lemma 4. For the rest of a sought 11-contraction sequence of H, we proceed by induction on $|V(H)| + |V(S)|$.

If the skeleton is empty $S = \emptyset$, then we pick the highest index i such that $V(H) \cap L_i \neq \emptyset$ and straightforwardly contract from layer i down by induction. If $S \neq \emptyset$, all faces of S are reduced (or empty), and the BFS tree $T \subseteq S$ from Definition 3.b has at most 2 leaves, we get that T consists of at most two T-vertical paths, and that S has at most two non-empty faces by Definition 3. b. Since the two faces are reduced, every layer of $\mathcal{L}$ contains at most $2+2=4$ vertices. So, $(H, S' = \emptyset, \mathcal{L})$ is also a splendid layered skeletal trigraph (note; no contraction happend) and we continue as before again by induction.

For all other cases, with a nonempty skeleton $S \neq \emptyset$, we branch as follows.

Case 1. The skeleton S has all faces empty. Then $S = H$ since S is a plane triangulation by Definition 3.a. Considering the BFS tree $T \subseteq S$ from Definition 3.b, we apply Lemma 6 and get e and cycle $D_e \subseteq T + e \subseteq H$. Let Q be the maximal T-vertical path starting in x and not hitting D_e. We set $S' := S - V(Q)$, using that the layered skeletal trigraph $(H, S', \mathcal{L})$ is splendid again, and we finish by induction.

Case 2. The skeleton S has a face ϕ which is neither empty nor reduced. Then ϕ is a rich face, and let j be the largest index such that $|U_\phi \cap L_j| > 1$ for $L_j \in \mathcal{L}$. We contract any two vertices of $U_\phi \cap L_j$ in H, creating a layered skeletal trigraph $(H', S, \mathcal{L}')$. For an illustration, see the face $\phi = \phi_3$ in Figure 1 in which the trigraph resulted by a contraction of two vertices from $U_{\phi_3} \cap L_6$ into the emphasized vertex r. We show that $(H', S, \mathcal{L}')$ conforms to Definition 3, and then apply induction.

Case 3. The skeleton S has all faces reduced (and some non-empty). As in Case 1, we apply Lemma 6 and get e and cycle $D_e \subseteq T + e \subseteq S$, and the path $Q \subseteq S$ in the interior of D_e. The interior of D_e contains at most two non-empty faces ϕ_1 and ϕ_2 of S. The considered case can be illustrated in Figure 1 (ignoring for now that the face ϕ_3 is not reduced) with the edge $e = e_0$. In general, there can be more than one empty faces of the skeleton S enclosed by D_{e_0}. We again set $S' := S - V(Q)$ and consider the layered skeletal trigraph $(H, S', \mathcal{L})$ with the (new) non-empty face ϕ bounded by D_e, which can be shown rich. Consequently, $(H, S', \mathcal{L})$ conforms to Definition 3, and we again finish by induction with it.

The whole proof, modulo straightforwad details regarding Definition 3, is now done. $\square$

4 Conclusion

We have provided a short self-contained proof of Theorem 1. While the proved bound is not the best currently possible, the proof given here is way simpler than those in [4, 5]. While sacrificing a bit of simplicity of the given proof, we can also give a better upper bound of 9 (thus matching [4]), but we are so far not sure whether a similarly simplified proof can be given for the upper bound of 8 as in [5].

References

[1] Michael A. Bekos, Giordano Da Lozzo, Petr Hliněný, and Michael Kaufmann. Graph product structure for h-framed graphs. *CoRR*, abs/2204.11495v1, 2022. `arXiv:2204.11495v1`.

[2] Édouard Bonnet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant. Twin-width I: tractable FO model checking. *J. ACM*, 69(1):3:1–3:46, 2022.

[3] Édouard Bonnet, O-joung Kwon, and David R. Wood. Reduced bandwidth: a qualitative strengthening of twin-width in minor-closed classes (and beyond). *CoRR*, abs/2202.11858, 2022. `arXiv:2202.11858`.

[4] Petr Hliněný. Twin-width of planar graphs is at most 9, and at most 6 when bipartite planar. *CoRR*, abs/2205.05378, 2022. `arXiv:2205.05378`.

[5] Petr Hliněný and Jan Jedelský. Twin-width of planar graphs is at most 8, and at most 6 when bipartite planar. *CoRR*, abs/2210.08620, 2022. Accepted to ICALP 2023. `arXiv:2210.08620`.

[6] Hugo Jacob and Marcin Pilipczuk. Bounding twin-width for bounded-treewidth graphs, planar graphs, and bipartite graphs. In *WG*, volume 13453 of *Lecture Notes in Computer Science*, pages 287–299. Springer, 2022.

[7] Daniel Král and Ander Lamaison. Planar graph with twin-width seven. *CoRR*, abs/2209.11537, 2022. `arXiv:2209.11537`.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Stack and Queue Numbers of Graphs Revisited

Petr Hliněný* Adam Straka†

Abstract

A long-standing question of the mutual relation between the stack and queue numbers of a graph, explicitly emphasized by Dujmović and Wood in 2005, was "half-answered" by Dujmović, Eppstein, Hickingbotham, Morin and Wood in 2022; they proved the existence of a graph family with the queue number at most 4 but unbounded stack number. We give an alternative very short, and still elementary, proof of the same fact.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-083`

1 Introduction

The graph parameters called stack and queue numbers relate to *linear layouts* (i.e, linear vertex orderings, usually of additional "nice" properties) of graphs, and have found numerous applications in theoretical computer science since then. The parameters were formally introduced by Heath, Leighton, and Rosenberg in [6,7], and their implicit question of whether the stack number of a graph is bounded in terms of its queue number, or vice versa, was subsequently emphasized by Dujmović and Wood in [3]. Quite recently, in 2022, Dujmović, Eppstein, Hickingbotham, Morin and Wood gave in [2] a negative answer to one half of the question; they proved the existence of a graph family with the queue number at most 4 but unbounded stack number (while it remains an open problem whether there exists a family of bounded stack number and unbounded queue number).

We give the basic definitions. Consider a graph G and a strict linear order $\prec$ on its vertex set $V(G)$. Two edges $xx', yy' \in E(G)$ with $x \prec x'$ and $y \prec y'$ are said to $\prec$-*cross* if

*Faculty of Informatics, Masaryk University, Brno, Czech republic. E-mail: `hlineny@fi.muni.cz`
†Faculty of Informatics, Masaryk University, Brno, Czech republic. E-mail: `adamstraka@mail.muni.cz`

$x \prec y \prec x' \prec y'$ or $y \prec x \prec y' \prec x'$, and to *$\prec$-nest* if $x \prec y \prec y' \prec x'$ or $y \prec x \prec x' \prec y'$. See Figure 1. The *stack number* $\mathrm{sn}(G)$ (*queue number* $\mathrm{qn}(G)$) of a graph G is the minimum integer k such that there exist a linear order $\prec$ of $V(G)$ and a colouring of the edges of G by k colours such that no two edges of the same colour $\prec$-cross ($\prec$-nest, resp.). The corresponding order $\prec$ together with the colouring is called a k-stack (k-queue) layout of G.

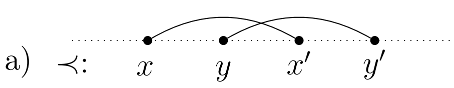


Figure 1: Edges xx' and yy' that (a) $\prec$-cross, and (b) $\prec$-nest.

In fact, a notion equal (modulo a negligible technical detail) to the stack number was known long before as the *book thickness* (or page number), see Persinger [8] and Atneosen [1].

To state the main result of [2], we define the following special graph H_n: the vertex set is $V(H_n) = \{1, \ldots, n\}^2$, and $uv \in E(H_n)$ where $u = [a, b] \in V(H_n)$ and $v = [c, d] \in V(H_n)$, if and only if $|a - c| + |b - d| = 1$ or $a - c = b - d \in \{-1, 1\}$. Note that H_n is the plane dual of the hexagonal ("honeycomb") grid, and see an illustration in Figure 2.

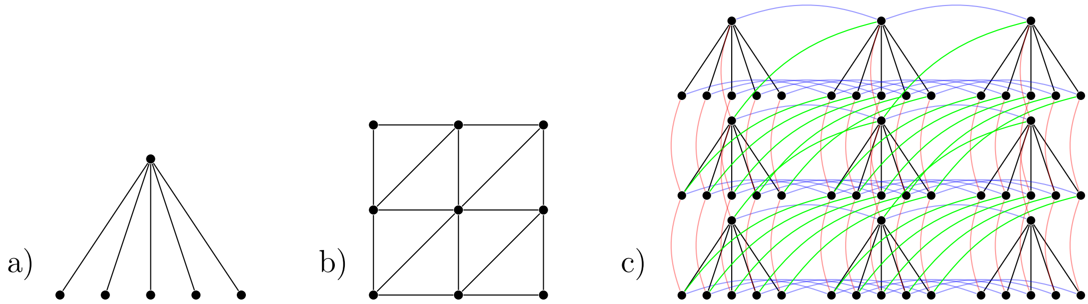


Figure 2: (a) The star S_5, (b) the graph H_3, and (c) their Cartesian product $S_5 \square H_3$. The four edge colours illustrate a queue layout for $S_5 \square H_3$.

Recall that S_n is the star with n leaves, and that $G_1 \square G_2$ denotes the Cartesian product of two graphs G_1 and G_2. Dujmović et al. [2] showed that, for all integers $a, n > 0$ and the Cartesian product $G = S_a \square H_n$, we have $\mathrm{qn}(G) \leq 4$. In fact, they noted that every H_n admits a so-called *strict* 3-queue layout, which "adds up" with a trivial 1-queue layout of S_a over Cartesian product by Wood [12]. Their main result reads:

Theorem 1 (Dujmović et al. [2])**.** *For every integer s, and for $a, n > 0$ which are sufficiently large with respect to s, the Cartesian product $G := S_a \square H_n$ is of stack number at least s.*

Our contribution is to give a very short simplified proof of Theorem 1 (based in parts on the ideas from [2], but also eliminating some rather long fragments of the former proof).

2 Proof of Theorem 1

We will use some classical results, the first two of which are truly folklore.

Proposition 2 (Ramsey [9]). *For all integers $r, s > 0$ there exists $R = R(r, s)$ such that for any assignment of two colours read and blue to the edges of the complete graph K_R, there is a red clique on r vertices or a blue clique on s vertices in it.*

Proposition 3 (Erdős–Szekeres [4]). *For given integers $r, s > 0$, any sequence of distinct elements of a linearly ordered set of length more than $r \cdot s$ contains an increasing subsequence of length $s + 1$ or a decreasing subsequence of length $r + 1$.*

Proposition 4 (Gale [5]). *Consider a dual hexagonal grid H_n as above. For any assignment of two colours to the vertices of H_n, there exists a monochromatic path on n vertices.*

Consider for the rest any fixed stack layout of the graph G of Theorem 1, with the linear order $\prec$ on the vertex set $V(G)$. Recall that $V(G) = \{(u, p) : u \in V(S_a), p \in V(H_n)\}$.

Lemma 5. *Let L be the set of leaves of S_a, and let $b = a^{-m}$ where $m = 2^{n^2-1}$. There is a subsequence $(u_1, \ldots, u_b)$ in the set L of length b such that for each $p \in V(H_n)$, either $(u_1, p) \prec (u_2, p) \prec \ldots \prec (u_b, p)$, or $(u_1, p) \succ (u_2, p) \succ \ldots \succ (u_b, p)$.*

Proof. Let $V(H_n) = \{p_1, \ldots, p_{n^2}\}$ be the vertices of H_n. Start with the permutation $\sigma_1 = (u_{i[1,1]}, \ldots, u_{i[1,a_1=a]})$ of L such that $(u_{i[1,1]}, p_1) \prec \ldots \prec (u_{i[1,a_1]}, p_1)$. By Proposition 3, for each $j \in \{2, \ldots, n^2\}$, the sequence σ_{j-1} contains a subsequence $\sigma_j = (u_{i[j,1]}, \ldots, u_{i[j,a_j]})$ such that $a_j \geq \sqrt{a_{j-1}}$, and $(u_{i[j,1]}, p_j) \prec \ldots \prec (u_{i[j,a_j]}, p_j)$ or $(u_{i[j,1]}, p_j) \succ \ldots \succ (u_{i[j,a_j]}, p_j)$. By simple calculus, we get $a_{n^2} \geq a_1^m = b$ which is the desired outcome. □

Let $S_b \subseteq S_a$ be the (specific) substar of S_a defined by the subset of leaves $\{u_1, \ldots, u_b\}$ (of Lemma 5). Colour every vertex $p \in V(H_n)$ red if $(u_1, p) \prec \ldots \prec (u_b, p)$, and colour p blue otherwise. From this and Proposition 4, we immediately obtain:

Corollary 6. *There is a subgraph $Q \subseteq H_n$, being a path on n vertices, such that, without loss of generality, $(u_1, q) \prec \ldots \prec (u_b, q)$ holds for every vertex $q \in V(Q)$.* □

Define $X \subseteq G$ to be the subgraph induced on the vertex set $V(S_b) \times V(Q)$, i.e., $X = S_b \square Q$, and denote by $\mathcal{R}$ the set of paths $R_i \subseteq X$ induced on $\{u_i\} \times V(Q)$ for $i = 1, \ldots, b$. We extend $\prec$ to a partial order on $\mathcal{R}$ as follows; for $R_i, R_j \in \mathcal{R}$, we have $R_i \prec R_j$, if and only if $u \prec w$ for all $u \in V(R_i)$ and $w \in V(R_j)$. We say that R_i and R_j are $\prec$-*separated* if $R_i \prec R_j$ or $R_i \succ R_j$, and that R_i and R_j are $\prec$-*crossing* if there exist edges $e \in E(R_i)$ and $f \in E(R_j)$ such that e, f $\prec$-cross. The following is simple but crucial:

Lemma 7. *Every two distinct paths $R_i, R_j \in \mathcal{R}$ are either $\prec$-crossing, or $\prec$-separated.*

Proof. Assume the contrary; up to symmetry meaning that all edges of R_i are nested in some edge $e_2 = \{(u_j, q), (u_j, q')\} \in E(R_j)$. Then, in particular, $e_1 = \{(u_i, q), (u_i, q')\} \in E(R_i)$ is nested in e_2, and so $(u_j, q) \prec (u_i, q)$ and $(u_j, q') \succ (u_i, q')$. This contradicts Corollary 6. □

Corollary 8. *For all integers c, d and n, and for $b = |\mathcal{R}|$ sufficiently large with respect to c, d, we have that $\mathcal{R}$ contains at least c pairwise $\prec$-separated or d pairwise $\prec$-crossing paths.*

Proof. Imagine a pair of paths $\{R_i, R_j\} \subseteq \mathcal{R}$ coloured red if R_i, R_j are $\prec$-crossing, and blue if they are $\prec$-separated. With respect to Lemma 7, we apply Proposition 2 with $b \geq R(c, d)$. ☐

We finish as follows.

Proof of Theorem 1. Respecting the above definition of the set of paths $\mathcal{R}$ in G, we branch into the two cases determined by Corollary 8.

Case I. There are c pairwise $\prec$-separated paths in $\mathcal{R}$.

Without loss of generality, let these paths be $R_1 \prec \ldots \prec R_c$. For the root t of S_b, label the n vertices of the set $\{t\} \times V(Q) \subseteq V(X)$ by $t_1 \prec \ldots \prec t_n$. There are two subcases.

- $R_{\lfloor c/2 \rfloor} \prec t_{\lceil n/2 \rceil}$. For each $i = 1, \ldots, \min(\lfloor c/2 \rfloor, \lceil n/2 \rceil)$, pick an edge of X from $t_{\lceil n/2 \rceil + i - 1}$ to $V(R_i)$ (which exist since R_i hits every copy of S_b in X by the definition). We have got $\min(\lfloor c/2 \rfloor, \lceil n/2 \rceil)$ edges in X that pairwise $\prec$-cross, as in Figure 3.

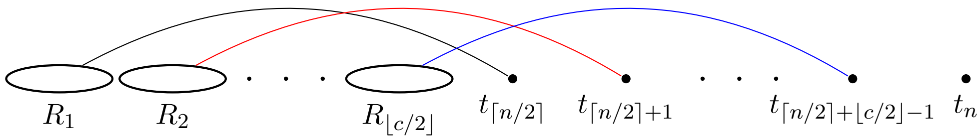


Figure 3: Case I, where $R_{\lfloor c/2 \rfloor} \prec t_{\lceil n/2 \rceil}$ and $\lceil n/2 \rceil > \lfloor c/2 \rfloor$.

- $t_{\lceil n/2 \rceil} \prec R_{\lfloor c/2 \rfloor + 1}$ (note that $t_{\lceil n/2 \rceil}$ may be "$\prec$-nested" in $R_{\lfloor c/2 \rfloor}$). This is symmetric to the previous, and we get $\min(\lceil c/2 \rceil, \lceil n/2 \rceil)$ pairwise $\prec$-crossing edges in X between vertices of $R_{\lfloor c/2 \rfloor + 1}, \ldots, R_c$ and $s_1, \ldots, s_{\lceil n/2 \rceil}$.

Case II. There are d pairwise $\prec$-crossing paths in $\mathcal{R}$.

Pick any path R_0 out of these d paths. In $Z := \bigcup_{R \in \mathcal{R}, R \neq R_0} E(R)$ there are at least $d - 1$ edges which $\prec$-cross some edge of R_0, and so at least $(d-1)/n$ of them cross the same edge $e \in E(R_0)$. Having $e = u_1 u_2$, $u_1 \prec u_2$, and $f = v_1 v_2 \in E(X)$ such that e and f $\prec$-cross, we say that v_1 is the *inside* vertex of f if $u_1 \prec v_1 \prec u_2$, and then v_2 is the *outside* vertex. By the pigeonhole principle, there is a set $Z' \subseteq Z$ of $d' = |Z'| \geq (d-1)/n^2$ edges $\prec$-crossing e such that their inside vertices belong to the same copy of S_b in X.

The outside vertices of the edges of Z' belong to at most two other copies of S_b in X (determined by a neighbourhood in the path Q), and each is before of after e in $\prec$. By the pigeonhole principle again, and without loss of generality, there is a set $Z'' \subseteq Z'$ of size $|Z''| \geq \frac{1}{2} \cdot \frac{1}{2} d' = d'/4$, such that also the outside vertices of the edges of Z'' belong to the same copy of S_b in X, and they all lie after e in $\prec$. See Figure 4. Moreover, by Corollary 6 (the ordering claimed therein), the edges in Z'' must pairwise $\prec$-cross.

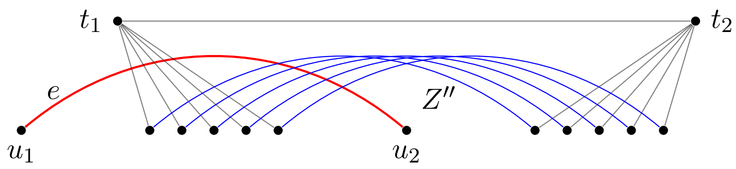


Figure 4: Case II, with emphasized edge e, blue parwise-crossing edges of Z'', and t_1, t_2 being two copies of the root of S_b.

To finish the proof, we set $n = 2s$ and $a = R(2s, 4n^2 s + 1)^m$ where $m = 2^{n^2-1}$. Then in Lemma 5 we get $b = R(2s, 4n^2 s + 1)$, and in Corollary 8 we have $c = 2s$ and $d = 4n^2 s + 1$. In Case I, we then obtain at least $\min(\lfloor c/2 \rfloor, \lceil n/2 \rceil) = s$ edges of $X \subseteq G$ that pairwise $\prec$-cross. In Case II, it is at least $d'/4 = (d-1)/(4n^2) = s$ such pairwise $\prec$-crossing edges, too. Edges that pairwise $\prec$-cross obviously must receive distinct colours. A valid stack layout based on $\prec$ hence needs at least s colours, and since $\prec$ has been arbitrary for the graph G, we finally conclude that $\mathrm{sn}(G) \geq s$. □

3 Conclusion

We have provided a short elementary proof of Theorem 1. Although the original proof in [2] is not very long or difficult, by carefully rearranging the arguments we have succeeded in eliminating some technical steps of the proof in [2] and, in particular, resolved the case of pairwise crossing paths in a direct short way. Briefly explaining, our proof skips initial technical parts of [2] preceding the use of Proposition 3 (Erdős–Szekeres) and readily applies Proposition 3 and Proposition 4 in a way similar to [2], and then it concludes by Proposition 2 (Ramsey) in which both outcomes straightforwardly lead to a large set of pairwise crossing edges, thus avoiding other technical steps needed in [2] mainly at the end of the arguments.

The presented proof is based on the Bachelor's thesis of the second author [10, 11].

References

[1] Gail Adele Atneosen. *On the embeddability of compacta in n-books: intrinsic and extrinsic properties.* Michigan State University, 1968.

[2] Vida Dujmović, David Eppstein, Robert Hickingbotham, Pat Morin, and David R Wood. Stack-number is not bounded by queue-number. *Combinatorica*, 42(2):151–164, 2022.

[3] Vida Dujmović and David R Wood. Stacks, queues and tracks: Layouts of graph subdivisions. *Discrete Mathematics and Theoretical Computer Science*, 7:155–202, 2005.

[4] Pál Erdös and G. Szekeres. A combinatorial problem in geometry. *Compositio Mathematica*, 2:463–470, 1935.

[5] David Gale. The game of hex and the Brouwer fixed-point theorem. *The American Mathematical Monthly*, 86(10):818–827, 1979.

[6] Lenwood S. Heath, Frank Thomson Leighton, and Arnold L. Rosenberg. Comparing queues and stacks as machines for laying out graphs. *SIAM Journal on Discrete Mathematics*, 5(3):398–412, 1992.

[7] Lenwood S. Heath and Arnold L. Rosenberg. Laying out graphs using queues. *SIAM Journal on Computing*, 21(5):927–958, 1992.

[8] CA Persinger. Subsets of n-books in E3. *Pacific Journal of Mathematics*, 18(1):169–173, 1966.

[9] F. P. Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, s2-30(1):264–286, 1930.

[10] Adam Straka. Stack number and queue number of graphs, 2023. Bachelor's thesis, Masaryk University, Faculty of Informatics. URL: `https://is.muni.cz/th/m5pwj/`.

[11] Adam Straka. Stack number and queue number of graphs, 2023. `arXiv:2305.09700`.

[12] David R Wood. Queue layouts of graph products and powers. *Discrete Mathematics & Theoretical Computer Science*, 7, 2005.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Crux, space constraints and subdivisions

(Extended abstract)

Seonghyuk Im [*] Jaehoon Kim[†] Younjin Kim [‡] Hong Liu [§]

Abstract

The existence of H-subdivisions within a graph G has deep connections with topological, structural and extremal properties of G. One prominent example of such a connection, due to Bollobás and Thomason and independently Komlós and Szemerédi, asserts that the average degree of G being d ensures a $K_{\Omega(\sqrt{d})}$-subdivision in G. Although this square-root bound is the best possible, various results showed that much larger clique subdivisions can be found in a graph for many natural classes. We investigate the connection between crux, a notion capturing the essential order of a graph, and the existence of large clique subdivisions.

Our main result gives an asymptotically optimal bound on the size of a largest clique subdivision in a generic graph G, which is determined by both its average degree and its crux size. As corollaries, we obtain

- a characterisation of extremal graphs for which the square-root bound above is tight: they are essentially a disjoint union of graphs each of which has the crux size linear in d;
- a unifying approach to find a clique subdivision of almost optimal size in graphs which do not contain a fixed bipartite graph as a subgraph;

[*]Department of Mathematical Sciences, KAIST, South Korea & Extremal Combinatorics and Probability Group(ECOPRO), Institute for Basic Science(IBS), South Korea. Email: seonghyuk@kaist.ac.kr.
Supported by Institute for Basic Science (IBS-R029-C4), the POSCO Science Fellowship of POSCO TJ Park Foundation, and the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIT) No. RS-2023-00210430.

[†]Department of Mathematical Sciences, KAIST, South Korea. Email: jaehoon.kim@kaist.ac.kr.
Supported by the POSCO Science Fellowship of POSCO TJ Park Foundation and the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIT) No. RS-2023-00210430.

[‡]Extremal Combinatorics and Probability Group (ECOPRO), Institute for Basic Science(IBS), South Korea. Email: mathyounjinkim@gmail.com.
Supported by Institute for Basic Science (IBS-R029-C4).

[§]Extremal Combinatorics and Probability Group (ECOPRO), Institute for Basic Science (IBS), Daejeon, South Korea, Email: hongliu@ibs.re.kr.
Supported by Institute for Basic Science (IBS-R029-C4).

- and that the clique subdivision size in random graphs $G(n,p)$ witnesses a dichotomy: when $p = \omega(n^{-1/2})$, the barrier is the space, while when $p = o(n^{-1/2})$, the bottleneck is the density.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-084`

1 Introduction

For a graph H, a subdivision of H (or an H-subdivision) is a graph obtained by replacing each edge of H by internally vertex-disjoint paths. Studies on the existence of certain subdivisions in a given graph G provide deep understandings on various aspects of G. For example, the cornerstone theorem of Kuratowski [12] in 1930 completely characterises planar graphs by proving that graphs are planar if and only if they do not contain a subdivision of either K_5, the complete graph on five vertices, or $K_{3,3}$, the complete bipartite graph with three vertices in each class.

What conditions on graphs G guarantee an H-subdivision in them? A fundamental result of Mader [14] in 1967 states that a large enough average degree always provides a desired subdivision. Namely, for every $t \in \mathbb{N}$, there exists a smallest integer $f(t)$ such that every graph G with average degree at least $f(t)$ contains a subdivision of K_t. He further conjectured that $f(t) = O(t^2)$. This conjecture was verified in the 90s by Bollobás and Thomason [3] and independently by Komlós and Szemerédi [10]. In fact, $f(t) = \Theta(t^2)$; the lower bound was observed by Jung [8] in 1970: consider the n-vertex graph which is a disjoint union of $n/5t^2$ copies of $K_{\frac{t^2}{10},\frac{t^2}{10}}$. A clique subdivision must be embedded in a connected graph; this example, though may have arbitrary large order n, is essentially the same as one copy of $K_{\frac{t^2}{10},\frac{t^2}{10}}$, which does not contain a K_t-subdivision. Indeed, at least $\binom{t/2}{2}$ many edges are subdivided in any K_t-subdivision in $K_{\frac{t^2}{10},\frac{t^2}{10}}$, which would require around $t^2/8 > t^2/10$ vertices on one side. In other words, apart from the obvious "degree constraint" from the average degree, there is also some "space constraint" forbidding a K_t-subdivision.

From the extremal example above, it is then natural to wonder if G does not structurally look like $K_{\frac{t^2}{10},\frac{t^2}{10}}$, can we find a larger clique subdivision? Indeed, Mader [15] conjectured that every C_4-free graph with average degree d contains a subdivision of $K_{\Omega(d)}$ and recently it was resolved by Liu and Montgomery [13]. Furthermore, they proved that for every $t \geq s \geq 2$, there exists a constant $c = c(s,t)$ such that if G is $K_{s,t}$-free and has average degree d, then G has a subdivision of a clique of order $cd^{s/2(s-1)}$.

Note that a C_4-free graph with average degree at least d must have at least $\Omega(d^2)$ vertices as the maximum number of edges of an n-vertex C_4-free graph is $O(n^{3/2})$, hence providing enough space to put a $K_{\Omega(d)}$-subdivision with $O(d^2)$-vertices. Similarly, the number $d^{s/2(s-1)}$ also matches with the conjectured extremal number of $K_{s,t}$. Thus, all these H-free conditions relax the "space constraints". Hence, this suggests that 'the essential order' of the graph G, rather than structural F-freeness, is an important factor for the size of the largest clique subdivision. Indeed, Liu and Montgomery [13] conjectured that every graph G with its 'essential order' n and average degree d contains a K_t subdivision with $t = \Omega(\min\{d, \frac{n}{\log n}\})$.

Such a notion of 'essential order', called crux, was recently introduced by Haslegrave, Hu, Kim, Liu, Luan and Wang [5]. We write $d(G)$ for the average degree of G.

Definition 1.1 (Crux)**.** Let $\alpha > 0$ and G be a graph. A subgraph $H \subseteq G$ is an *α-crux* if $d(H) \geq \alpha \cdot d(G)$. Let $c_\alpha(G)$ be the order of a smallest α-crux in G, that is:

$$c_\alpha(G) = \min\{n : \exists H \subseteq G \text{ s.t. } |H| = n, d(H) \geq \alpha d(G)\}.$$

We will write simply $c(G)$ when $\alpha = 1/100$; the choice of $1/100$ here is not special and can be replaced with any small number. Roughly speaking, the crux of a graph is large when the edges are relatively uniformly distributed.

Our main result reads as follows. It implies in particular that the space constraints, measured by the crux size, is a deciding factor for the size of largest clique subdivision in a graph.

Theorem 1.2. *There exists an absolute constant $\beta > 0$ such that the following is true. Let G be a graph with $d(G) = d$. Then G contains a $K_{\beta t/(\log\log t)^6}$-subdivision where*

$$t = \min\left\{d, \sqrt{\frac{c(G)}{\log c(G)}}\right\}.$$

Theorem 1.2 asymptotically confirms a conjecture of Liu and Montgomery [13]. The bound above is optimal up to the multiplicative $(\log\log t)^6$ factor: the d-blowup of a d-vertex $O(1)$-regular expander satisfies $c(G) = \Theta(d^2)$ and the largest clique subdivision has order $d/\sqrt{\log d}$ (see [13] for more details).

2 Applications

2.1 Characterisation of extremal graphs

The first consequence of our main result is a structural characterisation of extremal graphs G having the smallest possible clique subdivision size $\Theta(\sqrt{d(G)})$, showing that the *only obstruction* to get a larger than usual clique subdivision is a small crux. In other words, if $c(G) = \omega(d)$, then one can embed a K_t-subdivision with $t = \omega(\sqrt{d})$. Theorem 1.2 does not imply this result but along the way of proving it, we obtain this.

Theorem 2.1. *Given a graph G with average degree d, if the largest clique subdivision has order $\Theta(\sqrt{d})$, then its crux size is linear in d, i.e. $c(G) = O(d)$.*

Theorem 2.1 implies that the extremal graphs are *essentially disjoint union of dense small graphs whose crux size is linear in their average degrees.* This can be viewed as an analogous result of Myers [16] who studied the extremal graphs for embedding clique minors.

2.2 Graphs without a fixed bipartite graph

The next application provides a lowerbound on the largest clique subdivision size, which is optimal up to a polylog-factor, in a graph without a fixed copy of bipartite graph H. This generalises the result of Liu and Montgomery [13] on $K_{s,t}$-free graphs. We would like to remark that the proof of Liu and Montgomery makes heavy use of the structure of the forbidden graph $K_{s,t}$, hence their argument does not extend to general H-free graphs. Below, we write $x = \tilde{\Omega}(y)$ if there exists positive constants a, b such that $x \geq ay \log^{-b} y$.

Corollary 2.2. *Let H be a bipartite graph with extremal number $\mathrm{ex}(n, H) = O(n^{1+\tau})$ for some $0 < \tau < 1$ and let G be an H-free graph with average degree d. Then G contains a K_t-subdivision where*

$$t = \begin{cases} \tilde{\Omega}(d^{\frac{1}{2\tau}}) & \text{if } \tau > 1/2 \\ \tilde{\Omega}(d) & \text{if } \tau \leq 1/2. \end{cases}$$

Proof. Let $\alpha = 1/100$, and F be a smallest α-crux of G of order $c(G)$. As F is H-free, $e(F) \leq O(|F|^{1+\tau})$, hence

$$c(G) = |F| = \Omega(d(F)^{1/\tau}) = \Omega(d(G)^{1/\tau}).$$

Therefore, Theorem 1.2 implies that G has a clique subdivision of size $\tilde{\Omega}(\min\{d, c(G)^{\frac{1}{2}}\}) = \tilde{\Omega}(\min\{d, d^{\frac{1}{2\tau}}\})$. □

The bound above is best possible up to the polylogarithmic factor if $\mathrm{ex}(n, H) = \Theta(n^{1+\tau})$. To see this, let G be an n-vertex bipartite H-free graph G with $\Theta(n^{1+\tau})$ edges. If G has a K_t-subdivision, then at least $t/2$ core vertices are in the same part of a bipartition of G. For any two of them, a path connecting them uses at least one vertex of the other part of G so $\binom{t/2}{2} \leq n$ and therefore $t = O(\sqrt{n}) = O(d^{1/2\tau})$.

2.3 Dichotomy on Erdős-Rényi random graphs

The last application delas with the subdivisions in Erdős-Rényi random graphs. While the size of the largest K_t-minor in a random graph is widely studied (for example [2, 11, 4]), the only known results for clique subdivision is when p is a constant. More precisely, when $p \in (0, 1)$ is a constant, Bollobás and Catlin [1] proved in 1981 that the largest clique subdivision of $G(n, p)$ is $(\sqrt{2/(1-p)} + o(1))\sqrt{n}$ with high probability (w.h.p.).

We determine the size of the largest clique subdivision upto polylog-factor when $p = \omega(\frac{\log n}{n})$. We remark that when $p = o(\frac{\log n}{n})$, the clique subdivision is typically extremely small in $G(n, p)$: only logarithmic in n.

Corollary 2.3. *Suppose $p = \omega(\frac{\log n}{n})$ and $p = 1 - \Omega(1)$. Then w.h.p., the largest t that $G = G(n, p)$ has a K_t-subdivision is given by*

$$t = \tilde{\Theta}(\min\{np, \sqrt{n}\}).$$

The proof is obtained by showing $c(G) = \Omega(n)$ for a binomial ramdom graph using the standard concentration inequalities. Corollary 2.3 implies an interesting dichotomy on clique subdivision size in $G(n,p)$ above and below the density $1/\sqrt{n}$: when $p = \omega(n^{-1/2})$, then it is limited solely by the space constraints, while when $p = o(n^{-1/2})$, the degree constraint is the bottleneck.

3 Outline of the proof

In this section, we sketch the proof of our main theorem. For the detail version of the proof, see the online vertsion of our preprint [7].

3.1 Sublinear expander

A main tool we use in this paper is the sublinear expander notion introduced by Komlós and Szemerédi [9, 10]. Let $N^i_G(X)$ be the set of vertices which is distance exactly i from X. In particular, $N^0_G(X) = X$. We write $N_G(X)$ to denote $N^1_G(X)$ and we write $B^i_G(X) = \bigcup_{j\le i} N^i_G(X)$. For given ε, k, we define $\rho(x) = \rho(x,\varepsilon,k)$ as

$$\rho(x) = \begin{cases} 0 & \text{if } x < \frac{k}{5} \\ \frac{\varepsilon}{\log^2(15x/k)} & \text{if } x \ge \frac{k}{5} \end{cases}$$

Note that $\rho(x)$ is a decreasing function and $x\rho(x)$ is a increasing function for $x \ge \frac{k}{2}$. Komlós and Szemerédi introduced the notion of (ε,k)-expander, which is a graph in which every set of appropriate size has not too small external neighbourhood. Haslegrave, Kim and Liu [6] slightly generalised this notion to a robust version. Roughly speaking, a graph G is a robust-expander if every set of appropriate size has not too small external neighbourhood even after deleting a small number of vertices and edges. For an edge set $F \subseteq E(G)$, we write $G \setminus F$ to denote the graph with the vertex set $V(G)$ and the edge set $E(G) \setminus F$.

Definition 3.1 ([6])**.** For $\varepsilon > 0, k > 0$, a graph G is (ε,k)*-robust-expander* if for every subset $X \subseteq V(G)$ of size $\frac{k}{2} \le |X| \le \frac{|V(G)|}{2}$ and an edge set $F \subseteq E(G)$ with $|F| \le d(G)\rho(|X|)|X|$, we have $|N_{G-F}(X)| \ge \rho(|X|)|X|$.

This notion of sublinear expander is very useful in the following three aspects.

- Every graph contains a robust-expander subgraph with almost the same average degree.
- This provides a short connection between any two large sets while avoiding a relatively small set of vertices and edges.
- No metter which small set of vertices we delete, the remaining graph still has large average degree.

These three aspects are captured in the following three results.

Theorem 3.2. *[6, 9, 10] If $\varepsilon > 0$ is sufficiently small (independent from k) so that $\int_1^\infty \rho(x,\varepsilon,k)/x dx < \frac{1}{8}$, then every graph G contains an (ε,k)-robust-expander H with $d(H) \geq d(G)/2$ and $\delta(H) \geq d(H)/2$ as a subgraph.*

Lemma 3.3. *[6, 10] Let G be an n-vertex (ε,k)-robust-expander. Then for any two vertex sets X_1, X_2 of size at least $x \geq \frac{k}{2}$, and a vertex set W of size at most $\frac{x\rho(x)}{4}$, there exists a path between X_1 and X_2 in $G - W$ of length at most $\frac{2}{\varepsilon}\log^3(\frac{15n}{k})$.*

Lemma 3.4. *Suppose $0 < \frac{1}{n} \ll \varepsilon < 1$ and $k < \frac{n}{10}$. Let G be an n-vertex (ε,k)-robust-expander. Then for every $W \subseteq V(G)$ with $|W| \leq \frac{1}{20}\rho(n,\varepsilon,k)\cdot n$, we have $d(G-W) \geq \frac{1}{20}\rho(n,\varepsilon,k)\cdot d(G)$.*

3.2 Outlines of the main steps

Assume that G and t are as in Theorem 1.2 and let $s = \beta t/(\log\log t)^6$ for some small enough constant $\beta > 0$. Using Lemma 3.2, we can assume that our graph is an expander with an appropriate choice of ε and k. Our proof bacisally use the structures introduced in [13]. We find $\Omega(s)$ rooted trees with many branches and leaves (called units and webs) and build short vertex-disjoint paths between them to form a desired clique subdivision. On the other hand, directly mimicking such an argument does not provide a subdivision of the desired size. We need to find a clique subdivision of the size which almost matches the smaller bounds of the degree constraint d and the space constraint $\sqrt{c(G)/\log c(G)}$.

If the bound from the degree constraint is stronger (i.e., d is smaller than $\sqrt{c(G)/\log c(G)}$), the main goal is to collect many vertices of degree at least t. If the bound from the space constraint is stronger (i.e., $\sqrt{n/\log n}$ is much smaller than d), the main difficulty is to find short paths connecting the vertices. As we want to build $\Omega(s^2)$ vertex disjoint paths in G to form a K_s-subdivision, we need to be able to find $\Omega(s^2)$ paths of average length at most $O(n/s^2) = \log n(\log\log n)^{O(1)}$. However, Lemma 3.3 only guarantees a much longer path of length $O(\log^3(n/k))$. To obtain paths of desired length, we may take k to be $n/(\log n)^{O(1)}$, but such a choice of k does not allow us to obtain expansions of small sets, introducing another difficulty. Another issue from this approach is that we don't have any controls on the order of (ε,k)-expander we take. The order of such an expander puts additional space constraints for finding a desired subdivision as well as affect the length of paths we obtain from Lemma 3.3.

To overcome the above difficulties, we consider several cases and use (ε,k)-expander in each case with different choices of k. Let $d = d(G)$. By iteratively applying Theorem 3.2, we obtain the following graphs

$$G \supseteq G_1 \supseteq G_2 \supseteq H$$

where G_1 is an $(\varepsilon,\varepsilon d)$-expander, G_2 is an (ε,d^2)-expander, and H is an $(\varepsilon,c(G)/100)$-expander such that each graph has minimum degree at least $d/16$. Let $n_1 = |G_1|, n_2 = |G_2|$, and $n_H = |H|$. We now consider the following four cases depending on the values of t, n_1, n_2, and n_H.

Case 1: $d \leq \exp(\log^{1/6} n_1)$: In this case, we can adapt a theorem in [13] to obtain a desired $K_{\Omega(d)}$-subdivision.
Case 2: $t \leq \min\{\frac{\sqrt{n_1}}{(\log n_1)^{O(1)}}, \frac{d}{(\log d)^{O(1)}}\}$: In this case, as t is quite smaller than both d and $\sqrt{n_1/\log n_1}$, the degree constraint and the space constraint within G_1 are not strong obstacles to obtain a desired $K_{\Omega(t)}$-subdivision. Hence, by utilizing the properties of the expanders, we can construct many units and webs and connect them with short paths to obtain a desired $K_{\Omega(t)}$-subdivision.
Case 3: $d \leq \frac{\sqrt{n_2}}{(\log n_2)^{O(1)}}$: In this case, the space constraint within G_2 is much weaker than the degree constraint, so our main concern is to collect $\Omega(t)$ vertices of degree at least $d(\log n_2)^{O(1)}$. Although Lemma 3.4 provides a set of large degree vertices, it only gives vertices of degree $d/(\log n_2)^{O(1)}$, which is smaller than what we need. Considering two cases where the edge distribution of G_2 is close to uniform and skewed, a careful analysis provides a desired set of vertices of large degree in both cases.
Case 4: The remaining cases: In the remaining case, we will find a desired K_s-subdivision in H. Note that as we are not in case 1–3, we obtain the ineqaulity

$$\frac{n_H}{(\log c(G))^{O(1)}} \leq c(G) \leq n_H.$$

As H is $(\varepsilon, c(G)/100)$-expander, this ensures that $\rho(x, \varepsilon, c(G)/100) = (\log\log c(G))^{O(1)}$ for every $c(G) \leq x \leq n_H$. With this extra assumption on our hand, Lemma 3.4 now provides a set of $\Omega(t/(\log\log t)^{O(1)})$ vertices of degree $\Omega(t/(\log\log t)^{O(1)})$, which matches the bound from the degree constraint. Moreover, Lemma 3.3 also yields a path of length $(\log\log c(G))^{O(1)}$ between two large sets of size at least $c(G)/100$. Note that the definition of crux ensures some expansion of all vertex set smaller than $c(G)/100$. By utilizing this, we can show that the $O(\log c(G))$-th ball $B_H^{O(\log c(G))}(v)$ of a well-chosen vertex v has size at least $c(G)/100$. This together with Lemma 3.3 provides a desired path of length $(\log\log c(G))^{O(1)}$ between two balls $B_H^{O(\log c(G))}(v)$ and $B_H^{O(\log c(G))}(u)$ of well-chosen vertices u and v. Combining these ideas with further technical analysis, we obtain the desired K_s-subdivison. We omit further details.

References

[1] Bela Bollobás and Paul A. Catlin. Topological cliques of random graphs. *Journal of Combinatorial Theory. Series B*, 30(2):224–227, 1981.

[2] Bela Bollobás, Paul A. Catlin, and Paul Erdős. Hadwiger's conjecture is true for almost every graph. *European Journal of Combinatorics*, 1(3):195–199, 1980.

[3] Béla Bollobás and Andrew Thomason. Proof of a conjecture of Mader, Erdős and Hajnal on topological complete subgraphs. *European Journal of Combinatorics*, 19(8):883–887, 1998.

[4] Nikolaos Fountoulakis, Daniela Kühn, and Deryk Osthus. The order of the largest complete minor in a random graph. *Random Structures & Algorithms*, 33(2):127–141, 2008.

[5] John Haslegrave, Jie Hu, Jaehoon Kim, Hong Liu, Bingyu Luan, and Guanghui Wang. Crux and long cycles in graphs. *SIAM J. Discrete Math.*, 36(4):2942–2958, 2022.

[6] John Haslegrave, Jaehoon Kim, and Hong Liu. Extremal density for sparse minors and subdivisions. *International Mathematics Research Notices*, 2021.

[7] Seonghyuk Im, Jaehoon Kim, Younjin Kim, and Hong Liu. Crux, space constraints and subdivisions. arXiv preprint, arXiv:2207.06653, 2022.

[8] Heinz Adolf Jung. Eine Verallgemeinerung des n-fachen Zusammenhangs für Graphen. *Mathematische Annalen*, 187:95–103, 1970.

[9] János Komlós and Endre Szemerédi. Topological cliques in graphs. *Combinatorics, Probability and Computing*, 3(2):247–256, 1994.

[10] János Komlós and Endre Szemerédi. Topological cliques in graphs. II. *Combinatorics, Probability and Computing*, 5(1):79–90, 1996.

[11] Michael Krivelevich and Benjamin Sudakov. Minors in expanding graphs. *Geometric and Functional Analysis*, 19(1):294–331, 2009.

[12] Kazimierz Kuratowski. Sur le problème des courbes gauches en topologie. *Fundamenta Mathematica*, 15:271–283, 1930.

[13] Hong Liu and Richard Montgomery. A proof of Mader's conjecture on large clique subdivisions in C_4-free graphs. *Journal of the London Mathematical Society. Second Series*, 95(1):203–222, 2017.

[14] Wolfgang Mader. Homomorphieeigenschaften und mittlere Kantendichte von Graphen. *Mathematische Annalen*, 174:265–268, 1967.

[15] Wolfgang Mader. An extremal problem for subdivisions of K_5^-. *Journal of Graph Theory*, 30(4):261–276, 1999.

[16] Joseph Samuel Myers. Graphs without large complete minors are quasi-random. *Combinatorics, Probability and Computing*, 11(6):571–585, 2002.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Cops and Robber on Hyperbolic Manifolds

(Extended abstract)

Vesna Iršič * Bojan Mohar † Alexandra Wesolek ‡

Abstract

The Cops and Robber game on geodesic spaces is a pursuit-evasion game with discrete steps which captures the behavior of the game played on graphs, as well as that of continuous pursuit-evasion games. One of the outstanding open problems about the game on graphs is to determine which graphs embeddable in a surface of genus g have largest cop number. It is known that the cop number of genus g graphs is $O(g)$ and that there are examples whose cop number is $\tilde{\Omega}(\sqrt{g}\,)$. The same phenomenon occurs when the game is played on geodesic surfaces.

In this paper we obtain a surprising result when the game is played on a surface with constant curvature. It is shown that two cops have a strategy to come arbitrarily close to the robber, independently of the genus. For special hyperbolic surfaces we also give upper bounds on the number of cops needed to catch the robber. Our results generalize to higher-dimensional hyperbolic manifolds.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-085`

1 Introduction

The Cops and Robber game is a pursuit-evasion game. The game is commonly played on graphs [1, 5, 6, 7, 9, 12, 16, 19], and as a new variant on geodesic spaces [15, 21, 22]. The

*Faculty of Mathematics and Physics, University of Ljubljana, Slovenia. Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia. E-mail: `vesna.irsic@fmf.uni-lj.si`.

†Department of Mathematics, Simon Fraser University, Burnaby, BC, Canada. E-mail: `mohar@sfu.ca`.
Supported in part by the NSERC Discovery Grant R611450 (Canada) and by the Research Project J1-2452 of ARRS (Slovenia). On leave from IMFM, Ljubljana.

‡Université Lyon, CNRS, INSA Lyon, UCBL, LIRIS, UMR5205, F-69622 Villeurbanne, France. E-mail: `agwesole@sfu.ca`.

players in the Cops and Robber game are the robber r and k cops $c_1, \ldots, c_k$. On graphs the players occupy vertices while on geodesic spaces the players occupy points in space. The game is played in rounds. The robber chooses initial positions for the players $r^0, c_1^0, \ldots, c_k^0$ and in each round of the game the players can move to a new position[1]. Each round of the game has two turns, the first one for the robber and the second one for the cops. In particular, each of the cops can make a step at the cops' turn. When the game is played on a graph, the players are allowed to move to an adjacent vertex at their turn. When the game is played on a geodesic space X, the robber chooses an agility function $\tau : \mathbb{N} \to \mathbb{R}_+$ at the beginning of the game such that $\sum_{n\geq 1} \tau(n) = \infty$. In the n-th round, each player makes a step of length at most $\tau(n)$ (at their turn). The position of the robber r after round n is denoted as r^n, and the position of the cop c_i as c_i^n. In the following we give the robber the pronoun he, him, while the cops have the pronoun she, her. We say the cops *catch* the robber if at some point in the game the cop c_i occupies the same position as the robber r. If the robber is not caught, we say the robber *escapes*. The *cop number* $c(G)$ of a graph G is the minimum number of cops that can catch the robber (regardless of the robber's strategy and initial positions). For a geodesic space X we denote by $c_0(X)$ the *cop catch number*, which is the minimum number of cops that can catch the robber. Further, if the game is played on a geodesic space, we say that *the cops win* the game if

$$\inf_{n,i} d(c_i^n, r^n) = 0. \tag{1}$$

If the cops do not win the game, we say that *the robber wins* the game, which means that he can stay at distance at least ε away from the cops, for some $\varepsilon > 0$. For a geodesic space we denote by $c(X)$ the *cop win number*, that is the minimum number of cops that can win the game.

One of the first results about cop numbers given by Aigner and Fromme states that planar graphs have cop number at most 3 [1]. For graphs embeddable in a surface of genus g it is known that the cop number is at most linear in g and recent progress was made on improving the linear factor [8, 11, 24]. It is an outstanding open problem to determine which graphs embeddable in a surface of genus g have largest cop number. There are graphs of genus g with cop number at least $g^{\frac{1}{2}-o(1)}$, one such example are binomial random graphs $G_{n,p}$ with $p = \frac{2\log n}{n}$ [5, 20]. The gap between the upper and lower bound is large and it is conjectured that the lower bound gives the right order of the cop number.

Conjecture 1 ([20, 22])**.** *Let S be a a graph of genus g. Then $c(S) = O(\sqrt{g})$.*

Similarly, the cop win number for a surface of genus g is at most linear in g [21]. It was shown that for graphs of cop number at least 3 there exists a surface S of genus g with $c(S) \geq c(G)$ [21]. Therefore there are compact surfaces of genus g with cop win number at least $g^{\frac{1}{2}-o(1)}$. The following conjecture is a tough conjecture since it implies Conjecture 1.

Conjecture 2 ([22])**.** *Let S be a geodesic surface of genus g. Then $c(S) = O(\sqrt{g})$.*

[1]The rules of the Cops and Robber game on graphs we define here are slightly different to standard rules, but they do not affect the outcome of the game on connected graphs.

Surprisingly, the upper bound from Conjecture 2 can be significantly improved for surfaces of genus g which are hyperbolic, that means they have constant curvature -1. Since our results extend to higher-dimensional hyperbolic manifolds, we state the more general version.

Theorem 3. *If M is a compact hyperbolic manifold, then $c(M) = 2$.*

An important tool for determining upper bounds for cop numbers is the Isometric Path Lemma, which we use to establish Theorem 3 and to show the cop catch number for special surfaces of genus $g \geq 2$ is at most 6.

Lemma 4 ([1, 22])**.** *Let I be an isometric path starting at A and ending at B. Then one cop c can guard I after spending time equal to the length of I on the path to adjust himself.*

The Cops and Robber game on geodesic spaces is tightly related to continuous and discrete pursuit-evasion games on metric spaces. In continuous pursuit-evasion games the players make decisions at every point in the time interval $[0, \infty)$. For example, Besicovitch showed that in the Lion and Man game as introduced by Rado (see Littlewood's Miscellany [18]), the man can escape the lion when the game is played on a disk. Croft studied a variation of this game with multiple pursuers on higher dimensional balls [10] and Satimov and Kushkarov studied the game on the sphere [23].

The Discrete Lion and Man game is the Cops and Robber game where the agility function is constant, i.e. $\tau \equiv K$ for some constant K. It was shown that in the Discrete Lion and Man game the lion can catch the man on a disk, more generally, the lion can catch the man on any compact CAT(0)-space [4, 26]. While the Cops and Robber game is a discrete pursuit-evasion game, it captures the properties of the continuous game, in the sense that in the Cops and Robber game one cop cannot catch the robber when the game is played on a disk [15].

2 Proof Sketch of Theorem 3

We denote by $\mathbb{H}^n$ the n-dimensional hyperbolic space. By the Killing-Hopf Theorem [13, 17] any hyperbolic manifold arises from a tessellation of hyperbolic space, for an example see Figure 2. More precisely, any hyperbolic manifold is isometric to $\mathbb{H}^n/\Gamma$ where Γ is a group of isometries of $\mathbb{H}^n$ acting freely and properly discontinuously. In order to show Theorem 3, we play the game in the covering space $\mathbb{H}^n$ of the manifold. It was shown that if C is the covering space of a geodesic space X that locally preserves distances, then $c(X) \leq c(C)$ [15]. We will use the idea of the theorem in this proof. In order to simplify our exposition, we provide a sketch of the proof for Theorem 3 only for 2-dimensional manifolds which are surfaces. Let $s = \mathrm{sys}(S)$ be the systolic girth of the hyperbolic surface S, which is the length of the smallest non-contractible curve.

To show that $c(S) > 1$ we play the game with one cop c and the robber r. The robber chooses the agility function $\tau \equiv \frac{s}{8}$ and initial positions such that $d(c^0, r^0) > \frac{s}{8}$. If $d(c^k, r^k) \geq \frac{3s}{8}$, then the robber does not move and $r^{k+1} = r^k$. If $d(c^k, r^k) < \frac{3s}{8}$, the robber moves in the direction opposite to the cop's position, i.e. the shortest paths from r^k to c^k and r^k to r^{k+1} meet at r^k at angle π. To argue that such a position exists with the additional assumption that $d(r^{k+1}, c^k) = d(r^k, c^k) + \frac{s}{8}$, the Gauss-Bonnet Theorem can be applied, for details we refer to the full version of the paper [14]. In both cases the robber can stay at distance at least $\frac{s}{8}$ to the cop, which proves the lower bound.

We sketch the proof strategy for the upper bound. Let $D = \operatorname{diam}(S)$, which is the largest distance between two points in S. The rounds are grouped into blocks, that is, we consider integers representing time steps $1 = t_0 < t_1 < t_2 < \dots$ such that $\sum_{k=t_i}^{t_{i+1}-1} \tau(k) \geq 30D$. For each time step k we choose a representation C_i^k, R^k of c_i^k, r^k in the covering space $\mathbb{H}^2$. By definition of distance on a hyperbolic surface, $d_{\mathbb{H}^2}(C_i^k, R^k) \geq d_{\mathbb{H}^2}(c_i^k, r^k)$ $(i = 1, 2)$. We show that the cops have a strategy such that $\inf_k d_{\mathbb{H}^2}(C_i^k, R^k) = 0$. At each time step t_i, the cop c_2 chooses a new representation $C_2^{t_i}$ of its position $c_2^{t_i}$ in the covering space, which means $d(C_2^{t_i-1}, C_2^{t_i})$ is possibly greater than $\tau(t_i)$ but we maintain that $d(c_2^{t_i-1}, c_2^{t_i}) \leq \tau(t_i)$. In between the time steps t_i, t_{i+1} we choose the representation of the cops and the robber such that they are coherent with the agility function, which means $d_{\mathbb{H}^2}(R^{k-1}, R^k) \leq \tau(k)$ and $d(C_i^{k-1}, C_i^k) \leq \tau(k)$ for $t_i < k < t_{i-1}$. Let $R^{t_i}, C_1^{t_i}$ be a copy of the robber's and cop's position in the covering space, such that their distance in the covering space is the same as in the surface. We consider the geodesic g_0 through $R^{t_i}, C_1^{t_i}$. The choice of the position $C_2^{t_i}$ for cop c_2 is such that it is close to the geodesic g_0 but sufficiently far from R^{t_i}, which we make more precise in the following.

Let P be the point on g_0 at distance $10D$ to R^{t_i} that is further away from $C_1^{t_i}$. Note that there is a copy $C_2^{t_i}$ of $c_2^{t_i}$ in the covering space which is at distance at most $D = \operatorname{diam}(S)$ from P. We consider $h = o_{g_0}(C_2^{t_i})$, the orthogonal geodesic to g_0 through $C_2^{t_i}$.

Claim 1. *$h \cap g_0$ is at distance between $9D$ and $11D$ from R^{t_i} and at distance at most D from $C_2^{t_i}$.*

The strategy of cop c_2 is to chase the orthogonal projection of R^k on h. Note that by Claim 1 the distance from R^{t_i} to h is at least $9D$. Let B, B' be points on h at distance $8D$ from $B_0 := g_0 \cap h$, see Figure 1(a). The cop c_2 can guard the path from B to B' on h, since his distance to B_0 is at most D, so his distance to B, B' is at most $9D$, which is at least the distance from R^{t_i} to B, B'. Let $g_k = o_h(C_1^k)$ be the orthogonal geodesic to h through C_1^k, see Figure 1(b).

Suppose R^k, R^{k+1} are contained in the triangle defined by $B, h \cap g_k$ and C_1^k. Then we move cop c_1 towards B such that:

$$\text{The robber's position } R^{k+1} \text{ and } B \text{ are on the same side of } g_{k+1}. \tag{2}$$

The $(k+1)$-st position C_1^{k+1} of cop c_1 is

(a) the point between C_1^k and B s.t. $d(C_1^{k+1}, C_1^k) = \tau(k+1)$ if this step does not violate (2),

(b) otherwise, the closest point to R^{k+1} on the geodesic $o_h(R^{k+1})$ (the orthogonal geodesic to h through R^{k+1}) which satisfies $d(C_1^{k+1}, C_1^k) = \tau(k+1)$.

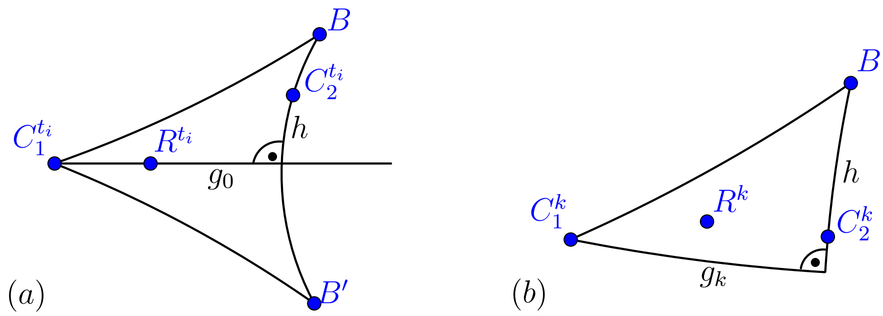


Figure 1: Figure (a) is a schematic drawing of the robber's and the cops' position at time step $t_i\,(1 \leq i)$. Figure (b) is a schematic drawing of the robber's and the cops' position at time step $t_i < k < t_{i+1}$.

The strategy is similar if the robber moves towards B'. If the robber crosses $\overline{C_1^k B}$ or $\overline{C_1^k B'}$ for some $t_i < k < t_{i+1}$, then the cops' strategy for the steps $t_i + 1, \ldots, t_{i+1} - 1$ is simply to walk towards the robber. At step t_{i+1} the strategies reset.

Claim 2. *Either the robber gets caught by the cops or he crosses $\overline{C_1^t B}$ or $\overline{C_1^t B'}$ for some t with $t_i < t < t_{i+1}$. Further, for each $\varepsilon > 0$ there exists some $\delta = \delta(\varepsilon)$, such that if $d(R^{t_i}, C_1^{t_i}) \geq \varepsilon$, then*

$$d(R^t, C_1^t) - d(R^{t_i}, C_1^{t_i}) \geq \delta\varepsilon.$$

Claim 2 shows that the cop c_1 comes eventually ε-close to the robber, which means $d(R^k, C_1^k) < \varepsilon$ for some k. This is enough to show that the cops can win the game on S. For the proofs of Claims 1 and 2 we refer the reader to the full version of the paper [14]. $\square$

3 Catching the Robber

We define by $P(k, \theta)$ the regular k-gon in the Poincaré disk $\mathcal{D}$ centred at $O = (0, 0)$ with angle θ at the vertices. We denote its vertices by $v_1, \ldots, v_k$ in counter-clockwise direction and let a_i be the (oriented) edge from v_i to v_{i+1} and a_i^{-1} be the reversed edge from v_{i+1} to v_i (we consider the indices modulo k). We are going to consider three standard hyperbolic surfaces for $g \geq 2$, where one of them is non-orientable.

- Let $S(g)$ be the orientable surface obtained from $P\left(4g, \frac{2\pi}{4g}\right)$ by identifying the (oriented) edges a_{4i-3} with a_{4i-1}^{-1} and a_{4i-2} with a_{4i}^{-1} for $i = 1, \ldots, g$. The surface $S(2)$ is depicted in Figure 2.
- Let $S'(g)$ be the orientable surface obtained from $P\left(4g+2, \frac{2\pi}{2g+1}\right)$ by identifying opposite (oriented) edges a_i, a_{i+2g+1}^{-1} for $i = 1, \ldots, 2g$.

- Let $N(g)$ be the non-orientable surface obtained from $P\left(2g, \frac{2\pi}{2g}\right)$ by identifying the (oriented) edge a_{2i} with a_{2i+1} for $i = 1, \dots, g$.

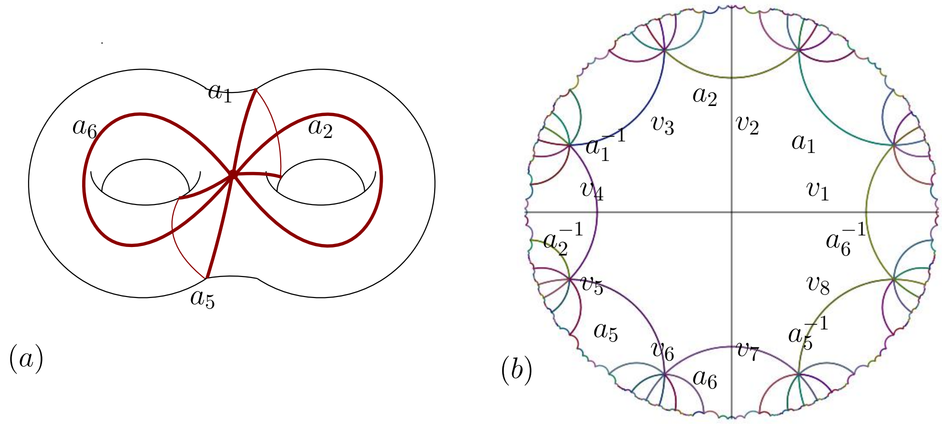


Figure 2: The surface $S(g)$ as a standard geometric model for the double torus with its fundamental domain $P\left(4g, \frac{2\pi}{4g}\right)$ in the Poincaré disk $\mathcal{D}$.

Lemma 5. *Suppose the robber is contained in a bounded convex polygon in the Poincaré disk $\mathcal{D}$ where n cops guard the boundary of the polygon. Then these n cops can catch the robber.*

We can deduce the following theorem from Lemma 5.

Theorem 6. *If $g \geq 2$, then (a) $c_0(S(g)) \leq 5$, (b) $c_0(S'(g)) \leq 6$ and (c) $c_0(N(g)) \leq 4$.*

Proof. We will give the proof only for (a), the proof for the other surfaces is similar. Let O be the midpoint of the fundamental polygon $P\left(4g, \frac{2\pi}{4g}\right)$. We will play the game in the covering space and choose the player's positions such that they are in $P\left(4g, \frac{2\pi}{4g}\right)$. We will first use the cops c_1, c_2, c_3 to guard isometric paths. We start by moving cop c_1 to the isometric path $\overline{Ov_1}$, cop c_2 to the isometric path $\overline{Ov_5}$ and cop c_3 to the isometric path $\overline{Ov_9}$. By Lemma 4 we can assume that after a finite amount of time the cops guard the respective isometric paths. Now if the robber is in one of the triangles Ov_jv_{j+1} for some $1 \leq j \leq 4$, then the robber's moves are restricted to the specified triangles since $a_1 = a_3^{-1}$ and $a_2 = a_4^{-1}$. Similarly if the robber is contained in one of the triangles Ov_jv_{j+1} for some $5 \leq j \leq 8$ his moves are restricted to these triangles. If the robber is outside Ov_jv_{j+1} for some $1 \leq j \leq 8$, we move cop c_2 to the isometric path $\overline{Ov_{13}}$ and wait until he is guarding it. If the robber is in one of the triangles Ov_jv_{j+1} for $9 \leq j \leq 12$, his moves are restricted to these triangles. If the robber is not contained in one of these triangles we keep going for $i = 3, 4, \dots$ in the same way, moving the cop currently guarding $\overline{Ov_{1+4i}}$ to guard $\overline{Ov_{1+4(i+2)}}$ unless $i+2 = g$, in which case the robber is contained in one of $Ov_{4g-3}v_{4g-2}, Ov_{4g-2}v_{4g-1}, Ov_{4g-1}v_{4g}$ or $Ov_{4g}v_1$, cop c_1 guards Ov_1 and one of c_2 or c_3 guards Ov_{4g-3}.

We assume without loss of generality that cop c_1 guards $\overline{Ov_1}$, cop c_2 guards $\overline{Ov_5}$ and the robber is in one of the triangles Ov_jv_{j+1} for some $1 \leq j \leq 8$. Cop c_3, c_4, c_5 will guard $\overline{Ov_2}, \overline{Ov_3}, \overline{Ov_4}$, respectively. Now the robber is captured in either $R_1 = Ov_1v_2 \cup Ov_3v_4$ or $R_2 = Ov_2v_3 \cup Ov_4v_5$. The regions R_1 and R_2 can be embedded in the covering space $\mathcal{D}$ such that they form a quadrilateral which is guarded by four of the cops. By Lemma 5 we can now catch the robber. $\square$

References

[1] M. Aigner and M. Fromme. A game of cops and robbers. *Discrete Appl. Math.*, 8(1):1–11, 1984.

[2] James W Anderson. Convexity, area, and trigonometry. In *Hyperbolic Geometry*, pages 111–151. Springer, 1999.

[3] Riccardo Benedetti and Carlo Petronio. *Lectures on hyperbolic geometry*. Springer Science & Business Media, 1992.

[4] Andrew Beveridge and Yiqing Cai. Two-dimensional pursuit-evasion in a compact domain with piecewise analytic boundary. *ArXiv e-prints*, May 2015.

[5] Béla Bollobás, Gábor Kun, and Imre Leader. Cops and robbers in a random graph. *Journal of Combinatorial Theory, Series B*, 103(2):226–236, 2013.

[6] Anthony Bonato. *An invitation to pursuit-evasion games and graph theory*, volume 97 of *Student Mathematical Library*. American Mathematical Society, Providence, RI, 2022.

[7] Anthony Bonato and Richard J. Nowakowski. *The game of cops and robbers on graphs*, volume 61 of *Student Mathematical Library*. American Mathematical Society, Providence, RI, 2011.

[8] Nathan Bowler, Joshua Erde, Florian Lehner, and Max Pitz. Bounding the cop number of a graph by its genus. *SIAM J. Discrete Math.*, 35(4):2459–2489, 2021.

[9] Peter Bradshaw. A proof of the meyniel conjecture for abelian cayley graphs. *Discrete Mathematics*, 343(1):111546, 2020.

[10] H. T. Croft. 'Lion and man': A postscript. *J. Lond. Math. Soc.*, 39:385–390, 1964.

[11] Joshua Erde and Florian Lehner. Improved bounds on the cop number of a graph drawn on a surface. In *Extended Abstracts EuroComb 2021*, pages 111–116. Springer, 2021.

[12] Peter Frankl. Cops and robbers in graphs with large girth and Cayley graphs. *Discrete Appl. Math.*, 17(3):301–305, 1987.

[13] Heinz Hopf. Zum Clifford-Kleinschen Raumproblem. *Math. Ann.*, 95(1):313–339, 1926.

[14] Vesna Iršič, Bojan Mohar, and Alexandra Wesolek. Cops and robber on hyperbolic manifolds. *In preparation.*

[15] Vesna Iršič, Bojan Mohar, and Alexandra Wesolek. Cops and robber game in higher-dimensional manifolds with spherical and Euclidean metric. *C. R. Math. Acad. Sci. Soc. R. Can.*, 44(3):50–68, 2022.

[16] Gwenaël Joret, Marcin Kamiński, and Dirk Oliver Theis. The cops and robber game on graphs with forbidden (induced) subgraphs. *Contrib. Discrete Math.*, 5(2):40–51, 2010.

[17] Wilhelm Killing. Ueber die Clifford-Klein'schen Raumformen. *Math. Ann.*, 39(2):257–278, 1891.

[18] J. E. Littlewood. *Littlewood's miscellany.* Cambridge University Press, Cambridge, 1986. Edited and with a foreword by Béla Bollobás.

[19] Tomasz Łuczak and Paweł Prałat. Chasing robbers on random graphs: zigzag theorem. *Random Structures & Algorithms*, 37(4):516–524, 2010.

[20] Bojan Mohar. Notes on cops and robber game on graphs. *ArXiv e-prints*, October 2017.

[21] Bojan Mohar. Min-max theorem for the game of cops and robber on geodesic spaces. *ArXiv e-prints*, December 2021.

[22] Bojan Mohar. The game of cops and robber on geodesic spaces. *ArXiv e-prints*, May 2022.

[23] N. Yu. Satimov and A. Sh. Kuchkarov. On the solution of a model differential pursuit-evasion game on a sphere. *Uzb. Mat. Zh.*, 2000(1):45–50, 2000.

[24] Bernd SW Schröder. The copnumber of a graph is bounded by [3/2 genus (g)]+ 3. In *Categorical perspectives*, pages 243–263. Springer, 2001.

[25] Stothers, Wilson. Hyperbolic geometry. http://www.maths.gla.ac.uk/wws/cabripages/hyperbolic/hyperbolic0.html. [Online; accessed 16-January-2023].

[26] Olga Yufereva. Lion and man game in compact spaces. *Dyn. Games Appl.*, 9(1):281–292, 2019.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

How connectivity affects the extremal number of trees

(Extended abstract)

Suyun Jiang* Hong Liu† Nika Salia†

Abstract

The Erdős-Sós conjecture states that the maximum number of edges in an n-vertex graph without a given k-vertex tree is at most $\frac{n(k-2)}{2}$. Despite significant interest, the conjecture remains unsolved. Recently, Caro, Patkós, and Tuza considered this problem for host graphs that are connected. Settling a problem posed by them, for a k-vertex tree T, we construct n-vertex connected graphs that are T-free with at least $(1/4 - o_k(1))nk$ edges, showing that the additional connectivity condition can reduce the maximum size by at most a factor of 2. Furthermore, we show that this is optimal: there is a family of k-vertex brooms T such that the maximum size of an n-vertex connected T-free graph is at most $(1/4 + o_k(1))nk$.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-086`

1 Introduction

In extremal graph theory, a central focus is determining the extremal number of various graphs. The extremal number, denoted by $\mathrm{ex}(n, F)$, is the maximum number of edges in an n-vertex graph that does not contain a graph F as a subgraph, not necessarily induced. While the asymptotic behavior of this function has been determined for all non-bipartite

*School of Artificial Intelligence, Jianghan University, Wuhan, Hubei, China, and Extremal Combinatorics and Probability Group (ECOPRO), Institute for Basic Science (IBS), Daejeon, South Korea. E-mail: `jiang.suyun@163.com`. Supported by National Natural Science Foundation of China (11901246) and China Scholarship Council and IBS-R029-C4.

†Extremal Combinatorics and Probability Group (ECOPRO), Institute for Basic Science (IBS), Daejeon, South Korea. E-mail: `hongliu@ibs.re.kr`, `salianika@gmail.com`. Supported by IBS-R029-C4.

graphs by Erdős, Stone, and Simonovits [6, 10], the behavior for bipartite graphs remains open with significant interest from the community.

The Erdős-Gallai theorem [9], established in 1959, studied the k-vertex path P_k, stating that $\mathrm{ex}(n, P_k) \leqslant \frac{n(k-2)}{2}$, and the maximum value is achieved by $\cup K_{k-1}$, the disjoint union of K_{k-1} when $k-1 | n$. Once the extremal number of the path is determined, extending the question to a tree is a natural next step. In 1962, Erdős and Sós [7] conjectured that for any k-vertex tree, its extremal number is at most $\frac{n(k-2)}{2}$, and again the disjoint union of K_{k-1} serves as an example for tightness.

Motivated by the fact that the conjectured maximizer $\cup K_{k-1}$ is not connected, a natural variant is to consider host graphs that are connected, see e.g [15, 4]. Formally, the connected extremal number $\mathrm{ex}_c(n, F)$ is the maximum number of edges in an n-vertex connected graph without a subgraph isomorphic to F. While the additional connectivity condition does not affect the asymptotics of the extremal number when the forbidden graph is non-bipartite or 2-edge-connected, Caro, Patkós, and Tuza [4] investigated what effect it has for trees. Notice that, in contrast with the classical extremal number, its connected relative is not even a monotone function of n. Indeed, for paths, it is known that for every $k \geqslant 10$, $\mathrm{ex}_c(k, P_k) = \binom{k-2}{2} + 2 < \binom{k-1}{2} = \mathrm{ex}_c(k-1, P_k)$.

Such connected variant for trees was in fact studied before and could date back to the work of Kopylov [19] in 1977, in which he resolved the problem for paths, showing that for $n \geqslant k \geqslant 4$,

$$\mathrm{ex}_c(n, P_k) = \max\left\{\binom{k-2}{2} + (n-(k-2)), \left\lfloor\frac{k-2}{2}\right\rfloor(n - \left\lceil\frac{k}{2}\right\rceil) + \binom{\lceil\frac{k}{2}\rceil}{2}\right\}. \quad (1)$$

Later, Balister, Győri, Lehel, Schelp [1] characterized extremal graphs for every n. There are also recent developments towards the stability version of this theorem by Füredi, Kostochka, Luo, Verstraëte [13, 14].

By Erdős-Gallai theorem and Kopylov's result (1), we see that the asymptotic of the maximum number of edges in an n-vertex P_k-free graph does not change by an additional connectivity constraint as $n \to \infty$ and $k \to \infty$. Caro, Patkós, and Tuza [4] studied how much smaller $\mathrm{ex}_c(n, T)$ can be for a k-vertex tree T, compared to $\frac{n(k-2)}{2}$. Formally, they defined

$$\gamma_k := \inf\left\{\limsup_{n\to\infty} \frac{\mathrm{ex}_c(n, T_k)}{\frac{n(k-2)}{2}} : T_k \text{ is a } k\text{-vertex tree}\right\} \text{ and } \gamma := \lim_{k\to\infty} \gamma_k. \quad (2)$$

It is not hard to see that this limit exists. From above, Caro, Patkós, and Tuza [4] found a family of trees whose connected extremal number is asymptotically smaller, yielding $\gamma \leqslant \frac{2}{3}$. From below, for every tree, they gave constructions showing that $\gamma \geqslant \frac{1}{3}$. They asked where the truth lies between $\frac{1}{3}$ and $\frac{2}{3}$.

Our main result settles this problem.

Theorem 1.1. *Let γ be as defined in* (2)*, we have* $\gamma = \frac{1}{2}$.

To obtain the lower bound $\gamma \geqslant \frac{1}{2}$, we provide several families of different constructions depending on the 'center of mass' of the forbidden tree (see Section 2.1), realizing the following.

Theorem 1.2. *For any k-vertex tree T_k, we have*

$$\mathrm{ex}_c(n, T_k) \geqslant \left(\frac{1}{4} - o_k(1)\right) kn.$$

On the other hand, we determine the exact connected extremal number of brooms with k vertices and diameter d, denoted by $B(k,d)$, for large enough n. In particular, $B(k,d)$ is the graph obtained from a path of $d+1$ vertices by blowing up a leaf to an independent set of size $k-d$. The following theorem is stated using graphs $\mathsf{G}_{n,\cdot,\cdot}$ and $\mathsf{F}_{n,\cdot,\cdot}$, which are defined in Sections 2.2 and 3.1. Some of these graphs (so-called edge blow-up of stars) have been studied before, see e.g. [5, 8, 27]. As the path is also a broom, the result below can be viewed as an extension of Kopylov's result (1).

Theorem 1.3. *For every integer k and d such that $k \geqslant d+2 \geqslant 8$, and $n \geqslant k^{dk}$ we have*

$$\mathrm{ex}_c(n, B(k,d)) = \begin{cases} e(\mathsf{G}_{n,d,\lfloor \frac{d-1}{2} \rfloor}) & \text{if } d \geqslant \frac{k+5}{2}, \\ \max\{e(\mathsf{G}_{n,d,\lfloor \frac{d-1}{2} \rfloor}), e(\mathsf{F}_{n,\frac{k+2}{2},1})\} & \text{if } d = \frac{k+2}{2} \text{ or } \frac{k+4}{2}, \\ \max\{e(\mathsf{G}_{n,d,\lfloor \frac{d-1}{2} \rfloor}), e(\mathsf{F}_{n,d,2}), e(\mathsf{F}_{n,d,3})\} & \text{if } d = \frac{k+3}{2}, \\ \left\lfloor \frac{(k-d)n}{2} \right\rfloor & \text{if } d \leqslant \frac{k+1}{2}. \end{cases}$$

As a corollary, we get the matching upper bound $\gamma \leqslant \frac{1}{2}$ as $\mathrm{ex}_c(n, B(k, \lceil \frac{k}{2} \rceil)) = \left(\frac{1}{4} + o_k(1)\right) kn$.

2 Overview of the proof of Theorem 1.2

In this section, we first introduce a key concept: the barycenter of a tree. We then provide three special graphs $\mathsf{G}_{n,\frac{k-c}{2},\frac{k-c}{4}}$, $\mathsf{S}_{n,x}$ and $\mathsf{P}_{n,x}$, each of which has $(\frac{1}{4} - o_k(1))kn$ edges, where c is a constant. By considering the degree of the barycenter vertex of the tree T_k, we can find a T_k-free graph G such that G is isomorphic to one of $\mathsf{G}_{n,\frac{k-c}{2},\frac{k-c}{4}}$, $\mathsf{S}_{n,x}$ and $\mathsf{P}_{n,x}$ for some x.

2.1 The Barycenter of a tree

For any tree T on k vertices, we call a vertex v of T *a barycenter* if v belongs to a largest connected component of $T - e$ for every edge e in T, that is, the vertex v belongs to the component of size at least $\lceil \frac{k}{2} \rceil$ in the graph obtained from T by removing an edge e.

Proposition 2.1. *Every tree has either a unique barycenter, or there are exactly two barycenters in the tree joined by an edge.*

2.2 Constructions of various classes of graphs

The family $\mathsf{G}_{n,\cdot,\cdot}$. We first define the family of extremal graphs for (1). Recall '$\cup$' denotes the disjoint union of graphs, '+' denotes the join of the graphs, and $\overline{K}_t$ denotes an independent set of size t. For $n \geqslant k \geqslant 2s$ let $\mathsf{G}_{n,k,s} := \left(K_{k-2s} \cup \overline{K}_{n-k+s}\right) + K_s$, see Figure 2.1. Note that for every n and k, there exists a constant a such that $k < a < 2k$ and the only extremal graphs achieving equality in (1) are $\mathsf{G}_{n,k-1,1}$ for $n \leqslant a$ and $\mathsf{G}_{n,k-1,\lfloor\frac{k-2}{2}\rfloor}$ for $n \geqslant a$. Clearly, $e(\mathsf{G}_{n,\frac{k-c}{2},\frac{k-c}{4}}) = (\frac{1}{4} - o_k(1))kn$.

The families $\mathsf{S}_{n,x}$ and $\mathsf{P}_{n,x}$. Let x be an integer such that $\frac{k}{2} < x < k$ or $x = \lfloor\frac{k-2}{2}\rfloor$. For the sake of simplicity of the write-up, we denote

$$a_x := \begin{cases} \left\lfloor \frac{2x^2}{k} \right\rfloor - 2 & \text{if } \frac{k}{2} < x < k, \text{ and} \\ x & \text{if } x = \lfloor \frac{k-2}{2} \rfloor. \end{cases}$$

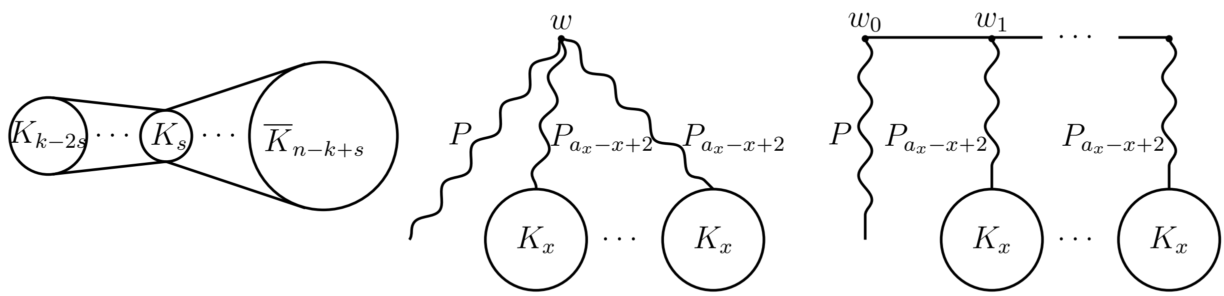


Figure 2.1: The graph $\mathsf{G}_{n,k,s}$ on the left, the graph $\mathsf{S}_{n,x}$ in the middle and the graph $\mathsf{P}_{n,x}$ on the right.

Let $\mathsf{S}_{n,x}$ be a graph consisting of $\left\lfloor \frac{n-1}{a_x} \right\rfloor$ vertex disjoint K_x with pendant paths of length $a_x - x$, a path of $n - 1 - a_x \left\lfloor \frac{n-1}{a_x} \right\rfloor$ vertices and a vertex w adjacent with an endpoint of each of these paths.

Let $\mathsf{P}_{n,x}$ be a graph consisting of $\left\lfloor \frac{n}{a_x+1} \right\rfloor$ vertex disjoint K_x with pendant paths of length $a_x - x + 1$ with the terminal leaf w_i ($i \geqslant 1$), a path of $n - (a_x + 1) \left\lfloor \frac{n}{a_x+1} \right\rfloor$ vertices with a terminal leaf w_0, such that $w_0 w_1 \dots w_{\lfloor \frac{n}{a_x+1} \rfloor}$ is a path.

It is easy to see that $e(\mathsf{S}_{n,x}) = (\frac{1}{4} - o_k(1))kn$ and $e(\mathsf{P}_{n,x}) = (\frac{1}{4} - o_k(1))kn$.

Now we give the overview of the proof of Theorem 1.2: Let v be a barycenter of the tree T_k, which exists by Proposition 2.1. Let $x_1 = \lfloor\frac{k-2}{2}\rfloor$ and $x_2 = \left\lfloor\frac{k}{\sqrt{2}}\right\rfloor$. We split the proof into three cases: $d(v) = 2$, $d(v) \geqslant 4$ and $d(v) = 3$. For the cases $d(v) = 2$ and $d(v) \geqslant 4$, we show that S_{n,x_1} and P_{n,x_1} are T_k-free, respectively. For the case $d(v) = 3$, we show that either T_k can not be embedded in S_{n,x_2} or P_{n,x_2}, or T_k contains two vertex disjoint sub-trees

S_1 and S_2, each of which is isomorphic to a spider[1] with the central vertex of degree at most three and $v(S_1)+v(S_2) \geqslant \frac{k-c}{2}$ for some constant c and thus T_k can not embedding in $\mathsf{G}_{n,2(\lfloor\frac{k-c}{4}\rfloor-5),\lfloor\frac{k-c}{4}\rfloor-5}$. In these three cases, we can always get the desired lower bound.

3 Overview of the proof of Theorem 1.3

In this section, we first construct some edge blow-up of stars $\mathsf{F}_{n,d,\cdot}$ in Section 3.1, which $e(\mathsf{F}_{n,d,\cdot})$ achieves the lower bound of $\mathrm{ex}_c(n, B(k,d))$. To prove the upper bound of $ex_c(n, B(k,d))$ we need some theorems, see Section 3.2.

3.1 The family $\mathsf{F}_{n,\cdot,\cdot}$

For $n > d \geqslant 2$, let $\mathsf{F}_{n,d,1}$ be the n-vertex connected graph such that every maximal 2-connected block is a clique of size $d-1$ except at most one clique of size $n - \lfloor\frac{n-1}{d-2}\rfloor(d-2)$, all sharing a common vertex. Thus if $n-1=(d-2)p_1+q_1$ for non-negative integers p_1 and q_1 such that $0 \leqslant q_1 < d-2$ then

$$e(\mathsf{F}_{n,d,1}) = p_1\binom{d-1}{2} + \binom{q_1+1}{2} = \frac{(d-1)(n-1)}{2} - \frac{q_1(d-2-q_1)}{2}.$$

Let $\mathsf{F}_{n,d,2}$ be the n-vertex connected graph such that every maximal 2-connected block is a clique with one of size $d-1$, the rest of size $d-2$ except at most one clique of size $n-1-\lfloor\frac{n-2}{d-3}\rfloor(d-3)$, all sharing a common vertex. Thus if $n-2=(d-3)p_2+q_2$ for integers p_2 and q_2 such that $p_2 \geqslant 1$ and $0 \leqslant q_2 < d-3$ then

$$e(\mathsf{F}_{n,d,2}) = p_2\binom{d-2}{2} + d - 2 + \binom{q_2+1}{2} = \frac{(d-2)n}{2} - \frac{q_2(d-3-q_2)}{2}.$$

For an even integer d, let $n-1=(d-3)p_3+q_3$ for integers p_3 and q_3 such that $0 \leqslant q_3 < d-3$. If $p_3 \geqslant q_3$, let $\mathsf{F}_{n,d,3}$ be the n-vertex connected graph such that it contains p_3 maximal 2-connected blocks $G_1,\ldots,G_{p_3}$ all sharing a common vertex v with p_3-q_3 of them being the cliques of size $d-2$. The rest of the maximal 2-connected blocks G_i are the cliques of size $d-1$ without a perfect matching of $G_i - v$. We have

$$e(\mathsf{F}_{n,d,3}) = \frac{(d-2)(n-1)}{2}.$$

If $p_3 < q_3$, let $\mathsf{F}_{n,d,3}$ be the n-vertex connected graph such that it contains p_3+1 maximal 2-connected blocks $G_1,\ldots,G_{p_3+1}$ all sharing a common vertex v with p_3 of them being the cliques of size $d-1$ without a perfect matching of $G_i - v$ and the remaining one is a clique of size q_3-p_3+1. We have

$$e(\mathsf{F}_{n,d,3}) = p_3\frac{(d-2)^2}{2} + \binom{q_3-p_3+1}{2} = \frac{(d-2)(n-1)}{2} - \frac{(q_3-p_3)(d-3-(q_3-p_3))}{2}.$$

[1]Spider is a tree with all vertices of degree at most two, except one vertex of any degree, referred to as the central vertex of the spider.

3.2 Other tools

Let $\mathcal{C}_{\geqslant k}$ denote the class of cycles of length at least k. For a class of graphs $\mathcal{F}$, the Turán number of $\mathcal{F}$ is the maximum number of edges in a graph not containing a subgraph F for all $F \in \mathcal{F}$, denoted by $\mathrm{ex}(n, \mathcal{F})$.

Woodall [25] and independently Kopylov [19] improved Erdős-Gallai theorem [9] for long cycles and obtained the following exact result for every n, see also [12].

Theorem 3.1 (Woodall [25], Kopylov [19]). *Let $n = p(k-2) + q + 1$, where $0 \leqslant q < k-2$ and $k \geqslant 3$, $p \geqslant 1$,*

$$\mathrm{ex}(n, \mathcal{C}_{\geqslant k}) = p\binom{k-1}{2} + \binom{q+1}{2} = \frac{(k-1)(n-1)}{2} - \frac{q(k-2-q)}{2}.$$

Theorem 3.2 (Kopylov [19], Woodall [25], Fan, Lv and Wang [11]). *Suppose $n \geqslant k \geqslant 5$, then every 2-connected n-vertex $\mathcal{C}_{\geqslant k}$-free graph contains at most*

$$\max\left\{\binom{k-2}{2} + 2(n-(k-2)), \left\lfloor\frac{k-1}{2}\right\rfloor\left(n - \left\lceil\frac{k+1}{2}\right\rceil\right) + \binom{\lceil\frac{k+1}{2}\rceil}{2}\right\} \textit{ edges.}$$

The extremal graphs are $\mathsf{G}_{n,k,2}$ and $\mathsf{G}_{n,k,\lfloor\frac{k-1}{2}\rfloor}$.

Now we give a overview of the proof of Theorem 1.3: For the lower bound of $\mathrm{ex}_c(n, B(k,d))$, we consider the following $B(k,d)$-free graphs: $\mathsf{G}_{n,d,\lfloor\frac{d-1}{2}\rfloor}$, an almost $(k-d)$-regular graph, $\mathsf{F}_{n,\frac{k+2}{2},1}$ if $d = \frac{k+2}{2}$ or $\frac{k+4}{2}$, $\mathsf{F}_{n,d,2}$ and $\mathsf{F}_{n,d,3}$ if $d = \frac{k+3}{2}$.

For the matching upper bound, let G be an n-vertex connected $B(k,d)$-free graph with $n \geqslant k^{dk}$ and let v be a vertex of G such that $d(v) = \Delta(G)$. We divide the proof into three cases depending on the value of $\Delta(G)$. If $\Delta(G) \leqslant k-d$ then we easily get the desired upper bound $\left\lfloor\frac{(k-d)n}{2}\right\rfloor$. If $k-d+1 \leqslant \Delta(G) \leqslant k-2$, then there exists a path of at least dk vertices starting at v by considering a breadth-first search tree from the vertex v, and then by pigeonhole principle we will find a copy of $B(k,d)$ in G resulting in a contradiction. For the last case $\Delta(G) \geqslant k-1$, if G is 2-connected then we get the desired upper bound by Theorem 3.2; otherwise, we assume G is not 2-connected and its maximal 2-connected blocks are $G_1, G_2, \ldots, G_s$. By Theorem 3.1, the circumference of G is either $d-2$ or $d-1$ or we have the desired upper bound. And by Theorem 3.2, each $e(G_i)$ is bounded for $i \in [s]$. By using the properties of G (the circumference and $B(k,d)$-free), we can get the desired upper bound.

4 Concluding remarks

For the Turán problem, we determine the maximum effect the additional connectivity condition could have over all trees. An interesting future direction of research would be to identify appropriate parameters (if exist) of a tree that determine the asymptotic behavior of its connected extremal number. The constructions in Sections 2 and 3 could be useful for this problem.

The reader interested in the problems and extensions related to Erdős-Sós conjecture we refer to the following papers [2, 22, 17, 20, 21, 23, 26, 3], extensions for Berge hypergraphs see [16, 18], extensions for colored graphs see [24].

Caro, Patkós, Tuza [4] asked whether the connected extremal number becomes monotone eventually. In particular, for every graph F, there exists a constant N_F, such that for every $n \geqslant N_F$, $\mathrm{ex}_c(n, F) \leqslant \mathrm{ex}_c(n+1, F)$. We observe that this is true when F contains a cycle.

Proposition 4.1. *For every graph F containing a cycle, there exists a constant N_F such that for every $n > N_F$ we have $\mathrm{ex}_c(n, F) < \mathrm{ex}_c(n+1, F)$.*

References

[1] Paul N. Balister, Ervin Győri, Jenő Lehel, and Richard H. Schelp. Connected graphs without long paths. *Discrete Mathematics*, 308(19):4487–4494, 2008.

[2] Guido Besomi, Matías Pavez-Signé, and Maya Stein. On the Erdős-Sós conjecture for trees with bounded degree. *Combinatorics, Probability and Computing*, 30(5):741–761, 2021.

[3] Stephan Brandt and Edward Dobson. The Erdős-Sós conjecture for graphs of girth 5. *Discrete Mathematics*, 150:411–414, 1996.

[4] Yair Caro, Balázs Patkós, and Zsolt Tuza. Connected Turán number of trees. *arXiv preprint arXiv:2208.06126*, 2022.

[5] Guantao Chen, Ronald J. Gould, Florian Pfender, and Bing Wei. Extremal graphs for intersecting cliques. *Journal of Combinatorial Theory, Series B*, 89:159–171, 2003.

[6] Paul Erdős and A. H. Stone. On the structure of linear graphs. *Bulletin of the American Mathematical Society*, 52:1087–1091, 1946.

[7] Paul Erdős. Extremal problems in graph theory. In *Theory of graphs and its applications (Proc. Sympos. Smolenice, 1963)*, pages 29–36. Publishing House of the Czechoslovak Academy of Science, Prague, 1964.

[8] Paul Erdős, Z. Füredi, Ronald J. Gould, and D. S. Gunderson. Extremal graphs for intersecting triangles. *Journal of Combinatorial Theory, Series B*, 64:89–100, 1995.

[9] Paul Erdős and Tibor Gallai. On maximal paths and circuits of graphs. *Acta Mathematica Academiae Scientiarum Hungarica*, 10(3-4):337–356, 1959.

[10] Paul Erdős and Miklós Simonovits. A limit theorem in graph theory. *Studia Scientiarum Mathematicarum Hungarica*, 1:51–57, 1966.

[11] Genghua Fan, Xuezheng Lv, and Pei Wang. Cycles in 2-connected graphs. *Journal of Combinatorial Theory, Series B*, 92(2):379–394, 2004.

[12] Ralph J. Faudree and Richard H. Schelp. Path Ramsey numbers in multicolorings. *Journal of Combinatorial Theory, Series B*, 19(2):150–160, 1975.

[13] Zoltán Füredi, Alexandr Kostochka, Ruth Luo, and Jacques Verstraëte. Stability in the Erdős–Gallai theorem on cycles and paths, II. *Discrete Mathematics*, 341(5):1253–1263, 2018.

[14] Zoltán Füredi, Alexandr Kostochka, and Jacques Verstraëte. Stability in the Erdős–Gallai theorems on cycles and paths. *Journal of Combinatorial Theory, Series B*, 121:197–228, 2016.

[15] Zoltán Füredi and Miklós Simonovits. The history of degenerate (bipartite) extremal graph problems. In *Erdős Centennial*, pages 169–264. Springer, 2013.

[16] Dániel Gerbner, Abhishek Methuku, and Cory Palmer. General lemmas for Berge–Turán hypergraph problems. *European Journal of Combinatorics*, 86:103082, 2020.

[17] Agnieszka Görlich and Andrzej Żak. On Erdős-Sós conjecture for trees of large size. *The Electronic Journal of Combinatorics*, 23(1):#P1.52, 2016.

[18] Ervin Győri, Nika Salia, Casey Tompkins, and Oscar Zamora. Turán numbers of Berge trees. *Discrete Mathematics*, 346(4):113286, 2023.

[19] G. N. Kopylov. On maximal paths and cycles in a graph. *Doklady Akademii Nauk SSSR*, 234(1):19–21, 1977.

[20] Bernard Lidický, Hong Liu, and Cory Palmer. On the Turán number of forests. *The Electronic Journal of Combinatorics*, 20(2):#P62, 2013.

[21] Andrew McLennan. The Erdős-Sós conjecture for trees of diameter four. *Journal of Graph Theory*, 49(4):291–301, 2005.

[22] Václav Rozhoň. A local approach to the Erdős-Sós conjecture. *SIAM Journal on Discrete Mathematics*, 33(2):643–664, 2019.

[23] Jean-François Saclé and Mariusz Woźniak. The Erdős-Sós conjecture for graphs without C_4. *Journal of Combinatorial Theory, Series B*, 70:367–372, 1997.

[24] Nika Salia, Casey Tompkins, and Oscar Zamora. An Erdős-Gallai type theorem for vertex colored graphs. *Graphs and Combinatorics*, 35(3):689–694, 2019.

[25] D. R. Woodall. Maximal circuits of graphs. I. *Acta Mathematica Academiae Scientiarum Hungarica*, 28(1-2):77–80, 1976.

[26] Mariusz Woźniak. On the Erdős-Sós conjecture. *Journal of Graph Theory*, 21(2):229–234, 1996.

[27] Long-Tu Yuan. Extremal graphs for edge blow-up of graphs. *Journal of Combinatorial Theory, Series B*, 152:379–398, 2022.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Semi-algebraic and semi-linear Ramsey numbers

(Extended abstract)

Zhihan Jin* István Tomon†

Abstract

An r-uniform hypergraph H is *semi-algebraic* of complexity $\mathbf{t} = (d, D, m)$ if the vertices of H correspond to points in $\mathbb{R}^d$, and the edges of H are determined by the sign-pattern of m degree-D polynomials. Semi-algebraic hypergraphs of bounded complexity provide a general framework for studying geometrically defined hypergraphs.

The much-studied *semi-algebraic Ramsey number* $R_r^{\mathbf{t}}(s, n)$ denotes the smallest N such that every r-uniform semi-algebraic hypergraph of complexity $\mathbf{t}$ on N vertices contains either a clique of size s, or an independent set of size n. Conlon, Fox, Pach, Sudakov and Suk proved that $R_r^{\mathbf{t}}(n, n) < \mathrm{tw}_{r-1}(n^{O(1)})$, where $\mathrm{tw}_k(x)$ is a tower of 2's of height k with an x on the top. This bound is also the best possible if $\min\{d, D, m\}$ is sufficiently large with respect to r. They conjectured that in the asymmetric case, we have $R_3^{\mathbf{t}}(s, n) < n^{O(1)}$ for fixed s. We refute this conjecture by showing that $R_3^{\mathbf{t}}(4, n) > n^{(\log n)^{1/3-o(1)}}$ for some complexity $\mathbf{t}$.

In addition, motivated by the results of Bukh-Matoušek and Basit-Chernikov-Starchenko-Tao-Tran, we study the complexity of the Ramsey problem when the defining polynomials are linear, that is, when $D = 1$. In particular, we prove that $R_r^{d,1,m}(n, n) \leq 2^{O(n^{4r^2m^2})}$, while from below, we establish $R_r^{1,1,1}(n, n) \geq 2^{\Omega(n^{\lfloor r/2 \rfloor - 1})}$.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-087`

*Department of Mathematics, ETH Zürich, 8092 Zürich Switzerland. E-mail: `zhihan.jin@ifor.math.ethz.ch`.

†Department of Mathematics and Mathematical Statistics, Umeå University, 901 87 Umeå, Sweden. E-mail: `istvan.tomon@umu.se`. Supported by SNSF grant 200021_196965.

1 Introduction

Given positive integers r, s, n, the *Ramsey number* $R_r(s,n)$ denotes the smallest N such that every r-uniform hypergraph on N vertices contains either a clique of size s, or an independent set of size n. For convenience, we write $R_r(n)$ instead of $R_r(n,n)$. In the case of graphs, that is $r=2$, classical results of Erdős and Szekeres [15] and Erdős [12] tell us that $R_2(n) = 2^{\Theta(n)}$, and in case s is fixed and n is sufficiently large, we have $R_2(s,n) = n^{\Theta(s)}$. However, in case $r \geq 3$, the Ramsey numbers are less understood. Erdős and Rado [11] and Erdős, Hajnal and Rado [14] show that

$$\mathrm{tw}_{r-1}(\Omega(n^2)) < R_r(n) < \mathrm{tw}_r(O(n)).$$

Also, in the asymmetric case, we have $R_r(s,n) = \mathrm{tw}_{r-1}(n^{\Theta_{r,s}(1)})$ for $s \geq r+2$ [14, 18], and $R_3(4,n) = 2^{n^{\Theta(1)}}$ [10]. Here, $\mathrm{tw}_k(x)$ is the *tower function* defined as $\mathrm{tw}_1(x) := x$ and $\mathrm{tw}_k(x) := 2^{\mathrm{tw}_{k-1}(x)}$. Hence, there is an almost exponential gap between the lower and upper bound for $R_r(n)$ in case $r \geq 3$, and it is a major open problem to close this gap. Note that, however, the rough order of the asymmetric Ramsey number $R_r(s,n)$ is more understood, at least up to the height of the required tower. See [7] for recent developments.

Yet the situation changes if we restrict our attention to hypergraphs that arise from geometric considerations. To this end, an r-uniform hypergraph H is *semi-algebraic of complexity* (d, D, m) if the following holds. There is an enumeration $v_1, \dots, v_N$ of the vertices of H, an assignment $v_i \mapsto p_i$ with $p_i \in \mathbb{R}^d$ for $i \in [N]$, and m polynomials $f_1, \dots, f_m : (\mathbb{R}^d)^r \mapsto \mathbb{R}$ of (total) degree at most D such that for $1 \leq i_1 < \cdots < i_r \leq N$, whether $\{v_{i_1}, \dots, v_{i_r}\}$ is an edge of H depends only on the sign-pattern of $(f_1(p_{i_1}, \dots, p_{i_r}), \dots, f_m(p_{i_1}, \dots, p_{i_r}))$. More precisely, there is a function $\Phi : \{+, -, 0\}^m \mapsto \{\mathrm{True}, \mathrm{False}\}$ such that $\{v_{i_1}, \dots, v_{i_r}\}$ is an edge if and only if

$$\Phi(\mathrm{sign}(f_1(p_{i_1}, \dots, p_{i_r})), \dots, \mathrm{sign}(f_m(p_{i_1}, \dots, p_{i_r}))) = \mathrm{True}.$$

Semi-algebraic graphs and hypergraphs of bounded complexity provide a general model to study certain geometric structures, such as intersection and incidence graphs of geometric objects, order types of point configurations, convex subsets of the plane, and so on. The semi-algebraic Ramsey number $R_r^{\mathbf{t}}(s,n)$ denotes the smallest N such that any r-uniform semi-algebraic hypergraph of complexity $\mathbf{t}$ on N vertices contains either a clique of size s or an independent set of size n. Alon, Pach, Pinchasi, Radoičić and Sharir [1] proved that $R_2^{\mathbf{t}}(n) = n^{\Theta(1)}$, which was extended by Conlon, Fox, Pach, Sudakov and Suk [6] to $R_r^{\mathbf{t}}(n) = \mathrm{tw}_{r-1}(n^{O(1)})$ for general r. In [6] and [9], matching lower bounds are provided in case the parameters d, D, m are sufficiently large with respect to r. Specifically, for every $r \geq 2$, there exists $\mathbf{t}$ such that $R_r^{\mathbf{t}}(n) = \mathrm{tw}_{r-1}(n^{\Theta(1)})$. Here and later, the constants hidden by the $O(.), \Omega(.), \Theta(.)$ notation might depend on $r, \mathbf{t}$ and s, unless specified otherwise.

1.1 Asymmetric Ramsey numbers

In contrast, asymmetric semi-algebraic Ramsey numbers appear to be more mysterious in case $r \geq 3$. For uniformity $r = 3$, in the special subcase $d = 1$, it was established in [6] that $R_3^{\mathbf{t}}(s,n) < 2^{(\log n)^{O(1)}}$. Furthermore, if $d \geq 2$, a result of Suk [19] shows that

$$R_3^{\mathbf{t}}(s,n) < 2^{2^{(\log n)^{1/2+o(1)}}} = 2^{n^{o(1)}}.$$

However, the best known lower bound constructions provide only polynomial growth, which leads to the natural conjecture that $R_3^{\mathbf{t}}(s,n) = n^{O(1)}$, formulated in both [6] and [19]. Our first main result refutes this conjecture.

Theorem 1.1. *There exists $\mathbf{t} = (d, D, m)$ such that*

$$R_3^{\mathbf{t}}(4,n) > n^{(\log n)^{1/3-o(1)}}.$$

1.2 Semi-linear hypergraphs

As discussed above, if d, D, m are sufficiently large with respect to r, then $R_r^{d,D,m}(n) = \mathrm{tw}_{r-1}(n^{\Omega(1)})$. In the constructions provided by both [6] and [9], the parameters d and D grow with r. In particular, [9] shows that one can take $d = r-3$ for $r \geq 4$. Furthermore, the Veronese mapping[1] implies that every r-uniform semi-algebraic hypergraph of complexity (d, D, m) is also of complexity (d', r, m) for some d' depending only on d and D. However, this raises the question whether the upper bound $R_r^{d,D,m}(n) < \mathrm{tw}_{r-1}(n^{O(1)})$ can be significantly improved if we assume that d or D are small compared to r. In support of this, Bukh and Matoušek [4] showed that if $d = 1$, that is, when the vertices of the hypergraph correspond to points on the real line, then any r-uniform semi-algebraic hypergraph of complexity $(1, D, m)$ containing no clique or independent set of size n has at most $2^{2^{O(n)}}$ vertices (in [4], the constant hidden by the $O(.)$ notation might depend on the defining polynomials, but a careful inspection of their proof yields that it can be bounded only by a function of D, m and r as well). Also, this bound is the best possible if D and m are sufficiently large. In this paper, we consider what happens if we bound the parameter D instead, that is, the degrees of the defining polynomials.

A semi-algebraic hypergraph of complexity (d, D, m) is *semi-linear*, if $D = 1$, that is, all defining polynomials are linear functions[2] The study of semi-linear hypergraphs was initiated by Basit, Chernikov, Starchenko, Tao and Tran [3], who considered these hypergraphs in the setting of Zarankiewicz's problem. There are many extensively studied families of graphs that are semi-linear of bounded complexity, for example intersection graphs of axis-parallel boxes in $\mathbb{R}^d$, circle graphs, and shift graphs. Motivated by the large literature (e.g. [2, 5, 8, 13, 17]) concerned with the Ramsey properties of such families, Tomon [22] studied the Ramsey properties of semi-linear graphs and showed that $R_2^{d,1,m}(s,n) \leq n^{1+o(1)}$ holds for every fixed s, d and m. This already shows a behavior unique to semi-linearity, as a construction of Suk and Tomon [20] shows that $R_2^{d,2,m}(3,n) = \Omega(n^{4/3})$

[1] A Veronese mapping sends $(x_1, \ldots, x_d) \in \mathbb{R}^d$ to some point whose coordinates are monomials of $x_1, \ldots, x_d$. E.g. $(x_1, x_2, x_3) \mapsto (x_1^2 x_2, x_2^2 x_3^2, x_1 x_2 x_3, x_3^3)$.

[2] To clarify, e.g. $(x_1, x_2) \mapsto 2x_1 + 3x_2 + 5$ is a linear function, but $(x_1, x_2) \mapsto x_1 x_2 + 3x_1 + 3$ is *not* linear, it is only multi-linear.

for some d and m. Tomon [22] also proposed the problem of determining the Ramsey numbers of r-uniform semi-linear hypergraphs for $r \geq 3$. Our second main result settles this problem.

Theorem 1.2. *For every triple of positive integers r, d, m, there exists $c = c(r, m) > 0$ such that*

$$R_r^{d,1,m}(n) \leq 2^{cn^{4r^2m^2}}.$$

Let us highlight that the bound in Theorem 1.2 does not depend on the dimension d, only on the uniformity r and the number of polynomials m. From below, in case $r \geq 3$, the semi-linear Ramsey number grows at least exponentially, showing that Theorem 1.2 is sharp up to the value of c and the exponent $4r^2m^2$. Indeed, let H be the 3-uniform hypergraph on vertex set $\{1, \dots, N\}$ in which for $x < y < z$, $\{x, y, z\}$ is an edge if $x + z < 2y$. Then H is semi-linear of complexity $(1, 1, 1)$, and it is easy to show that $\omega(H), \alpha(H) \leq \lceil \log_2 N \rceil + 1$. Thus, $R_3^{1,1,1}(n) \geq 2^{\Omega(n)}$. We show that even faster growth can be achieved by examining certain more convoluted constructions of higher uniformity.

Theorem 1.3. *For every $r \geq 4$, there exists a constant $c > 0$ such that*

$$R_r^{1,1,1}(n) \geq 2^{cn^{\lfloor r/2 \rfloor - 1}}.$$

2 A lower bound for $R_3^{\mathbf{t}}(4, n)$

In this section, we outline the construction for Theorem 1.1, which builds on a variant of the famous stepping-up lemma of Erdős and Hajnal (see [16]).

Given distinct $\alpha, \beta \in \{0, 1\}^N$, let $\delta(\alpha, \beta) := \min\{i : \alpha(i) \neq \beta(i)\}$. Let $\prec$ be the lexicographical order over $\{0, 1\}^N$, i.e. $\alpha \prec \beta \Leftrightarrow \alpha(\delta(\alpha, \beta)) < \beta(\delta(\alpha, \beta))$. An important property of $\delta(\cdot, \cdot)$ is that for any $\alpha_1 \prec \cdots \prec \alpha_\ell$, there is a unique i which achieves the minimum of $\delta(\alpha_i, \alpha_{i+1})$.

Now we define our notion of the step-up.

Definition 1. The *step-up* of a graph G is the 3-uniform hypergraph H on vertex set $\{0, 1\}^N$ defined as follows. For $\alpha, \beta, \gamma \in \{0, 1\}^N$ with $\alpha \prec \beta \prec \gamma$, we have $\{\alpha, \beta, \gamma\} \in E(H)$ if and only if $\delta(\alpha, \beta) < \delta(\beta, \gamma)$ and $\{\delta(\alpha, \beta), \delta(\beta, \gamma)\} \in E(G)$.

The next lemma relates the clique and independence numbers of both graphs.

Lemma 2.1. $\omega(H) \leq \omega(G) + 1$ *and* $\alpha(H) \leq N^{\alpha(G)} + 1$.

By a construction of Suk and Tomon [20], there exists a semi-algebraic graph G on $\Theta(m^{4/3})$ vertices with $\omega(G) = 2$ and $\alpha(G) \leq 2m$ for all $m \in \mathbb{N}$. The methods in [6, 9] show that the step-up of G, denoted by H, remains semi-algebraic. Pick m such that $n = \Theta((m^{4/3})^{2m})$, i.e. $m = \Omega(\log n / \log \log n)$. Then, $|V(H)| = 2^{\Theta(m^{4/3})} = n^{\Omega((\log n)^{1/3 - o(1)})}$, $\omega(H) = 3$, $\alpha(H) \leq |V(G)|^{2m} + 1 < n$. This finishes the proof.

3 Semi-linear hypergraphs

In this section, we outline the proof of Theorem 1.2, i.e. $R_r^{d,1,m}(n) \leq 2^{O(n^{4r^2m^2})}$. Let H be an r-uniform semi-linear hypergraph on vertex set $[N]$ of complexity $(d,1,m)$. We observe that H is the Boolean combination of $2m$ semi-linear hypergraphs $H_1,\dots,H_{2m}$, where H_i is defined by a matrix $P_i \in \mathbb{R}^{r\times N}$ as follows: for $1 \leq q_1 < \dots < q_r \leq N$, we have $\{q_1,\dots,q_r\} \in H$ if and only if $\sum_{i=1}^r P(i,q_i) < 0$. Therefore, our goal is to find $C \subseteq [N]$ of size $(\log N)^{\Omega_{r,m}(1)}$ such that for each $i \in [2m]$, C is either a clique or an independent set in H_i. We will find such a C by trimming and transforming our matrices in several steps.

For $\delta \in \mathbb{R}$, the *shift of a sequence* $x_1,\dots,x_N$ *by* δ is the sequence $x_1+\delta,\dots,x_N+\delta$. Given $\Delta > 1$ and $\tau \in \{-,+\}\times\{\searrow,\nearrow\}$, a sequence $x_1,\dots,x_N$ is called (Δ,τ)-*exponential* if for all $i \in [N-1]$, we have $0 < x_i < x_{i+1}/\Delta$ in the case $\tau = (+,\nearrow)$; $0 < x_{i+1} < x_i/\Delta$ in the case $\tau = (+,\searrow)$; $0 < (-x_i) < (-x_{i+1}/\Delta)$ in the case $\tau = (-,\nearrow)$; $0 < (-x_{i+1}) < (-x_i/\Delta)$ in the case $\tau = (-,\searrow)$. Also, say that a sequence is Δ-*exponential* if it is (Δ,τ)-exponential for some $\tau \in \{-,+\}\times\{\searrow,\nearrow\}$.

Lemma 3.1. *For every q there exists $c = c(q) > 0$ such that the following holds. Let $M \in \mathbb{R}^{q\times N}$ be a matrix such that no row contains repeated elements. Then M contains a $q \times N'$ sized submatrix for $N' = c(\log N)^{1/q}$ such that every row of M is $2q$-exponential.*

Applying Lemma 3.1 to the concatenation of the matrices P_ℓ for $\ell \in [2m]$, we find a subset $I \subset [N]$ of size $N' = c(\log N)^{1/2mr}$ (say $I' = [N']$) such that for each ℓ, the submatrix P'_ℓ of P_ℓ induced by columns in I is $2r$-exponential. Let H' be the subgraph of H induced by I, and H'_i be the subgraph of H'_i induced by I. It is easy to see that every H'_ℓ is defined by P'_ℓ, and that H' is a Boolean combination of $H'_1,\dots,H'_{2m}$. Theorem 1.2 is then an easy consequence of the following key lemma, which guarantees $C \subset I$ of size $(N')^{\Omega_{r,m}(1)}$ such that for all $i \in [2m]$, C is either a clique or an independent set in H'_i.

Lemma 3.2. *For every r and k, there exists $c = c(r,k) > 0$ such that the following holds. Let $P_1,\dots,P_k$ be $r \times N$ matrices where all rows are $2r$-exponential. Then there exists $C \subset [N]$ such that $|C| \geq cN^{\frac{1}{rk-k+1}}$ and C is a clique or an independent set in H_i for every $i \in [k]$.*

The proof of this lemma builds on the following observation. Using that each row of P_ℓ is $(2r)$-exponential, whether $\{q_1,\dots,q_r\}$ is an edge of H_ℓ depends (essentially) on the maximum of $H_\ell(1,q_1),\dots,H_\ell(r,q_r)$. Further details are omitted.

References

[1] Noga Alon, János Pach, Rom Pinchasi, Radoš Radoičić, and Micha Sharir. Crossing patterns of semi-algebraic sets. *Journal of Combinatorial Theory, Series A*, 111(2):310–326, 2005.

[2] Edgar Asplund and B Grünbaum. On a coloring problem. *Mathematica Scandinavica*, 8(1):181–188, 1960.

[3] Abdul Basit, Artem Chernikov, Sergei Starchenko, Terence Tao, and Chieu-Minh Tran. Zarankiewicz's problem for semilinear hypergraphs. *Forum Math. Sigma*, 9:Paper No. e59, 23, 2021.

[4] Boris Bukh and Jiří Matoušek. Erdős-Szekeres-type statements: Ramsey function and decidability in dimension 1. *Duke Mathematical Journal*, 163(12):2243–2270, 2014.

[5] Parinya Chalermsook and Bartosz Walczak. Coloring and maximum weight independent set of rectangles. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 860–868. SIAM, 2021.

[6] David Conlon, Jacob Fox, János Pach, Benny Sudakov, and Andrew Suk. Ramsey-type results for semi-algebraic relations. *Transactions of the American Mathematical Society*, 366(9):5043–5065, 2014.

[7] David Conlon, Jacob Fox, and Benny Sudakov. Hypergraph ramsey numbers. *Journal of the American Mathematical Society*, 23(1):247–266, 2010.

[8] James Davies and Rose McCarty. Circle graphs are quadratically χ-bounded. *Bull. London Math. Society*, 53(3):673–679, 2021.

[9] Marek Eliáš, Jiří Matoušek, Edgardo Roldán-Pensado, and Zuzana Safernová. Lower bounds on geometric ramsey functions. *SIAM Journal on Discrete Mathematics*, 28(4):1960–1970, 2014.

[10] Paul Erdős and András Hajnal. On ramsey like theorems. problems and results. In *Combinatorics (Proc. Conf. Combinatorial Math., Math. Inst., Oxford, 1972)*, pages 123–140. Citeseer, 1972.

[11] Paul Erdős and Richard Rado. Combinatorial theorems on classifications of subsets of a given set. *Proceedings of the London mathematical Society*, 3(1):417–439, 1952.

[12] Paul Erdős. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53(4):292–294, 1947.

[13] Paul Erdős and András Hajnal. Some remarks on set theory. ix. combinatorial problems in measure theory and set theory. *Michigan Mathematical Journal*, 11(2):107–127, 1964.

[14] Paul Erdős, András Hajnal, and Richard Rado. Partition relations for cardinal numbers. *Acta Math. Acad. Sci. Hungar*, 16:93–196, 1965.

[15] Paul Erdős and George Szekeres. A combinatorial problem in geometry. *Compositio mathematica*, 2:463–470, 1935.

[16] Ronald L Graham, Bruce L Rothschild, and Joel H Spencer. *Ramsey theory*, volume 20. John Wiley & Sons, 1991.

[17] David Larman, Jiří Matoušek, János Pach, and Jenő Törőcsik. A Ramsey-type result for convex sets. *Bull. London Math. Soc.*, 26(2):132–136, 1994.

[18] Dhruv Mubayi and Andrew Suk. New lower bounds for hypergraph ramsey numbers. *Bulletin of the London Mathematical Society*, 50(2):189–201, 2018.

[19] Andrew Suk. Semi-algebraic ramsey numbers. *Journal of Combinatorial Theory, Series B*, 116:465–483, 2016.

[20] Andrew Suk and István Tomon. Hasse diagrams with large chromatic number. *Bulletin of the London Mathematical Society*, 53, 2021.

[21] Endre Szemerédi and William T. Trotter. Extremal problems in discrete geometry. *Combinatorica*, 3:381–392, 1983.

[22] István Tomon. Ramsey properties of semilinear graphs. *Israel Journal of Mathematics*, pages 1–27, 2022.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Perfect-matching covers of cubic graphs with colouring defect 3

(Extended abstract)

Ján Karabáš* Edita Máčajová† Roman Nedela‡ Martin Škoviera§

Abstract

The colouring defect of a cubic graph is the smallest number of edges left uncovered by any set of three perfect matchings. While 3-edge-colourable graphs have defect 0, those that cannot be 3-edge-coloured have defect at least 3. We show that every bridgeless cubic graph with defect 3 can have its edges covered with at most five perfect matchings, which verifies a long-standing conjecture of Berge for this class of graphs. Moreover, we determine the extremal graphs.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-088`

1 Introduction

A strong form of Petersen's Perfect Matching Theorem [15] states that each edge of a bridgeless cubic graph G is contained in a perfect matching. The minimum number of perfect matchings needed to cover all the edges of G is its *perfect matching index*, denoted by $\pi(G)$. In 1970's, Berge conjectured (unpublished, see [3, 10, 16]) that $\pi(G) \leq 5$ for

*Faculty of Natural Sciences, Matej Bel University, Banská Bystrica, SK and Mathematical Institute of Slovak Academy of Sciences, Banská Bystrica, SK. E-mail: `jan.karabas@umb.sk`. Supported by grants APVV-19-0308 and VEGA 2/0078/20.

†Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, SK. E-mail: `macajova@dcs.fmph.uniba.sk`. Supported by grants APVV-19-0308 and VEGA 1/0743/21.

‡Faculty of Applied Sciences, University of West Bohemia, Pilsen, CZ and Mathematical Institute of Slovak Academy of Sciences, Banská Bystrica, SK. E-mail: `nedela@savbb.sk`. Supported by grants APVV-19-0308 and VEGA 2/0078/20.

§Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, SK. E-mail: `skoviera@dcs.fmph.uniba.sk`. Supported grants APVV-19-0308 and VEGA 1/0727/22.

every bridgeless cubic graph G. After more than 50 years, this conjecture remains widely open. In fact, very little is known unless the graph in question has a very specific structure, see for example [1, 4, 5, 9]).

In this paper we investigate perfect-matching covers of bridgeless cubic graphs that are close to 3-edge-colourable cubic graphs. If a cubic graph G can be 3-edge-coloured, then obviously $\pi(G) = 3$, and $\pi(G) \geq 4$ otherwise. If G cannot be 3-edge-coloured, then any set $\{M_1, M_2, M_3\}$ of three perfect matchings of G leaves some edges uncovered. The minimum number of uncovered edges is the *colouring defect* of G, denoted by $\mathrm{df}(G)$. This concept was introduced and extensively studied by Steffen et al. in [8, 17]. Together with oddness, resistance, perfect matching index, and other similar invariants it serves as one of measures of uncolourability of cubic graphs [2].

Steffen [17] showed that the colouring defect of every non-3-edge-colourable cubic graph (henceforth just *defect*, for short) is at least 3. Cubic graphs with defect 3 thus constitute a class of cubic graphs that is in a certain sense closest to 3-edge-colourable graphs. The purpose of this paper is to show that Berge's conjecture holds for this class of cubic graphs and to characterise the extremal graphs where five perfect matchings are actually necessary. Our main result reads as follows.

Theorem 1.1. *Every bridgeless cubic graph G of defect 3 can have its edges covered with at most five perfect matchings; that is, $4 \leq \pi(G) \leq 5$. If G is 3-connected, then $\pi(G) = 5$ if and only if G arises from the Petersen graph by inflating any number of vertices of a fixed vertex-star (possibly zero) by quasi-bipartite cubic graphs in a correct way.*

For cubic graphs with defect 3 this result significantly improves the result of Steffen [17, Theorem 2.14] which states that every cyclically 4-edge-connected cubic graph with defect 3 or 4 satisfies Berge's conjecture.

2 Auxiliary results

The proof of Theorem 1.1 will be executed in several steps and will use a number of tools. One of key ingredients, applied several times and at various stages of the proof, is the following theorem which explores 6-edge-cuts in cubic graphs. Given a subgraph H of a graph G, let $\delta_G(H)$ denote the edge-cut comprising all edges with exactly one end in H.

Theorem 2.1. *Let G be a bridgeless cubic graph and let $H \subseteq G$ be a subgraph with $|\delta_G(H)| = 6$. Then H has a perfect matching, or else H contains an independent set S of trivalent vertices such that*

(i) *the number of components of $H - S$ equals $|S| + 2$, and*

(ii) *every component L of $H - S$ has $|\delta_G(L)| = 3$.*

A bridgeless cubic graph Q will be called *quasi-bipartite* if it contains an independent set of vertices U such that the graph obtained by the contraction of each component of

$Q - U$ to a vertex is a cubic bipartite graph where U is one of the partite sets. Roughly speaking, a quasi-bipartite cubic graph arises from a bipartite cubic graph by inflating certain vertices in one of the partite sets to larger subgraphs, while preserving the edges between the partite sets. The previous theorem thus implies that if we add two new vertices u and v to H and create a cubic graph H^+ from H by attaching the edges of $\delta_G(H)$ to u and v, then H^+ becomes a quasi-bipartite with the independent set $U = S \cup \{u, v\}$.

The second auxiliary result is also related to bipartite graphs. A cubic graph G is said to be *almost bipartite* if it is bridgeless, not bipartite, and contains two edges e and f such that $G - \{e, f\}$ is a bipartite graph. The edges e and f are the *surplus edges* of G. Observe that if a cubic graph G is almost bipartite, then it has a component such that e connects vertices within one partite set and f connects vertices within the other partite set. Moreover, it can be shown that G has a perfect matching that contains both surplus edges. As a consequence, we obtain the following.

Theorem 2.2. *Every almost bipartite cubic graph is* 3*-edge-colourable.*

The *bipartite index* of a graph G is defined to be the smallest number of edges that must be deleted in order to make the graph bipartite. The previous theorem implies that every bridgeless cubic graph with bipartite index at most 2 is 3-edge-colourable. On the other hand, there exist infinitely many snarks whose bipartite index equals 3, for example the Isaacs flower snarks [6]. In this sense, Theorem 2.2 is best possible.

3 Berge covers for cubic graphs of defect 3

A *Berge cover* of a cubic graph G is a collection of five perfect matchings that cover all the edges of G. To find such a cover for a graph of defect 3 we employ a structure created by three perfect matchings. For a bridgeless cubic graph G we define an *optimal* 3*-array* of perfect matchings to be any set $\mathcal{M} = \{M_1, M_2, M_3\}$ of three perfect matchings such that the number of edges not covered by $M_1 \cup M_2 \cup M_3$ equals the defect of G. The *core* of $\mathcal{M}$ is the subgraph of G induced by the set of all edges that are not simply covered by $\mathcal{M}$. It is not difficult to see that if $\mathrm{df}(G) = 3$, then the core of $\mathcal{M}$ is a chordless hexagon which alternates the uncovered edges with the doubly covered ones [17]. If G is the Petersen graph, then any hexagon can be taken as the core of a suitable optimal 3-array. In particular, the defect of the Petersen graph equals 3.

To prove the first statement of Theorem 1.1 we show that every optimal 3-array $\mathcal{M}$ for a cubic graph G with $\mathrm{df}(G) = 3$ extends to a Berge cover. The key step towards the proof is the next lemma. At the crucial moment of its proof we apply Theorem 2.1 to the 6-edge-cut $\delta_G(W)$ where W is a suitable path of length 3 lying in the core of $\mathcal{M}$.

Lemma 3.1. *Let G be a bridgeless cubic graph of defect* 3 *and let $\mathcal{M}$ be an optimal* 3*-array of perfect matchings of G. Then G has a fourth perfect matching which covers at least two of the three edges left uncovered by $\mathcal{M}$.*

With the help of Lemme 3.1 we can prove the following.

Theorem 3.2. *Every bridgeless cubic graph with defect* 3 *has a Berge cover.*

Proof. Assume that $\mathrm{df}(G) = 3$, and let $\mathcal{M}$ be an arbitrary optimal array for G. Let M_4 be a perfect matching guaranteed by Lemma 3.1, which covers at least two of the uncovered edges. Since G has perfect matching that covers any preassigned edge, we can take a perfect matching M_5 that covers the third uncovered edge. Clearly, $\mathcal{M} \cup \{M_4, M_5\}$ is a Berge cover of G. □

4 Cyclically 4-edge-connected graphs

Our main result restricted to cyclically 4-edge-connected graphs reads as follows.

Theorem 4.1. *Let G be a cyclically* 4*-edge-connected cubic graph with defect* 3*. Then* $\pi(G) = 4$*, unless G is the Petersen graph.*

Proof (sketch). We prove that if $\pi(G) \geq 5$, then G is the Petersen graph. Take an optimal 3-array $\mathcal{M}$ for G whose core is a 6-cycle $C = (v_0v_1 \dots v_5)$, and set $H = G - V(C)$. If H had a perfect matching, we could extend it to a perfect matching M_4 of the entire G in such a way that $\mathcal{M} \cup \{M_4\}$ covers all the edges of G, implying that $\pi(G) = 4$. Therefore H has no perfect matching, and we can apply Theorem 2.1 to the edge-cut $\delta_G(H)$. Let $S \subseteq V(H)$ be the independent set of trivalent vertices stated in Theorem 2.1. Then each component of $H - S$ is a single vertex, because G is 4-edge-connected. It follows that H is bipartite, and therefore 3-edge-colourable.

We now investigate the 6-tuples of colours on $\delta_G(H)$ induced by 3-edge-colourings of H, ordered cyclically around C. It is easy to see that all three colours must always occur, otherwise the missing colour could be extended to a perfect matching M_4 of G, yielding a contradiction as before. There remain 15 colouring types for $\delta_G(H)$ of which 7 are excluded because they would enable a 3-edge-colouring of G.

For $i \in \{0, \dots, 5\}$ let u_i be the neighbour of v_i lying in H. We claim that $u_i = u_j$ whenever $j \equiv i + 3 \pmod 6$. If $u_i = u_j$, then indeed $j \equiv i + 3 \pmod 6$, otherwise G would have a triangle or a quadrilateral intersecting C. The former possibility cannot occur due to cyclic connectivity. In the latter case, the quadrilateral would share two edges with C, in which case $\mathcal{M}$ could be modified to a 3-edge-colouring of G, a contradiction. Suppose that there exist vertices u_i and u_j such that $u_i \neq u_j$ and $j \equiv i + 3 \pmod 6$, say $u_2 \neq u_5$. Create a cubic graph $H^\sharp$ from H as follows: add two new vertices s and t, connect them between themselves and to $\{u_0, u_1, u_3, u_4\}$, and finally join u_2 to u_5. This can be done in such a way that no 3-edge-colouring of H extends to $H^\sharp$, implying that $H^\sharp$ is not 3-edge-colourable. However, $H^\sharp$ is almost bipartite, which contradicts Theorem 2.2. Therefore $u_0 = u_3$, $u_1 = u_4$, and $u_2 = u_5$. It follows that $\delta_G(C \cup \{u_0, u_1, u_2\})$ is a 3-edge-cut, which must be trivial due to cyclic connectivity. Hence G has $6 + 3 + 1$ vertices, and this means that G is the Petersen graph. □

5 General case of Theorem 1.1

To move away from cyclically 4-edge-connected graphs we modify the classical method of snark reduction to cubic graphs of defect 3. By a *snark* we mean a 2-connected cubic graph that admits no 3-edge-colouring. A snark is *nontrivial* if it is cyclically 4-edge-connected with girth at least 5. It is well known that every snark can be transformed to a nontrivial snark by a sequence of certain simple reductions (like contracting a triangle). Performing a reduction of a snark G means to identify an edge-cut R in G whose removal leaves a component H which is not 3-edge-colourable. By adding a small number of vertices or edges it is possible to extend H to a snark G', a *reduction* of G along R.

A reduction of a snark G with defect 3 to a nontrivial snark G' of defect 3 may not always be possible. Such a situation occurs, for example, when G contains an *essential triangle*, one whose contraction produces a snark with defect greater than 3. It can be shown that the increase of defect by contracting an essential triangle can be arbitrarily large. Nevertheless, a snark with defect 3 can have at most one essential triangle, and if so, then it is the only obstruction to reduction.

Theorem 5.1. *Every snark G with $\mathrm{df}(G) = 3$ admits a reduction to a snark G' with $\mathrm{df}(G') = 3$ such that either G' is nontrivial or G' arises from a nontrivial snark K with $\mathrm{df}(K) \geq 4$ by inflating a vertex to a triangle; the triangle is essential in both G and G'.*

The proof of this theorem is quite involved and requires a careful analysis of Fano flows associated with 3-arrays (for the definition of a Fano flow see [7]).

Reductions can be conveniently handled with the help of two well-known operations. Let G and H be cubic graphs with distinguished edges e and f, respectively. We define a 2-*sum* $G \oplus_2 H$ to be a cubic graph obtained by deleting e and f and connecting the 2-valent vertices of G to those of H. If instead of distinguished edges we have distinguished vertices u and v of G and H, respectively, we can similarly define a 3-sum $G \oplus_3 H$. Note that $G \oplus_3 H$ can be regarded as being obtained from G by *inflating* the vertex u to $H - v$.

A cubic graph G containing a cycle-separating 2-edge-cut or 3-edge-cut can be expressed as $G \oplus_2 H$ or $G \oplus_3 H$ uniquely, only depending on the chosen edge-cut. It is easy to see that if two 2-cuts or 3-cuts intersect, the result of decomposition does not depend on the order in which the cuts are taken. As a consequence, we have the following.

Theorem 5.2. *Every 2-connected cubic graph G admits a decomposition into a collection $\{G_1, \ldots, G_m\}$ of cyclically 4-edge-connected cubic graphs such that G can be reconstructed from them by a repeated application of 2-sums and 3-sums. Moreover, this collection is unique up to ordering and isomorphism.*

The first step in the proof of the general case of Theorem 1.1 is to show that, somewhat surprisingly, cubic graphs with defect 3 containing an essential triangle behave nicely.

Theorem 5.3. *If a cubic graph G with defect 3 has an essential triangle, then $\pi(G) = 4$.*

Proof (sketch). Let T be an essential triangle of G and let $\mathcal{M}$ be an optimal 3-array for G with hexagonal core C. Let w be the unique neighbour of T not lying on C. First assume that G/T is cyclically 4-edge-connected. We claim that G has a perfect matching M_4 such that $\mathcal{M} \cup \{M_4\}$ covers all the edges of G. If not, we apply Theorem 2.1 to the 6-cut $\delta_G(C \cup T \cup \{w\})$ and with the help of Theorem 2.2 we derive a contradiction in a similar manner as in the proof of Theorem 4.1. If G is not cyclically 4-edge-connected, we contract T to a vertex t and decompose G/T to $K_1, \ldots, K_m$ according to Theorem 5.2. Exactly one K_i, say K_1, contains the vertex t. We inflate t back to T, transforming K_1 to a graph L. Note that C survives the decomposition of G intact, so T is an essential triangle of L. As L/T is cyclically 4-edge-connected, Theorem 4.1 implies that $\pi(L/T) = 4$ or L/T is the Petersen graph. In both cases L has a cover with four perfect matchings. By using 2-sums and 3-sums this cover can be extended to a cover of the entire G. □

Let G and H be 2-connected cubic graphs where H is quasi-bipartite with independent set U. We say that a 3-sum $G \oplus_3 H$ is *correct* if the distinguished vertex of H forms a trivial component of $H - U$. Note that the result of a correct 3-sum is again quasi-bipartite.

Theorem 5.4. *Let G and H be 2-connected cubic graphs where $\pi(G) \geq 5$ and H is 3-edge-colourable. Then $\pi(G \oplus_3 H) \geq 5$ if and only if H is quasi-bipartite and the 3-sum is correct.*

Our intention is to characterise all 2-connected cubic graphs G with $\mathrm{df}(G) = 3$ and $\pi(G) = 5$. If G has a 2-edge-cut, then G can be expressed as $G' \oplus_2 H$. Since no hexagonal core can intersect a 2-edge-cut, the core stays within one summand, say G'. We conclude that $\mathrm{df}(G') = 3$, $\pi(G') = 5$, and that H is 3-edge-colourable. It follows that it is enough to characterise 3-connected cubic graphs with $\mathrm{df}(G) = 3$ and $\pi(G) = 5$. This is done in the next theorem, whose proof concludes that of Theorem 1.1.

Theorem 5.5. *Let G be a 3-connected cubic graph with $\mathrm{df}(G) = 3$. Then $\pi(G) = 5$ if and only if G arises from the Petersen graph by inflating any number of vertices of a fixed vertex-star by quasi-bipartite cubic graphs in a correct way.*

Proof (sketch). Assume that $\pi(G) = 5$. The statement is clearly true if G is cyclically 4-edge-connected, so we may assume that $H_0 = G$ can be expressed as a 3-sum $H_1 \oplus_3 H_1'$. By Theorem 5.3, G has no essential triangle, so every hexagonal core of G survives in one of the summands, say H_1. By Theorem 5.4, $\pi(H_1) = 5$, H_1' is quasi-bipartite, and the 3-sum is correct. We now continue with the decomposition by applying Theorem 5.4 to H_1, and so forth. Eventually, we obtain a collection $\{G_1, \ldots, G_m\}$ of cyclically 4-edge-connected cubic graphs exactly one of which, say G_1, is a snark, which has $\mathrm{df}(G_1) = 3$ and $\pi(G_2) = 5$. By Theorem 4.1, G_1 is the Petersen graph. It means that G arises from G_1 by a repeated correct 3-sum with a number of quasi-bipartite graphs. Since a fixed hexagon $C \subseteq G_1$ must survive the summation as a core, only the four vertices of $G - V(C)$, forming a vertex-star complementary to C, are eligible as distinguished vertices for 3-sums. Thus G has the structure which is described in Theorem 1.1. The reverse implication proceeds along similar lines. □

6 Final remarks

This paper summarises results presented in several papers at various stages of writing. Full proofs of Theorems 2.1-2.2, Theorem 3.2, and Theorem 4.1 can be found in [12], which is available on arXiv. Theorem 5.1 and the fact that the contraction of an essential triangle can increase defect arbitrarily are proved in [13]. The latter result heavily depends on results proved in [14, Theorems 5.1-5.2]. Finally, Theorems 5.3-5.5 and Theorem 1.1 will be proved in [11].

References

[1] F. Chen and G. Fan. Fulkerson-covers of hypohamiltonian graphs. *Discrete Appl. Math.*, 186:66–73, 2015.

[2] M. A. Fiol, G. Mazzuoccolo, and E. Steffen. Measures of edge-uncolorability of cubic graphs. *Electron. J. Combin.*, 25(P4.54), 2018.

[3] J.-L. Fouquet and J.-M. Vanherpe. On Fulkerson conjecture. *Discuss. Math. Graph Theory*, 31:253–272, 2011.

[4] R.-X. Hao, X. Wang J. Niu, C.-Q. Zhang, and T. Zhang. A note on Berge-Fulkerson coloring. *Discrete Math.*, 309:4235–4240, 2009.

[5] R.-X. Hao, C.-Q. Zhang, and T. Zheng. Berge-Fulkerson coloring for $C_{(8)}$-linked graphs. *J. Graph Theory*, 88:46–60, 2018.

[6] R. Isaacs. Infinite families of non-trivial trivalent graphs which are not Tait colorable. *Amer. Math. Monthly*, 82:221–239, 1975.

[7] L. Jin, G. Mazzuoccolo, and E. Steffen. Cores, joins and the Fano-flow conjectures. *Discuss. Math. Graph Theory*, 38:165–175, 2018.

[8] L. Jin and E. Steffen. Petersen cores and the oddness of cubic graphs. *J. Graph Theory*, 84:109–120, 2017.

[9] S. Liu, R.-X. Hao, and C.-Q. Zhang. Rotation snark, Berge-Fulkerson conjecture and Catlin's 4-flow reduction. *Appl. Math. Comput.*, 410:126441, 2021.

[10] G. Mazzuoccolo. The equivalence of two conjectures of Berge and Fulkerson. *J. Graph Theory*, 68:125–128, 2011.

[11] J. Karabáš, E. Máčajová, R. Nedela, and M. Škoviera. Berge covers of cubic graphs with colouring defect 3. in preparation, 2022.

[12] J. Karabáš, E. Máčajová, R. Nedela, and M. Škoviera. Berge's conjecture for cubic graphs with small colouring defect. `arXiv:2210.13234 [math.CO]`, 2022.

[13] J. Karabáš, E. Máčajová, R. Nedela, and M. Škoviera. Cubic graphs with colouring defect 3. manuscript, 2022.

[14] J. Karabáš, E. Máčajová, R. Nedela, and M. Škoviera. Girth, oddness, and colouring defect of snarks. *Discrete Math.*, 345:113040, 2022.

[15] T. Schönberger. Ein Beweis des Petersenschen Graphensatzes. *Acta Litt. Sci. Szeged*, 7:51–57, 1934.

[16] P. D. Seymour. On multicolourings of cubic graphs, and conjectures of Fulkerson and Tutte. *London. Math. Soc.*, 38:423–460, 1979.

[17] E. Steffen. 1-Factor and cycle covers of cubic graphs. *J. Graph Theory*, 78:195–206, 2015.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

High-rank subtensors of high-rank tensors

(Extended abstract)

Thomas Karam *

Abstract

Let $d \geq 2$ be a positive integer. We show that for a class of notions R of rank for order-d tensors, which includes in particular the tensor rank, the slice rank and the partition rank, there exist functions $F_{d,R}$ and $G_{d,R}$ such that if an order-d tensor has R-rank at least $G_{d,R}(l)$ then we can restrict its entries to a product of sets $X_1 \times \cdots \times X_d$ such that the restriction has R-rank at least l and the sets $X_1, \ldots, X_d$ each have size at most $F_{d,R}(l)$. Furthermore, our proof methods allow us to show that under a very natural condition we can require the sets $X_1, \ldots, X_d$ to be pairwise disjoint.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-089`

The results described below are proved and discussed further in the paper [5].

1 Main results

The last few years have seen a sequence of successes in using notions of ranks for higher-dimensional tensors to solve combinatorial problems. A central idea from the breakthrough solution to the cap-set problem by Ellenberg and Gijswijt [2], which was based on a technique of Croot, Lev, and Pach [1], was reformulated by Tao [11] in terms of the notion of slice rank for tensors, leading to what is now known as the slice rank polynomial method. The slice rank was further studied by Sawin and Tao [10], and bounds shown there on

*Mathematical Institute, University of Oxford. E-mail: `thomas.karam@maths.ox.ac.uk`. Supported by ERC grant 883810.

the slice rank involving orderings on the coordinates were later used by Sauermann [9] to prove under suitable conditions the existence of solutions with pairwise distinct variables to systems of equations in subsets of $\mathbb{F}_p^n$ that are not exponentially sparse. Another fruitful generalisation of the idea underlying the slice rank has been the partition rank, which was defined by Naslund [8] in order to prove a polynomial upper bound on the size of subsets of $\mathbb{F}_{p^r}^n$ not containing any k-right corners (with p a prime integer and $r \geq 1$ a positive integer) and very recently used again by Naslund [7] to prove exponential lower bounds on the chromatic number of $\mathbb{R}^n$ with multiple forbidden distances.

We will focus on high-rank subtensors of tensors: it is a standard fact from linear algebra that if A is a matrix of rank k then A has a $k \times k$ submatrix with rank k, and we will study here the extent to which this statement can be generalised to notions of rank for higher-order tensors, in particular to the tensor rank, to the slice rank and to the partition rank. The results that we obtain in this direction as well as the methods that we use in their proofs will also allow us to prove that under a very natural assumption we can find a subtensor such that the coordinates take values in pairwise disjoint sets. As we explain in a few paragraphs, the formulation of this second result also arises naturally as an analogue of the standard inequality that every oriented graph has a bipartition such that at least a quarter of the edges go from the first part to the second.

We now define the relevant notions of higher-dimensional ranks for tensors and state our main theorems.

Definition 1. *Let $d \geq 2$ be a positive integer and let $\mathbb{F}$ be a field. An* order-d tensor *over $\mathbb{F}$ is a function $T : Q_1 \times \cdots \times Q_d \to \mathbb{F}$ for some finite subsets $Q_1, \dots, Q_d$ of $\mathbb{N}$.*

Throughout we shall use the following notation. We write $\mathbb{F}$ for an arbitrary field, and all our statements will hold uniformly in $\mathbb{F}$. If $d \geq 2$ is a positive integer, then $Q_1, \dots, Q_d$ will always stand for finite subsets of $\mathbb{N}$. Given an order-d tensor $T : Q_1 \times \cdots \times Q_d \to \mathbb{F}$ and subsets $X_1 \subset Q_1, \dots, X_d \subset Q_d$, we shall write $T(X_1 \times \cdots \times X_d)$ for the restriction $X_1 \times \cdots \times X_d \to \mathbb{F}$ of T. For each positive integer n we write $[n]$ for the set $\{1, 2, \dots, n\}$. Given $x \in Q_1 \times \cdots \times Q_d$, and $I \subset [d]$, we write $x(I)$ for the restriction $(x_\alpha : \alpha \in I)$ of x to its coordinates in I.

Definition 2. *Let $d \geq 2$ be a positive integer, and let T be an order-d tensor. We say that T has* tensor rank at most 1 *if there exist functions $a_\alpha : Q_\alpha \to \mathbb{F}$ for each $\alpha \in [d]$ such that*

$$T(x_1, \dots, x_d) = a_1(x_1) \dots a_d(x_d)$$

for every $(x_1, \dots, x_d) \in Q_1 \times \cdots \times Q_d$.

We say that T has slice rank at most 1 *if there exist $\alpha \in [d]$ and functions $a : Q_\alpha \to \mathbb{F}$ and $b : \prod_{\alpha' \in [d], \alpha' \neq \alpha} Q_{\alpha'} \to \mathbb{F}$ such that we can write*

$$T(x_1, \dots, x_d) = a(x_\alpha) b(x_1, \dots, x_{\alpha-1}, x_{\alpha+1}, \dots, x_d)$$

for every $(x_1, \dots, x_d) \in Q_1 \times \cdots \times Q_d$.

We say that T has partition rank at most 1 *if there exist a bipartition $\{I, J\}$ of $[d]$ with I, J both non-empty and functions $a : \prod_{\alpha \in I} Q_\alpha \to \mathbb{F}$ and $b : \prod_{\alpha \in J} Q_\alpha \to \mathbb{F}$ such that we can write*

$$T(x_1, \dots, x_d) = a(x(I))b(x(J))$$

for every $(x_1, \dots, x_d) \in Q_1 \times \cdots \times Q_d$.

We say that the tensor rank *(resp.* slice rank*, resp.* partition rank*) of T is the smallest nonnegative integer k such that there exist tensors $T_1, \dots, T_k$ each of tensor rank at most 1 (resp. slice rank at most 1, resp. partition rank at most 1) and such that $T = T_1 + \cdots + T_k$. We denote by $\operatorname{tr} T$ the tensor rank of T, by $\operatorname{sr} T$ the slice rank of T, and by $\operatorname{pr} T$ the partition rank of T.*

We will begin by showing the fact that every matrix of rank k has a $k \times k$ subtensor with rank k generalises in the best way one could hope for to the tensor rank for all $d \geq 2$: every order-d tensor T with tensor rank k has a $k \times k \times \cdots \times k$ (d times) subtensor with tensor rank k. However, that becomes false for the order-3 slice rank: we thank Timothy Gowers for constructing a counterexample. It will nonetheless be true that if an order-3 tensor is such that all its subtensors with size at most $48l^3$ have slice rank at most l then the whole tensor has slice rank at most $51l^3$. Finally we will show that such an asymptotic subtensors property holds for the slice and partition rank for all $d \geq 2$ as well as for a more general class of notions of rank which we now define before stating this asymptotic result.

Definition 3. *Let $d \geq 2$ be a positive integer, and let R be a non-empty family of partitions of $[d]$. We say that an order-d tensor T has R-*rank at most 1 *if there exist a partition $P \in R$ and for each $I \in P$ a function $a_I : \prod_{\alpha \in I} Q_\alpha \to \mathbb{F}$ such that we can write*

$$T(x_1, \dots, x_d) = \prod_{I \in P} a_I(x(I))$$

*for every $(x_1, \dots, x_d) \in Q_1 \times \cdots \times Q_d$. We say that the R-*rank *of T is the smallest nonnegative integer k such that there exist order-d tensors $T_1, \dots, T_k$ with R-rank at most 1 such that $T = T_1 + \cdots + T_k$.*

We will denote by $R\mathrm{rk}\, T$ the R-rank of T. We can check that for every $d \geq 2$, the R-rank specialises to the tensor rank, to the slice rank, and to the partition rank.

We are now in a position to state our first main theorem.

Theorem 4. *Let $d \geq 2$ be a positive integer, and let R be a non-empty family of partitions of $[d]$. There exist functions $F_{d,R} : \mathbb{N} \to \mathbb{N}$ and $G_{d,R} : \mathbb{N} \to \mathbb{N}$ such that if T is an order-d tensor with $R\mathrm{rk}\, T \geq G_{d,R}(l)$ then there exist $X_1 \subset Q_1, \dots, X_d \subset Q_d$ each with size at most $F_{d,R}(l)$ such that $R\mathrm{rk}\, T(X_1 \times \cdots \times X_d) \geq l$.*

Another independent starting point is the following standard statement.

Proposition 5. *Let G be an oriented graph with vertex set V. There exists an ordered bipartition (X, Y) of V such that the number of edges $(u, v) \in X \times Y$ of G is at least a quarter of the total number of edges of G.*

This statement can be seen to be equivalent to the following: given a matrix $A : [n] \times [n] \to \mathbb{F}$ there exist disjoint subsets X, Y of $[n]$ such that the restriction $A(X \times Y)$ has at least a quarter as many support elements as A has outside the diagonal. A first step will be to obtain an analogue of this statement for ranks of matrices. We thank Lisa Sauermann for a sketch that led to the proof of that statement. We will then generalise this analogue to higher-order tensors. We note that Proposition 5 and its generalisation to uniform hypergraphs will themselves be involved in the proof of the general higher-order tensor case.

Let E be the set of points $(x_1, \dots, x_d) \in Q_1 \times \cdots \times Q_d$ that do *not* have pairwise distinct coordinates. The following definition will be central to our second main result.

Definition 6. *Let $d \geq 2$ be a positive integer, let R be a non-empty family of partitions of $[d]$. For $T : Q_1 \times \cdots \times Q_d \to \mathbb{F}$ an order-d tensor we define the* essential R-rank

$$\mathrm{e}R\mathrm{rk}\, T = \min_V R\mathrm{rk}(T + V)$$

where the minimum is taken over all order-d tensors $V : Q_1 \times \cdots \times Q_d \to \mathbb{F}$ with support contained inside E, and the disjoint R-rank

$$\mathrm{d}R\mathrm{rk}\, T = \max_{X_1, \dots, X_d} R\mathrm{rk}(T(X_1 \times \cdots \times X_d))$$

where the maximum is taken over all $X_1 \subset Q_1, \dots, X_d \subset Q_d$ with $X_1, \dots, X_d$ pairwise disjoint.

It seems worthwhile to compare the essential R-rank with the disjoint R-rank, as it is straightforward to show that a tensor has essential R-rank equal to 0 if and only if it has disjoint R-rank equal to 0: the corresponding tensors are the tensors supported inside E. Moreover, we can show that the disjoint R-rank is at most the essential R-rank.

Lemma 7. *Let $d \geq 2$ be a positive integer, and let R be a non-empty family of partitions of $[d]$. For every order d tensor $T : Q_1 \times \cdots \times Q_d \to \mathbb{F}$ we have*

$$\mathrm{d}R\mathrm{rk}\, T \leq \mathrm{e}R\mathrm{rk}\, T.$$

Our second main result is a weak converse to this last inequality.

Theorem 8. *Let $d \geq 2$ be a positive integer, and let R be a non-empty family of partitions of $[d]$. There exists a function $G'_{d,R} : \mathbb{N} \to \mathbb{N}$ such that if T is an order-d tensor such that $\mathrm{e}R\mathrm{rk}\, T \geq G'_{d,R}(l)$ then we have $\mathrm{d}R\mathrm{rk}\, T \geq l$.*

Theorem 8 is also an essential ingredient to the proof of the main result of the paper [3], where in joint work with Timothy Gowers we generalise a theorem of Green and Tao ([4], Theorem 1.7) on the approximate equidistribution of polynomials with high rank over finite prime fields to the case where the variables are chosen (uniformly and independently) at random in an arbitrary non-empty subset of the field rather than in the whole field. However, we will not focus on this application.

The methods involved in our proofs of Theorem 4 and of Theorem 8 are similar in several ways: those that we will use to prove the latter can be viewed as a moderate complication of those that we will use to prove the former.

2 Proof example

As a simple representative example of our proof techniques, let us explain how we show Theorem 4 in the case of the order-3 slice rank, assuming that it is already proved in the case of the order-3 tensor rank. We begin by proving a lemma showing that having a large separated set of slices guarantees a high slice rank. For $T : Q_1 \times Q_2 \times Q_3 \to \mathbb{F}$ and $x \in Q_1$ we write $T_x : Q_2 \times Q_3 \to \mathbb{F}$ for the matrix defined by $T_x(y, z) = T(x, y, z)$, and similarly define the notations T_y and T_z.

Lemma 9. *Let $T : Q_1 \times Q_2 \times Q_3 \to \mathbb{F}$ be an order-3 tensor, and $l \geq 1$ be an integer. If there exist $x_1 \dots, x_l \in Q_1$ such that*

$$\mathrm{rk}(\sum_{i=1}^{l} a_i T_{x_i}) \geq l$$

for every $(a_1, \dots, a_l) \in \mathbb{F}^l \setminus \{0\}$, then $\mathrm{sr}\, T \geq l$.

We next show a partial converse to the inequality $\mathrm{sr}\, T \leq \mathrm{tr}\, T$ which holds in the situation where all slices of T of all three kinds have bounded rank.

Lemma 10. *Let $T : Q_1 \times Q_2 \times Q_3 \to \mathbb{F}$ be an order-3 tensor. Let $m \geq 1$ be a positive integer. Assume that for all $x \in Q_1, y \in Q_2, z \in Q_3$ we have $\mathrm{rk}\, T_x, \mathrm{rk}\, T_y, \mathrm{rk}\, T_z \leq m$. Then $\mathrm{tr}\, T \leq m(\mathrm{sr}\, T)^2$.*

We are now ready to finish the proof.

Proposition 11. *Let $T : Q_1 \times Q_2 \times Q_3 \to \mathbb{F}$ be an order-3 tensor, and let $l \geq 1$ be a positive integer. If $\mathrm{sr}\, T \geq 51l^3$ then there exist $X \subset Q_1$, $Y \subset Q_2$, $Z \subset Q_3$ with size at most $48l^3$ such that $\mathrm{sr}\, T(X \times Y \times Z) \geq l$.*

Let $T : Q_1 \times Q_2 \times Q_3 \to \mathbb{F}$ be an order-3 tensor. If T satisfies the assumption of Lemma 9 then we can conclude using a multidimensional version of the standard statement on submatrices. If T satisfies the assumption of Lemma 10 then we conclude by reducing to the tensor rank. Furthermore, these two lemmas can be viewed to some extent as representing two extreme cases, to which we can always reduce: if T is an order-3 tensor with high slice rank but which is not in the first situation, then we can always decompose it as a sum $S + U$ where S has bounded slice rank and U is in the second situation, a decomposition which hence allows us to prove Proposition 11 for all order-3 tensors.

References

[1] E. Croot, V. Lev and P. Pach, *Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small*, Ann of Math. **185** (2017), 331-337.

[2] J. Ellenberg and D. Gijswijt, *On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression* Ann of Math. **185** (2017), 339-343.

[3] W. T. Gowers and T. Karam, *Equidistribution of high rank polynomials with variables restricted to subsets of* $\mathbb{F}_p$, arXiv:2209.04932 (2022).

[4] B. Green and T. Tao, *The distribution of polynomials over finite fields, with applications to the Gowers norms.* Contrib. Discrete Math. **4** (2009), no. 2, 1-36.

[5] T. Karam, *High-rank subtensors of high-rank tensors*, arXiv:2207.08030v2 (2022).

[6] S. Lovett, *The analytic rank of tensors and its applications*, Discrete Anal. **7** (2019), 1-10.

[7] E. Naslund, *The Chromatic Number of* $\mathbb{R}^n$ *with Multiple Forbidden Distances*, arxiv:2205.12312 (2022).

[8] E. Naslund, *The partition rank of a tensor and k-right corners in* $\mathbb{F}_q^n$, Jour. Combin. Th, A **174** (2020), 105190.

[9] L. Sauermann, *Finding solutions with distinct variables to systems of linear equations over* $\mathbb{F}_p$, Math. Ann. (2022).

[10] W. Sawin and T. Tao, *Notes on the "slice rank" of tensors*, https://terrytao.wordpress.com/2016/08/24/notes-on-the-slice-rank-of-tensors.

[11] T. Tao, *A symmetric formulation of the Croot-Lev-Pach-Ellenberg-Gijswijt capset bound*, https://terrytao.wordpress.com/2016/05/18/a-symmetric-formulation-of-the-croot-lev-pach-ellenberg-gijswijt-capset-bound.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Rainbow spanning trees in uniformly coloured perturbed graphs

(Extended abstract)

Kyriakos Katsamaktsis* Shoham Letzter† Amedeo Sgueglia‡

Abstract

We consider the problem of finding a copy of a rainbow spanning bounded-degree tree in the uniformly edge-coloured randomly perturbed graph.

Let G_0 be an n-vertex graph with minimum degree at least δn, and let T be a tree on n vertices with maximum degree at most d, where $\delta \in (0,1)$ and $d \geq 2$ are constants. We show that there exists $C = C(\delta, d) > 0$ such that, with high probability, if the edges of the union $G_0 \cup \mathbf{G}(n, C/n)$ are uniformly coloured with colours in $[n-1]$, then there is a rainbow copy of T.

Our result resolves in a strong form a conjecture of Aigner-Horev, Hefetz and Lahiri.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-090`

1 Introduction

Given $\delta \in (0,1)$, we define $\mathcal{G}_{\delta,n}$ to be the family of graphs on $[n]$ with minimum degree at least δn, and we let $\mathbf{G}(n,p)$ be the binomial random graph on $[n]$ with edge probability p. One of the central themes in extremal combinatorics is understanding how large δ

*Department of Mathematics, University College London, London WC1E 6BT, UK. E-mail: kyriakos.katsamaktsis.21@ucl.ac.uk. Research supported by the Engineering and Physical Sciences Research Council [grant number EP/W523835/1].

†Department of Mathematics, University College London, London WC1E 6BT, UK. E-mail: s.letzter@ucl.ac.uk. Research supported by the Royal Society.

‡Department of Mathematics, University College London, London WC1E 6BT, UK. E-mail: a.sgueglia@ucl.ac.uk. Research supported by the Royal Society.

needs to be so that, for each $G \in \mathcal{G}_{\delta,n}$, G contains a copy of a given graph. Similarly, probabilistic combinatorics aims to determine how large p needs to be for a given graph to appear in $\mathbf{G}(n,p)$ with high probability[1]. As an interpolation between the two graph models, Bohman Frieze, and Martin [4] introduced the *perturbed graph* model. For a given $\delta \in (0,1)$, this is defined as $G_0 \cup \mathbf{G}(n,p)$ where $G_0 \in \mathcal{G}_{\delta,n}$, i.e. as the n-vertex graph on $[n]$ whose edge set is the union of the edges of G_0 and the edges of $\mathbf{G}(n,p)$. Since [4], there has been a sizeable body of research extending and adapting results from the extremal and the probabilistic to the perturbed setting.

Another flourishing trend is to investigate the emergence of rainbow structures in uniformly edge-coloured graphs. Given an edge-coloured graph G, a subgraph H of G is *rainbow* if each edge of H has a distinct colour. A graph G is *uniformly coloured* in a set of colours $\mathcal{C}$ if each edge of G gets a colour independently and uniformly at random from $\mathcal{C}$. For example, for $G = \mathbf{G}(n, \omega(1)/n)$ uniformly coloured in $\mathcal{C} = [n]$, Aigner-Horev, Hefetz and Lahiri [1] showed that with high probability G admits a rainbow copy of any fixed almost-spanning bounded-degree tree. Other instances of similar problems in random graphs can be found in [3, 5–7]. Here we consider rainbow spanning bounded-degree trees in uniformly coloured perturbed graphs.

Theorem 1.1. *Let $\delta \in (0,1)$ and let $d \geq 2$ be a positive integer. Then there exists $C > 0$ such that the following holds. Let G_0 be a graph on n vertices with minimum degree at least δn. Suppose that T is a tree on n vertices with maximum degree at most d, and that $G \sim G_0 \cup \mathbf{G}(n, C/n)$ is uniformly coloured in $[n-1]$. Then, with high probability, G contains a rainbow copy of T.*

Theorem 1.1 provides a rainbow variant of a result of Krivelevich, Kwan and Sudakov [12], who showed that, under the same assumptions, with high probability, $G_0 \cup \mathbf{G}(n, C/n)$ contains a copy of T.

Aigner-Horev, Hefetz and Lahiri [1] already considered the question of embedding rainbow spanning trees in uniformly coloured perturbed graphs, and they proved that the same conclusion holds when the edges are uniformly coloured with $(1+\varepsilon)n$ colours (for an arbitrary constant ε) and C/n is replaced by $\omega(1)/n$. Moreover, Theorem 1.1 proves in a strong form Conjecture 1.4 of [1].

In the next section, Section 2, we consider the problem of finding rainbow almost-spanning bounded-degree trees in uniformly coloured random graphs. In Section 3, we sketch how to prove our main result.

2 Almost-spanning rainbow trees in random graphs

The first ingredient in our proof is Theorem 2.1, which says that we can embed almost-spanning trees with bounded degree in a rainbow fashion in random subgraphs of uniformly coloured pseudorandom graphs. The reason we need to consider random subgraphs of

[1] Formally, we say that a sequence of events $(A_n)_{n\in\mathbb{N}}$ holds *with high probability* if $\mathbb{P}[A_n] \to 1$ as $n \to \infty$.

pseudorandom graphs, as opposed to standard random graphs, is explained in Section 3.1. We do not define what we mean by *pseudorandom* here.

For $p \in [0,1]$, the *p-random subgraph* of a graph G, denoted by G_p, is the random graph resulting from sampling each edge of G independently with probability p.

Theorem 2.1. *Let $\varepsilon \in (0,1)$ and let $d \geq 2$ be a positive integer. Then there exists $C > 0$ such that the following holds. Let T be a tree on $(1-\varepsilon)n$ vertices, with maximum degree d, let G be a pseudorandom graph on n vertices, and write $p = C/n$. Suppose that G_p is coloured uniformly in $[n]$. Then, with high probability, G_p contains a rainbow copy of T.*

Theorem 2.1 resolves Conjecture 1.2 of [1].

The proof of Theorem 2.1 uses two previous results. The first, due to Alon, Krivelevich and Sudakov [2, Thm. 1.4], says that sparse expander graphs contain a copy of every almost-spanning bounded-degree tree. Because with $p \geq C/n$, for a large constant $C > 0$, in the p-random subgraph of a pseudorandom graph sufficiently large subsets of vertices expand, this result from [2] implies the uncoloured version of Theorem 2.1.

The second result we use is a simple consequence of a general result of Ferber and Krivelevich [6, Thm. 1.2] for binomial random subgraphs of uniformly edge-coloured hypergraphs. This allows us to deduce Theorem 2.1 from its uncoloured version.

Theorem 2.2 (Consequence of [6, Thm. 1.2])**.** *Let $\varepsilon, p, q \in (0,1)$ satisfy $q = \varepsilon^{-1}p$. Suppose that $\mathcal{H}$ is a collection of subgraphs of K_n with at most $(1-\varepsilon)n$ edges. Then*

$$\mathbb{P}\left[\begin{array}{c}\mathbf{G}(n,p)\ \textit{contains}\\ \textit{some}\ H \in \mathcal{H}\end{array}\right] \leq \mathbb{P}\left[\begin{array}{c}\textit{a uniformly edge-coloured}\ \mathbf{G}(n,q),\\ \textit{with colours in}\ [n],\ \textit{contains a rainbow}\ H \in \mathcal{H}\end{array}\right].$$

3 Rainbow spanning trees in randomly perturbed graphs

Let $G \sim G_0 \cup \mathbf{G}(n, C/n)$ and suppose G is uniformly coloured in $[n-1]$. Let T be the spanning tree of maximum degree at most d that we wish to embed in a rainbow fashion in G. Our proof splits into two cases, according to the structure of the tree T: when T has $\Omega(n)$ leaves; and when T has $\Omega(n)$ disjoint, not-too-short bare paths (where a *bare path* is a path whose interior vertices have degree 2 in T). An observation of Krivelevich [11] shows that each tree falls into at least one of these categories.

3.1 Embedding trees with long bare paths

Suppose that T has $\Omega(n)$ not-too-short disjoint bare paths. Consider r such paths of length ℓ (where $r = \Omega(n)$ and ℓ is a constant which is not too small), and denote the ends of the i-th path by x_i, y_i. Let F be the forest resulting from removing the interior vertices of these bare paths from T.

We will use Theorem 2.1 to embed F in G [2]. However, in order to be able to turn this into a rainbow embedding of T (by embedding a rainbow collection of r paths of length ℓ,

[2]Observe this still follows from Theorem 2.1 despite F being a forest. In fact we can find a rainbow embedding of the almost spanning tree which consists of F and the edges x_iy_i.

with the i-th path having endpoints x_i, y_i), we first prepare an absorbing structure, which is an adaptation of such a structure of Montgomery [13]. The building block of our absorber is given by the so-called (v, c)-*gadget*. These have been introduced by Gould, Kelly, Kühn and Osthus [8] in the context of random optimal proper colourings of the complete graph, and have already been used for perturbed graphs by the first two authors [9].

Given a vertex v and a colour c, a (v, c)-*gadget* $A_{v,c}$ is a graph on 11 vertices with the following property (the notation refers to Figure 1). $A_{v,c}$ contains two rainbow paths P and P' with the same end points, such that P uses all vertices in $A_{v,c}$ and has a c-coloured edge, and P' uses all vertices apart from v and all colours of P except for c.

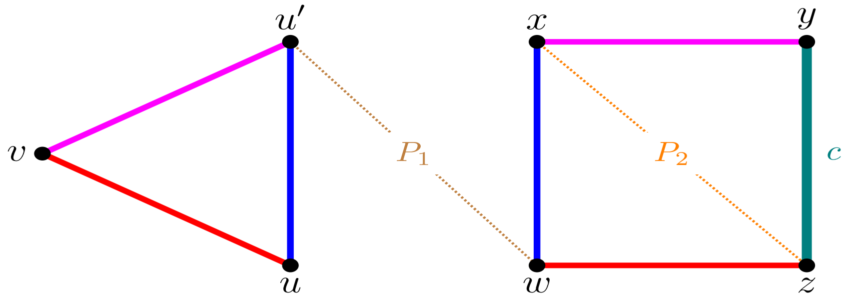


Figure 1: The (v, c)-gadget $A_{v,c}$, where the paths P_1 and P_2 have length three and are rainbow (with colours distinct from those already appearing). The path P (resp. P') is $uvu'P_1wxP_2zy$ (resp. $uu'P_1wzP_2xy$).

3.2 Embedding trees with many leaves

Suppose now that T has $\Omega(n)$ leaves. Roughly speaking, here is what we do. We first remove a constant proportion of the leaves, one leaf per parent, and embed the resulting almost-spanning tree in a rainbow fashion in G using Theorem 2.1. Completing this to a rainbow embedding of T amounts to finding a rainbow perfect matching between the removed leaves and their parents (since we removed one leaf for each parent), using all remaining colours. With some work, this follows from a forthcoming result of the authors [10], which in turn is an adaptation of a recent preprint of the first two authors [9].

Let L be a maximal collection of leaves with distinct parents. By the maximum degree assumption, $|L| = \Omega(n)$. Let L' be the collection of parents of the leaves in L, so $|L| = |L'|$. Let $T' = T \setminus L$. Let $\mathbf{G}_1 \sim \mathbf{G}(n, C/n)$ and colour $\mathbf{G}_1$ uniformly in $[n-1]$. By Theorem 2.1, with high probability, we can find a rainbow embedding of T' in $\mathbf{G}_1$[3]. Then, observe that the image of L' in V under the embedding, and the complement of $V(T')$ in the embedding, are distributed uniformly at random among all disjoint subsets of V of size $|L'|$.

[3] Theorem 2.1 applies when the number of vertices equals the number of colours, so formally it applies on a subgraph of $\mathbf{G}_1$ on $n-1$ vertices, which will be a binomial random graph with edge probability $C'/(n-1)$.

Draw a new copy of the random graph $\mathbf{G}_2 \sim \mathbf{G}(n, C/n)$. For each edge $e \in E(G_0) \cup E(\mathbf{G}_2)$, reveal whether its colour lies in $\mathcal{C}(T')$, the set of colours in the rainbow embedding of T'. Let G_0' be the subgraph of G_0 consisting of the edges which are disjoint from $E(\mathbf{G}_1)$ and have colours in $\mathcal{C}' := [n-1] \setminus \mathcal{C}(T')$. Then, from Chernoff's bound, it follows that $G_0'[L, L']$ has minimum degree $\Omega(n)$. Let $\mathbf{G}_2'$ be the subgraph of $\mathbf{G}_2$ whose edges are coloured in $\mathcal{C}'$. Then $\mathbf{G}_2'$ is a copy of the random graph $\mathbf{G}(n, C'/n)$, coloured uniformly in $\mathcal{C}'$, for an appropriate (but still large) constant C'.[4] Let $H = (G_0' \cup \mathbf{G}_2')[L, L']$. So H is a balanced bipartite graph, with bipartition $\{L, L'\}$, each of whose edges is coloured uniformly in $\mathcal{C}'$, a set of size $|L|$. It suffices to show that, with high probability, H has a rainbow perfect matching.

We now show that this reduces to finding a rainbow directed Hamilton cycle in a uniformly coloured directed perturbed graph. This can be proved as follows. Pick an arbitrary bijection $\pi : L' \to L$ and let D be the edge-coloured digraph on vertex set L with the following edges: for each $xy \in E(H)$, with $x \in L$ and $y \in L'$, add the directed edge $x\pi(y)$ and colour it by the colour of xy in H. It is straightforward to check that, if D has a rainbow directed Hamilton cycle, then H has a rainbow perfect matching. Indeed, suppose $x_1, \ldots, x_{|L|}$ is a rainbow Hamilton cycle in D. Then $x_1\pi^{-1}(x_2), x_2\pi^{-1}(x_3), \ldots, x_{|L|}\pi^{-1}(x_1)$ is a rainbow perfect matching in H.

It is also easy to check that D is distributed according to the *directed* perturbed model: this is the union of a digraph with linear minimum in- and out-degree, and $\mathbf{D}(n, p)$, the *random directed graph*, where each ordered pair of distinct vertices is an edge with probability p, independently. Moreover, D is uniformly coloured in $\mathcal{C}'$. The proof thus follows from the next theorem.

Theorem 3.1 ([10]). *Let $\delta \in (0, 1)$. Then there exists $C > 0$ such that the following holds. Let D_0 be a directed graph on vertex set $[n]$ with minimum in- and out-degree at least δn, and let $D \sim D_0 \cup \mathbf{D}(n, C/n)$ be uniformly coloured in $[n]$. Then, with high probability, D has a rainbow directed Hamilton cycle.*

References

[1] E. Aigner-Horev, D. Hefetz, and A. Lahiri, *Rainbow trees in uniformly edge-colored graphs*, Random Structures & Algorithms **62** (2023), no. 2, 287–303.

[2] N. Alon, M. Krivelevich, and B. Sudakov, *Embedding nearly-spanning bounded degree trees*, Combinatorica **27** (2007), no. 6, 629–644.

[3] D. Bal and A. Frieze, *Rainbow matchings and Hamilton cycles in random graphs*, Random Structures & Algorithms **48** (2016), no. 3, 503–523.

[4] T. Bohman, A. Frieze, and R. Martin, *How many random edges make a dense graph Hamiltonian?*, Random Structures & Algorithms **22** (2003), no. 1, 33–42.

[4] Actually, edges of $\mathbf{G}_2'$ which are also in $\mathbf{G}_1$ are not uniformly coloured, but there are very few of them ($O(\log n)$ typically), so we ignore this issue for the rest of the section.

[5] A. Ferber, *Closing gaps in problems related to Hamilton cycles in random graphs and hypergraphs*, Electronic Journal of Combinatorics **22** (2015), no. 1, Paper 1.61, 7 pp.

[6] A. Ferber and M. Krivelevich, *Rainbow Hamilton cycles in random graphs and hypergraphs*, Recent trends in combinatorics, vol. 159, Springer, 2016, pp. 167–189.

[7] A. Frieze and P.-S. Loh, *Rainbow Hamilton cycles in random graphs*, Random Structures & Algorithms **44** (2014), no. 3, 328–354.

[8] S. Gould, T. Kelly, D. Kühn, and D. Osthus, *Almost all optimally coloured complete graphs contain a rainbow Hamilton path*, Journal of Combinatorial Theory, Series B **156** (2022), 57–100.

[9] K. Katsamaktsis and S. Letzter, *Rainbow Hamiltonicity in uniformly coloured perturbed graphs*, arXiv:2304.09155 (2022), https://arxiv.org/abs/2304.09155.

[10] K. Katsamaktsis, S. Letzter, and A. Sgueglia, *Rainbow directed Hamilton cycles in uniformly coloured perturbed digraphs*, In preparation.

[11] M. Krivelevich, *Embedding spanning trees in random graphs*, SIAM Journal on Discrete Mathematics **24** (2010), no. 4, 1495–1500.

[12] M. Krivelevich, M. Kwan, and B. Sudakov, *Bounded-degree spanning trees in randomly perturbed graphs*, SIAM Journal on Discrete Mathematics **31** (2017), no. 1, 155–171.

[13] R. Montgomery, *Spanning trees in random graphs*, Advances in Mathematics **356** (2019), 106793.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Trees of Intermediate Volume Growth

(Extended abstract)

George Kontogeorgiou* Martin Winter†

Abstract

For every sufficiently well-behaved function $g\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ that grows at least linearly and at most exponentially we construct a tree T of uniform volume growth g, that is,

$$C_1 \cdot g(r/4) \leq |B_G(v,r)| \leq C_2 \cdot g(4r), \quad \text{for all } r \geq 0 \text{ and } v \in V(T),$$

with $C_1, C_2 > 0$ and where $B_G(v,r)$ denotes the ball of radius r centered at a vertex v. In particular, this yields examples of trees of uniform intermediate (*i.e.*, super-polynomial and sub-exponential) volume growth.

We use this construction to provide first examples of unimodular random rooted trees of uniform intermediate growth, answering a question by Itai Benjamini. We find a peculiar change in structural properties for these trees at growth $r^{\log\log r}$.

Our results can be applied to obtain triangulations of $\mathbb{R}^2$ with varied growth behaviours and a Riemannian metric on $\mathbb{R}^2$ for the same wide range of growth behaviors.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-091`

1 Introduction

For a graph G, a vertex $v \in V(G)$ and $r \geq 0$, the set $B_G(v,r) := \{w \in V(G) \mid d_G(v,w) \leq r\}$ is the *ball* of radius r around v. The growth of these balls as r increases is the *growth behavior* or *volume growth* of G at the vertex v. The two extreme cases of growth are the regular trees (of exponential growth) and the lattice graphs (of polynomial growth).

*Department of Mathematics, University of Warwick, CV4 7AL Coventry, UK. E-mail: `George.Kontogeorgiou@warwick.ac.uk`. Supported by EPSRC.

†Department of Mathematics, University of Warwick, CV4 7AL Coventry, UK. E-mail: `martin.h.winter@warwick.ac.uk`. Supported by EPSRC.

It is an ongoing endeavor to map the growth behaviors in various graph classes, the most famous example being Cayley graphs of finitely generated groups (see *e.g.* [8]). Major results in this regard are the existence of Cayley graphs of *intermediate* growth (that is, super-polynomial but sub-exponential) [7], and the proof that vertex-transitive graphs only have polynomial growth for integer exponents [11, Theorem 2].

Vertex-transitive graphs have the same growth at every vertex. In other graph classes this must be imposed more explicitly: following [5], a graph G is of *uniform* growth $g\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ if there are constants $c_1, C_1, c_2, C_2 \in \mathbb{R}_{>0}$ so that

$$C_1 \cdot g(c_1 r) \;\leq\; |B_G(v,r)| \;\leq\; C_2 \cdot g(c_2 r) \quad \text{for all } r \geq 0 \text{ and } v \in V(G). \tag{1.1}$$

In this article we construct trees for a wide range of growth behaviors, including intermediate and polynomial with non-integer exponents. The question of uniform intermediate growth for trees was initially posed by Itai Benjamini (private communication) in the context of unimodular random trees. However, even the existence of deterministic trees of such growth was unknown at that time.

We firstly verify the existence in the deterministic case for various growth behaviors. We then demonstrate that our construction extends to unimodular random rooted trees with the same wide range of growth behaviors, answering Benjamini's question in the positive. Finally, we probe the structure of these unimodular trees and find a threshold phenomenon happening at growth roughly $r^{\log\log r}$. As an application, in Section 4.1 we obtain triangulations of the plane, as well as Riemannian metrics on $\mathbb{R}^2$, both with the same wide range of growth behaviors.

Our work follows a history of studies on the growth rate of graphs, and particularly of trees. The first (unimodular) trees of uniform polynomial growth were constructed by Benjamini and Schramm [6]. Special attention to exponential growth for trees was given by Timár [10], focusing on the existence of a well-defined exponential rate. Recent advancement in this regard was made by Abert, Fraczyk and Hayes [1]. Intermediate but not necessarily uniform growth in trees has been studied by Amir and Yang [3] as well as the references given therein.

1.1 Motivation

The interest in such trees originates in the observation, made by physicists, that planar triangulations can have non-quadratic uniform growth [2, 4]. In their landmark paper [6] Benjamini and Schramm explained this curious phenomenon by constructing trees of every polynomial growth and then demonstrating how any tree of a particular growth can be turned into such a triangulation with a similar growth:

Construction 1.1. Suppose T is a tree of maximum degree $\bar{\Delta}$. Fix a triangulated sphere with at least $\bar{\Delta}$ pairwise disjoint triangles. Take copies of this sphere, one for each vertex of T, and identify two spheres along a triangle when the associated vertices are adjacent in T. This yields a planar triangulation. If T is of uniform growth g, so is this triangulation.

As we explain at the concluding remarks, our trees can be similarly adapted to produce triangulations of the plane.

1.2 Main results

We show the existence of deterministic and unimodular random rooted trees with growth $g\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ for various functions between polynomial and exponential growth:

Theorem 1. *If $g\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is super-additive and (eventually) log-concave, then there exists a deterministic tree T of uniform growth g.*

Theorem 2. *If $g\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is super-additive and (eventually) log-concave, then there exists a unimodular random rooted tree $(\mathcal{T}, \omega)$ of uniform growth g.*

Super-additivity and log-concavity formalize the constraints on prescribed growth to be "at least linear" and "at most exponential" and prevent certain pathologies, such as unbounded degree or too strong oscillations in the growth behavior. Concretely, a function g is *super-additive* if $g(x+y) \geq g(x)+g(y)$ for every x, y in its domain, and it is *log-concave* if $g(tx + (1-t)y) \geq g(x)^t g(y)^{1-t}$ for every x, y in its domain and every $t \in (0, 1)$.

We also prove a structure theorem (Theorem 3) that describes the structure of our unimodular trees depending on the prescribed growth rate. We show that the growth rate $r^{\log\log r}$ acts as a threshold, with "faster-growing" trees being *apocentric* (not unlike the classical canopy tree) and "slower-growing" trees being *balanced* (with a precise definition of these terms in Construction 2.3). In both cases, the trees are a.s. 1-ended for most prescribed growths.

2 The construction

For each integer sequence $\delta_1, \delta_2, \delta_3, \ldots \in \mathbb{N}$ with $\delta_n \geq 1$ we construct a tree $T = T(\delta_1, \delta_2, \ldots)$. The choice of sequence will determine the growth rate of T.

Construction 2.1. The trees T_n are defined recursively. In each tree we distinguish two special types of vertices: a *center*, and the *apocentric vertices*. Both will be defined alongside the trees:

(i) T_0 is a single vertex, which is the center of T_0 and an apocentric vertex.

(ii) T_n is built from $\delta_n + 1$ disjoint copies $\tau_0, \tau_1, \ldots, \tau_{\delta_n}$ of T_{n-1} that we join into a tree by adding the following edges: for each $i \in \{1, \ldots, \delta_n\}$ add an edge between the center of τ_i and some apocentric vertex of τ_0.

There is a choice in selecting these apocentric vertices of τ_0 (and we can choose the same apocentric vertex more than once), but we require that these adjacencies be distributed in a uniform way among the apocentric vertices of τ_0.

The center of T_n is that of τ_0; the apocentric vertices of T_n are those of $\tau_1, \ldots, \tau_{\delta_n}$.

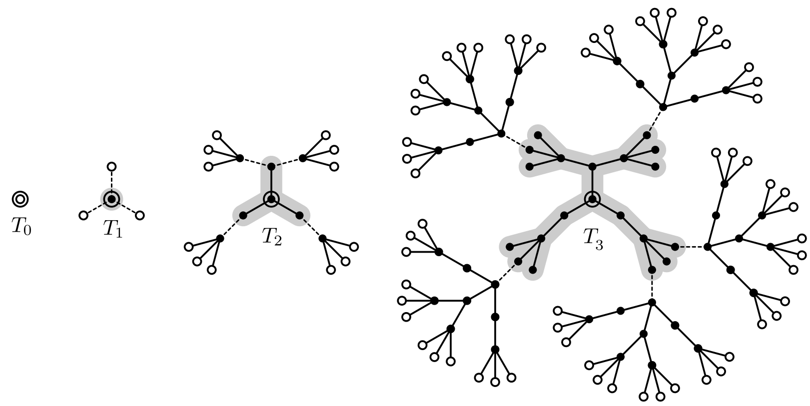


Figure 1: The first four trees $T_0, ..., T_3$ defined by $\delta_n := n + 2$. The ringed vertex is the center, and the white vertices are the apocentric vertices in the respective tree. The highlighted subgraph is the central copy τ_0 in T_n. The dashed lines are the new edges connecting the copies to form a single tree.

Observation 2.2.

(i) T_n has exactly $(\delta_1 + 1) \cdots (\delta_n + 1)$ vertices;

(ii) T_n has exactly $\delta_1 \cdots \delta_n$ apocentric vertices, all of which are leaves of the tree;

(iii) the distance from the center of T_n to any of its apocentric vertices is $2^n - 1$.

Construction 2.3. For each $n \geq 1$ identify T_n with one of its copies $\tau_0, \tau_1, ..., \tau_{\delta_{n+1}}$ in T_{n+1}. In this way we obtain an inclusion chain $T_0 \subset T_1 \subset T_2 \subset \cdots$ and the union $T = T(\delta_1, \delta_2, ...) := \bigcup_{n \geq 0} T_n$ is an infinite tree.

For later use we distinguish three natural types of limits:

- the *centric limit* always identifies T_n with the "central copy" τ_0 in T_{n+1}. This limit comes with a designated vertex $x^* \in V(T_0) \subset V(T)$, the *global center*.
- *apocentric limits* always identify T_n with an "apocentric copy" τ_i in T_{n+1}.
- *balanced limits* make infinitely many central and apocentric identifications.

We show that for a suitable sequence $\delta_1, \delta_2, \delta_3, ... \in \mathbb{N}$ and independently of the type of the limit, the tree T has a uniform volume growth, and that with a deliberate choice of the sequence we can model a wide range of growth behaviors, including polynomial, intermediate and exponential. This proves Theorem 1.

The following example computation gives an idea of the connection between the sequence $\delta_1, \delta_2, \delta_3, ... \in \mathbb{N}$ and the growth of T. Let T be the centric limit with global center $x^* \in V(T)$. By Observation 2.2 (iii) the ball of radius $r = 2^n - 1$ in T, centered at x^*, is exactly $T_n \subset T$. By Observation 2.2 (i) it follows that

$$|B_T(x^*, r)| = |T_n| = (\delta_1 + 1) \cdots (\delta_n + 1). \tag{2.1}$$

So, if we aim for $B_T(x^*, r) \approx g(r)$ with a given growth function $g\colon \mathbb{R}_0 \to \mathbb{R}_0$, then (2.1) suggests to use a sequence $\delta_1, \delta_2, \delta_3, \ldots \in \mathbb{N}$ for which $(\delta_1 + 1) \cdots (\delta_n + 1)$ approximates $g(2^n - 1)$. It turns out to be more convenient to approximate $g(2^n)$, so

$$\delta_n + 1 \approx \frac{g(2^n)}{g(2^{n-1})}, \tag{2.2}$$

where we introduce an error when rounding the right side to an integer.

To establish uniform growth with a prescribed growth rate g it remains to prove:

- the error introduced by rounding the right side of (2.2) is manageable;
- an estimation close to (2.1) holds for radii r that are not of the form $2^n - 1$;
- an estimation close to (2.1) holds for general limit trees and around vertices other than a designated "global center".

We address each of these points in our paper for g super-additive and log-concave.

We close with three examples demonstrating the versatility of Construction 2.3.

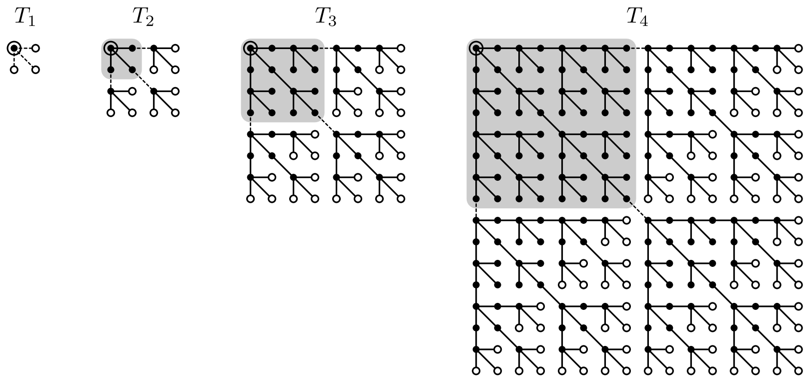


Figure 2: $T(3, 3, \ldots)$ embedded in the square of the 2D lattice.

Example 2.4 (Polynomial growth)**.** If we aim for polynomial growth $g(r) = r^\alpha, \alpha \in \mathbb{N}$ then the heuristics (2.2) suggests to use a constant sequence $\delta_n := 2^\alpha - 1$.

The corresponding trees T_n embed in the α-th powerof the α-dimensional lattice.

More generally, for any constant sequence $\delta_n := c$ we expect to find polynomial volume growth, potentially with a non-integer exponent $\log(c + 1)$.

Example 2.5 (Exponential growth)**.** For $\delta_n := d^{2^{n-1}}, d \in \mathbb{N}$ the centric limit T is the d-ary tree,

of exponential volume growth. Using (2.1) for $r = 2^n$ we find

$$|B_T(x^*, r-1)| = (\delta_1 + 1) \cdots (\delta_n + 1) = \prod_{k=1}^{n} \left(d^{2^{k-1}} + 1 \right) = \sum_{i=0}^{2^n - 1} d^i = \frac{d^{2^n} - 1}{d - 1} = \frac{d^r - 1}{d - 1}.$$

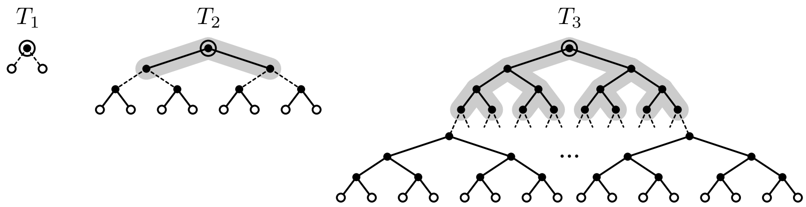


Figure 3: The binary tree constructed from Construction 2.1 using the sequence $\delta_n = 2^{2^{n-1}}$.

Extrapolating from Example 2.4 and Example 2.5, it seems reasonable that unbounded sequences $\delta_1, \delta_2, \delta_3, \ldots$ with a growth sufficiently below doubly exponential result in intermediate volume growth.

Example 2.6 (Intermediate growth)**.** For $\delta_n := (n+3)^\alpha - 1$, $\alpha \in \mathbb{N}$ we can compute this explicitly (see Figure 1 for the case $\alpha = 1$). If T is the centric limit with global center $x^* \in V(T)$ and $r = 2^n$, then:

$$\begin{aligned}|B_T(x^*, r-1)| = (\delta_1 + 1) \cdots (\delta_n + 1) = (\tfrac{1}{6}(n+3)!)^\alpha \sim (n! n^3)^\alpha \sim (n^n e^{-n} n^{7/2})^\alpha \\ = r^{\alpha \log\log r} r^{-\alpha/\ln 2} (\log r)^{7\alpha/2}\end{aligned}$$

Indeed, by Theorem 1, this choice of sequence leads to a tree of uniform intermediate volume growth. Trees constructed from $\delta_n \sim n^\alpha$ present an interesting boundary case in Section 3 when we discuss unimodular random trees (see also Theorem 3).

3 Passing to unimodular random trees

A *rooted graph* is a pair of the form (G, o), where G is a graph and $o \in V(G)$ is "a root". For a definition of *random rooted graphs* we follow [6]: firstly, there is a natural topology on the set of rooted graphs – the *local topology* – induced by the metric

$$\mathrm{dist}\big((G,o),(G',o')\big) := 2^{-R} \quad \begin{aligned} &\text{if } B_G(o,r) \cong B_{G'}(o',r) \text{ for all } 0 \le r \le R \\ &\text{and } B_G(o,R+1) \not\cong B_{G'}(o',R+1), \end{aligned}$$

where it is understood that $B_G(o,r)$ is rooted at o and that isomorphisms between rooted graphs preserve roots.

A random rooted graph (G,o) is a Borel probability measure (for the local topology) on the set of locally finite, connected rooted graphs. We call (G,o) *finite* if the set of infinite rooted graphs has (G,o)-measure zero. If in addition the conditional distribution of the root in (G,o) over each finite graph is uniform, then (G,o) is called *unbiased*.

Given a sequence (G_n, o_n) of unbiased random rooted graphs, a random rooted graph (G,o) is said to be the *Benjamini-Schramm limit* of (G_n, o_n) if for every finite rooted graph (H, ω) and natural number $r \ge 0$ we have

$$\lim_{n\to\infty} P\big(B_{G_n}(o_n, r) \cong (H,\omega)\big) = P\big(B_G(o,r) \cong (H,\omega)\big).$$

If it exists, (G, o) is the unique limit. If a random rooted graph is the Benjamini-Schramm limit of some sequence, we say that it is *sofic*.

One can show that a set of graphs of uniformly bounded degree is compact in the local topology, and thus, a sequence (G_n, o_n) of uniformly bounded degree always has a convergent subsequence.

We say that a random rooted graph (G, o) is of uniform growth $g\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, if there are constants $c_1, C_1, c_2, C_2 \in \mathbb{R}_{>0}$ so that a.s.

$$C_1 \cdot g(c_1 r) \leq |B_G(o, r)| \leq C_2 \cdot g(c_2 r), \quad \text{for all } r \geq 0.$$

A random rooted graph (G, o) is *unimodular* if it obeys the *mass transport principle*, *i.e.*,

$$\mathbb{E}\Big[\sum_{x \in V(G)} f(G, o, x)\Big] = \mathbb{E}\Big[\sum_{x \in V(G)} f(G, x, o)\Big]$$

for every *transport function* f, which, for our purpose, are sufficiently defined as Borel functions over doubly-pointed graphs that output non-negative real numbers (for a precise definition we direct the reader to [9]).

The function f simulates mass transport between vertices, and the mass transport principle states, roughly, that the root o sends, on average, as much mass to other vertices as it receives from them. Unimodular graphs are significant in the theory of random graphs and encompass some important classes, notably, all sofic graphs.

Proposition 3.1. *Let $g\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be super-additive and log-concave. Let T_n be as in Construction 2.1. Then the sequence (T_n, o_n) has a subsequence that converges in the Benjamini-Schramm sense to a unimodular random rooted tree of uniform growth g.*

This proves Theorem 2. We examine the structure of the Benjamini-Schramm limit $\mathcal{T}$:

Theorem 3 (Structure Theorem)**.**

(i) If $g = \Omega(r^{\alpha \log\log r})$, $\alpha > 1$, then $(\mathcal{T}, \omega)$ is a.s. an apocentric limit. In particular, it is 1-ended.

(ii) If $g = O(r^{\log\log r})$, then $(\mathcal{T}, \omega)$ is a.s. a balanced limit and it is a.s. 1-ended or 2-ended. In particular, if $\delta_n \neq 1$ eventually, then $(\mathcal{T}, \omega)$ is a.s. 1-ended and the probability for being isomorphic to any particular tree is 0.

The significance of the distinction worked out in Theorem 3 becomes more apparent with an example: the Benjamini-Schramm limit of (T_n, o_n) for a sufficiently fast growing function g (including intermediate) can be a.s. isomorphic to a single deterministic tree. Such a limit can be seen as deterministic trees with a randomly chosen root.

Example 3.2. Define recursively $\delta_1 := 1, \delta_2 := 2$ and $\delta_{n+1} := \delta_n \delta_{n-1}$. Hence, $\delta_n = 2^{F_n}$, for all $n \geq 0$, where F_n denotes the n-th Fibonacci number.

Let $T_n, n \geq 0$ be the sequence of trees according to Construction 2.1.

Then (T_n, o_n) has a subsequence that converges in the Benjamini-Schramm sense to a random rooted tree $(\mathcal{T}, \omega)$ of uniform volume growth.

From $|T_n| = (\delta_1 + 1) \cdots (\delta_n + 1)$ we have, for $r := 2^n$,

$$\tfrac{1}{2} D^{r^\alpha} \leq |T_n| \leq \tfrac{1}{2} r \cdot D^{r^\alpha}, \quad \text{with } D := 2^{\varphi^2/\sqrt{5}} \approx 2.251 \text{ and } \alpha := \log \varphi \approx 0.6942.$$

Here $\varphi \approx 1.618$ denotes the golden ratio. The growth is therefore *intermediate*.

We claim that $(\mathcal{T}, \omega)$ is a.s. isomorphic to a particular deterministic tree: from Theorem 3 we can see that $(\mathcal{T}, \omega)$ is an apocentric limit. We show that the T_n are highly symmetric in that any two apocentric copies $\tau, \tau' \prec_{n-1} T_n$ are in fact indistinguishable by symmetry. In consequence, there exists only one possible inclusion chain leading to an apocentric limit, and $\mathcal{T}$ is the unique tree obtained in this way.

Unimodular random rooted trees that are a.s. isomorphic to a unique tree of smaller uniform growth can be constructed by setting $\delta_{n+1} := \delta_n \delta_{n-1}$ for only some n, and $\delta_{n+1} := \delta_n$ otherwise.

In contrast, Benjamini-Schramm limits for g of growth below $r^{\log \log r}$ have measure zero on every countable set of trees, hence this approach cannot yield examples with uniform growth. It remains to ask whether this is an artifact of our construction or a general phenomenon.

4 Concluding remarks and open questions

4.1 Planar triangulations

Having established the existence of trees for various growth rates, we can use Construction 1.1 to conclude the existence of planar triangulations with the same range of growth behaviors. In fact, we can say more: previously known triangulations of polynomial growth are planar, but not necessarily triangulations of the plane, *i.e.*, they are not necessarily homeomorphic to $\mathbb{R}^2$. For this to be true, the tree T needs to be *1-ended*, which is the case *e.g.* for apocentric limits obtained from a sequence of Construction 2.1. Choosing a suitable metric on each triangle then also yields a Riemannian metric on $\mathbb{R}^2$ with the respective growth behavior.

4.2 Subgraphs of uniform growth

At the early stages of our research, the approach for constructing trees of uniform intermediate growth was to start from just any graph of intermediate growth (such as a Cayley graph of the Grigorchuk group [7]), and extract a spanning tree that inherits this growth in some way. Ironically, working out the details of this extraction led to an understanding of the desired trees that allowed us constructing them without a need for the ambient graph. Still, we ask:

Question 4.1. Given a graph G of uniform growth g, is there a spanning tree (or just any embedded tree) of the same uniform growth?

4.3 Beyond the construction

The unimodular random rooted trees of uniform volume growth constructed in Section 3 were obtained as Benjamini-Schramm limits of the sequence T_n. We found a threshold at growth $r^{\log\log r}$ and it remains open whether this is an artifact of our construction or whether it points to a fundamental phase change phenomenon in unimodular trees of uniform growth.

Question 4.2. To what extent are unimodular trees with growths on either side of the threshold $r^{\log\log r}$ structurally different?

References

[1] Miklós Abert, Mikolaj Fraczyk, and Ben Hayes. Co-spectral radius, equivalence relations and the growth of unimodular random rooted trees. *arXiv preprint arXiv:2205.06692*, 2022.

[2] Jan Ambj, Jan Ambjørn, Bergfinnur Durhuus, Thordur Jonsson, Orur Jonsson, et al. *Quantum geometry: a statistical field theory approach.* Cambridge University Press, 1997.

[3] Gideon Amir and Shangjie Yang. The branching number of intermediate growth trees. *arXiv preprint arXiv:2205.14238*, 2022.

[4] Omer Angel. Growth and percolation on the uniform infinite planar triangulation. *Geometric And Functional Analysis*, 13(5):935–974, 2003.

[5] Laszlo Babai. The growth rate of vertex-transitive planar graphs. *SODA '97: Proceedings of the eighth annual ACM-SIAM symposium on Discrete algorithms*, pages 564–573, 1997.

[6] Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. In *Selected Works of Oded Schramm*, pages 533–545. Springer, 2011.

[7] Rostislav Ivanovich Grigorchuck. Degrees of growth of finitely generated groups and the theory of invariant means. *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya*, 48(5):939–985, 1984.

[8] Clara Löh. *Geometric group theory.* Springer, 2017.

[9] R. Lyons and D. Aldous. Processes on unimodular random networks. *Electronic Journal of Probability*, 12(54):1454–1508, 2007.

[10] Ádám Timár. A stationary random graph of no growth rate. In *Annales de l'IHP Probabilités et statistiques*, volume 50, pages 1161–1164, 2014.

[11] Vladimir Ivanovich Trofimov. Graphs with polynomial growth. *Mathematics of the USSR-Sbornik*, 51(2):405, 1985.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Finding pairwise disjoint vector pairs in $\mathbb{F}_2^n$ with a prescribed sequence of differences

(Extended abstract)

Benedek Kovács*

Abstract

We consider the following question by Balister, Győri and Schelp: given 2^{n-1} nonzero vectors in $\mathbb{F}_2^n$ with zero sum, is it always possible to partition $\mathbb{F}_2^n$ into pairs such that the difference between the two elements of the i-th pair is equal to the i-th given vector? An analogous question in $\mathbb{F}_p$ was resolved by Preissmann and Mischler in 2009. In this paper, we prove the conjecture in $\mathbb{F}_2^n$ in the case when there are at most $n - 2\log n - 1$ distinct values among the given differences, and also in the case when at least a fraction $\frac{28}{29}$ of the differences are equal.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-092`

1 Introduction

We consider the following conjecture of Balister, Győri and Schelp [2] from 2008:

Main conjecture 1.1. Let $n \geq 2$ be an integer and $m = 2^{n-1}$. If the nonzero difference vectors $\mathbf{d_1}$, $\mathbf{d_2}$, $\ldots$, $\mathbf{d_m}$ are given in $\mathbb{F}_2^n$ such that $\sum_{i=1}^{m} \mathbf{d_i} = \mathbf{0}$ (and the $\mathbf{d_i}$'s are not necessarily distinct), then $\mathbb{F}_2^n$ can be partitioned into disjoint pairs $\{\mathbf{a_i}, \mathbf{b_i}\}$ $(1 \leq i \leq m)$ such that $\mathbf{a_i} - \mathbf{b_i} = \mathbf{d_i}$ holds for every i.

*ELTE Linear Hypergraphs Research Group, Eötvös Loránd University, Budapest, Hungary. E-mail: `benoke98@student.elte.hu`. Supported by the ÚNKP, New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund.

In 2008, Bacher [1] has independently posed another analogous version of this conjecture, where instead of $\mathbb{F}_2^n$, we partition the elements of $\mathbb{F}_p \setminus \{0\}$ into pairs, where p is an odd prime, and there is no restriction on the sum of the given (nonzero) differences. For this case, Preissmann and Mischler gave a positive answer [9]; their method relies on summing the values of an appropriate multivariate polynomial over $\mathbb{F}_p$.

Theorem 1.2 (Preissmann, Mischler). *Let p be an odd prime and $M = \frac{p-1}{2}$. If in $\mathbb{F}_p$, the nonzero differences $d_1, d_2, \dots, d_M$ are given, then $\mathbb{F}_p \setminus \{0\}$ can be partitioned into disjoint pairs $\{a_i, b_i\}$ $(1 \le i \le M)$ such that for each i, $a_i - b_i = d_i$ holds.*

Later, Kohen and Sadofschi [6] gave a new proof of this claim using the Combinatorial Nullstellensatz.

The statement can also be investigated for other cyclic groups as well. The following conjecture of Adamaszek pertaining to cyclic groups of even order has been proven by Kohen and Sadofschi [7]:

Theorem 1.3 (Kohen, Sadofschi). *Let $n = 2M$ be even. If the elements $d_1, d_2, \dots, d_M \in (\mathbb{Z}/n\mathbb{Z})^\times$ are arbitrarily given, then $\mathbb{Z}/n\mathbb{Z}$ can be partitioned into disjoint pairs $\{a_i, b_i\}$ such that for each i, we have $a_i - b_i = d_i$.*

Another way to generalize Theorem 1.2 is if we consider the problem for $\mathbb{F}_p^n \setminus \{\mathbf{0}\}$ instead of $\mathbb{F}_p \setminus \{0\}$. Karasev and Petrov showed that in this case, the same statement does not hold (by considering the case when every $\mathbf{d_i}$ is equal to the same nonzero vector $\mathbf{d}$). However they have shown the following claim [5, Theorem 3]:

Theorem 1.4 (Karasev, Petrov). *Let p be an odd prime and $M = \frac{p^n-1}{2}$. If the sets $\{\mathbf{d_{1,1}}, \dots, \mathbf{d_{1,n}}\}, \{\mathbf{d_{2,1}}, \dots, \mathbf{d_{2,n}}\}, \dots, \{\mathbf{d_{M,1}}, \dots, \mathbf{d_{M,n}}\}$ are given in $\mathbb{F}_p^n$ such that each set is a basis of $\mathbb{F}_p^n$, then there exists a function $g : [M] \to [n]$ such that $\mathbb{F}_p^n \setminus \{\mathbf{0}\}$ can be subdivided into disjoint pairs $\{\mathbf{a_i}, \mathbf{b_i}\}$, $1 \le i \le M$ with $\mathbf{a_i} - \mathbf{b_i} = \mathbf{d_{i,g(i)}}$ for every i.*

If we investigate the statement in $\mathbb{F}_2^n$ instead of $\mathbb{F}_p^n$, then to obtain a perfect matching, we also need to include the zero vector in the set of elements to be matched. Even in this case, the claim does not hold for arbitrary nonzero differences, as the sum of differences has to be equal to the sum of all elements of the vector space, which is zero. By the main conjecture, this would be a sufficient condition for an adequate perfect matching to exist.

The authors of [2] have also verified this conjecture for the case $n \le 5$, and they have proved the main conjecture in the following special case [2, Theorem 4]:

Theorem 1.5 (Balister, Győri, Schelp). *The main conjecture is true in the case when the vectors $\mathbf{d_1}, \mathbf{d_2}, \dots, \mathbf{d_{\frac{m}{2}}}$ are all equal, and for every integer $1 \le i \le \frac{m}{2}$ we have $\mathbf{d_{2i-1}} = \mathbf{d_{2i}}$.*

In 2021, Correia, Pokrovskiy and Sudakov [3] published the following result:

Theorem 1.6 (Correia, Pokrovskiy, Sudakov). *Let G be a multigraph whose edges are t-coloured, so that each colour class is a matching of size at least $t + 20t^{15/16}$. Then there exists a rainbow matching (that is, a matching with t edges of all distinct colours).*

Applying this result to the graph on the vertex set $\mathbb{F}_2^n$ with colour class i consisting of the edges between pairs of difference $\mathbf{d_i}$, we get that for any $M \leq \frac{1}{2}N - C \cdot N^{15/16}$ nonzero differences $\mathbf{d_i}$, where $N = 2^n$, we can find disjoint pairs $\{\mathbf{a_i}, \mathbf{b_i}\}$ such that $\mathbf{a_i} - \mathbf{b_i} = \mathbf{d_i}$. However this method does not result in perfect matchings.

Gao, Ramadurai, Wanless and Wormald [4] conjectured that Theorem 1.6 holds for $t+2$ in place of $t+20t^{15/16}$, which would resolve this problem for any $M \leq \frac{1}{2}N - 2$ nonzero difference vectors.

In this paper, we prove the main conjecture in the following two special cases:

- when the number of distinct values among the $\frac{1}{2}N$ difference vectors is at most $n - 2\log n - 1$;
- and when n is sufficiently large and at least a fraction $\frac{28}{29}$ of the difference vectors are all equal.

2 Perfect matching in the case of few difference classes

Let the nonzero differences $\mathbf{d_1}, \mathbf{d_2}, \ldots, \mathbf{d_m}$ be given such that $\sum_{i=1}^{m} \mathbf{d_i} = \mathbf{0}$, where $m = 2^{n-1}$. The collections containing all differences equal to a fixed vector $\mathbf{d}$ will be called *difference classes*. For a given configuration $\{\mathbf{d_1}, \mathbf{d_2}, \ldots, \mathbf{d_m}\}$, let t denote the number of nonempty difference classes. We would like to give a value $T(n)$ as large as possible, for which we can guarantee the existence of a suitable perfect matching of $\mathbb{F}_2^n$ in the case $t \leq T(n)$.

In the case $t = 1$ the task is trivial: take the $\langle \mathbf{d} \rangle$-cosets of $\mathbb{F}_2^n$ for the difference $\mathbf{d}$.

In the case $t = 2$, the task can be solved using Theorem 1.5, as $\sum \mathbf{d_i} = \mathbf{0}$ means that both difference classes have even size. So we have the structure that half of the differences are the same and the rest of the differences can be partitioned into equal-valued pairs.

Theorem 2.1. *The main conjecture is true in the case when the number of difference classes is at most $n - 2\log n - 1$.*

Lemma 2.2. *Let $n \geq 4$, and let P_n denote the power set of $[n]$ as a poset ordered by containment. Let H be a subset of P_n of size at most $n+1$, for which $\emptyset \notin H$ and $[n] \notin H$. Moreover assume that H does not contain all of the one-element sets and does not contain all of the $n-1$ element sets either. Then $P_n \setminus H$ contains a chain of size $n+1$.*

Proof sketch. This can be proved using the fact that P_n admits a decomposition into disjoint symmetric chains (i.e. chains containing one set of each integer cardinality between k and $n-k$ for some k); the proof of this fact can be found in [8, Proposition 2]. □

Proof sketch of theorem 2.1. Let the distinct values of the given differences be $\mathbf{u_1}$, $\mathbf{u_2}$, …, $\mathbf{u_t}$ where for each $1 \leq i \leq t$, $\mathbf{u_i}$ appears n_i times with $n_1 \geq n_2 \geq \cdots \geq n_t$.

Let $U = \langle \mathbf{u_1}, \ldots, \mathbf{u_t} \rangle$ and $k = \dim U$. (Then $k \leq t$.) We can assume that $k \geq 2$, as otherwise $t = 1$, a case already seen. Call the U-cosets of $\mathbb{F}_2^n$ *layers*. We create perfect matchings of each layer separately, and will not modify any finished layers later. Our algorithm consists of 3 phases.

Phase 1: We create perfect matchings in some (less than t) layers in such a way that an even number of vectors will remain in each difference class.

Phase 2: We create perfect matchings in some (less than t) layers in such a way that in each difference class, the number of remaining vectors will be divisible by 2^{k-1}.

Phase 3: All of the remaining differences are used to create homogeneous layers (i.e. layers consisting of differences from only one class).

Phase 1. Let $H = \{\mathbf{u_i} : 2 \leq i \leq t, n_i \equiv 1 \pmod 2\}$, and we use the notation $\mathbf{u} = \mathbf{u_1}$.

We call a subset S of H a *circuit* if its elements are linearly dependent mod $\mathbf{u}$, and this property does not hold for any proper subset of S. A circuit S has *good parity* if the sum of its elements is $|S|\mathbf{u}$, and *bad parity* if the sum of its elements is $(|S|+1)\mathbf{u}$.

As we have $\sum n_i \mathbf{u_i} = \mathbf{0}$, the sum of the elements of H is equal to $n_1\mathbf{u}$, so it is $\mathbf{0}$ mod $\mathbf{u}$. We will apply the following step repeatedly: as long as H is nonempty, we select some of its elements (at least three of them), and we will create a perfect matching of a full layer using one copy of each of the selected vectors, and a suitable number of copies of $\mathbf{u}$. The vectors of H used in this process will be removed from H. When all elements of H are depleted, we move on to Phase 2.

A sequence $(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_i})$ consisting of nonzero vectors in H is *diverse* if $\mathbf{0}$, $\mathbf{v_1}$, $\mathbf{v_1}+\mathbf{v_2}$, $\ldots$, $\mathbf{v_1}+\mathbf{v_2}+\cdots+\mathbf{v_{i-1}}$ are all distinct mod $\mathbf{u}$. If $(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_i})$ is a diverse sequence for which $\mathbf{v_1}+\mathbf{v_2}+\cdots+\mathbf{v_i} = i\mathbf{u}$ holds, then we can make a layer with one copy of each of $\mathbf{v_1}$, $\ldots$, $\mathbf{v_i}$ as differences, and all remaining differences being $\mathbf{u}$.

In each step, take A to be a subset of H of minimal size with $\mathbf{0}$ sum mod $\mathbf{u}$. If A has good parity, then any ordering of it is diverse, and a layer can be created. If it has bad parity, then from $H \setminus A$ we can select another minimal-size subset B with $\mathbf{0}$ sum mod $\mathbf{u}$, which can also be assumed to have bad parity. Then $A \cup B$ can be put into a diverse order (by applications of Lemma 2.2), which will be used to create the layer.

Phase 2. For each $2 \leq i \leq n$, if the number of remaining copies of $\mathbf{u_i}$ has a remainder m_i mod 2^{k-1}, then a single layer containing m_i copies of $\mathbf{u_i}$ and $2^{k-1} - m_i$ copies of $\mathbf{u}$ is made, which is possible by the main conjecture for two difference classes.

Phase 3. As the number of remaining vectors in each class is divisible by 2^{k-1}, this phase can be trivially performed, completing the required perfect matching of $\mathbb{F}_2^n$.

By calculation, it can be seen that altogether we used $\frac{4}{3}(t-1) \cdot 2^{k-1}$ copies of $\mathbf{u}$ during the first two phases, which is less than $\frac{2^{n-1}}{t}$, and so less than n_1, as required. □

3 Perfect matching in the case of many equal vectors

In this chapter, we resolve the main conjecture (for sufficiently large n) in the special case when at least a fraction $\frac{28}{29}$ of the difference vectors are all equal, and the others are arbitrary. So in contrast to the theorem of Balister, Győri and Schelp (see Theorem 1.5), here we do not require that all differences appear an even number of times.

Lemma 3.1. *Let G be a finite abelian group, and let $X \subseteq G$. Then in G, we can select at least $\frac{|G|}{|X|(|X|-1)+1}$ pairwise disjoint translates of X.*

Proof idea. Keep choosing translates of X greedily which are disjoint from the previously chosen ones. $\square$

Remark 3.2. If the group G has exponent 2, the lemma can be improved to say that at least $\frac{|G|}{\binom{|X|}{2}+1}$ pairwise disjoint translates of X can be selected.

Lemma 3.3. *Let $n \geq 2$ and $a \geq t \geq 2$ be integers, for which $\sum_{i=0}^{\lfloor t/2 \rfloor} \binom{a}{i} > 2^n$. Then in $\mathbb{F}_2^n$, among any a vectors one can find at most t which are linearly dependent.*

Proof idea. From the assumption, there exist two distinct subsets of the given vectors of size $\leq \left\lfloor \frac{t}{2} \right\rfloor$ with the same sum. Take the symmetric difference of these two subsets. $\square$

Theorem 3.4. *The main conjecture is true in the case when at least a fraction $\frac{28}{29}$ of the differences are all equal, and n is sufficiently large.*

Proof sketch. Let $\mathbf{u} \in \mathbb{F}_2^n$ appear more than $\frac{28}{29}m$ times among the given differences $\mathbf{d_1}$, $\mathbf{d_2}$, $\ldots$, $\mathbf{d_m}$. (Here $m = \frac{1}{2} \cdot 2^n$.) Let H denote the multiset of vectors $\mathbf{d_i}$ not equal to $\mathbf{u}$. Then $|H| < \frac{1}{29}m$.

We will partition $\mathbb{F}_2^n$ into pairs in the following way. In each step we select some elements $(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_i})$ of the multiset H, and among the elements of $\mathbb{F}_2^n$ not yet used, for each $1 \leq j \leq i$ we select a pair of elements with difference $\mathbf{v_j}$ (so that these pairs are disjoint from each other). The set of elements used in each step will be a union of some $\langle \mathbf{u} \rangle$-cosets; so at the end of the process, after having used all elements of H, all the remaining differences will be equal to $\mathbf{u}$, and these can be assigned to one coset each.

Similarly to the notions used in the proof of Theorem 2.1, define diverse sequences, and circuits and their parity.

If the nonzero vectors $(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_i})$ form a diverse sequence, and $\mathbf{v_1}+\mathbf{v_2}+\cdots+\mathbf{v_i} = i\mathbf{u}$, then we can take i pairwise disjoint vector pairs which use each vector in the given sequence as a difference precisely once, and whose union is equal to the union of some $\langle \mathbf{u} \rangle$-cosets. In each step of the partitioning of $\mathbb{F}_2^n$, we will use such a pattern.

We partition H into circuits by always removing the smallest circuit from it, and then in this partition, we pair up bad circuits according to increasing order of their size. Then similarly to the proof of Theorem 2.1, each good circuit, or pair of bad circuits can be arranged in a diverse order. We will use these diversely-ordered classes in decreasing order of size, always trying to find a translate of the corresponding vector set in $\mathbb{F}_2^n$ that does not contain any previously-used vectors. Classes of size greater than 8 will be called *large*, and otherwise a class is called *small*.

When we selected the circuits C_i in H (always selecting the smallest possible circuit within the remaining vectors), then because of Lemma 3.3, as long as the number of remaining vectors (a) fulfilled the inequality $\binom{a}{0} + \binom{a}{1} + \binom{a}{2} > 2^n$, we always found a circuit of size at most 4, leading to small classes. Therefore the total size of large classes is at most $4 \cdot 2^{n/2}$.

For sets of vectors X corresponding to large classes, by a calculation via Remark 3.2, we can find more disjoint translates of X in $\mathbb{F}_2^n$ then there are previously-used points, hence there will always be a translate of X which is completely unused.

For X corresponding to small classes, the total size of previous classes is less than $\frac{1}{58} \cdot 2^n$, and calculating by Remark 3.2, using the fact that $\frac{|X|}{2} \leq 8$, there will again be a sufficient number of pairwise disjoint translates of X. $\square$

References

[1] Question posed by R. Bacher on MathLinks in 2008, currently available here: https://artofproblemsolving.com/community/c6h183554

[2] Balister, P. N., Győri, E., & Schelp, R. H. (2011). Coloring vertices and edges of a graph by nonempty subsets of a set. European Journal of Combinatorics, 32(4), 533-537.

[3] Correia, D. M., Pokrovskiy, A., & Sudakov, B. (2021). Short proofs of rainbow matching results. arXiv preprint arXiv:2108.07734.

[4] Gao, P., Ramadurai, R., Wanless, I. M., & Wormald, N. (2021). Full rainbow matchings in graphs and hypergraphs. Combinatorics, Probability and Computing, 1-19.

[5] Karasev, R. N., & Petrov, F. V. (2012). Partitions of nonzero elements of a finite field into pairs. Israel Journal of Mathematics, 192(1), 143-156.

[6] Kohen, D., & Sadofschi, I. (2010). A New Approach on the Seating Couples Problem. arXiv:1006.2571

[7] Kohen, D., & Sadofschi, I. (2016). On a generalization of the seating couples problem. Discrete Mathematics, 339(12), 3017-3019.

[8] Mička, O. (2017). Hypercube problems. `http://ktiml.mff.cuni.cz/ gregor/hypercube/lecture17.pdf`

[9] Preissmann, E., & Mischler, M. (2009). Seating Couples Around the King's Table and a New Characterization of Prime Numbers. The American Mathematical Monthly, 116(3), 268-272.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Avoiding intersections of given size in finite affine spaces $\mathrm{AG}(2,q)$

(Extended abstract)

Benedek Kovács* Zoltán Lóránt Nagy†

Abstract

We study the set of intersection sizes of a k-dimensional affine subspace and a point set of size $m \in [0, 2^n]$ of the n dimensional binary affine space $\mathrm{AG}(n, 2)$.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-093`

1 Introduction

In this paper, we address the following general problem. Let S be a subset of size m of the affine space $\mathrm{AG}(n, q)$. Does there always exist a k-dimensional affine subspace which contains exactly t points of S? A k-dimensional affine subspace will be referred to as a *k-flat*, and a k-flat which contains exactly t points of a given set S will be called a *$[k, t]$-flat (induced by S)*. If for every m-element set S in the n-dimensional affine space, there is a k-flat containing exactly t points of S, we say that the pair $[n, m]$ forces the pair $[k, t]$, or that t-sets are unavoidable in k-flats. We use the notation $[n, m]_q \to [k, t]$ for this concept.

Our main focus will be the case $q = 2$, and we will omit the index q except when we wish to refer to arbitrary finite fields. The graph theoretic analogue of this problem was initiated by Erdős, Füredi, Rothschild and T. Sós [7]. Let $G(n, m)$ denote a graph on n vertices and m edges. Fix a positive integer k and a pair of integers (n, m) such that $0 \le m \le \binom{n}{2}$. For

*ELTE Linear Hypergraphs Research Group, Eötvös Loránd University, Budapest, Hungary. The author is partially supported by the ÚNKP, New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund.

†ELTE Linear Hypergraphs Research Group, Eötvös Loránd University, Budapest, Hungary. The author is supported by the Hungarian Research Grant (NKFI) No. PD 134953.

which t does it hold that any n-vertex graph with m edges contains an induced subgraph on k vertices having exactly t edges? Equivalently, we are seeking pairs (k, t) such that k-vertex induced subgraphs with t edges are unavoidable in graphs of the form $G(n, m)$. Erdős, Füredi, Rothschild and T. Sós introduced the notation $(n, m) \to (k, t)$ for the case when this is true.

Their main result showed that forced pairs (k, t) are rare in the following sense. Consider the set $Sp(n; k, t)$ of all edge cardinalities m such that $(n, m) \to (k, t)$. Its density is the ratio $\frac{|Sp(n;k,t)|}{\binom{n}{2}}$. They proved that the limit superior of this density is bounded from above by $2/3$ apart from a handful cases, and is 0 for the majority of the pairs (k, t) for fixed and large enough k. Erdős, Füredi, Rothschild and T. Sós conjectured that in fact it is bounded from above by $1/2$ apart from finitely many pairs (k, t). This was confirmed recently by He, Ma and Zhao [9]. Several related problems have been studied in the last couple of years [1, 4]. Our problem can be viewed as the q-analogue of this problem, when we investigate whether all subspaces of given dimension k can avoid to have an intersection of size t with an m-element set of the space AG(n, q), which corresponds to $\mathbb{F}_q^n$.

Considering the case $q = 3$, $k = 1$ and $t = 3$, we in turn get the famous cap set problem, which asks for the maximum number of points in AG$(n, 3)$ without creating a line, or in other words, without containing a 3-term arithmetic progression. There has been a recent breakthrough due to Ellenberg and Gijswijt [6], building upon the ideas of Croot, Lev and Pach [5], which showed that to avoid complete lines, $|S|/3^n$ has to be exponentially small. This connection highlights the complexity of the problem.

There has been significant interest in the case when we want to forbid each intersection of size larger than or equal to f instead of avoiding only f-sets in k-flats. If there exists a set $S \in AG(n, q)$ for which all k-dimensional affine subspaces contain at most c points of S, then S is called (k, c)-subspace evasive.The importance of such sets relies on its connections to explicit contructions to bipartite Ramsey graphs by Pudlák and Rödl [10] and with list-decodable codes by Guruswami [8].

By a standard application of the first moment method, Guruswami obtained random subsets of $\mathbb{F}^n$ of large size which are (k, c)-subspace evasive. ($\mathbb{F}$ denotes a finite field.)

Theorem 1.1 (Guruswami, [8])**.** *For any fixed pair (k, c), there exists a (k, c)-evasive set in $\mathbb{F}^n$ of size at least $C \cdot |\mathbb{F}|^{n(1-\frac{2k}{c})}$, where $C > 0$ is a constant independent of n.*

Corollary 1.2. *For any fixed pair (k, t) with $t > 1$, there exists a constant $C > 0$ for which the following holds: if $m \le C \cdot 2^{n(1-\frac{2\cdot k}{t-1})}$, then t-sets are avoidable in the k-flats of* AG$(n, 2)$*, i.e., in this case $[n, m] \not\to [k, t]$.*

Surprisingly it turned out that the obtained bounds are sharp in a weak sense.

Theorem 1.3 (Sudakov and Tomon, [11])**.** *Let $\mathbb{F}$ be a field, $k \in \mathbb{Z}^+$ and $\varepsilon \in (0, 0.05)$. If n is sufficiently large with respect to k, and $S \subseteq \mathbb{F}^n$ has size $m \ge |\mathbb{F}|^{n(1-\varepsilon)}$, then S is not $(k, \frac{k-\log_2(1/\varepsilon)}{8\varepsilon})$-subspace evasive.*

Note however that S being non-(k, c)-evasive does not necessarily imply that S contains a k-flat with $c+1$ points, except when $c = |\mathbb{F}|^n - 1$. The latter case on the other hand is not covered by the theorem, as $\frac{k-\log_2(1/\varepsilon)}{8\varepsilon} < 2^{k-3}$ holds for every $\varepsilon \in (0, 0.05)$.

2 Our main results

Similarly to the investigation in [7], we define the set of forcing sizes m with respect to $[k, t]$, which are the sizes for which a $[k, t]$-flat is unavoidable, as follows.

Definition 2.1.

$$Sp(n; k, t) := \{m : [n, m]_q \to [k, t]\}$$

is called the *set of forcing sizes with respect to* n, k, t, and we refer to it as the $(n; k, t)$-*spectrum.*

$$\rho(n; k, t) := \frac{Sp(n; k, t)}{|\mathbb{F}|^n}$$

is the *density of the spectrum.*

Our aim is to characterize the spectra or at least bound the density of the spectra for various values of k and f. Note that from now on, $\mathbb{F}$ is considered to be the binary field. When it is not confusing, we use the notation $[a, b]$ for the integers in the interval. For a set of integers $H \subseteq \mathbb{Z}$ and $c \in \mathbb{Z}$, let $c - H = \{c - h : h \in H\}$. For any set $X \in \mathbb{F}_2^n$, $\overline{X}$ will denote the complement of X. Observe first that determining the spectrum $Sp[n, k, f]$ or its density is essentially the same problem as determining $Sp[n, k, 2^k - f]$ or its density.

Lemma 2.2. *If* $n, k \geq 1$ *are integers and* $0 \leq f \leq 2^k$ *then* $Sp[n, k, f] = 2^n - Sp[n, k, 2^k - f]$.

2.1 Some exact results on the spectrum

We start with investigating $Sp[n, k, f]$ in cases where k is small.

Proposition 2.3. *(i)* $Sp[n, 1, 0] = [0, 2^n - 2]$, *(ii)* $Sp[n, 1, 1] = [1, 2^n - 1]$,
(iii) $Sp[n, 2, 1] = [0, 2^n] \setminus (\{2^n\} \cup \{2^n - 2^d : d \in [0, n]\})$,
(iv) $Sp[n, 2, 2] = [2, 2^n - 2]$, *(v)* $Sp[n, 2, 3] = [0, 2^n] \setminus (\{0\} \cup \{2^d : d \in [0, n]\})$.

The determination of the spectrum $Sp[n, 2, 4]$ for the full 2-dimensional flat is still a challenging problem. This has been studied under the name of Sidon sets in binary vector spaces. A subset S of an Abelian group is a *Sidon set* if the only solutions to the equation $a + b = c + d$ with $a, b, c, d \in A$ are the trivial solutions when (a, b) is a permutation of (c, d). Observe that for $A = \mathbb{F}_2^n$, S contains a $(2, 4)$-flat if and only if it is not a Sidon set. There are known results on Sidon sets in this setting which imply the following.

Proposition 2.4 (([3], also see [2, 12]))**.**

1. *There exists a constant* $C > 0$ *such that* $[n, m] \to [2, 4]$ *for every* $m \geq C \cdot 2^{\frac{1}{2}n}$.
2. *The explicit construction* $\{(x, x^3) : x \in \mathbb{F}_{2^{n/2}}\}$ *shows that* $[n, 2^{\frac{1}{2}n}] \not\to [2, 4]$.

The complete characterization of the spectrum for case $k = 3, f = 4$ requires a combination of various tools, including probabilistic methods.

Proposition 2.5. *For every $n \geq 3$, $Sp[n, 3, 4] = [4, 2^n - 4]$.*

2.2 Bounds on the density of the spectrum - unavoidable elements

Our main results are concerning the case when f is a power of 2.

Theorem 2.6. *Suppose that $k > \ell > 0$. Then there exist absolute constants $C, D > 0$ such that $[n, m] \to [k, 2^{k-\ell}]$ for $m \in \left(C \cdot 2^{n\left(1-\frac{1}{2^{k-\ell-1}}\right)}, D \cdot 2^n\right)$. Moreover, for each $\varepsilon > 0$ and sufficiently large n, $\rho(n; k, 2^{k-\ell}) \geq \frac{1-\varepsilon}{2^{\ell-1}}$.*

If ℓ is small, we can prove even stronger results.

Theorem 2.7. *Suppose that $\ell \in \{0, 1\}$. Then there exists an absolute constant C such that $[n, m] \to [k, 2^{k-\ell}]$ for $m \in \left(C \cdot 2^{n\left(1-\frac{1}{2^{k-\ell-1}}\right)}, 2^n - C \cdot 2^{n\left(1-\frac{1}{2^{k-\ell-1}}\right)}\right)$.*

Finally we discuss a case when t is the sum of two consecutive powers of 2. This case is significantly more involved compared to the case of $t = 2^{k-\ell}$.

Theorem 2.8. *For every pair (k, ℓ) of integers with $2 \leq \ell \leq k - 1$, the density of integer values m within the interval $\left[0, \frac{1}{2^{\ell-1}} \cdot 2^n\right]$ for which $[n, m] \to [k, 3 \cdot 2^{k-\ell}]$ holds, tends to 1 as $n \to \infty$. Hence $\rho(n; k, 3 \cdot 2^{k-\ell}) \geq \frac{1}{2^{\ell-1}}$.*

The above described results are relying on bounds on the size of cut in hypercube, induction and supersaturation results, and bounds on the additive energy.

If f is not a power of 2, then we must expect that the spectrum is *scattered*, several values of m are missing from it.

2.3 Missing elements from the spectrum

The random construction of Guruswami showed that we can obtain a (k, c)-evasive set on at least $2^{n(1-\frac{2k}{c})}/2^{k+1}$ points. Below we refine his argument using alteration.

Theorem 2.9. *Let $k, c \in \mathbb{Z}^+$ with $c \geq k + 1$. Then for $n > k$, there exists a (k, c)-evasive set in $\mathbb{F}_2^n$ of size at least*

$$\lfloor K \cdot 2^{n\left(1-\frac{k}{c}\right)} \rfloor - 1 \quad \textit{where} \quad K = K(k, c) := \frac{c}{c+1} \cdot 2^{k(k+1)/c} \cdot \left(2e^{2/3}(c+1)\binom{2^k}{c+1}\right)^{-\frac{1}{c}}.$$

Thus if $t > k$, then the smallest element in the spectrum of $Sp[n, k, t]$ is exponential in n. There is another reason why some elements are missing from a spectrum. The lexicographic construction below shows that if m can be obtained as a sum of few powers of 2, then the intersections with k-flats avoids many possible sizes.

Definition 2.10. Let m be an integer in $[0,2^n]$. The *lexicographic construction of m points* is defined as follows. Take every point $P=(x_1,x_2,\ldots,x_n)$ in $\mathbb{F}_2^n$ for which the binary representation of the n-digit binary number $\overline{x_1x_2\ldots x_n}$ represents a number smaller than m.

Proposition 2.11. *Let m be a positive integer and let $M(m)$ denote the number of nonzero digits in the binary representation of m. If $M(t)>M(m)$ then $[n,m]\not\to[k,t]$.*

Corollary 2.12. *The lexicographic construction guarantees that for a given pair of integers k,t, the number of avoidable intersection sizes with respect to k-flats is at least a polynomial in n having degree $M(t)$.*

The theorem above can be combined with the former results on (k,c)-subspace-evasive sets. Indeed, suppose that we have a point set S_0 which avoids k-flats having intersection of size r for each $r\in[t_1,t_2]$ and $c<t_2-t_1$. Then the union of S_0 and a (k,c)-subspace-evasive set S_1 will avoid k-flats with intersections of size $t\in[t_1+c,t_2]$.

In fact the lexicographic construction shows further values of m which are avoidable. Suppose that m is a difference of two powers of 2. While $M(m)$ might be large, the lexicographic construction of m points shows that $[k,t]$ is avoidable if t cannot be written as a power of 2 or a difference of two powers of 2. It is easy to deduce a statement similar to Proposition 2.11. Still, we believe that in general the following conjecture holds.

Conjecture 2.13. *Let $t>k$ integers. Then* $\lim_{n\to\infty}\rho[n,k,t]=1$.

References

[1] Axenovich, M., Weber, L. (2021). Absolutely avoidable order-size pairs for induced subgraphs. arXiv preprint arXiv:2106.14908.

[2] Bennett, M. (2018). Bounds on sizes of general caps in $AG(n,q)$ via the Croot-Lev-Pach polynomial method. arXiv preprint arXiv:1806.05303.

[3] Bose, R. C., Ray-Chaudhuri, D.K. (1960). On a Class of Error Correcting Binary Group Codes. Information and Control 3, pp. 68–79.

[4] Caro, Y., Lauri, J., Zarb, C. (2021). The feasibility problem for line graphs. arXiv preprint arXiv:2107.13806.

[5] Croot, E., Lev, V. F., Pach, P. P. (2017). Progression-free sets in are exponentially small. Annals of Mathematics, 331-337.

[6] Ellenberg, J. S., Gijswijt, D. (2017). On large subsets of with no three-term arithmetic progression. Annals of Mathematics, 339-343.

[7] Erdős, P., Füredi, Z., Rothschild, B. L., Sós, V. T. (1999). Induced subgraphs of given sizes. Discrete mathematics, 200(1-3), 61-77.

[8] Guruswami, V. (2011). Linear-algebraic list decoding of folded reed-solomon codes, in Proceedings of the 26th IEEE Conference on Computational Complexity.

[9] He, J., Ma, J., Zhao, L. (2021). Improvements on induced subgraphs of given sizes. arXiv preprint arXiv:2101.03898.

[10] Pudlák, P., Rödl, V. (2004). Pseudorandom sets and explicit constructions of Ramsey graphs. Quaderni di Matematica, 13, 327–346.

[11] Sudakov, B., Tomon, I. (2022). The extremal number of tight cycles. International Mathematics Research Notices, 2022(13), 9663-9684.

[12] Tait, M., Won, R. (2021). Improved Bounds on Sizes of Generalized Caps in AG(n,q). SIAM Journal on Discrete Mathematics, 35(1), 521-531.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Graph Covers and Generalized Snarks

(Extended abstract)

Jan Kratochvíl* Roman Nedela†

Abstract

The notion of graph cover, also known as locally bijective homomorphism, is a discretization of covering spaces known from general topology. It is a pair of incidence-preserving vertex- and edge-mappings between two graphs, the edge-component being bijective on the edge-neighborhoods of every vertex and its image. In line with the current trends in topological graph theory and its applications in mathematical physics, graphs are considered in the most relaxed form and as such they may contain multiple edges, loops and semi-edges.

Nevertheless, simple graphs (binary structures without multiple edges, loops, or semi-edges) play an important role. The Strong Dichotomy Conjecture of Bok et al. [2022] states that for every fixed graph H, deciding if an input graph covers H is either polynomial time solvable for arbitrary input graphs, or NP-complete for simple ones. These authors introduced the following quasi-order on the class of connected graphs: A connected graph A is called *stronger than* a connected graph B if every simple graph that covers A also covers B. Witnesses of A not being stronger than B are *generalized snarks* in the sense that they are simple graphs that cover A but do not cover B. Bok et al. conjectured that if A has no semi-edges, then A is stronger than B if and only if A covers B. We prove this conjecture for cubic one-vertex graphs, and we also justify it for all cubic graphs A with at most 4 vertices.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-094`

*Department of Applied Mathematics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic. E-mail: `honza@kam.mff.cuni.cz`. Supported by Czech Science Foundation through research grant GAČR 20-15576S.

†Faculty of Applied Sciences, University of West Bohemia, Pilsen, Czech Republic. E-mail: `nedela@savbb.sk`. Supported by Czech Science Foundation through research grant GAČR 20-15576S.

1 Introduction

Combinatorial treatment of graph coverings had its primary incentive in the solution of Heawood's Map Colour Problem due to Ringel, Youngs and others [18, 17]. That coverings underlie the techniques that led to the eventual solution of the problem was recognized by Alpert and Gross [9]. These ideas further crystallized in 1974 in the work of Gross [8] where voltage graphs were introduced as a means of a purely combinatorial description of regular graph coverings. In parallel, motivated by the effort to construct infinite families of highly symmetrical graphs, Biggs came with the very same idea which appeared in monograph [2]. Much of the theory of combinatorial graph coverings in its own right was subsequently developed by Gross and Tucker in the 1970's. We refer the reader to [10, 20] and the references therein.

In [13] the combinatorial theory of graph coverings and voltage assignments was established and extended onto a more general class of graphs which include edges with free ends (called semi-edges). The new concept of a graph proved to be useful in applications as well as in theoretical considerations.

Nowadays, the construction of a graph covering over a prescribed graph is established as a useful technique allowing to construct effectively infinite families of graphs sharing prescribed properties. In particular, it was used to construct extremal regular graphs with fixed degree and diameter [15], to construct cages and their approximations [11], and in investigation of flows on graphs [16] .

From the computational complexity point of view, Bodlaender [3] showed that deciding if a given graph G covers a given graph H (both graphs are part of the input) is NP-complete. Abello et al. [1] asked for the complexity of this question when parameterized by the target graph H. They gave the first examples of graphs for which the problem, referred to as H-Cover, is NP-complete or polynomial time solvable. It should be noted that in this seminal paper, both the parameter and the input graphs are allowed to have loops and multiple edges, but not semi-edges. The impact of semi-edges for the complexity issues is first discussed in Bok et al. [5]. It is perhaps somewhat surprising that in all cases where the complexity of the H-Cover problem is known to be NP-complete, it has been proved NP-complete even for simple graphs on input. This has been now conjectured to hold true in general, as the Strong Dichotomy Conjecture (cf. Conjecture 1 below) in [6].

Bok et al. [4] have proved that the Strong Dichotomy Conjecture holds true for all graphs, provided it holds true for connected ones. The curiosity of the NP-hardness reduction is its non-constructiveness. For two graphs A and B, they use a simple graph A' which covers A and does not cover B, if such a simple graph exists (and an arbitrary simple cover of A otherwise). However, there is no clue how to decide if such an A' exists or not. (This paradox is not undermining the reduction, because the approach is used for fixed graphs A and B, two of the connected components of the target graph H, to prove the existence of a polynomial time reduction between computational problems.) As a consequence, they defined a binary relation between connected graphs, saying that a graph A is stronger than a graph B if such a graph A' does not exist. Our aim is to contribute to the study of this relation, which shows a surprising connection to the well studied area of edge-colorings of graphs. In order to describe this connection, let us remind the reader of the (graph-theoretical) notion of a *snark*.

A snark is a simple 2-connected cubic graph which is not 3-edge-colorable. The interest in snarks was boosted by an observation by Heawood (1890) that the Four Color Theorem is equivalent to the statement: there are no planar snarks. For about 80 years only few examples of non-trivial snarks were constructed until Isaacs (1975) introduced the infinite family of flower snarks and the operation of dot-product allowing to construct a new nontrivial snark from two given ones. Investigation of snarks is an active area of research due to the fact that many long-standing conjectures on graphs (such as the 5-flow conjecture or the cycle double cover conjecture) can be reduced to problems on snarks, see [19, 7, 14] and the references therein.

If we denote by $F(3,0)$ the one-vertex graph with 3 semi-edges, and by $F(1,1)$ the one-vertex graph with 1 semi-edge and 1 loop, a snark is a simple cubic graph that covers $F(1,1)$ but does not cover $F(3,0)$, i.e., a witness for the fact that $B = F(1,1)$ is not stronger than $A = F(3,0)$. By the Petersen theorem, a 2-connected cubic graph always contains a perfect matching and hence it covers $F(1,1)$. In this sense every witness A' for A not being stronger than B can be viewed as a *generalized snark*.

It is easy to see that A is stronger than B whenever A covers B. For all known pairs A, B such that A is stronger than B and A does not cover B, the graph A contains semi-edges. In [12] it is conjectured that this is always the case (cf. Conjecture 2 below). In this paper, we justify these conjectures in several general situations, namely for cubic graphs.

2 Preliminaries

Definition 1. *A* graph *is a finite set of vertices accompanied with a finite set of* edges*, where an edge is either a* loop*, or a* semi-edge*, or a* normal edge*. A normal edge is incident with two distinct vertices and adds 1 to the degree of each of them. A loop is incident with a single vertex and adds 2 to its degree. A semi-edge is also incident with a single vertex, but adds only 1 to its degree.*

As defined, we only consider undirected graphs. However, graphs may have multiple loops and/or multiple semi-edges incident with the same vertex, and also multiple normal edges incident with the same pair of vertices. A graph is called *simple* if it has no loops, no semi-edges and no multiple normal edges. The *edge-neighborhood* $E_G(u)$ of a vertex u is the set of edges of G incident with u.

Definition 2. *A* covering projection *from a graph G to a connected graph H is a pair of surjective mappings $f_V\colon V(G) \to V(H)$ and $f_E\colon E(G) \to E(H)$ such that*

- f_E maps semi-edges onto semi-edges and loops onto loops, respectively, (normal edges may be mapped onto normal edges, loops, and semi-edges),

- f_E is incidence preserving (i.e., if $e \in E(G)$ is incident with vertices $u, v \in V(G)$, then $f_E(e)$ is incident with $f_V(u)$ and $f_V(v)$, which may of course be the same vertex),

- f_E is a local bijection on the edge-neighborhoods of any vertex and its image.

The last condition implies that f_V is degree preserving and, together with the other conditions, that the preimage of a normal edge incident with vertices $u, v \in V(H)$ is a disjoint union of normal edges, each incident with one vertex in $f_V^{-1}(u)$ and with one vertex in $f_V^{-1}(v)$, spanning $f_V^{-1}(u) \cup f_V^{-1}(v)$; the preimage of a loop incident with vertex $u \in V(H)$ is a disjoint union of cycles spanning $f_V^{-1}(u)$ (a loop is a cycle of length 1, and two parallel edges form a cycle of length 2); and the preimage of a semi-edge incident with a vertex $u \in V(H)$ is a disjoint union of semi-edges and normal edges spanning $f_V^{-1}(u)$.

If a graph G allows a covering projection onto a graph H, we say that G *covers* H, and we write $G \to H$.

It is well known that in a covering projection to a connected graph, the sizes of preimages of all vertices of the target graph are the same. This implies that $|V(H)|$ divides $|V(G)|$ whenever $G \to H$ for a connected graph H. We say that G is a k-fold cover of H, with $k = \frac{|V(G)|}{|V(H)|}$ in such a case. It follows that both vertex- and edge-component of a covering projection into a connected graph are surjective mappings.

We are interested in the following computational problem, parameterized by the target graph H.

PROBLEM:	H-COVER
INPUT:	A graph G.
QUESTION:	Does G cover H?

Abello et al. [1] raised the question of characterizing the complexity of H-COVER for simple graphs H. Despite intensive effort and several general results, the complete characterization and even a conjecture on what are the easy and hard cases is not in sight. Bok et al. [5] was the first paper that studied this question for (multi)graphs with semi-edges. A polynomial/NP-completeness dichotomy is believed in, and it has been conjectured in a stronger form in [6]:

Conjecture 1 (Strong Dichotomy Conjecture). For every graph H, the H-COVER problem is either polynomial-time solvable for general input graphs, or NP-complete for simple graphs on input.

In this connection, the following relation among graphs introduced in [4] seems to play quite an important role.

Definition 3. *A connected graph A is said to be* stronger than *a connected graph B, denoted by $A \triangleright B$, if it holds true that any simple graph covers A only if it also covers B. Formally,*

$$A \triangleright B \quad \Leftrightarrow \quad \forall G \text{ simple graph} : ((G \to A) \Rightarrow (G \to B)).$$

It follows from the definition (and from the fact that the composition of covering projections is also a covering projection) that $A \triangleright B$ whenever $A \to B$. Moreover, if A is a simple graph, then $A \triangleright B$ if and only if $A \to B$. The graphs $F(3,0)$ and $F(1,1)$ defined in the Introduction provide an example of graphs such that $F(3,0) \triangleright F(1,1)$ though $F(3,0) \not\to F(1,1)$.

It would certainly be too ambitious a goal to try to understand the complexity of the "being stronger" relation, as understanding the complexity of the $\triangleright$ relation would require a full understanding of covering graphs by simple graphs, which is known to be NP-complete for many instances of the target graphs. However, there may be a hope for understanding $A \triangleright B$ for those pairs of graphs A, B such that $A \not\to B$. In the problem session of GROW 2022 workshop in Koper, September 2022, we have conjectured that the presence of semi-edges in A is vital in this sense (cf. [12]).

Conjecture 2. If A has no semi-edges, then $A \triangleright B$ if and only if $A \to B$.

3 Our results

The goal of this paper is to justify the above mentioned conjectures for several general situations. We first show that A cannot be much smaller than B in order to be stronger than it. Then we prove the conjectures for bipartite two-vertex graphs A.

Theorem 1. *Let A and B be connected graphs such that $A \triangleright B$. Then $|V(B)|$ divides $2|V(A)|$. If, moreover, A has no semi-edges, then $|V(B)|$ divides $|V(A)|$.*

Theorem 2. *Let A be a dipole, i.e., a graph with two vertices joined by d parallel edges. Then for every graph B, $A \triangleright B$ implies $A \to B$.*

We further pay a closer attention to cubic graphs. By a technical case analysis, which involves construction of several generalized snarks, we prove that Conjectures 2 and 3 hold true for all graphs A with at most 4 vertices (and all graphs B). Our last two results prove the conjectures for all one-vertex cubic graphs B (i.e., for $B = F(3,0)$ and $B = F(1,1)$) and arbitrary A by actually completely describing the graphs A (even those with semi-edges) such that $A \triangleright B$.

Theorem 3. *For any connected graph A, it holds true that $A \triangleright F(3,0)$ if and only if $A \to F(3,0)$.*

For the last result, we need to introduce a new notion. A *semi-covering projection* from a graph G to a graph H is a pair of vertex and edge mappings which are incidence preserving (like covering projections), but semi-edges are allowed to be mapped on loops and the preimage of a loop in H is allowed to be any 2-regular subgraph of G spanning the preimage of the vertex incident with the loop in H (i.e., a disjoint union of cycles, digons, loops and open paths). We write $G \rightsquigarrow H$ when G allows a semi-covering projection onto H. With the help of this notion we can describe the graphs that are stronger than $F(1,1)$. (Note that a graph without semi-edges which semi-covers $F(1,1)$ also covers $F(1,1)$.)

Theorem 4. *For any connected graph A, it holds true that $A \triangleright F(1,1)$ if and only if $A \rightsquigarrow F(1,1)$.*

4 Acknowledgment

The authors thank Petr Hliněný for useful discussions.

References

[1] James Abello, Michael R. Fellows, and John C. Stillwell. On the complexity and combinatorics of covering finite complexes. *Australian Journal of Combinatorics*, 4:103–112, 1991.

[2] Norman Biggs. *Algebraic graph theory*. Cambridge University Press, 1974.

[3] Hans L. Bodlaender. The classification of coverings of processor networks. *Journal of Parallel Distributed Computing*, 6:166–182, 1989.

[4] Jan Bok, Jiří Fiala, Nikola Jedličková, Jan Kratochvíl, and Michaela Seifrtová. Computational complexity of covering disconnected multigraphs. In Evripidis Bampis and Aris Pagourtzis, editors, *Fundamentals of Computation Theory*, volume 12867 of *Lecture Notes in Computer Science*, pages 85–99. Springer, 2021.

[5] Jan Bok, Jiří Fiala, Petr Hliněný, Nikola Jedličková, and Jan Kratochvíl. Computational complexity of covering multigraphs with semi-edges: Small cases. In Filippo Bonchi and Simon J. Puglisi, editors, *46th International Symposium on Mathematical Foundations of Computer Science, MFCS 2021, August 23-27, 2021, Tallinn, Estonia*, volume 202 of *LIPIcs*, pages 21:1–21:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

[6] Jan Bok, Jiří Fiala, Nikola Jedličková, Jan Kratochvíl, and Paweł Rzążewski. List covering of regular multigraphs. In Cristina Bazgan and Henning Fernau, editors, *Combinatorial Algorithms - 33rd International Workshop, IWOCA 2022, Trier, Germany, June 7-9, 2022, Proceedings*, volume 13270 of *Lecture Notes in Computer Science*, pages 228–242. Springer, 2022.

[7] Miquel Angel Fiol, Giuseppe Mazzuoccolo, and Eckhard Steffen. Measures of edge-uncolorability of cubic graphs. *Electronic journal of combinatorics*, 25(4):# P4.54, 2018.

[8] Jonathan L. Gross. Voltage graphs. *Discrete mathematics*, 9(3):239–246, 1974.

[9] Jonathan L. Gross and Seth R. Alpert. The topological theory of current graphs. *Journal of Combinatorial Theory, Series B*, 17(3):218–233, 1974.

[10] Jonathan L. Gross and Thomas W. Tucker. *Topological graph theory*. Courier Corporation, 2001.

[11] Robert Jajcay and Jozef Širáň. Small vertex-transitive graphs of given degree and girth. *Ars Mathematica Contemporanea*, 4(2):375–384, 2011.

[12] Jan Kratochvíl. Towards strong dichotomy of graph covers. In *GROW 2022 - Book of Open Problems*, page 10, 2022.

[13] Aleksander Malnič, Roman Nedela, and Martin Škoviera. Lifting graph automorphisms by voltage assignments. *European Journal of Combinatorics*, 21(7):927–947, 2000.

[14] Jan Mazák, J. Rajník, and Martin Škoviera. Morphology of small snarks. *Electronic J. Combin.*, 29(4), 2022.

[15] Mirka Miller and Jozef Širáň. Moore graphs and beyond: A survey of the degree/diameter problem. *The electronic journal of combinatorics*, 20(2):# DS14v2, 2013.

[16] Roman Nedela and Martin Škoviera. Cayley snarks and almost simple groups. *Combinatorica*, 21:583–590, 2001.

[17] Gerhard Ringel. *Map color theorem*, volume 209. Springer Science & Business Media, 2012.

[18] Gerhard Ringel and John W. T. Youngs. Solution of the heawood map-coloring problem. *Proceedings of the National Academy of Sciences*, 60(2):438–445, 1968.

[19] John J. Watkins. Snarks. *Annals of the New York Academy of Sciences*, 576(1):606–622, 1989.

[20] Arthur T. White. *Graphs, groups and surfaces.* Elsevier, 1985.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

On edge-ordered graphs with linear extremal functions

(Extended abstract)

Gaurav Kucheriya* Gábor Tardos†

Abstract

The systematic study of Turán-type extremal problems for edge-ordered graphs was initiated by Gerbner et al. in 2020. Here we characterize connected edge-ordered graphs with linear extremal functions. This characterization is similar in spirit to results of Füredi et al. (2020) about vertex-ordered and convex geometric graphs.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-095`

1 History

Turán-type extremal graph theory asks how many edges an n-vertex simple graph can have if it does not contain a subgraph isomorphic to a *forbidden graph.* We introduce the relevant notation here.

Definition 1.1. *We say that a simple graph G* avoids *another simple graph H, if no subgraph of G is isomorphic to H. The Turán number* $\mathrm{ex}(n, H)$ *of a* forbidden *finite simple graph H (having at least one edge) is the maximum number of edges in an n-vertex simple graph avoiding H.*

*Department of Applied Mathematics, Charles University, Prague, Czechia. E-mail: gaurav@kam.mff.cuni.cz. Supported by GAČR grant 22-19073S.

†Alfréd Rényi Institute of Mathematics, Budapest, Hungary. E-mail: tardos@renyi.hu. Supported by the ERC advanced grants ERMiD and GeoScape and the National Research, Development and Innovation Office (NKFIH) grants K-132696 and SNN-135643.

This theory has proved to be useful and applicable in combinatorics, as well as in combinatorial geometry, number theory and other parts of mathematics and theoretical computer science. Turán-type extremal graph theory was later extended in several directions, including hypergraphs, geometric graphs, convex geometric graphs, vertex-ordered graphs, etc. Here we work with *edge-ordered graphs* as introduced by Gerbner, Methuku, Nagy, Pálvölgyi, Tardos and Vizer in [4]. The several extension of extremal graph theory each proved useful and applicable in different parts of mathematics and this also holds for the (still new) edge-ordered version discussed here, see e.g. [2]. Let us recall the basic definitions.

Definition 1.2. *An* edge-ordered graph *is a finite simple graph G together with a linear order on its edge set E. We often give the edge-order with an injective labeling $L : E \to \mathbb{R}$. We denote the edge-ordered graph obtained this way by G^L, in which an edge e precedes another edge f in the edge-order (denoted by $e < f$) if $L(e) < L(f)$. We call G^L the* labeling *or* edge-ordering *of G and call G the simple graph* underlying G^L.

An isomorphism between edge-ordered graphs must respect the edge-order. A subgraph of an edge-ordered graph inherits the edge-order and so it is also an edge-ordered graph. We say that the edge-ordered graph G contains *another edge-ordered graph H, if H is isomorphic to a subgraph of G otherwise we say that G* avoids *H.*

For a positive integer n and an edge-ordered graph H, let the Turán number $\mathrm{ex}_<(n, H)$ *be the maximal number of edges in an edge-ordered graph on n vertices that avoids H. Fixing the* forbidden edge-ordered graph *H, $\mathrm{ex}_<(n, H)$ is a function of n and we call it the* extremal function *of H. Note that this definition does not make sense if H has no edges, so we insist that H is* non-trivial*, that is, it has at least one edge.*

Braß, Károlyi and Valtr, [1] introduced *convex geometric graphs* while Pach and Tardos, [7] introduced *vertex-ordered graphs* and studied their extremal theories. In both cases a simple graph is given extra structure by specifying an order on their vertices (a cyclic order for convex geometric graphs and a linear order for vertex-ordered graphs). Characterizing the convex geometric or vertex-ordered graphs with a linear extremal function seems to be beyond reach (so far), but Füredi, Kostochka, Mubayi and Verstraëte, [3] found such a characterization for *connected* convex geometric graphs and also for *connected* vertex-ordered graphs. The situation seems to be similar for edge-ordered graphs: while we could not give a general characterization of edge-ordered graphs with linear extremal functions, in Section 2 we characterize when connected edge-ordered graphs have linear extremal functions. This characterization is also a dichotomy result: we show that whenever the extremal function of a connected edge-ordered graph is not linear, it must be $\Omega(n \log n)$.

2 Main result

In classical (unordered) extremal theory the following dichotomy is immediate:

Observation 1. *If H is a forest, then* $\mathrm{ex}(n, H) = O(n)$*, otherwise* $\mathrm{ex}(n, H) = \Omega(n^c)$ *for some $c = c(H) > 1$.*

The analogous statement fails for edge-ordered graphs: the paper [4] exhibits several edge-ordered paths with extremal functions $\Theta(n \log n)$. Therefore, when looking for an analogous result for edge-ordered graphs, we have a choice to make. Either we want to characterize the edge-ordered graphs with linear extremal functions, or the ones with extremal functions that are *almost linear*, i.e., $n^{1+o(1)}$. In the latter direction the authors of [4] formulated a conjecture that we recently verified, see [6]. The former problem seems to be considerably more difficult as there is not even a reasonable conjecture characterizing all edge-ordered graphs with a linear extremal function.

The first result in this direction appeared in the MSc thesis of the first author, [5]: he gave a simple characterization of edge-ordered *paths* with linear extremal functions. In this section we generalize this result and provide a characterization for *connected* edge-ordered graphs with linear extremal functons, see Theorem 1. This is the same restriction considered in [3] with respect to vertex-ordered and convex geometric graphs. Our theorem also states that if the extremal function of a connected edge-ordered graph is not linear, then it is $\Omega(n \log n)$. Such a dichotomy does not hold for edge-ordered graphs in general as [4] exhibits a (necessarily disconnected) edge-ordered graph whose extremal function is $\Theta(n\alpha(n))$, where α is the inverse of the Ackermann function.

In order to formulate the characterization and dichotomy in Theorem 1 we need to introduce some terminology. The *reverse* G^R of an edge-ordered graph G is obtained from G by reversing its edge-order. The *order chromatic number* $\chi_<(G)$ of an edge-ordered graph G is the smallest chromatic number $\chi(H)$ of a simple graph H such that all edge-orderings of H contain G. (If no such H exists we write $\chi_<(G) = \infty$). The order chromatic number was introduced in the paper [4] to play the role of the (ordinary) chromatic number in a version of the Erdős-Stone-Simonovits theorem for edge-ordered graphs, see Theorem 2.3 in [4]. For the purposes of our Theorem 1, one does not even have to apply this definition, it is enough to apply Lemma 2.1 below that gives a simple characterization when the order chromatic number of an edge-ordered forest is two. We call a vertex v of an edge-ordered graph *close* if the edges adjacent to v form an interval in the edge-order.

Lemma 2.1 ([4]). *A non-trivial edge-ordered forest has order chromatic number 2 if and only if it has a proper 2-coloring such that all vertices in one of the color classes are close.*

We call the edges $e_1 < e_2$ *consecutive* in an edge-ordered graph G if no edge e of G satisfies $e_1 < e < e_2$. An edge-ordered graph G is a *semi-caterpillar* if the underlying simple graph is a non-trivial tree and any pair of consecutive edges in G are either adjacent in G or they are directly connected by an edge larger than both of them.

Theorem 1 (Dichotomy). *If G or its reverse G^R is a semi-caterpillar of order chromatic number 2, then* $\mathrm{ex}_<(n, G) = O(n)$. *For any other non-trivial connected edge-ordered graph G we have* $\mathrm{ex}_<(n, G) = \Omega(n \log n)$.

Neither direction of the above dichotomy seems to follow from earlier results. For lack of space we do not give the full proofs, just sketch the main concepts involved. We start with saying a few words on semi-caterpillars.

Recall that the definition of semi-caterpillars insists that each pair of consecutive edges must either be adjacent or they are connected directly by a single larger edge. If we insist that all pair of consecutive edges are adjacent in an edge-ordered tree, we obtain a sub-class of semi-caterpillars, let us call these *basic caterpillars*. Note that the underlying simple graphs of basic caterpillars are (conventional) caterpillars: each vertex is at distance at most one from a single path. It is easy to prove that the order chromatic number of all basic caterpillars is 2. Neither statement generalizes to all semi-caterpillars, but it is not hard to prove that all vertices of a semi-caterpillar are at distance at most *two* of a single path. See Figure 1 for an example of an order chromatic number 2 semi-caterpillar whose underlying graph is not a caterpillar. The fact that basic caterpillars have linear extremal functions follows easily from the following two observations about a concept we call basic extension: A *basic extension* of a non-trivial edge-ordered graph G is an edge-ordered graph obtained by adding a single new edge to G that connects one end of the smallest edge of G to a new vertex outside G and making this new edge smaller than any edge in G.

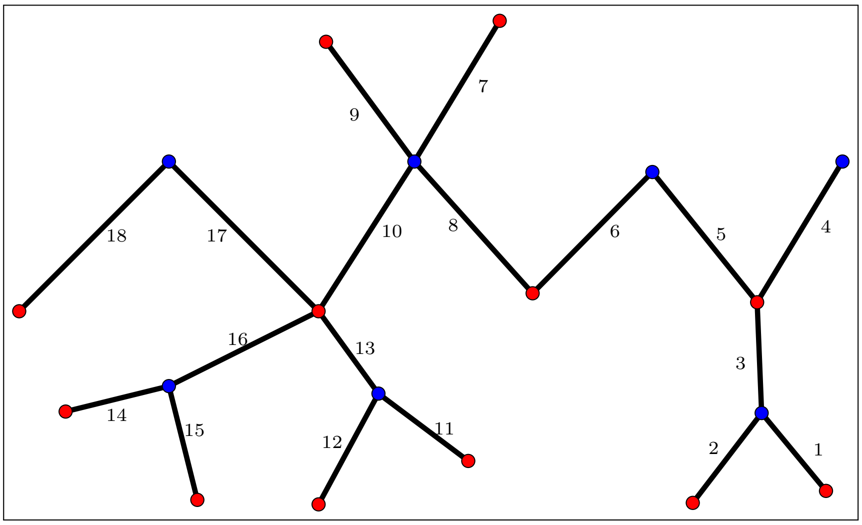


Figure 1: A semi-caterpillar with a linear extremal function

Lemma 2.2. *A non-trivial edge-ordered graph G without isolated vertices is a basic caterpillar if and only if G is obtained (up to isomorphism) from the (only) edge-ordering of the graph K_2 by a sequence of basic extensions.*

Lemma 2.3. *If G' is a basic extension of the edge-ordered graph G, then* $\mathrm{ex}_<(n, G') = \mathrm{ex}_<(n, G) + O(n)$.

Lemma 2.2 is very easy to prove and Lemma 2.3 was already implicit in [4]. We use a similar approach for proving the first statement of Theorem 1, but we will have to resolve several complications on the way. To formulate a version of Lemma 2.2 for semi-caterpillars of order chromatic number 2 we will introduce a generalization of basic extensions we call extensions. We deal with edge-ordered trees, so the underlying simple graphs are bipartite.

We will have to break symmetry and distinguish the two sides. This is largly motivated by Lemma 2.1.

An *edge-ordered bigraph* G is an edge-ordered graph G_0 together with a proper 2-coloring to *left* and *right* vertices, so each edge has a left end and a right end. We call G_0 the edge-ordered graph *underlying* G. Note that we use many terms, like edge-ordered forest, edge-ordered tree, edge-ordered path in a simpler sense meaning an edge-ordered graph whose underlying simple graph is a forest, a tree, or a path, respectively. Our use of edge-ordered bigraph as explained above is more than an edge-ordered graph whose underlying simple graph is bipartite. The notions of *isomorphism*, *subgraph*, *contain* and *avoid* naturally extend to edge-ordered bigraphs.

The paper [4] introduced edge-ordered bigraphs in order to break symmetry. Using them one can distinguish the two ways a connected edge-ordered graph may be embedded in another edge-ordered graph if both underlying simple graphs happen to be bipartite: after making them into edge-ordered bigraphs by designating left and right vertices in both graphs either all left vertices map to left vertices and the mapping ensures containment between the edge-ordered bigraphs or all left vertices map to right vertices in which case it does not.[1]

Let G be a non-trivial edge-ordered bigraph and let e be the smallest edge in G. We call the edge-ordered bigraph G' an *extension* of G if G' is obtained from G by adding new edges to it, such that

1. every new edge connects one end of e to a new degree 1 vertex;
2. all new edges are smaller than the edge e;
3. all new edges incident to the left end of e are smaller than any new edge incident to the right end of e.

Let T_0 denote the unique edge-ordered bigraph with a single edge and two vertices. We are now ready to formulate our analogue of Lemma 2.2 for semi-caterpillars of order chromatic number 2.

Lemma 2.4. *An edge-ordered graph is a semi-caterpillar of order chromatic number 2 if and only if it is isomorphic to the underlying edge-ordered graph of an edge-ordered bigraph obtained by a sequence of extensions from T_0.*

The proof of this lemma uses among other things the characterization in Lemma 2.1. If we could complement Lemma 2.4 with an appropriate analogue of Lemma 2.3, that would finish the proof of the first statement of Theorem 1. This analogue should state that if G' is an extension of the edge-ordered bigraph G and their underlying edge-ordered graphs are G'_0 and G_0, respectively, then $\mathrm{ex}_<(n, G'_0) - \mathrm{ex}_<(n, G_0) = O(n)$ or—at least—that $\mathrm{ex}_<(n, G'_0)$ is linear if $\mathrm{ex}_<(n, G_0)$ is linear. Unfortunately, neither statement holds.

This makes our proof of the first statement of Theorem 1 necessarily more involved: instead of being able to concentrate on a single extension step, we have to argue about the entire sequence of extensions that produces a certain edge-ordered bigraph.

[1]The paper [4] used the terms *edge-ordered bipartite graph* instead of edge-ordered bigraph and the terms *left-contain* and *right-contain* for the two ways an edge-ordered bigraph can contain a path.

And now we say a few words on the proof of the second statement of Theorem 1 which states that the extremal functions of connected edge-ordered graphs not covered by the first statement are $\Omega(n \log n)$. Its main ingredient is the following lemma. We denote the simple path on k vertices by P_k and denote its labeling by listing the labels along the path in the upper index. For example, P_4^{213} mentioned in the lemma below is the edge-ordered 3-edge path whose middle edge is the smallest.

Lemma 2.5. *Let G be a non-trivial edge-ordered tree. G is a semi-caterpillar if and only if it does not contain any of the edge-ordered paths P_4^{213}, P_5^{1342} or P_5^{1432}.*

Using this lemma one can finish the proof of the second statement of Theorem 1 as follows. If the order chromatic number of G is not 2, then by the edge-ordered version of the Erdős-Stone-Simonovits theorem (see [4]) $\mathrm{ex}_<(n, G) = \Theta(n^2)$. Recall that G is assumed to be connected, so if it is not an edge-ordered tree, then its underlying simple graph contains a cycle C_k and therefore $\mathrm{ex}_<(n, G) \geq \mathrm{ex}(n, C_k) = \Omega(n^{1+1/k})$. If G is a non-trivial edge-ordered tree, but not a semi-caterpillar, then it contains one of the three edge-ordered paths listed in Lemma 2.5 and therefore the extremal function $\mathrm{ex}_<(n, G)$ is at least the extremal function of the corresponding edge-ordered path. The extremal functions of these edge-ordered paths were studied in the paper [4] and we know that $\mathrm{ex}_<(n, P_5^{1432}) = \Theta(n \log n)$ and $\mathrm{ex}_<(n, P_5^{1342}) = \Omega(n \log n)$, so we are done in these cases. In the only remaining case G contains P_4^{213}.

The paper [4] calculates the extremal function of P_4^{213} also, but unfortunately it is linear. Now we apply the same argument to the reverse G^R of G, which is also connected. We obtain that if G^R is not a semi-caterpillar of order chromatic number 2, then its extremal function is $\Omega(n \log n)$ or else it contains P_4^{213}. Note that the extremal function of G and G^R coincide, so we are done unless both edge-ordered graphs G and G^R contain P_4^{213} or, in other words, G contains both P_4^{213} and P_4^{132}.

There are edge-ordered graphs with linear extremal functions containing both P_4^{213} and P_4^{132}, for example the disjoint union of P_4^{213} and P_4^{465}. But recall that G is connected. We finish the proof by showing that the extremal function of any *connected* edge-ordered graph G containing both P_4^{213} and P_4^{132} is $\Omega(n \log n)$. The proof uses the construction in Lemma 4.11 of [4].

3 Concluding remarks

Studying the extremal functions of edge-ordered graphs, especially those at the lower end of the spectrum seems very interesting. In particular, Gerbner, Methuku, Nagy, Pálvölgyi, Tardos and Vizer in [4] studied the extremal functions of many edge-ordered graphs, among them all edge-ordered paths consisting of up to four edges. In the journal version of this paper we extend their research to paths of five edges and beyond. Unfortunately, we do not have enough space to include the highlights of this line of research here.

References

[1] P. Braß, Gy. Károlyi, P. Valtr, A Turán-type extremal theory of convex geometric graphs, *Discrete and Computational Geometry—The Goodman-Pollack Festschrift* (B. Aronov et al., eds.), Springer-Verlag, Berlin, 2003, 277–302.

[2] B. Keszegh, D. Pálvölgyi, The number of tangencies between two families of curves, manuscript, arXiv:2111.08787.

[3] Z. Füredi, A. Kostochke, D. Mubayi, J. Verstraëte, Ordered and convex geometric trees with linear extremal function, *Discrete & Computational Geometry* **64** (2020), 324–338.

[4] D. Gerbner, A. Methuku, D. Nagy, D. Pálvölgyi, G. Tardos, M. Vizer, Edge-ordered Turán problems, *Journal of Combinatorial Theory, Ser. B* **160** (2023), 66–113.

[5] G. Kucheriya, Master's thesis, Moscow Institute of Physics and Technology, 2021.

[6] G. Kucheriya, G. Tardos, A characterization of edge-ordered graphs with almost linear extremal functions, *Combinatorica* (to appear), arXiv:2206.12979.

[7] J. Pach, G. Tardos, Forbidden paths and cycles in ordered graphs and matrices, *Israel Journal of Mathematics* **155** (2006), 359–380.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

A POLYNOMIAL REMOVAL LEMMA FOR POSETS

(EXTENDED ABSTRACT)

Panna Tímea Fekete* Gábor Kun†

Abstract

We prove a removal lemma with polynomial bound for posets. Alon and Shapira proved that every class of undirected graphs closed under the removal of edges and vertices is strongly testable. However, their bounds on the queries are not very effective, since they heavily rely on Szemerédi's regularity lemma. The case of posets turns out to be simpler: we show that every class of posets closed under the removal of edges is easily testable, that is, strongly testable with a polynomial bound on the queries. We also give a simple classification: for every class of posets closed under the removal of edges and vertices there is an h such that the class is indistinguishable from the class of posets without chains of length h (by testing with a constant number of queries). The analogous results hold for comparability graphs.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-096`

1 Introduction

The relationship between local and global properties of structures is a central theme in combinatorics and computer science. A classical result of this kind is the triangle removal

*Institute of Mathematics, Eötvös Loránd University, POB 120, H-1518 Budapest, Hungary and Alfréd Rényi Institute of Mathematics, POB 127, H-1364 Budapest, Hungary. E-mail: `fekete.panna.timea@renyi.hu`. Supported by the doctoral student scholarship program of the Co-operative Doctoral Program of the Ministry of Innovation and Technology financed from the National Research, Development and Innovation Fund and the ERC Synergy Grant No. 810115.

†Institute of Mathematics, Eötvös Loránd University, POB 120, H-1518 Budapest, Hungary and Alfréd Rényi Institute of Mathematics, POB 127, H-1364 Budapest, Hungary. E-mail: , `kungabor@renyi.hu`. Supported by Hungarian Academy of Sciences Momentum Grant no. 2022-58 and ERC Advanced Grant ERMiD.

lemma by Ruzsa and Szemerédi [21] usually stated in the form that if a graph G admits $\delta|V(G)|^3$ triangles then it can be made triangle-free by the removal of $\varepsilon|V(G)|^2$ edges, where δ depends only on ε. This can be applied to give a combinatorial proof of Roth's theorem [19] on 3-term arithmetic progressions, while the hypergraph removal lemma has been used to prove Szemerédi's theorem. Removal lemmas were proved for abelian groups by Green [13], for linear systems of equations by Král, Serra and Vena [16] and for permutations by Klimošová and Král [15], and by Fox and Wei [10], as well. These have led to applications in computer science allowing to test many properties by sampling a constant number of element from a structure [20].

We say that a property of (di)graphs is a set of finite (di)graphs. A (di)graph G is *ε-far* from having a property Φ if any (di)graph G' on the vertex set $V(G)$ that differs by at most $\varepsilon|V(G)|^2$ edges from G does not have the property Φ either. A property Φ is *strongly testable* if for every $\varepsilon > 0$ there exists an $f(\varepsilon)$ such that if G is ε-far from having the property Φ then for a (di)graph G the induced (directed) subgraph on $f(\varepsilon)$ vertices chosen uniformly at random does not have the property Φ with probability at least one half. Alon and Shapira [3] proved that every property of undirected graphs closed under the removal of edges and vertices is strongly testable, see Lovász and Szegedy for an analytic approach [17], while Rödl and Schacht generalized this to hypergraphs [18], see also Austin and Tao [7].

Unfortunately, the dependence on ε can be quite bad already in the case of undirected graphs: the known upper bounds in the Alon-Shapira theorem are wowzer functions due to the iterated involvement of Szemerédi's regularity lemma. We call a property *easily testable* if $f(\varepsilon)$ can be bounded by a polynomial of ε. Even triangle-free graphs are hard to test, i.e., this class is not easily testable: Behrend's construction [8] on sets of integers without 3-term arithmetic progression leads to a lower bound of magnitude $\varepsilon^{c\log(\varepsilon)}$. Alon proved that H-freeness is easily testable in case of undirected graphs if and only if H is bipartite. For forbidden induced subgraphs Alon and Shapira gave a characterization [4], where there are very few easy cases, and ordered graphs studied by Gishboliner and Tomon are similar [11]. On the other hand, 3-colorability and, in general, "partition problems" surprisingly turned out to be easily testable, see Goldreich, Goldwasser and Ron [12]. Even a conjecture to draw the borderline between easy and hard properties seems beyond reach.

The goal of this paper is to study testability of posets as special digraphs. By a poset we mean a set equipped with a partial order that is anti-reflexive, asymmetric and transitive.

One can show that every property of posets closed under the removal of edges and vertices is strongly testable, similarly to the proof of Alon and Shapira [3], using the poset version of Szemerédi's regularity lemma proved by Hladky, Máthé, Patel and Pikhurko [14]). We show that properties of posets defined by forbidden subposets are easily testable. This is equivalent to the following removal lemma with polynomial bound. The height of a finite poset P is defined as the length of its longest chain, while the width is the size of the largest antichain, denoted by $h(P)$ and $w(P)$, respectively. The chain of length h is denoted by C_h.

Given two finite posets P, Q a mapping $f : Q \to P$ is a homomorphism if it is order-preserving, i.e., $f(x) \prec f(y)$ for every $x \prec y$. The probability that a uniform random mapping from Q to P is a homomorphism is denoted by $t(Q,P)$.

Theorem 1.1. *For every $\varepsilon > 0$ and positive integers h, w there exists $\delta > 0$ such that for every finite poset Q of height h and width w and an arbitrary finite poset P, if $t(Q,P) < \delta$ then there exists a Q-free poset P' on the base set of P obtained by deleting at most $\varepsilon|P|^2$ edges of P. Moreover, P' is C_h-free and the dependence of δ on ε is polynomial.*

We use this theorem to prove testability for (not necessarily finite) classes of finite posets. The height and width of $\mathcal{P}$ of a set of finite posets are

$$h(\mathcal{P}) = \min_{P \in \mathcal{P}} h(P) \qquad w(\mathcal{P}) = \min_{\substack{P \in \mathcal{P}: \\ h(P) = h(\mathcal{P})}} w(P).$$

Theorem 1.2. *For every family of finite posets $\mathcal{P}$ the property of not containing any member of $\mathcal{P}$ as a subposet is easily testable. Moreover, the number of queries depends only on $h(\mathcal{P})$ and $w(\mathcal{P})$.*

We say that two properties Φ_1 and Φ_2 of posets are not distinguishable if for every $\varepsilon > 0$ and $i = 1, 2$ there exists N such that for every poset P on at least N elements with property Φ_i there exists a poset P' with property Φ_{3-i} such that P' is obtained by deleting at most $\varepsilon|P|^2$ edges of P.

Theorem 1.3. *For every family of finite posets $\mathcal{P}$ there exists an h such that the property of not containing any member of $\mathcal{P}$ as a subposet is not distinguishable from the property of not containing the chain C_h as a subposet.*

Note that in our case it is meaningless to allow adding edges to the original poset, since adding edges will not change whether the poset is $\mathcal{P}$-free.

The comparability graph G corresponding to a poset P has vertex set $V(G) = P$ and edge set $E(G) = \{(x,y) : x \prec y \text{ or } y \prec x\}$. Alon and Fox proved that it is hard to test if a given graph is a comparability graph (or if it is perfect) [6]. Besides posets our results apply to comparability graphs, too. Given a set of (possibly infinitely many) finite graphs $\mathcal{F}$ we define the chromatic number $\chi(\mathcal{F})$ and the independence number $\alpha(\mathcal{F})$ as follows.

$$\chi(\mathcal{F}) = \min_{F \in \mathcal{F}} \chi(P) \qquad \alpha(\mathcal{F}) = \min_{\substack{F \in \mathcal{F}: \\ \chi(F) = \chi(\mathcal{F})}} \alpha(F).$$

Theorem 1.4. *For every family of finite graphs $\mathcal{F}$ the property of a given comparability graph not containing any member of $\mathcal{F}$ as a subgraph is easily testable. Moreover, the number of queries depends only on $\chi(\mathcal{F})$ and $\alpha(\mathcal{F})$.*

Theorem 1.5. *For every family of finite graphs $\mathcal{F}$ the property of being a comparability graph and not containing any member of $\mathcal{F}$ as a subgraph is not distinguishable from the property of being a comparability graph and having chromatic number at most $\chi(\mathcal{F}) - 1$.*

The proofs are based on the same ideas as in case of posets, we do not include them.

2 Density bounds

The complete h-partite poset with antichains of size w will be denoted by $K_{h\times w}(= K_{w,w,\dots,w})$. In particular, $K_{h\times 1}$ is a chain of length h, but for this we will use the shorthand notation C_h.

The next lemma provides a lower bound on the density of the complete h-partite poset $K_{h\times w}$ in terms of the density of the chain of length h. The proof uses standard techniques appearing in the solution to the Zarankiewicz problem. We will use the notation $[n] := \{1, 2, \dots, n\}$.

Lemma 2.1. *For every poset P and positive integers h, w the inequality*

$$t(K_{h\times w}, P) \geq t^{w^2}(C_h, P)$$

holds.

Proof. The following two claims imply the lemma.

Claim 2.2.

$$t(K_{h\times w}, P) \geq t^w(K_{w,1,w,1,\dots}, P)$$

Proof. Let $(x_{i,j})_{i\in[h],j\in[w]}$ be chosen uniformly and independently at random.

$$t(K_{h\times w}, P) = \mathbb{P}_{(x_{i,j})_{i\in[h],j\in[w]}}\Big(\forall i' \in [h-1], j, j' \in [w] \quad x_{i',j} \prec x_{i'+1,j'}\Big)$$
$$= \mathbb{E}_{\substack{(x_{i,j})_{i\in[h],j\in[w]}\\ i \text{ odd}}}\left[\mathbb{P}_{\substack{(x_{i,j})_{i\in[h],j\in[w]}\\ i \text{ even}}}\left(\forall i' \in [h-1], j, j' \in [w] \quad x_{i',j} \prec x_{i'+1,j'} \,\middle|\, (x_{i,j})_{i\in[h],j\in[w]}, i \text{ odd}\right)\right].$$

Here we split $K_{h\times w}$ into w edge-disjoint copies of $K_{w,1,w,1,\dots}$. Since the events corresponding to elements in the same even layer are independent we obtain that this equals

$$\mathbb{E}_{\substack{(x_{i,j})_{i\in[h],j\in[w]}\\ i \text{ odd}}}\left[\mathbb{P}^w_{\substack{(x_{i,1})_{i\in[h]}\\ i \text{ even}}}\left(\forall i' \in [h-1], j' \in [w] \quad \substack{\text{if } i' \text{ odd then } x_{i',j'} \prec x_{i'+1,1}\\ \text{if } i' \text{ even then } x_{i',1} \prec x_{i'+1,j'}} \,\middle|\, (x_{i,j})_{i\in[h],j\in[w]}, i \text{ odd}\right)\right]$$
$$\geq \mathbb{E}^w_{\substack{(x_{i,j})_{i\in[h],j\in[w]}\\ i \text{ odd}}}\left[\mathbb{P}_{\substack{(x_{i,1})_{i\in[h]}\\ i \text{ even}}}\left(\forall i' \in [h-1], j' \in [w] \quad \substack{\text{if } i' \text{ odd then } x_{i',j'} \prec x_{i'+1,1}\\ \text{if } i' \text{ even then } x_{i',1} \prec x_{i'+1,j'}} \,\middle|\, (x_{i,j})_{i\in[h],j\in[w]}, i \text{ odd}\right)\right]$$
$$= \mathbb{P}^w_{\substack{(x_{i,1})_{i\in[h]} \text{ for } i \text{ even}\\ (x_{i,j})_{i\in[h],j\in[w]} \text{ for } i \text{ odd}}}\left(\forall i' \in [h-1], j' \in [w] \quad \substack{\text{if } i' \text{ odd then } x_{i',j'} \prec x_{i'+1,1}\\ \text{if } i' \text{ even then } x_{i',1} \prec x_{i'+1,j'}}\right)$$
$$= t^w(K_{w,1,w,1,\dots}, P),$$

where we have applied Jensen's inequality. □

Claim 2.3.

$$t(K_{w,1,w,1,\dots}, P) \geq t^w(C_h, P)$$

The proof of this Claim follows the same lines as the previous one, we do not include it. The lemma follows. □

3 Removal lemmas with polynomial bounds

First we prove a removal lemma for chains.

Lemma 3.1. *For every $\varepsilon > 0$ and positive integer h there exists a $\delta > 0$ such that for every finite poset P if $t(C_h, P) < \delta$ then there exists a C_h-free poset P' on the base set of P obtained by deleting at most $\varepsilon|P|^2$ edges of P. Moreover, the dependence of δ on ε is polynomial: $\delta = \left\lceil \frac{3}{\varepsilon} \right\rceil^{-h}$.*

Remark 3.2. For a fixed h the dependence of δ on ε is similar to that in the lemma. Consider a random h-partite poset with classes $S_1, \ldots, S_h$ of equal size large enough, where for two element $x \in S_i, y \in S_{i+1}$ with probability $h^2\varepsilon < 1$ we have $x \prec y$. The expected value of $t(C_h, P)$ is $\varepsilon^{h-1}h^{h-2}$. On the other hand, it is not hard to see that we need to remove essentially $\varepsilon|P|^2$ edges to make the poset C_h-free. (This is the expected number of edges between two consecutive classes.)

Proof. (of Lemma 3.1) Set $\gamma = \left\lceil \frac{3}{\varepsilon} \right\rceil^{-1}$. We partition the poset P into classes $S_1, \ldots S_{1/\gamma}$ of size differing by at most one such that if $x \prec y$ holds for $x \in S_i$ and $y \in S_j$ then $i \le j$. This is possible, since every finite poset has a linear extension.

Now we will delete edges in order to get a C_h -free poset.

First, delete edges inside the classes – this way we delete at most $\gamma|P|^2$ edges. The remaining digraph is still a poset, denote it by P_1.

We define a function $r : P_1 \to [1/\gamma]$. Given an element $x \in S_i$ the integer $r(x)$ will be the largest integer such that x is the maximal element of "many" chains with length $r(x)$.

Set $r(x) = 1$ for every $x \in S_1$.

Assume that for $i < 1/\gamma$ the function r is defined on $\cup_{j=1}^{i} S_j$.

Given $x \in S_{i+1}$ let $r(x)$ be the largest integer such that $\big|\{y : y \prec x \text{ in } P_1, r(y) = r(x) - 1\}\big| \ge \gamma|P|$, and 1 if there is no such integer. Note that $r(x) \le r(y)$ holds for every $x \prec y$. There are at least $(\gamma|P|)^{r(x)-1}$ chains of length $r(x)$ ending at x for every $x \in S_{i+1}$ such that r is strictly increasing on these chains.

Once the function r is defined we delete every edge (x, y) in P_1 for $y \prec x$ if $r(x) = r(y)$. This concerns at most $\gamma|P|^2$ edges, otherwise $r(x)$ would be larger. Note that the remaining digraph P_2 is still a poset and for every x there are still at least $(\gamma|P|)^{r(x)-1}$ chains of length $r(x)$ ending at x such that r is strictly increasing on these chains.

There is no element, where r takes value $(h+1)$, since every such element would be the end of at least $(\gamma|P|)^h$ chains of length at least $(h+1)$, but we do not have that many different chains of length h. By the same reason the number of elements, where r takes value h, is at most $\gamma|P|$. We delete every edge adjacent to these elements: this way we delete at most $\gamma|P|^2$ edges, denote the remaining poset by P'.

The total number of edges deleted is at most $3\gamma|P|^2 < \varepsilon|P|^2$.

The poset P' does not contain any chain of length at least h, since edges where the value of r at the end-vertex is at least h has been deleted, while edges where the value of

r at the end-vertex is not greater than at the starting vertex have also been deleted. This finishes the proof of the lemma. $\square$

Proof. (of Theorem 1.1) Set $\delta = \lceil \frac{3}{\varepsilon} \rceil^{-hw^2}$. The poset Q is a subposet of $K_{h\times w}$, hence Lemma 2.1 implies $\delta > t(Q,P) \geq t(K_{h\times w},P) \geq t^{w^2}(C_h,P)$. By Lemma 3.1 there exists a C_h-free subposet P' of P obtained by deleting at most $\varepsilon|P|^2$ edges. $\square$

Corollary 3.3. *For every $\varepsilon > 0$ and positive integers h,w there exists $\delta > 0$ such that for every finite graph F of chromatic number h and independence number w and an arbitrary finite comparability graph G if $t(F,G) < \delta$ then there exists an F-free comparability graph G' on the vertex set of G obtained by deleting at most $\varepsilon|V(G)|^2$ edges of G. Moreover, G' is K_h-free and the dependence on ε is polynomial: $\delta = \left\lceil \frac{3}{\varepsilon} \right\rceil^{-hw^2}$.*

Proof. The graph F is a subgraph of the multipartite Turán graph T with h classes each of size w, hence $t(F,G) \geq t(T,G)$. Let P be one of the posets whose comparability graph is G. Note that $t(T,G) \geq t(K_{h\times w},P)$, since we may assume that T is the comparability graph of $K_{h\times w}$, hence every homomorphism of $K_{h\times w}$ to P is a comparability-preserving map from T to G, i.e., a graph homomorphism.

We obtain by Lemma 2.1 that $\delta > t(F,G) \geq t(K_{h\times w},P) \geq t^{w^2}(C_h,P)$.

Lemma 3.1 implies that there exists a C_h-free subposet P' of P obtained by deleting at most $\varepsilon|P|^2$ edges, and its comparability graph G' satisfies the conditions of the corollary. $\square$

4 Property testing

Proof. (of Theorem 1.2) Set $h = h(\mathcal{P})$ and $w = w(\mathcal{P})$. Consider an $\varepsilon > 0$ and a poset P such that after removing $\varepsilon|P|^2$ the resulting poset still contains a subposet in $\mathcal{P}$. By Corollary 3.3 the probability that hw elements chosen uniformly at random contain $K_{h\times w}$, and hence a poset in $\mathcal{P}$ as a subposet is at least $\left\lceil \frac{3}{\varepsilon} \right\rceil^{-hw^2}$. If we choose $hw \left\lceil \frac{3}{\varepsilon} \right\rceil^{hw^2}$ elements uniformly at random then the probability of finding a poset in $\mathcal{P}$ as subposet is more than a half. $\square$

Proof. (of Theorem 1.3) If a poset does not contain the chain $C_{h(\mathcal{P})}$ as a subposet then it does not contain any poset from $\mathcal{P}$.

In order to prove the other direction consider a poset $Q \in \mathcal{P}$ with height $h(\mathcal{P})$. If a poset P does not contain Q as a subposet then there is no injective homomorphism from Q to P, hence $t(Q,P) \leq |P|^{-1}|Q|^2$. Theorem 1.1 shows that by the removal of $3|P|^{-1/(h(Q)w(Q)^2)}|Q|^{1/(2h(Q)w(Q)^2)}|P|^2$ edges from P one obtains a $C_{h(\mathcal{P})}$-free poset. $\square$

References

[1] Alon, Noga and Fischer, Eldar and Krivelevich, Michael and Szegedy, Mario (2000). Efficient testing of large graphs. Combinatorica, 20(4), 451-476.

[2] Alon, Noga and Shapira, Asaf (2003). Testing subgraphs in directed graphs. In Proceedings of the thirty-fifth annual ACM symposium on Theory of computing (pp. 700-709).

[3] Alon, Noga and Shapira, Asaf (2005). Every monotone graph property is testable. In Proceedings of the thirty-seventh annual ACM symposium on Theory of computing (pp. 128-137).

[4] Alon, Noga and Shapira, Asaf (2006). A characterization of easily testable induced subgraphs. Combinatorics, Probability and Computing, 15(6), 791-805.

[5] Alon, Noga and Fox, Jacob (2011). Testing perfection is hard. arXiv preprint arXiv:1110.2828.

[6] Alon, Noga and Fox, Jacob (2015). Easily testable graph properties. Combinatorics, Probability and Computing, 24(4), 646-657.

[7] Austin, Tim and Tao, Terence (2010). Testability and repair of hereditary hypergraph properties. Random Structures & Algorithms, 36(4), 373-463.

[8] Behrend, Felix A (1946). On sets of integers which contain no three terms in arithmetical progression. Proceedings of the National Academy of Sciences, 32(12), 331-332.

[9] Fox, Jacob (2011). A new proof of the graph removal lemma. Annals of Mathematics, 561-579.

[10] Fox, Jacob and Wei, Fan (2018). Fast property testing and metrics for permutations. Combinatorics, Probability and Computing, 27(4), 539-579.

[11] Gishboliner, Lior and Tomon, István (2021). Polynomial removal lemmas for ordered graphs. arXiv preprint arXiv:2110.03577.

[12] Goldreich, Oded and Goldwasser, Shari and Ron, Dana (1998). Property testing and its connection to learning and approximation. Journal of the ACM (JACM), 45(4), 653-750.

[13] Green, Ben (2005). A Szemerédi-type regularity lemma in abelian groups, with applications. Geometric & Functional Analysis GAFA, 15(2), 340-376.

[14] Hladkỳ, Jan and Máthé, András and Patel, Viresh and Pikhurko, Oleg (2015). Poset limits can be totally ordered. Transactions of the American Mathematical Society, 367(6),4319–4337.

[15] Klimošová, Tereza and Král', Daniel (2014). Hereditary properties of permutations are strongly testable. In Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms (pp. 1164-1173). Society for Industrial and Applied Mathematics.

[16] Král', Daniel and Serra, Oriol and Vena, Lluís (2012). A removal lemma for systems of linear equations over finite fields. Israel Journal of Mathematics, 187, 193-207.

[17] Lovász, László and Szegedy, Balázs (2008). Testing properties of graphs and functions. arXiv preprint arXiv:0803.1248.

[18] Rödl, Vojtěch and Schacht, Mathias (2009). Generalizations of the removal lemma. Combinatorica, 29(4), 467-501.

[19] Roth, Klaus F (1953). On certain sets of integers. J. London Math. Soc, 28(1), 104-109.

[20] Rubinfeld, Ronitt and Sudan, Madhu (1996). Robust characterizations of polynomials with applications to program testing. SIAM Journal on Computing, 25(2), 252-271.

[21] Ruzsa, Imre and Szemerédi, Endre (1978). Triple systems with no six points carrying three triangles. Combinatorics (Keszthely, 1976), Coll. Math. Soc. J. Bolyai, 18(939-945), 2.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Tiling Dense Hypergraphs

(Extended abstract)

Richard Lang*

Abstract

There are three essentially necessary conditions for perfect tilings in hypergraphs, which correspond to barriers in space, divisibility and covering. It is natural to ask when these conditions are asymptotically sufficient. Our main result confirms this for hypergraph families that are approximately closed under taking a typical induced subgraph of constant order. As an application, we recover and extend a series of well-known results for perfect tilings in hypergraphs and related settings involving vertex orderings and rainbow structures.

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-097

1 Introduction

A basic question in combinatorics is whether a combinatorial object on a ground set of vertices contains a particular substructure that spans all vertices. Since the corresponding decision problems are typically computationally intractable, the 'extremal' approach has focused on identifying easily verifiable sufficient conditions, a classic example being minimum degree conditions in the graph setting. Over the past decades, a robust literature has developed around these problems [9, 13, 21, 22], yet many questions remain widely open.

More recently, efforts have increasingly been dedicated to formulating an axiomatic approach. The idea is to identify a set of 'simple' conditions that are essentially necessary for the existence of the desired substructure. One then aims to show that satisfying these properties in a robust manner guarantees the substructure in question. Important

*Fachbereich Mathematik, Universität Hamburg, 20146 Hamburg, Germany. E-mail: richard.lang@uni-hamburg.de. This research was supported by DFG (450397222), FAPESP (21/11020-9) and H2020-MSCA (101018431).

milestones in this line of research are due to Keevash and Mycroft [10], Han [5] (perfect matchings) and Keevash [8] (designs) in the hypergraph setting. For graphs, analogous results have been obtained by Kühn, Osthus and Treglown [15] (Hamilton cycles), Knox and Treglown [11], Lang and Sanhueza-Matamala [16] (easily separable graphs) as well as Hurley, Joos and Lang [7] (perfect mixed tilings).

We continue this branch of research by introducing a framework for perfect tilings in hypergraphs. The literature on this subject has identified three natural barriers that prevent perfect tilings, which correspond to obstructions in space, divisibility and covering. Freschi and Treglown [4] raised the 'meta question' whether this already includes all relevant obstacles. We answer this in the affirmative for hypergraphs families (and related structures) whose features are approximately replicated by a typical subgraph of constant order. Our main contribution states that any hypergraph which robustly overcomes each of the obstructions must already contain a perfect tiling.

As an application, we recover and extend the milestones for perfect tilings under minimum degree conditions in graphs [12, 14] and hypergraphs [6, 19] as well as recent breakthroughs in the ordered setting [4], quasirandom setting [2] and the rainbow setting [1, 18].

2 A framework for hypergraph tiling

In the following, we formulate a simplified version of our main result. A *k-uniform hypergraph* (or *k-graph* for short) G consists of vertices $V(G)$ and edges $E(G)$, where each edge is a set of k vertices. Given another k-graph F, our goal is to find a *perfect F-tiling* in G, which is a collection of pairwise disjoint copies of F that together cover all vertices of G. Note that the special case, when F is a single edge, corresponds to a *perfect matching.* We denote by $\mathrm{Hom}(F;G)$ the set of *homomorphisms* from F to G, meaning the functions $\phi\colon V(F) \to V(G)$ that map edges of F to edges of G.

Obstacles for perfect tilings

Let us review three essentially necessary conditions for perfect tilings in hypergraphs.

Space. A first obstruction to perfect tilings involves space. For example, a simple instance of the *space barrier* is obtained by taking a complete graph and deleting the edges within a subset of more than half of the vertices. We formalise the corresponding space property via a linear programming relaxation. A *perfect fractional F-tiling* G is a function $\omega\colon \mathrm{Hom}(F;G) \to [0,1]$ such that for all $v \in V(G)$, we have $\sum_{\phi\in\mathrm{Hom}(F;G)} \omega(\phi)|\phi^{-1}(v)| = 1$. Let Spa_F be the set of k-graphs with a perfect fractional F-tiling.

Divisibility. Another type of obstacle for perfect tilings arises from divisibility. For instance, it is not possible to find a perfect matching in the union of two disjoint odd cliques — a basic example of the *divisibility barrier.* We can capture this phenomenon as follows. For a homomorphism $\phi \in \mathrm{Hom}(F;G)$, denote by $\mathbf{1}_\phi \in \mathbb{N}^{V(G)}$ the indicator vector of the image of ϕ, which satisfies $\mathbf{1}_\phi(v) = |\phi^{-1}(v)|$ for each $v \in V(G)$. Similarly, for a vertex $u \in V(G)$, denote by $\mathbf{1}_u$ the indicator vector with $\mathbf{1}_u(u) = 1$ and zero otherwise. The *F-lattice* of G is the additive subgroup $\mathcal{L}(F;G) \subseteq \mathbb{Z}^{V(G)}$ generated by the vectors $\mathbf{1}_\phi$ with $\phi \in \mathrm{Hom}(F;G)$. We say that $\mathcal{L}(F;G)$ is *complete* if it contains all *transferrals* $\mathbf{1}_v - \mathbf{1}_u$ with $u, v \in V(G)$. Denote by Div_F the set of k-graphs with complete F-lattice.

Cover. There are hypergraphs which satisfy the space and divisibility condition, but do not contain a perfect tiling simply because some vertices are not on any copy of F at all. Such a configuration is called a *cover barrier*. Motivated by this, we say that G is *F-covered* if for every vertex $v \in V(G)$, there is a homomorphism $\phi \in \mathrm{Hom}(F;G)$ such that $|\phi^{-1}(v)| = 1$. We denote by Cov_F the set of F-covered k-graphs.

Necessity. The next claim confirms that the space, divisibility and cover properties are essentially necessary for the existence of a perfect tiling. We abbreviate $m = |V(F)|$.

Observation 2.1. *If G has more than m vertices and contains a perfect F-tiling after deleting any choice of m vertices. Then G satisfies* Spa_F, Div_F *and* Cov_F.

Proof. The cover property follows trivially. The space property can be obtained by averaging over all fractional perfect F-tilings obtained after deleting m vertices. For the completeness of the lattice $\mathcal{L}(F;G)$, let $G' \subseteq G$ be obtained by deleting $m-1$ arbitrary vertices, and let $u, v \in V(G')$. By assumption, $G'-u$ has a perfect F-tiling $\mathcal{F}_u$, and $G'-v$ has a perfect F-tiling $\mathcal{F}_v$. We identify these tilings with the corresponding elements of $\mathrm{Hom}(F;G)$. It follows that $\mathcal{L}(F;G)$ contains the transferral $\mathbf{1}_v - \mathbf{1}_u = \sum_{\phi\in\mathcal{F}_v} \mathbf{1}_\phi - \sum_{\phi\in\mathcal{F}_u} \mathbf{1}_\phi$, as desired. □

Sufficient conditions for perfect matchings. Now we are ready to formulate our main result, which inverts the implication of Observation 2.1. It states that every hypergraph which robustly overcomes the space, divisibility and cover barrier has a perfect tiling. Our notion of robustness is formalised with the following key definition.

Definition 2.2 (Property graph)**.** For a k-graph G and property $\mathcal{P}$, the *property graph*, denoted by $P^{(s)}(G;\mathcal{P})$, is the s-uniform hypergraph on vertex set $V(G)$ with an edge $S \subseteq V(G)$ whenever the induced subgraph $G[S]$ satisfies $\mathcal{P}$.

Informally, we regard G as 'robustly' satisfying $\mathcal{P}$ if $P^{(s)}(G;\mathcal{P}')$ has minimum degree vertex $1 - \exp(-\Omega(s))$ where $\mathcal{P} \approx \mathcal{P}'$. However, in practise a lower degree condition suffices due the possibility of 'boosting'. Let $\delta(s)$ be the *minimum vertex degree threshold for perfect s-uniform matchings*, that is the least $\delta \in [0,1]$ such that for all $\mu > 0$ and n large enough and divisible by s, every n-vertex s-graph P with $\delta_1(P) \geq (\delta+\mu)\binom{n-1}{s-1}$ admits a perfect matching.

Theorem 2.3. *For every k-graph F on m vertices, $s \geq 1$ and $\mu > 0$ there is n_0 such that for all $n \geq n_0$ divisible by m the following holds. Let G be a k-graph on n vertices with*

$$\delta_1\big(P^{(s)}(G;\mathrm{Spa}_F \cap \mathrm{Div}_F \cap \mathrm{Cov}_F)\big) \geq (\delta(s)+\mu)\binom{n-1}{s-1}.$$

Then G has a perfect F-tiling.

Keevash and Mycroft [10] as well as Han [5] investigated similar phenomena in the setting of perfect matchings. These results differ in their notion of robustness and in their proof techniques. In particular, Keevash and Mycroft [10] introduced the concept of completeness for lattices and used it to find a suitable allocation for the Hypergraph Blow-up Lemma. Independently, Lo and Markström [17] developed an absorption-based approach to hypergraph tiling using a (more restrictive) form of lattice completeness. Han [5] combined and extended these ideas to give a simpler proof of the Keevash–Mycroft Theorem avoiding the (Strong) Hypergraph Regularity Lemma.

Our main framework contributes to this line of research in two ways. Firstly, the interface is simple but practical. For host graph families that are approximately closed under taking typical induced subgraphs of constant order, Theorem 2.3 practically decomposes the problem of finding perfect tilings into verifying the space, divisibility and cover properties separately, which greatly simplifies the analysis. The fact that the building blocks of these properties are formulated in terms of homomorphisms adds a lot of flexility to this approach. A more general result, which also applies to structures beyond hypergraphs is proved in the full version of the paper. In combination, we obtain short and insightful proofs of many old and new results.

The second important point about Theorem 2.3 is that the proof itself is quite short. The argument is self-contained, after discounting classic insights from combinatorics, and it does not involve the Regularity Lemma. The techniques can easily be extended to other configurations involving exceptional vertices and the partite setting. Finally, our framework can also be used to derive stability results via the theory of property testing.

Proof outline. Let us sketch the proof of Theorem 2.3. Consider a k-graph G which robustly satisfies $\mathcal{P} := \mathrm{Spa}_F \cap \mathrm{Div}_F \cap \mathrm{Cov}_F$. Our goal is to find a perfect F-tiling in G. For perfect matchings, this has been done by considering a partition $\mathcal{V}$ of $V(G)$ together with a reduced k-graph Γ on the clusters of $\mathcal{V}$, whose edges track the local edge densities of G. Under the right notion of robustness, this implies that Γ also satisfies $\mathcal{P}$. This framework allows to find a perfect matching in G either via a Hypergraph Blow-up Lemma [10] or via an absorption argument plus some classic insights on matchings in sparse graphs [5]. The main idea of our proof is to replace the reduced graph $\Gamma \in \mathcal{P}$, which approximates the whole structure of G, with a family of reduced k-graphs $\mathcal{R} \subseteq \mathcal{P}$, that describe parts of the local structure of G with higher accuracy.

To illustrate this, let us outline why G contains for some k-graph $R \in \mathcal{P}$ the blow-up R^*.[1] Recall that by assumption the property s-graph $P := P^{(s)}(G; \mathcal{P})$ is quite dense. Thus, by an old result of Erdős [3], we may find a complete s-partite s-graph $K \subseteq P$ with parts of size b where b is much larger than s. Note that each edge $S \in E(K)$ corresponds to an element $G[S]$ of $\mathcal{P}$, but for distinct edges these elements might differ or not have their vertices in the same parts of K. To deal with this, we give a colour to each of these

[1] Meaning that R^* is obtained by replacing each vertex of R by constant number of vertices and replacing the edges with complete partite subgraphs.

configurations and apply Ramsey's theorem. This results in a subgraph $R^* \subseteq K$ with (somewhat smaller) parts of size b' and an s-vertex k-graph $R \in \mathcal{P}$ such that every edge $S \in E(R^*)$ induces an k-graph isomorphic to R with its vertices in the 'same' parts; just as desired. We informally call R^* a '$\mathcal{P}$-blow-up' with 'local reduced graph' R.

Given this observation, the proof of Theorem 2.3 proceeds in two steps implementing the Absorption Method [20]. First we match most of the vertices, then we incorporate the leftover vertices. For the first step, we show that under the assumption that the property graph has large minimum vertex degree one can partition most of the vertices of G with $\mathcal{P}$-blow-ups. We then find an almost perfect tiling in each of these blow-ups. For the second step, we show that every set of m vertices is anchored in many $\mathcal{P}$-blow-ups. This allows us to reserve a small set of $\mathcal{P}$-blow-ups beforehand to host a special structure, which can be used to absorb the leftover vertices.

The remaining challenge of the proof then consists in spelling out the embedding arguments into the blow-ups. This step is equivalent to an allocation in the context of a Blow-up Lemma applied to a 'global reduced graph' Γ. However, in our context we allocate to local reduced graph R. Since its blow-up R^* is complete partite this immediately results in the desired embedding. So in particular, we may avoid the technical details of using a (Hypergraph) Blow-up Lemma.

References

[1] Y. Cheng, J. Han, B. Wang, and G. Wang, *Rainbow spanning structures in graph and hypergraph systems*, arXiv:2105.10219 (2021).

[2] L. Ding, J. Han, S. Sun, G. Wang, and W. Zhou, *F-factors in quasi-random hypergraphs*, arXiv:2108.10731 (2021).

[3] P. Erdős, *On extremal problems of graphs and generalized graphs*, Israel Journal of Mathematics **2** (1964), 183–190.

[4] A. Freschi and A. Treglown, *Dirac-type results for tilings and coverings in ordered graphs*, Forum Math. Sigma **10** (2022), e104.

[5] J. Han, *On perfect matchings in k-complexes*, Int. Math. Res. Not. **2021** (2021), no. 11, 8741–8762.

[6] J. Han, C. Zang, and Y. Zhao, *Minimum vertex degree thresholds for tiling complete 3-partite 3-graphs*, J. Combin. Theory Ser. A **149** (2017), 115–147.

[7] E. Hurley, F. Joos, and R. Lang, *Sufficient conditions for perfect tilings*, arXiv:2201.03944 (2022).

[8] P. Keevash, *The existence of designs*, arXiv:1401.3665 (2014).

[9] ———, *Hypergraph matchings and designs*, arXiv:1807.05752 (2018).

[10] P. Keevash and R. Mycroft, *A geometric theory for hypergraph matching*, vol. 233, Amer. Math. Soc., 2015.

[11] F. Knox and A. Treglown, *Embedding spanning bipartite graphs of small bandwidth*, Comb. Probab. Comput. **22** (2013), no. 1, 71–96.

[12] J. Komlós, *Tiling Turán theorems*, Combinatorica **20** (2000), no. 2, 203–218.

[13] D. Kühn and D. Osthus, *Embedding large subgraphs into dense graphs*, arXiv:0901.3541 (2009).

[14] ———, *The minimum degree threshold for perfect graph packings*, Combinatorica **29** (2009), no. 1, 65–107.

[15] D. Kühn, D. Osthus, and A. Treglown, *Hamiltonian degree sequences in digraphs*, J. Combin. Theory Ser. B **100** (2010), no. 4, 367–380.

[16] R. Lang and N. Sanhueza-Matamala, *On sufficient conditions for spanning substructures in dense graphs*, to appear in Proc. Lond. Math. Soc. (2023).

[17] A. Lo and K. Markström, *F-factors in hypergraphs via absorption*, Graphs Combin. **31** (2015), no. 3, 679–712.

[18] R. Montgomery, A. Müyesser, and Y. Pehova, *Transversal factors and spanning trees*, Adv. Comb. (2022), Paper No. 3, 25.

[19] R. Mycroft, *Packing k-partite k-uniform hypergraphs*, J. Combin. Theory Ser. A **138** (2016), 60–132.

[20] V. Rödl, A. Ruciński, and E. Szemerédi, *Perfect matchings in large uniform hypergraphs with large minimum collective degree*, J. Combin. Theory Ser. A **116** (2009), no. 3, 613–636.

[21] M. Simonovits and E. Szemerédi, *Embedding Graphs into Larger Graphs: Results, Methods, and Problems*, Building Bridges II: Mathematics of László Lovász, Bolyai Society Mathematical Studies, 28. Springer, 2019, pp. 445–592.

[22] Y. Zhao, *Recent advances on Dirac-type problems for hypergraphs*, Recent trends in combinatorics, 2016, pp. 145–165.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

On Perfect Subdivision Tilings

(Extended abstract)

Hyunwoo Lee*

Abstract

For a given graph H, we say that a graph G has a perfect H-subdivision tiling if G contains a collection of vertex-disjoint subdivisions of H covering all vertices of G. Let $\delta_{sub}(n, H)$ be the smallest integer k such that any n-vertex graph G with minimum degree at least k has a perfect H-subdivision tiling. For every graph H, we asymptotically determined the value of $\delta_{sub}(n, H)$. More precisely, for every graph H with at least one edge, there is a constant $1 < \xi^*(H) \leq 2$ such that $\delta_{sub}(n, H) = \left(1 - \frac{1}{\xi^*(H)} + o(1)\right) n$ if H has a bipartite subdivision with two parts having different parities. Otherwise, the threshold may depend on the parity of n.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-098`

1 Introduction

Embedding a large sparse subgraph into a dense graph is one of the most central problems in extremal graph theory. It is well-known that any graph G with minimum degree at least $\lfloor \frac{v(G)}{2} \rfloor$ has a Hamiltonian cycle, hence also a perfect matching if the number of vertices $v(G)$ of G is even. A natural generalization of a perfect matching is a perfect H-tiling, for a general graph H. We say G has a perfect H-tiling if G contains a collection of vertex-disjoint copies of H, whose union covers all vertices of G. For a positive integer n divisible by $v(H)$, we denote by $\delta(n, H)$ the minimum integer k such that any n-vertex

*Department of Mathematical Sciences, KAIST, South Korea and Extremal Combinatorics and Probability Group (ECOPRO), Institute for Basic Science (IBS). E-mail: `hyunwoo.lee@kaist.ac.kr`. Supported by the POSCO Science Fellowship of POSCO TJ Park Foundation and the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIT) No. RS-2023-00210430, and the Institute for Basic Science (IBS-R029-C4).

graph G with minimum degree at least k has a perfect H-tiling. For any integer $r \geq 2$, the Hajnal-Szemerédi [11] theorem states that the number $\delta(n, K_r)$ is equal to $\left(1 - \frac{1}{r}\right) n$.

The minimum degree threshold of perfect tiling for general graph H was first proved by Alon and Yuster [2]. They showed that if n is divisible by $v(H)$, then $\delta(n, H) \leq \left(1 - \frac{1}{\chi(H)}\right) n + o(n)$, where $\chi(H)$ is a chromatic number of H. Komlós, Sárközy, and Szemerédi [18] improved the $o(n)$ term in Alon-Yuster theorem to some constant $C = C(H)$, which settled the conjecture of Alon and Yuster [2]. Another direction for an asymptotic extension of Hajnal-Szemerédi theorem was proved by Komlós [17]. We write the critical chromatic number of H as $\chi_{cr}(H)$, which is defined as $\chi_{cr}(H) = \frac{(\chi(H)-1)v(H)}{v(H)-\sigma(H)}$, where $\sigma(H)$ is the minimum possible size of color class in the optimal proper coloring of H. Komlós showed that for any $\gamma > 0$, there exists $n_0 = n_0(\gamma, H)$ such that if $n \geq n_0$, then for any n-vertex graph G whose minimum degree is at least $\left(1 - \frac{1}{\chi_{cr}(H)}\right) n$ contains an H-tiling which covers at least $(1-\gamma)n$ vertices of G. Komlós [17] conjectured that the number of uncovered vertices can be reduced to a constant and this conjecture was confirmed by Shokoufandeh and Zhao [30]. More precisely, the following holds.

Theorem 1.1 (Shokoufandeh and Zhao [30])**.** *Let H be a graph. Then there exists a constant $C = C(H)$, which only depends on H such that any graph G on n-vertices with minimum degree at least $\left(1 - \frac{1}{\chi_{cr}(H)}\right) n$ contains a H-tiling which covers all but at most C vertices of G.*

The almost exact value of $\delta(n, H)$ for every graph H was determined by Kühn and Osthus [23] up to an additive constant depending only on H.

Theorem 1.2 (Kühn and Osthus [23])**.** *Let H be a graph and n be a positive integer which divisible by $v(H)$. Then there exist a constant $C = C(H)$ and $\chi(H) - 1 < \chi^*(H) \leq \chi(H)$ depending only on H such that*

$$\left(1 - \frac{1}{\chi^*(H)}\right) n - 1 \leq \delta(n, H) \leq \left(1 - \frac{1}{\chi^*(H)}\right) n + C.$$

Indeed, in [23], the authors stated how we can compute $\chi^*(H)$ for a given graph H.

1.1 Main results

Motivated by the Kühn-Osthus theorem on perfect H-tilings, several variations of Theorem 1.2 were considered. For instances, see [5, 10, 12, 13, 14, 22, 26].

We consider a problem related to the concept of perfect H-tilings and subdivision embeddings. Consider graphs G and H. We say for a graph H' is a subdivision of H if H' is obtained from H by replacing edges of H to vertex-disjoint paths. Let H and G be graphs. An *H-subdivision tiling* is a collection of disjoint union of subdivisions of H. We say that G has a *perfect H-subdivision tiling* if G has an H-subdivision tiling which covers all vertices of G. A natural question would be to determine the minimum degree threshold which ensures the existence of perfect H-subdivision tiling in any n-vertex graph G. We define this minimum degree threshold as the following.

Definition 1.3. Let H be a graph. We denote the minimum degree threshold for perfect H-subdivision tilings by $\delta_{sub}(n, H)$, which is the smallest integer k such that any n-vertex graph G with minimum degree at least k has a perfect H-subdivision tiling.

If H has no edges, then perfect H-subdivision tiling exists if and only if $v(G)$ is divisible by $v(H)$, regardless of the minimum degree $\delta(G)$ of G. Thus, from now on, we only consider graphs with at least one edge.

Since embedding bipartite graphs generally requires less minimum degree than non-bipartite graphs, we want to cover most of the vertices of the host graph with subdivisions of H that are bipartite. Suppose every bipartite subdivision of H is in some sense balanced. In that case, one cannot perfectly tile them in a highly unbalanced complete bipartite graph which has a smaller minimum degree than a balanced complete bipartite graph. For this reason, we need to measure how unbalanced bipartite subdivisions of H can be, as it poses some space barriers on the problem.

For this purpose, we introduce the following two definitions.

Definition 1.4. Let H be a graph and $X \subseteq V(H)$. We define a function $f_H : 2^{V(H)} \to \mathbb{R}$ as $f_H(X) = \frac{v(H)+e(H[X])+e(H[Y])}{|X|+e(H[Y])}$ where $Y = V(H) \setminus X$.

Definition 1.5. Let H be a graph. We define $\xi(H) \coloneqq \min\{f_H(X) : X \subseteq V(H)\}$.

Note that we always have $1 < \xi(H) \leq 2$. Another crucial factor for perfect subdivision tiling problem is the divisibility issue. Assume that all bipartite subdivisions of H have bipartitions with both parts having the same parity. If G is a complete bipartite graph $K_{a,b}$ with a, b having different parities, then we cannot find a perfect H-subdivision tiling in G, as it poses some divisibility barriers on the problem. Hence, we need to introduce the following definitions concerning the difference between two parts in bipartitions of subdivisions of H and their highest common factor.

Definition 1.6. Let H be a graph. We define $\mathcal{C}(H) \coloneqq \{(|X|+e(H[Y]))-(|Y|+e(H[X])) : X \subseteq V(H), Y = V(H) \setminus X\}$. We denote by $hcf_\xi(H)$ the highest common factor of all integers in $\mathcal{C}(H)$. (If $\mathcal{C}(H) = \{0\}$, we define $hcf_\xi(H) = \infty$.)

By considering the space and divisibility barriers, we introduce the following parameter measuring both obstacles for the problem. We will show that this is the determining factor for $\delta_{sub}(n, H)$.

Definition 1.7. Let H be a graph. We define

$$\xi^*(H) \coloneqq \begin{cases} \xi(H) & \text{if } hcf_\xi(H) = 1, \\ \max\{\frac{3}{2}, \xi(H)\} & \text{if } hcf_\xi(H) = 2, \\ 2 & \text{otherwise.} \end{cases}$$

We are now ready to state our main theorems. The following theorem gives an asymptotically exact value for $\delta_{sub}(n, H)$ except only one case, that is when $hcf_\xi(H) = 2$.

Theorem 1.8. *Let H be a graph with $hcf_\xi(H) \neq 2$. For every $\gamma > 0$, there exists an integer $n_0 = n_0(\gamma, H)$ such that the following holds. For every integer $n \geq n_0$,*

$$\left(1 - \frac{1}{\xi^*(H)}\right) n - 1 \leq \delta_{sub}(n, H) \leq \left(1 - \frac{1}{\xi^*(H)} + \gamma\right) n$$

This theorem asymptotically determine $\delta_{sub}(n, H)$ as long as $hcf_\xi(H) \neq 2$. If $hcf_\xi(H) = 2$, then the parity of n is also important. The following theorem asymptotically determines $\delta_{sub}(n, H)$ for this case.

Theorem 1.9. *Let H be a graph with $hcf_\xi(H) = 2$. For every $\gamma > 0$, there exists an integer $n_0 = n_0(\gamma, H)$ such that the following holds. For every integer $n \geq n_0$,*

$$\begin{aligned} \frac{1}{2}n - 1 \leq \delta_{sub}(n, H) &\leq \left(\frac{1}{2} + \gamma\right) n && \text{if } n \text{ is odd,} \\ \left(1 - \frac{1}{\xi^*(H)}\right) n - 1 \leq \delta_{sub}(n, H) &\leq \left(1 - \frac{1}{\xi^*(H)} + \gamma\right) n && \text{if } n \text{ is even.} \end{aligned}$$

One consequence of Theorems 1.8 and 1.9 is that the value of $\delta_{sub}(n, K_r)$ behaves unpredictably when r is small. Indeed, $\delta_{sub}(n, K_2) = \left(\frac{1}{3} + o(1)\right) n$ and for each $r \in \{3, 4, 5\}$, we have $\delta_{sub}(n, K_r) = \left(\frac{2}{r+1} + o(1)\right) n$. For the case $r = 7$, if n is even, we have $\delta_{sub}(n, K_7) = \left(\frac{1}{3} + o(1)\right) n$ otherwise, we have $\delta_{sub}(n, K_7) = \left(\frac{1}{2} + o(1)\right) n$. Finally, for every $r \geq 8$ and $r = 6$, we have $\delta_{sub}(n, K_r) = \left(\frac{1}{2} + o(1)\right) n$. This is contrasting to normal H-tiling problem. This means determining factors for minimum degree thresholds of perfect H-tilings and perfect H-subdivision tilings are essentially different. Probably, the most interesting difference between the perfect H-tiling and the perfect H-subdivision tiling is that the monotonicity does not hold for subdivision tiling. For a perfect tiling, if H_2 is a spanning subgraph of H_1, then obviously $\delta(n, H_2) \leq \delta(n, H_1)$. However, for perfect subdivision tiling, this does not hold in many cases. For example, our results implies $\delta_{sub}(n, K_4) = \frac{2}{5}n + o(n) < \delta_{sub}(n, C_4) = \frac{1}{2}n + o(n)$.

As $\xi^*(H)$ is the determining factor for the minimum degree threshold, it is convenient for us to specify a bipartite subdivision achieving the value $\xi^*(H)$. We introduce the following definition.

Definition 1.10. Let H be a graph. We denote by X_H a subset of $V(H)$, where $f_H(X_H) = \xi(H)$. If there are multiple choices of X_H, we fix one choice arbitrarily. We define a graph H^* obtained from H by replacing all edges in $H[X_H]$ and $H[V(H) \setminus X_H]$ to paths of length two.

Note that H^* is a subdivision of H, which is a bipartite graph and $v(H^*) = v(H) + e(H[X_H]) + e(H[V(H) \setminus X_H])$.

We observe that the inequality $\xi(H) \geq \chi^*(H^*)$ holds. Hence, if $hcf_\xi(H) \leq 2$, we may use Theorem 1.1 to find an H^*-tiling that covers all but at most constant number of vertices of G in a graph G with $\delta(G) \geq \left(1 - \frac{1}{\xi(H)} + \gamma\right) n$. In order to cover the leftover vertices, we use the absorption method. The absorption method was introduced in [28], and since then, it has been used to solve various crucial problems in extremal combinatorics. The main difficulty to apply the absorption method in our setting is that in many cases, the host graph is not sufficiently dense to guarantee that any vertices can be absorbed in the final step. To overcome this difficulty, we use the regularity lemma and an extremal result on the domination number to obtain some control over the vertices that can be absorbed.

2 Proof overview

2.1 Lower bounds

It is easy to check that the following observation holds.

Observation 2.1. *Let H be a graph and let F be a bipartite subdivision of H with bipartition (A, B). Then $\frac{|B|}{|A|} \leq \frac{1}{\xi(H)-1}$.*

This observation allows us to obtain the following proposition, which poses a lower bound for $\delta_{sub}(n, H)$ when we do not care about $hcf_\xi(H)$. Since $\xi(H)$ measures how unbalanced a bipartition of a subdivision of H can be, if the given host graph is a sufficiently unbalanced complete bipartite graph, then we cannot perfectly tile it with the subdivisions of H. Thus, we can deduce the following.

Proposition 2.2. *For every integer $n > 0$ and every graph H, there is an n-vetex graph G with minimum degree at least $\lfloor \left(1 - \frac{1}{\xi(H)}\right) n \rfloor - 1$ such that G does not have a perfect H-subdivision tiling.*

Now we cause the divisibility issue to construct a lower bound example. To obtain the lower bound in Theorem 1.8 and the first case of Theorem 1.9, we prove that the following proposition holds.

Proposition 2.3. *Let H be a graph with $hcf_\xi(H) \neq 1$. Then for every integer $n > 0$, there is an n-vertex graph G with minimum degree at least $\lfloor \frac{n}{2} \rfloor - 1$ which does not have a perfect H-subdivision tiling except for $hcf_\xi(H) = 2$ and n is even.*

The remaining case is when $hcf_\xi(H) = 2$ and n is even. The lower bound of this case can be obtained from the following.

Proposition 2.4. *For every graph H with $hcf_\xi(H) = 2$ and for every even number n, there is an n-vertex graph G with minimum degree at least $\lfloor \frac{1}{3} n \rfloor - 1$ such that G does not contain a perfect H-subdivision tiling.*

Both the proof of Propositions 2.3 and 2.4 rely on the observation that if the host graph is a complete bipartite graph with the difference between two bipartitions are not divisible by $hcf_\xi(H)$, then there is no perfect H-subdivision tiling. This can be verified by the definition of $hcf_\xi(H)$.

2.2 Upper bounds

We now sketch the proof of our main results. We first start with the following observation.

Observation 2.5. $\delta_{sub}(n, H) \leq \left(\frac{1}{2} + o(1)\right) n$.

Indeed, for every graph H, there is at least one bipartite subdivision of H. By using Erdős-Stone-Simonovits theorem and Theorem 1.2, we can deduce that $\delta_{sub}(n, H) \leq \left(\frac{1}{2} + o(1)\right) n$. As Propositions 2.2 to 2.4 provides desired lower bounds, Observation 2.5 implies that it suffices to prove the two following lemmas.

Lemma 2.6. *Let $hcf_\xi(H) = 1$ and n be a sufficiently large number. If $\delta(G) \geq \left(1 - \frac{1}{\xi(H)} + o(1)\right) n$, then G has a perfect H-subdivision tiling.*

Lemma 2.7. *Let $hcf(\xi)(H) = 2$ and n be a sufficiently large even number. If $\delta(G) \geq \left(\max\{\frac{1}{3}, 1 - \frac{1}{\xi(H)}\} + o(1)\right) n$, then G has a perfect H-subdivision tiling.*

In order to prove the above lemmas, we use the absorption method. Since we are dealing with a subdivision embedding problem, we define our absorber as follows.

Definition 2.8. Let H and G be graphs and take two subsets $A \subseteq V(G)$ and $X \subseteq V(G) \setminus A$. We say A is a *$Sub(H)$-absorber* for X if the both $G[A]$ and $G[A \cup X]$ have perfect H-subdivision tilings. If $X = \{v\}$, we say A is a $Sub(H)$-absorber for v.

For example, we consider an appropriate subdivision of H with an edge xy in it and add two edges vx and vy to obtain a graph H'. Then a copy of H' ensures that $V(H') - \{v\}$ is a $Sub(H)$-absorber for v. In order to establish robust absorption structures, we wish to collect many vertices that belong to many copies of such graphs H'. We will ensure this using the concept of ε-regularity.

The following is the proof outline of Lemmas 2.6 and 2.7. We omit the details of the argument as we provide them in the full version [25] of the paper.

Step 1: Preprocessing. In order to find many copies of H' containing a given vertex v, we plan to utilize the concept of ε-regularity. For this, we apply the regularity lemma and use it to obtain many disjoint ε-regular pairs covering almost all vertices. Note that those ε-regular pairs are allowed to be somewhat asymmetric. By deleting a small number of vertices, we can further ensure some minimum degree condition on every ε-regular pair. Let Z be the small set of vertices not covered by the obtained ε-regular pairs with the minimum degree condition.

Step 2: Place the absorber. In each regular pair, the ε-regularity and the minimum degree condition ensure that every vertex v in it belongs to many copies of H'. Using this property, we can find a small subset $A \subseteq (V(G) \setminus Z)$ such that A is $Sub(H)$-absorber for any small set $X \subseteq V(G) \setminus (Z \cup A)$.

Step 3: Cover almost all vertices. By considering a suitable bipartite subdivision H'' of H and applying Erdős-Stone-Simonovits theorem, we find copies of H'' disjoint from A to cover all vertices of Z as well as a small set of additional vertices. Denote the set of such vertices as W_1. As $|A \cup W_1|$ is small, the remaining graph $G \setminus (A \cup W_1)$ still has almost the same minimum degree as G. By applying Theorem 1.1, we can find $W_2 \subseteq V(G) \setminus (A \cup W_1)$ such that $G[W_2]$ has a perfect H-subdivision tiling and $|V(G) \setminus (A \cup W_1 \cup W_2)|$ is small.

Step 4: Absorb the uncovered vertices. Let $X = V(G) \setminus (A \cup W_1 \cup W_2)$. Since X is small, by our choice of A, the set A is $Sub(H)$-absorber for X. This means $G[A \cup X]$ has a perfect H-subdivision tiling. Since A, W_1, W_2 and X are vertex-disjoint sets and $A \cup W_1 \cup W_2 \cup X = V(G)$, we obtain a perfect H-subdivision tiling of G by combining $G[A \cup X]$, $G[W_1]$ and $G[W_2]$.

References

[1] Noga Alon. Transversal numbers of uniform hypergraphs. *Graphs and Combinatorics*, 6(1):1–4, 1990.

[2] Noga Alon and Raphael Yuster. H-factors in dense graphs. *Journal of Combinatorial Theory, Series B*, 66(2):269–282, 1996.

[3] B Andrásfal. Graphentheoretische extremalprobleme. *Acta Mathematica Hungarica*, 15(3-4):413–438, 1964.

[4] Vladimir I Arnautov. Estimation of the exterior stability number of a graph by means of the minimal degree of the vertices. *Prikl. Mat. i Programmirovanie*, 11(3-8):126, 1974.

[5] József Balogh, Lina Li, and Andrew Treglown. Tilings in vertex ordered graphs. *Journal of Combinatorial Theory, Series B*, 155:171–201, 2022.

[6] Julia Böttcher, Mathias Schacht, and Anusch Taraz. Proof of the bandwidth conjecture of bollobás and komlós. *Mathematische Annalen*, 343(1):175–205, 2009.

[7] Paul Erdős and Miklós Simonovits. A limit theorem in graph theory. In *Studia Sci. Math. Hung.* Citeseer, 1965.

[8] Paul Erdős and Miklós Simonovits. Supersaturated graphs and hypergraphs. *Combinatorica*, 3(2):181–192, 1983.

[9] Paul Erdös and Arthur H Stone. On the structure of linear graphs. *Bulletin of the American Mathematical Society*, 52(12):1087–1091, 1946.

[10] Andrea Freschi and Andrew Treglown. Dirac-type results for tilings and coverings in ordered graphs. In *Forum of Mathematics, Sigma*, volume 10, page e104. Cambridge University Press, 2022.

[11] András Hajnal and Endre Szemerédi. Proof of a conjecture of p. erdos. *Combinatorial theory and its applications*, 2:601–623, 1970.

[12] Jie Han, Patrick Morris, and Andrew Treglown. Tilings in randomly perturbed graphs: Bridging the gap between hajnal-szemerédi and johansson-kahn-vu. *Random Structures & Algorithms*, 58(3):480–516, 2021.

[13] Eoin Hurley, Felix Joos, and Richard Lang. Sufficient conditions for perfect mixed tilings. *arXiv preprint arXiv:2201.03944*, 2022.

[14] Joseph Hyde, Hong Liu, and Andrew Treglown. A degree sequence komlós theorem. *SIAM Journal on Discrete Mathematics*, 33(4):2041–2061, 2019.

[15] GOH Katona, T Nemetz, and M Simonovits. On a graph-problem of turán in the theory of graphs. *Matematikai Lapok*, 15:228–238, 1964.

[16] Jaehoon Kim, Daniela Kühn, Deryk Osthus, and Mykhaylo Tyomkyn. A blow-up lemma for approximate decompositions. *Transactions of the American Mathematical Society*, 371(7):4655–4742, 2019.

[17] János Komlós. Tiling turán theorems. *Combinatorica*, 20(2):203–218, 2000.

[18] János Komlós, Gábor Sárközy, and Endre Szemerédi. Proof of the alon–yuster conjecture. *Discrete Mathematics*, 235(1-3):255–269, 2001.

[19] János Komlós, Gábor N Sárközy, and Endre Szemerédi. Blow-up lemma. *Combinatorica*, 17(1):109–123, 1997.

[20] János Komlós, Ali Shokoufandeh, Miklós Simonovits, and Endre Szemerédi. The regularity lemma and its applications in graph theory. *Summer school on theoretical aspects of computer science*, pages 84–112, 2002.

[21] János Komlós and Miklós Simonovits. Szemerédi's regularity lemma and its applications in graph theory. 1996.

[22] Daniela Kühn and Deryk Osthus. Embedding large subgraphs into dense graphs. *arXiv preprint arXiv:0901.3541*, 2009.

[23] Daniela Kühn and Deryk Osthus. The minimum degree threshold for perfect graph packings. *Combinatorica*, 29(1):65–107, 2009.

[24] Daniela Kühn, Deryk Osthus, and Anusch Taraz. Large planar subgraphs in dense graphs. *Journal of Combinatorial Theory, Series B*, 95(2):263–282, 2005.

[25] Hyunwoo Lee. On perfect subdivision tilings. *arXiv preprint arXiv:2302.09393*, 2023.

[26] Allan Lo and Klas Markström. f-factors in hypergraphs via absorption. *Graphs and Combinatorics*, 31(3):679–712, 2015.

[27] Charles Payan. Sur le nombre d'absorption d'un graphe simple. 1975.

[28] Vojtěch Rödl, Andrzej Ruciński, and Endre Szemerédi. A dirac-type theorem for 3-uniform hypergraphs. *Combinatorics, Probability and Computing*, 15(1-2):229–251, 2006.

[29] Vojtěch Rödl and Mathias Schacht. Regularity lemmas for graphs. In *Fete of combinatorics and computer science*, pages 287–325. Springer, 2010.

[30] Ali Shokoufandeh and Yi Zhao. Proof of a tiling conjecture of komlós. *Random Structures & Algorithms*, 23(2):180–205, 2003.

[31] Endre Szemerédi. Regular partitions of graphs. Technical report, Stanford Univ Calif Dept of Computer Science, 1975.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Cycle Partition of Dense Regular Digraphs and Oriented Graphs

(Extended abstract)

Allan Lo* Viresh Patel† Mehmet Akif Yıldız‡

Abstract

Magnant and Martin [24] conjectured that every d-regular graph on n vertices can be covered by $n/(d+1)$ vertex-disjoint paths. Gruslys and Letzter [11] verified this conjecture in the dense case, even for cycles rather than paths. We prove the analogous result for directed graphs and oriented graphs, that is, for all $\alpha > 0$, there exists $n_0 = n_0(\alpha)$ such that every d-regular digraph on n vertices with $d \geq \alpha n$ can be covered by at most $n/(d+1)$ vertex-disjoint cycles. Moreover if G is an oriented graph, then $n/(2d+1)$ cycles suffice. This also establishes Jackson's long standing conjecture [14] for large n that every d-regular oriented graph on n vertices with $n \leq 4d+1$ is Hamiltonian.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-099`

1 Introduction

As one of the most intensely studied structures in the graph theory, *a Hamilton cycle* in a (directed) graph is a (directed) cycle that visits every vertex. There are numerous

*School of Mathematics, University of Birmingham, UK. E-mail: `s.a.lo@bham.ac.uk`. Partially supported by EPSRC, grant no. EP/V002279/1 and EP/V048287/1. There are no additional data beyond that contained within the main manuscript.

†School of Mathematical Sciences, Queen Mary University of London, UK. E-mail: `viresh.patel@qmul.ac.uk`. Partially supported by the Netherlands Organisation for Scientific Research (NWO) Gravitation project NETWORKS (grant no. 024.002.003).

‡Korteweg de Vries Instituut voor Wiskunde, Universiteit van Amsterdam, The Netherlands. E-mail: `m.a.yildiz@uva.nl`. Supported by a Marie Skłodowska-Curie Action from the EC (COFUND grant no. 945045) and by the NWO Gravitation project NETWORKS (grant no. 024.002.003).

results that establish (best-possible) conditions guaranteeing the existence of Hamilton cycles in (directed) graphs. The seminal result of Dirac [5] states that every graph on $n \geq 3$ vertices with minimum degree at least $n/2$ is Hamiltonian. Ghouila-Houri [10] showed the corresponding version in directed graphs (*digraph* for short), that is, every digraph on $n \geq 3$ vertices with minimum semidegree at least $n/2$ (i.e. every vertex has in- and outdegree at least $n/2$) is Hamiltonian. These bounds are tight by taking e.g. the disjoint union of two cliques (a regular extremal example) or a slightly imbalanced complete bipartite graph (an irregular extremal example). Recall that an oriented graph is a digraph that can have at most one edge between each pair of vertices (whereas a digraph can have up to two, one in each direction). For oriented graphs, a more recent result of Keevash, Kühn and Osthus [15] establishes a (tight) minimum degree threshold of $\lceil (3n-4)/8 \rceil$ for Hamiltonicity. In contrast to graphs and digraphs, there are no *regular* extremal examples in the case of oriented graphs. Jackson [14] conjectured in 1981 that regularity actually reduces the degree threshold significantly for oriented graphs:

Conjecture 1.1 (Jackson [14])**.** *For each $d > 2$, every d-regular oriented graph on $n \leq 4d+1$ vertices has a Hamilton cycle.*

The disjoint union of two regular tournaments shows that Jackson's conjecture is best possible.[1] We note that the approximate version of Jackson's conjecture was recently verified by current authors in [23], that is for every $\varepsilon > 0$, there exists $n_0(\varepsilon)$ such that every d-regular oriented graph on $n \geq n_0(\varepsilon)$ vertices with $d \geq (1/4+\varepsilon)n$ is Hamiltonian. Here, we verify the exact version for large n.

Theorem 1.2. *There exists an integer n_0 such that every d-regular oriented graph on $n \geq n_0$ vertices with $n \leq 4d+1$ has a Hamilton cycle.*

Generalizing questions about Hamilton cycles, one can consider the question of covering the vertices of a (di)graph by as few vertex-disjoint cycles as possible. Indeed, we prove Theorem 1.2 by showing a more general result about covering regular digraphs and oriented graphs with few vertex-disjoint cycles.

Theorem 1.3. *For all $\alpha > 0$, there exists $n_0 = n_0(\alpha)$ such that every d-regular digraph G on n vertices with $d \geq \alpha n$ can be covered by at most $n/(d+1)$ vertex-disjoint cycles. Moreover if G is an oriented graph, then $n/(2d+1)$ cycles suffice.*

This is best possible by considering the disjoint union of complete digraphs of order $d+1$ for digraphs and the disjoint union of regular tournaments of order $2d+1$ for oriented graphs. Notice that we have $n/(2d+1) < 2$ when $n \leq 4d+1$, so that Theorem 1.3 implies Theorem 1.2. We also note that Theorem 1.3 generalizes the following result of Gruslys and Letzter [11] from regular graphs to regular digraphs and oriented graphs.

Theorem 1.4 (Gruslys and Letzter [11])**.** *For all $\alpha > 0$, there exists $n_0 = n_0(\alpha)$ such that every d-regular graph G on $n \geq n_0$ vertices with $d \geq \alpha n$ can be covered by at most $n/(d+1)$ vertex-disjoint cycles.*

[1] This example works for $n \equiv 2 \pmod 4$. Similar examples can also be constructed when $n \not\equiv 2 \pmod 4$.

Theorem 1.3 implies Theorem 1.4 by making every edge into a directed 2-cycle. Theorem 1.3 also has connections with several well-studied problems in extremal graph theory: here we mention some of them.

1.1 Path Cover

A weaker version of cycle cover is path cover. The *path cover number* $\pi(G)$ of a (di)graph G is the minimum number of vertex-disjoint (directed) paths needed to cover $V(G)$. This was introduced by Ore [25], and he showed that $\pi(G) \le n - \sigma_2(G)$ holds where $\sigma_2(G)$ denotes the minimum sum of the degrees over all non-adjacent vertices. Magnant and Martin [24] conjectured that *regularity* significantly reduces the upper bound for $\pi(G)$:

Conjecture 1.5 (Magnant and Martin [24])**.** *If G is a d-regular graph on n vertices, then $\pi(G) \le n/(d+1)$.*

It is known that Conjecture 1.5 holds for small values of d (see [24] for $d \le 5$ and see [7] for $d = 6$). Han [13] showed that, for dense graphs, it is enough to use $1 + n/(d+1)$ paths to cover almost all vertices. Also, Theorem 1.4 verifies Conjecture 1.5 in the dense case. It is worth noting that the Linear Arboricity Conjecture [2] implies Conjecture 1.5 for odd values of d, and gives $\pi(G) \le 2n/(d+2)$ for general d (see [7] for a detailed discussion).

For digraphs, the classical result of Gallai and Milgram [9] states that $\pi(G)$ can be bounded above by the size of maximum independent set (and Dilworth's [4] theorem says that the equality holds for the special case of posets). As our Theorem 1.3 generalizes Theorem 1.4 from graphs digraphs and oriented graphs, we believe the following stronger version of Conjecture 1.5 holds, which Theorem 1.3 establishes in the dense case.

Conjecture 1.6. *If G is a d-regular digraph on n vertices, then $\pi(G) \le n/(d+1)$. Moreover, $\pi(G) \le n/(2d+1)$ holds if G is oriented.*

Conjecture 1.6 implies Conjecture 1.5 by making every edge into a directed 2-cycle.

1.2 Edge-Disjoint Cycles

In a weaker version of the problem we consider, one is interested in finding *edge-disjoint* cycles whose union covers all the vertices. As a generalization of Dirac's theorem, it is conjectured [6] that if a graph G on n vertices has minimum degree n/k, then $V(G)$ can be covered by $k-1$ edge-disjoint cycles. The case $k = 3$ was also proved in [6]. The conjecture was proved for 2-connected graphs [16], and has been completely resolved in [17]. Later, Balogh, Mousset and Skokan [3] obtained a stability result, showing that every graph on n vertices with minimum degree nearly n/k has a special structure if it does not have $k-1$ edge-disjoint cycles covering all vertices. One can ask the same question for digraphs as a generalization of Ghoulia-Houri's theorem, and Theorem 1.3 answers it affirmatively for regular graphs:

Conjecture 1.7. *If G is a digraph with minimum semi-degree n/k, then $V(G)$ can be covered by $k-1$ edge-disjoint cycles.*

1.3 Extending Perfect Matchings

Gruslys and Letzter [11], as well as proving Theorem 1.4, proved that every large d-regular *bipartite graph* G on n vertices with d linear in n can be covered by at most $n/2d$ vertex-disjoint paths. They mentioned that one should be able to replace paths by cycles. Indeed, as a corollary of Theorem 1.3, the result below shows that those cycles can be found in such a way that they even contain any prescribed perfect matching.

Corollary 1.8. *For all $\alpha > 0$, there exists $n_1 = n_1(\alpha)$ such that, for every d-regular bipartite graph on $n \geq n_1$ vertices with $d \geq \alpha n$, any perfect matching can be extended to a union of at most $n/2d$ vertex-disjoint cycles.*

Note that Corollary 1.8 is tight by considering the disjoint union of $n/2d$ many $K_{d,d}$'s. It also shows that d-regular bipartite graphs on n (sufficiently large) vertices with $d > n/4$ are examples of graphs in which every perfect matching can be extended into a Hamilton cycle. This property is called the PMH-property in [1]. Häggkvist [12] initiated the study of sufficient conditions for the PMH-property (using the name F-Hamiltonian where F is a perfect matching) by showing $\sigma_2(G) \geq n+1$ is sufficient. Las Vergnas [21] proved a similar condition for bipartite graphs, and Yang [27] gave minimum edge density conditions to guarantee the PMH property in graphs and bipartite graphs. In the sparse setting, as a special case of a conjecture of Ruskey and Savage [26], Fink [8] proved that the hypercube has the PMH-property.

2 Sketch Proof

In this section, we sketch the proof of our main result Theorem 1.3. One of the key ingredients of the proof is a structural result that allows us to partition dense regular digraphs into (bipartite) robust expanders, which will be discussed in Section 2.1. In Section 2.2 we explain how a weaker version of Theorem 1.3 can be quickly derived from the structural result. In Sections 2.3, we briefly discuss some of the ingredients required for the full version of Theorem 1.3.

2.1 Robust Expanders

Robust expansion is a notion introduced and used by Kühn and Osthus together with several coauthors to obtain a number of breakthrough results on (di)graph decompositions and Hamiltonicity (see [20, 19, 18]). Here we present only the aspects relevant to the sketch proof and will suppress most parameters to ease exposition.

Informally, robust expanders are dense (di)graphs that are highly connected in some sense, and one of their key properties is that they are Hamiltonian under suitable (mild) degree conditions. If we could show that every d-regular digraph can be partitioned into at most $n/(qd + 1)$ robust expanders where $q = 2$ if G is oriented and $q = 1$ otherwise (we use this definition of q throughout the rest of the sketch proof), it would be enough to prove Theorem 1.3. Unfortunately, it is not true, but a similar result can be obtained by generalizing a structural result of Kühn, Lo, Osthus and Staden [18] about partitioning undirected graphs into robust expanders.

In order to give the reader some sense of robust (bipartite) expansion, we give the formal definition below but note that it will not be used in the sketch proof. Also the definition we give is slightly different but equivalent to that used in other work. Let $0 < \nu \le \tau < 1$ and suppose G is a digraph with subsets of vertices $A, B \subseteq V(G)$ (not necessarily disjoint) and $N := |A| + |B|$. We define $G[A, B]$ as the undirected bipartite graph on N vertices with bipartition A, B where, for each $a \in A$ and $b \in B$, ab is an (undirected) edge of $G[A, B]$ if and only if ab is a directed edge in $E(G)$. We say that $G[A, B]$ is a *bipartite robust* (ν, τ)-*expander* if for every $S \subseteq A$ with $\tau|A| \le |S| \le (1-\tau)|A|$, the set of vertices in B having at least νN inneighbours in A (in the graph G) has size at least $|S| + \nu N$. Henceforth, we suppress the parameters ν and τ and say simply that $G[A, B]$ is a bipartite robust expander.

For any d-regular n-vertex digraph G with d linear in n, we show that it is possible to give two vertex partitions $V(G) = V_{1*} \cup \cdots \cup V_{k*}$ and $V(G) = V_{*1} \cup \cdots \cup V_{*k}$ with $k \le 1 + n/(qd+1)$ such that for each i, $G[V_{i*}, V_{*i}]$ is a bipartite robust expander. Letting $V_{ij} = V_{i*} \cap V_{*j}$ for all $i, j \in [k]$, note that we actually give a k^2-partition $\{V_{ij} : i, j \in [k]\}$ of $V(G)$. The following is a simplified informal version of our structural result.

Theorem 2.1. *For any $\alpha > 0$, there exists an integer $n_0 = n_(\alpha)$ such that for all d-regular digraph graphs G on $n \ge n_0$ vertices with $d \ge \alpha n$, there is a partition $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ of $V(G)$ satisfying, for all $i \in [k]$,*

(i) *$G[V_{i*}, V_{*i}]$ is a bipartite robust expander with linear minimum degree;*
(ii) *$|V_{i*}| \approx |V_{*i}|$;*
(iii) *$k \le 1 + n/(qd+1)$.*

2.2 A Weaker Version of Theorem 1.3

Let G be as in Theorem 1.3, i.e. an n-vertex d-regular digraph with $d \ge \alpha n$ and n sufficiently large. In this subsection we describe how Theorem 2.1 can be used to show that almost all vertices of G can be covered by at most $1 + n/(qd+1)$ vertex-disjoint cycles (so one more cycle than stated in Theorem 1.3).

We apply Theorem 2.1 and obtain a partition $\{V_{ij} : i, j \in [k]\}$ of $V(G)$ satisfying (i)–(iii). By (ii), one can delete a small number of vertices in G so that $|V_{i*}| = |V_{*i}|$ holds for each $i \in [k]$ (for notational simplicity, we still write G and V_{ij} after deletion). A key property of bipartite robust expanders is that deleting any a small number of vertices only slightly weakens the bipartite robust expansion and minimum degree properties of (i). The following crucial observation shows that we can partition G into at most k vertex-disjoint cycles, which establishes the weaker version of Theorem 1.3 since $k \le 1+n/(qd+1)$ by (iii).

Fix $i \in [k]$, and assume $V_{ij} = \emptyset$ for all $j \in [k] \setminus \{i\}$, i.e. $V_{i*} = V_{ii} = V_{*i}$. In this case, one can use (i) to show that $G[V_{ii}]$ is a robust expander [2] with linear minimum degree; the

[2] Roughly speaking, G is a robust expander if $G[V(G), V(G)]$ is a bipartite robust expander.

result of [20] (see also [22]) then implies that $G[V_{ii}]$ is Hamiltonian. Now, assume $V_{ij} \neq \emptyset$ for at least one $j \in [k] \setminus \{i\}$. As $|V_{i*}| = |V_{*i}|$, we have $|V_{i*} \setminus V_{ii}| = |V_{*i} \setminus V_{ii}| > 0$, so write $V_{i*} \setminus V_{ii} = \{y_1, \dots, y_s\}$ and $V_{*i} \setminus V_{ii} = \{z_1, \dots, z_s\}$. Let $\phi : V_{i*} \to V_{*i}$ be given by $\phi(x) = x$ for all $x \in V_{ii}$ and $\phi(y_r) = z_r$ for all $r \in [s]$. Define $G(i, \phi)$ to be the digraph whose vertices are V_{i*} and with directed edge uw present in $G(i, \phi)$ if and only if $u\phi(w)$ is present in G. In other words, $G(i, \phi)$ is the digraph obtained from $G[V_{i*}, V_{*i}]$ by identifying each $u \in V_{i*} \setminus V_{ii}$ with $\phi(u) \in V_{*i} \setminus V_{ii}$ and deleting loops. By (i), one can show that $G(i, \phi)$ is a robust expander with linear minimum degree and hence (again by [20, 22]) contains a Hamilton cycle C. Without loss of generality, assume $y_1, \dots, y_s$ lie on C in this order. One can easily check that each path $y_r C y_{r+1}$ along C corresponds to a path in $G[V_{i*} \cup V_{*i}]$ from y_r to z_{r+1} and that these are vertex-disjoint and their union is $V_{i*} \cup V_{*i}$.

As a result, for each $i \in [k]$, we can cover $V_{i*} \cup V_{*i}$ by either a cycle (if $V_{ij} = \emptyset$ for all $j \in [k] \setminus \{i\}$) or a set of vertex-disjoint paths $\mathcal{Q}_i$ from $V_{i*} \setminus V_{ii}$ to $V_{*i} \setminus V_{ii}$. Note that the union of all these cycles and path systems gives a vertex-disjoint union of cycles covering G; indeed the path systems $\mathcal{Q}_i$ and $\mathcal{Q}_j$ only intersect in V_{ij} (where paths in $\mathcal{Q}_i$ start and paths in $\mathcal{Q}_j$ end) and in V_{ji} (where paths in $\mathcal{Q}_j$ start and paths in $\mathcal{Q}_i$ end). Our freedom to choose ϕ gives us some control over which pairs of endpoints are connected by paths in the $\mathcal{Q}_i$, and by choosing the ϕ's carefully, we can guarantee that the number of cycles in the union is at most k.

2.3 Balancing the Partition

We say $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ is a *balanced* partition of G if $|V_{i*}| = |V_{*i}|$ holds for each $i \in [k]$. Here we explain how to balance the partition $\mathcal{P}$ given in Theorem 2.1 in order that we can apply the methods described in the previous subsection. We use the idea of path contraction: consider a directed path in a digraph G and contract it so that the in- and outneighbours of the new vertex are respectively the inneighbours of the path's start vertex and the outneighbours of the path's end vertex. If the resulting graph can be partitioned into ℓ vertex-disjoint cycles, then so can G by simply uncontracting the path. Therefore, we seek a path system $\mathcal{Q}$ such that the contraction of $\mathcal{Q}$ makes the partition $\mathcal{P}$ balanced but also does not destroy the other properties of Theorem 2.1; the latter can (almost) be guaranteed by ensuring the number of edges of $\mathcal{Q}$ is small. It turns out (and is not difficult to show) that it suffices to find path systems $\mathcal{Q}_{ij}$ using edges of G from V_{i*} to V_{*j} satisfying

$$\sum_{j \neq i} e(\mathcal{Q}_{ij}) - \sum_{j \neq i} e(\mathcal{Q}_{ji}) = |V_{i*}| - |V_{*i}|.$$

We use a flow argument to find such a path system $\mathcal{Q}$.

Our argument up this point gives a collection of at most $k \leq 1 + n/(qd+1)$ vertex-disjoint cycles that cover G, which is one more than stated in Theorem 1.3. In fact, we only get $1 + n/(qd+1)$ cycles if Theorem 2.1 gives us a partition with $k = 1 + n/(qd+1)$ and $V_{ij} = \emptyset$ for all $i \neq j$. By carefully making use of the additional structure in this situation, we can reduce the number of cycles by 1.

References

[1] M. Abreu, J. B. Gauci, D. Labbate, G. Mazzuoccolo, and J. P. Zerafa. Extending perfect matchings to Hamiltonian cycles in line graphs. *Electron. J. Combin.* **28(1)**, (2021), Paper P1.7.

[2] J. Akiyama, G. Exoo, and F. Harary. Covering and packing in graphs IV: Linear arboricity. *Networks*, **11(1)**, (2006), 69–72.

[3] J. Balogh, F. Mousset, and J. Skokan. Stability for vertex cycle covers. *Electron. J. Combin.* **24** (2017), Paper P3.56.

[4] P. Dilworth. A decomposition theorem for partially ordered sets. *Ann. of Math.* **51(1)**, (1950), 161–166.

[5] G.A. Dirac. Some theorems on abstract graphs. *Proc. Lond. Math. Soc.* **2** (1952), 69–81.

[6] H. Enomoto, A. Kaneko, and Z. Tuza. P_3-factors and covering cycles in graphs of minimum degree $n/3$. *Combinatorics (Proc. Coll. Math. Soc. János Bolyai Eger, 1987)*, North Holland (1988), 213–220.

[7] U. Feige and E. Fuchs. On the path partition number of 6-regular graphs. *Journal of Graph Theory*, **101**, 345–378.

[8] J. Fink. Perfect matchings extend to Hamilton cycles in hypercubes. *J. Combin. Theory Ser. B*, **97** (2007) 1074–1076.

[9] T. Gallai and A. N. Milgram. Verallgemeinerung eines graphentheoretischen satzes von Redei. *Acta Sci. Math*, **21**, (1960), 181–186.

[10] A. Ghouila-Houri. Une condition suffisante d'existence d'un circuit hamiltonien. *C.R. Acad. Sci. Paris* **25** (1960), 495–497.

[11] V. Gruslys and S. Letzter. Cycle partitions of regular graphs. *Combin. Probab. Comput.*, **30(4)** (2021), 526–549.

[12] R. Häggkvist. On F-Hamiltonian graphs. In *Graph Theory and Related Topics* by J. A. Bondy, U.S.R. Murty, Academic Press, New York, (1979), 219–231.

[13] J. Han. On vertex-disjoint paths in regular graphs. *Electron. J. Combin.* **25** (2018), Paper P2.12.

[14] B. Jackson. Long paths and cycles in oriented graphs. *J. Graph Theory* **5** (1981), 245–252.

[15] P. Keevash, D. Kühn, and D. Osthus. An exact minimum degree condition for Hamilton cycles in oriented graphs. *J. Lond. Math. Soc.* **79** (2009), 144–166.

[16] M. Kouider. Covering vertices by cycles. *J. Graph Theory*, **18** (1994) 757–776.

[17] M. Kouider and Z. Lonc. Covering cycles and k-term degree sums. *Combinatorica*, **16** (1996), 407–412.

[18] D. Kühn, A. Lo, D. Osthus, and K. Staden. The robust component structure of dense regular graphs and applications. *Proc. Lond. Math. Soc.* **110(1)** (2015), 19–56.

[19] D. Kühn and D. Osthus. Hamilton decompositions of regular expanders: a proof of Kelly's conjecture for large tournaments. *Adv. Math.*, **237** (2013), 62–146.

[20] D. Kühn, D. Osthus, and A. Treglown. Hamiltonian degree sequences in digraphs. *J. Combin. Theory Ser. B*, **100** (2010), 367–380.

[21] M. Las Vergnas. Problèmes de couplages et problèmes hamiltoniens en Théorie des graphes. Ph.D. Thesis, University of Paris, (1972).

[22] A. Lo and V. Patel. Hamilton cycles in sparse robustly expanding digraphs. *Electron. J. Comb.*, **25(3)** (2018), P3.44

[23] A. Lo, V. Patel, and M.A. Yıldız. Hamilton cycles in dense regular digraphs and oriented graphs. arXiv: https://arxiv.org/abs/2203.10112 (2022).

[24] C. Magnant and D. Martin. A note on the path cover number of regular graphs. *Australas. J. Combin.* **43** (2009), 211–217.

[25] O. Ore. Arc coverings of graphs. *Annali di Matematica Pura ed Applicata*, **55**, (1961), 315–321.

[26] F. Ruskey and C. D. Savage. Hamilton cycles that extend transposition matchings in Cayley graphs of S_n. *SIAM Journal on Discrete Mathematics* **6(1)** (1993) 152–166.

[27] Z. Yang. On F-Hamiltonian graphs. *Discrete Math.*, **196**, (1999), 281–286.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Almost partitioning every 2-edge-coloured complete k-graph into k monochromatic tight cycles

(Extended abstract)

Allan Lo* Vincent Pfenninger†

Abstract

A k-uniform tight cycle is a k-graph with a cyclic order of its vertices such that every k consecutive vertices from an edge. We show that for $k \geq 3$, every red-blue edge-coloured complete k-graph on n vertices contains k vertex-disjoint monochromatic tight cycles that together cover $n - o(n)$ vertices.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-100`

1 Introduction

Monochromatic partitioning is an area of combinatorics that has its origin in a remark of Gerencsér and Gyárfás [9] that any 2-edge-colouring of the complete graph K_n contains a spanning path that consists of a red[1] path followed by a blue path. In particular, every 2-edge-coloured K_n admits a partition of its vertex set into a red path and a blue path. In a subsequent paper, Gyárfás [10] proved the stronger result that every 2-edge-coloured K_n contains a red cycle and a blue cycle that share at most one vertex and together cover

*School of Mathematics, University of Birmingham, UK. E-mail: `s.a.lo@bham.ac.uk`. The research leading to these results was supported by EPSRC, grant no. EP/V002279/1 and EP/V048287/1 (A. Lo). There are no additional data beyond that contained within the main manuscript.

†School of Mathematics, University of Birmingham, UK, E-mail: `v.pfenninger@bham.ac.uk` This project has received partial funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 786198, V. Pfenninger).

[1]We assume that the colours in a 2-edge-colouring are always red and blue.

all vertices. Lehel conjectured that every 2-edge-coloured K_n can be partitioned into a red cycle and a blue cycle. Here the empty set, a single vertex, and a single edge are considered to be cycles. Lehel's conjecture was first proved for large enough n by Łuczak, Rödl and Szemerédi [15] using the regularity lemma. Allen [1] later gave a different proof of this that does not use the regularity lemma, thus improving the bound on n. Bessy and Thomassé [3] finally gave a short and elegant proof of Lehel's conjecture for all n.

Here our interest lies in generalisations of Lehel's conjecture to hypergraphs and tight cycles. For related problems in hypergraphs about other types of cycles see [12, 16, 6]. See [11] and [17, Sections 7.4 and 9.7] for surveys on monochromatic partitioning.

For $k \geq 2$, a *k-graph* (or *k-uniform hypergraph*) H is a pair of sets $(V(H), E(H))$ such that $E(H) \subseteq \binom{V(H)}{k}$ (where for a set X, $\binom{X}{k}$ denotes the set of all subsets of X of size k). A *k-uniform tight cycle* is a k-graph with a cyclic order of its vertices such that its edges are exactly the sets of k consecutive vertices in the order. From now on, any set of at most k vertices is also considered a tight cycle.

The problem we want to consider is how to cover almost all vertices of every 2-edge-coloured complete k-graph with as few vertex-disjoint monochromatic tight cycles as possible. For $k = 3$, Bustamante, Hàn and Stein [5] proved that in every 2-edge-coloured complete 3-graph almost all vertices can be partitioned into a red and a blue tight cycle. Subsequently, Garbe, Mycroft, Lang, Lo, and Sanhueza-Matamala [8] showed that in fact there is a partition of all the vertices into two monochromatic tight cycles. Here it is necessary to allow the two monochromatic tight cycles to possibly have the same colour. Indeed for every $k \geq 3$, there exists a complete graph on arbitrarily many vertices that does not admit a partition of its vertices into a red and a blue tight cycle [13, Proposition 1]. In an earlier paper [13], the authors proved that for $k = 4$ it is also possible to almost partition the vertices of every 2-edge-coloured 4-graph into a red and a blue tight cycle. In the same paper, the authors also showed a weaker result for $k = 5$, that 4 vertex-disjoint monochromatic tight cycles suffice to cover almost all vertices of every 2-edge-coloured complete 5-graph. The only bound for general k that was known is given by a result of Bustamante, Corsten, Frankl, Pokrovskiy and Skokan [4]. They showed that every r-edge-coloured complete k-graph on n vertices can be partitioned into $c(r, k)$ monochromatic tight cycles. However, the constant $c(r, k)$ that can be obtained from their proof is very large. Indeed to cover almost all vertices they simply repeatedly find a monochromatic tight cycle using the fact that the Ramsey number for the k-uniform tight cycle on N vertices is linear in N.

Our aim here is to show a reasonable general bound on the number of tight cycles that are needed to almost partition a 2-edge-coloured complete k-graph. Indeed we show that k tight cycles suffice. We remark that this result is probably not tight. The only known lower bound on the number of tight cycles needed is the trivial lower bound of 2.

Theorem 1.1. *For every $\varepsilon > 0$ and $k \geq 3$, there exists an integer n_0 such that every 2-edge-coloured complete k-graph on $n \geq n_0$ vertices contains k vertex-disjoint monochromatic tight cycles covering at least $(1 - \varepsilon)n$ vertices.*

2 Sketch proof of Theorem 1.1

Our proof is based on a hypergraph version of Łuczak's connected matching method (see [14] for the original method). Roughly speaking the idea is as follows. Let K be a 2-edge-coloured complete k-graph on n vertices. We apply the Regular Slice Lemma (a form of the Hypergraph Regularity Lemma that is due to Allen, Böttcher, Cooley and Mycroft [2]) to K^{red}.[2] Since any regularity partition for a k-graph is also a regularity partition for its complement, this gives rise to a reduced k-graph $\mathcal{R}$ that is a 2-edge-coloured almost complete k-graph. So a red edge $i_1 \dots i_k$ in $\mathcal{R}$ means that K^{red} is regular with respect to the corresponding clusters $V_{i_1}, \dots, V_{i_k}$ and at least half the edges of K with one vertex in each V_{i_j}, $j \in [k]$ are red.[3] It can be shown that there is a red tight cycle in K that contains almost all of the vertices in $\bigcup_{j\in[k]} V_{i_j}$. The idea is now to combine a matching M of red edges in $\mathcal{R}$ into an even longer red tight cycle that covers almost all the vertices in the clusters that are covered by M. However, in order for this to work we will need to be able to construct a red tight path[4] that goes from the clusters corresponding to one edge of M to the clusters corresponding to another edge of M. So we require our red matching M in $\mathcal{R}$ to be 'connected' in some sense. To this end, we need the following definitions. A *tight pseudo-walk (of length m from e to e')* in a k-graph H is a sequence of edges $e_1 \dots e_m$ in H such that $|e_i \cap e_{i+1}| = k-1$ (and $(e_1, e_m) = (e, e')$). A k-graph H is *tightly connected* if for every pair of edges $e, e' \in H$, there is a tight pseudo-walk from e to e' in H. A *tight component* in a k-graph H is a maximal tightly connected subgraph of H. Let H be a 2-edge-coloured k-graph. A *red* or *blue tight component* in H is a tight component in H^{red} or in H^{blue}, respectively. A *monochromatic tight component* in H is a red or a blue tight component in H. The hypergraph version of Łuczak's connected matching method that we need now roughly says the following. If $\mathcal{R}$ contains a matching that covers almost all vertices and uses edges from at most k monochromatic tight components, then there exists k monochromatic tight cycles in K that are vertex-disjoint and together cover almost all vertices. The proof of Theorem 1.1 is now reduced to proving that every almost complete 2-edge-coloured k-graph contains a matching that covers almost all vertices and only uses edges from at most k monochromatic tight components.

For a set $S \subseteq V(H)$ with $|S| \le k-1$, we let $N_H(S) = \{S' \in \binom{V(H)}{k-|S|} : S \cup S' \in E(H)\}$ and $d_H(S) = |N_H(S)|$. A k-graph H on n vertices is called (μ, α)-*dense* if, for each $i \in [k-1]$, we have $d_H(S) \ge \mu\binom{n}{k-i}$ for all but at most $\alpha\binom{n}{i}$ sets $S \in \binom{V(H)}{i}$ and $d_H(S) = 0$ for all other sets $S \in \binom{V(H)}{i}$. Note that the reduced graph $\mathcal{R}$ will typically be $(1-\varepsilon, \varepsilon)$-dense.

Our discussion on a hypergraph version of Łuczak's connected matching method is encapsulated in the following corollary from our previous work. When we say that a statement holds for constants a and b with $0 < a \ll b$, we mean that the statement holds provided that a is chosen sufficiently small in terms of b. Moreover, if $1/n$ appears in such

[2]For a 2-edge-coloured k-graph H, we denote by H^{red} and H^{blue} the subgraph induced by the red edges and the subgraph induced by the blue edges of H, respectively.

[3]For $n \in \mathbb{N}$, $[n] = \{1, \dots, n\}$.

[4]A *tight path* is a k-graph with a linear order of its vertices such that every k-consecutive vertices form an edge. Or alternatively a tight path is a k-graph obtained by deleting a single vertex from a tight cycle.

a hierarchy then we implicitly assume that $n \in \mathbb{N}$.

Corollary 2.1 ([13, Corollary 20]). *Let $1/n \ll 1/m \ll \varepsilon \ll \eta \ll \gamma, 1/k, 1/s$ with $k \geq 3$. Suppose that every 2-edge-coloured $(1-\varepsilon, \varepsilon)$-dense k-graph H on m vertices contains a matching in H that covers all but at most ηm vertices of H and only contains edges from at most s monochromatic tight components of H. Then any 2-edge-coloured complete k-graph on n vertices contains s vertex-disjoint monochromatic tight cycles covering at least $(1-\gamma)n$ vertices.*

3 A large matching using edges from few monochromatic tight components

By Corollary 2.1, to prove Theorem 1.1, it suffices to prove the following lemma.

Lemma 3.1. *Let $1/n \ll \varepsilon \ll \eta \ll 1/k \leq 1/2$. Let H be a 2-edge-coloured $(1-\varepsilon, \varepsilon)$-dense k-graph on n vertices. Then there exists a matching in H that covers all but at most ηn vertices of H and only contains edges from at most k monochromatic tight components of H.*

The cases when $k = 3$ is already proved in [5] (in which they showed that a red and a blue tight component suffice). The first step of the proof is to find a red and a blue tight component R and B, respectively, of H such that almost all 2-subsets of $V(H)$ are contained in some edge of $R \cup B$. One then finds a large matching in $R \cup B$. Thus a natural first step of proving Lemma 3.1 is to find a constant number of monochromatic tight components of H, such that almost all $(k-1)$-subsets of $V(H)$ are contained in some edge of these tight components. However this is not possible when $k \geq 4$ as shown by the following example. Let $V_1, \dots, V_\ell$ be an equipartition of a set of n vertices. Consider the 2-edge-coloured complete k-graph on $\bigcup_{i \in [\ell]} V_i$ such that an edge e is red if $|e \cap V_i| > k/2$, and blue otherwise. Observe that there are ℓ blue tight components. Moreover, each $\binom{V_i}{k-1}$ is "covered" by a distinct blue tight component.

Instead, we will enlarge our maximal matching as we choose tight components as follows. Consider a monochromatic tight component F_* in H and let $\mathcal{G}_1 = \{F_*\}$. We say that two monochromatic tight components F_1 and F_2 of H are *adjacent* if they have opposite colours and there are edges $e_1 \in F_1$ and $e_2 \in F_2$ such that $|e_1 \cap e_2| = k-1$. Now for each $i \geq 2$ in turn, let $\mathcal{G}_i$ be the set of monochromatic tight components that are adjacent to a monochromatic tight component in $\mathcal{G}_{i-1}$ and not already in $\bigcup_{j \in [i-1]} \mathcal{G}_j$. Moreover, for each $i \geq 1$, we let $\mathcal{E}(\mathcal{G}_i) = \bigcup_{F \in \mathcal{G}_i} F$, that is, $\mathcal{E}(\mathcal{G}_i)$ is the set of edges that are in some monochromatic tight component $F \in \mathcal{G}_i$. It is easy to see that all edges of H are in $\bigcup_{j \in [2k]} \mathcal{E}(\mathcal{G}_j)$. In fact, if our initial monochromatic tight component F_* spans almost all vertices (such an F^* exists), then almost all edges of H are in $\bigcup_{j \in [k]} \mathcal{E}(\mathcal{G}_j)$. For simplicity, we assume that $H = \bigcup_{j \in [k]} \mathcal{E}(\mathcal{G}_j)$.

We now set $W_0 = V(H)$ and for each $i = 1, \dots, k$ in turn, we let M_i be a maximal matching in $H[W_{i-1}] \cap \bigcup_{j\in[i]} \mathcal{E}(\mathcal{G}_j)$ and $W_i = W_{i-1} \setminus V(M_i)$. Since $\bigcup_{j\in[k]} \mathcal{E}(\mathcal{G}_j) = H$ is $(1-\varepsilon, \varepsilon)$-dense, $\bigcup_{j\in[k]} M_j$ is a maximal matching in H covering almost all vertices of H.

It remains to show that $\bigcup_{j\in[k]} M_j$ is contained in at most k monochromatic tight components. Note that, for each $i \in [k+1]$,

$$H[W_{i-1}] \cap \bigcup_{j\in[i-1]} \mathcal{E}(\mathcal{G}_j) = \varnothing \tag{3.1}$$

by our choices of M_{i-1} and W_{i-1}. Hence $M_i \subseteq H[W_{i-1}] \cap \mathcal{E}(\mathcal{G}_i)$. Therefore, it suffices to show that $H[W_{i-1}] \cap \mathcal{E}(\mathcal{G}_i)$ (and so M_i) consists of edges from one monochromatic tight component.

Suppose for a contradiction that there are two edges e_1 and e_2 in $H[W_{i-1}] \cap \mathcal{E}(\mathcal{G}_i)$ that are in different monochromatic tight components. Suppose further that $|W_{i-1}| \geq \eta n$ (or else $\bigcup_{j\in[i-1]} M_j$ is already an almost perfect matching). In $H[W_{i-1}]$, there exists a tight pseudo-walk P from e_1 to e_2. Recall that $\bigcup_{j\in[i]} \mathcal{E}(\mathcal{G}_j)$ is tightly connected, so $\bigcup_{j\in[i]} \mathcal{E}(\mathcal{G}_j)$ contains a tight pseudo-walk P' from e_2 to e_1. Thus PP' (the concatenation of P and P') is a closed tight pseudo-walk. We then define a nearly triangulated plane graph[5] D such that every vertex of D corresponds to an edge in H, PP' is on the outer face and any walk in D corresponds to a tight pseudo-walk in H. We colour each vertex of D with the same colour of its corresponding edge in H. All edges in $\mathcal{E}(\mathcal{G}_i)$ (including e_1 and e_2) have the same colour, say red. Since e_1 and e_2 are not in the same red tight component, there is no red walk in D from e_1 to e_2. By adapting the proof of Gale [7] of the fact that the Hex game cannot end in a draw, one deduces that H contains a blue tight pseudo-walk P^* from an edge of P to an edge of P'. Since P' is contained in $\bigcup_{j\in[i]} \mathcal{E}(\mathcal{G}_j)$, we have $P^* \subseteq \bigcup_{j\in[i-1]} \mathcal{E}(\mathcal{G}_j)$. Therefore, $\emptyset \neq P \cap P^* \subseteq H[W_{i-1}] \cap \bigcup_{j\in[i-1]} \mathcal{E}(\mathcal{G}_j)$ contradicting (3.1).

References

[1] P. Allen. Covering two-edge-coloured complete graphs with two disjoint monochromatic cycles. *Combin. Probab. Comput.*, 17(4):471–486, 2008.

[2] P. Allen, J. Böttcher, O. Cooley, and R. Mycroft. Tight cycles and regular slices in dense hypergraphs. *J. Combin. Theory Ser. A*, 149:30–100, 2017.

[3] S. Bessy and S. Thomassé. Partitioning a graph into a cycle and an anticycle, a proof of lehel's conjecture. *J. Combin. Theory Ser. B*, 100(2):176–180, 2010.

[4] S. Bustamante, J. Corsten, N. Frankl, A. Pokrovskiy, and J. Skokan. Partitioning edge-colored hypergraphs into few monochromatic tight cycles. *SIAM J. Discrete Math.*, 34(2):1460–1471, 2020.

[5] A nearly triangulated plane graph is a plane graph such that all the faces except the outer face are triangles.

[5] S. Bustamante, H. Hàn, and M. Stein. Almost partitioning 2-colored complete 3-uniform hypergraphs into two monochromatic tight or loose cycles. *J. Graph Theory*, 91(1):5–15, 2019.

[6] S. Bustamante and M. Stein. Partitioning 2-coloured complete k-uniform hypergraphs into monochromatic ℓ-cycles. *European J. Combin.*, 71:213–221, 2018.

[7] D. Gale. The game of Hex and the Brouwer fixed-point theorem. *Amer. Math. Monthly*, 86(10):818–827, 1979.

[8] F. Garbe, R. Mycroft, R. Lang, A. Lo, and N. Sanhueza-Matamala. Partitioning 2-coloured complete 3-uniform hypergraphs into two monochromatic tight cycles. *In preparation.*

[9] L. Gerencsér and A. Gyárfás. On Ramsey-type problems. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.*, 10:167–170, 1967.

[10] A. Gyárfás. Vertex coverings by monochromatic paths and cycles. *J. Graph Theory*, 7(1):131–135, 1983.

[11] A. Gyárfás. Vertex covers by monochromatic pieces — a survey of results and problems. *Discrete Math.*, 339(7):1970–1977, 2016.

[12] A. Gyárfás and G. N. Sárközy. Monochromatic path and cycle partitions in hypergraphs. *Electron. J. Combin.*, 20(1):18, 2013.

[13] A. Lo and V. Pfenninger. Towards Lehel's conjecture for 4-uniform tight cycles. *Electron. J. Combin.*, 30(1):P1.13, 2023.

[14] T. Łuczak. $R(C_n, C_n, C_n) \leq (4 + o(1))n$. *J. Combin. Theory Ser. B*, 75(2):174–187, 1999.

[15] T. Łuczak, V. Rödl, and E. Szemerédi. Partitioning two-coloured complete graphs into two monochromatic cycles. *Combin. Probab. Comput.*, 7(4):423–436, 1998.

[16] G. N. Sárközy. Improved monochromatic loose cycle partitions in hypergraphs. *Discrete Math.*, 334:52–62, 2014.

[17] M. Simonovits and E. Szemerédi. Embedding graphs into larger graphs: results, methods, and problems. In *Building bridges II—mathematics of László Lovász*, volume 28 of *Bolyai Soc. Math. Stud.*, pages 445–592. Springer, Berlin, [2019] ©2019.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Kneser graphs are Hamiltonian*

(Extended abstract)

Arturo Merino[†] Torsten Mütze[‡] Namrata[§]

Abstract

For integers $k \geq 1$ and $n \geq 2k+1$, the Kneser graph $K(n,k)$ has as vertices all k-element subsets of an n-element ground set, and an edge between any two disjoint sets. It has been conjectured since the 1970s that all Kneser graphs admit a Hamilton cycle, with one notable exception, namely the Petersen graph $K(5,2)$. This problem received considerable attention in the literature, including a recent solution for the sparsest case $n = 2k+1$. The main contribution of this paper is to prove the conjecture in full generality. We also extend this Hamiltonicity result to all connected generalized Johnson graphs (except the Petersen graph). The generalized Johnson graph $J(n,k,s)$ has as vertices all k-element subsets of an n-element ground set, and an edge between any two sets whose intersection has size exactly s. Clearly, we have $K(n,k) = J(n,k,0)$, i.e., generalized Johnson graph include Kneser graphs as a special case. Our results imply that all known families of vertex-transitive graphs defined by intersecting set systems have a Hamilton cycle, which settles an interesting special case of Lovász' conjecture on Hamilton cycles in vertex-transitive graphs from 1970. Our main technical innovation is to study cycles in Kneser graphs by a kinetic system of multiple gliders that move at different speeds and that interact over time, reminiscent of the gliders in Conway's Game of Life, and to analyze this system combinatorially and via linear algebra.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-101`

*This work was supported by Czech Science Foundation grant GA 22-15272S. An extended abstract has been accepted for presentation at the 55th Annual ACM Symposium on the Theory of Computing (STOC 2023).

[†]Department of Mathematics, TU Berlin, Germany. E-mail: `merino@math.tu-berlin.de`

[‡]Department of Computer Science, University of Warwick, United Kingdom & Department of Theoretical Computer Science and Mathematical Logic, Charles University, Prague, Czech Republic. E-mail: `torsten.mutze@warwick.ac.uk`

[§]Department of Computer Science, University of Warwick, United Kingdom. E-mail: `namrata@warwick.ac.uk`

1 Introduction

For integers $k \geq 1$ and $n \geq 2k+1$, the *Kneser graph* $K(n,k)$ has as vertices all k-element subsets of $[n] := \{1,2,\ldots,n\}$, and an edge between any two sets A and B that are disjoint, i.e., $A \cap B = \emptyset$. Kneser graphs were introduced by Lovász [Lov78] in his celebrated proof of Kneser's conjecture. Using the Borsuk-Ulam theorem, he proved that the chromatic number of $K(n,k)$ equals $n-2k+2$. Observe also that the maximum independent set in $K(n,k)$ has size $\binom{n-1}{k-1}$ by the famous Erdős-Ko-Rado [EKR61] theorem. Furthermore, the graph $K(n,k)$ is vertex-transitive, i.e., it 'looks the same' from the point of view of any vertex, and all vertices have degree $\binom{n-k}{k}$. Lastly, note that when $n < ck$, the Kneser graph $K(n,k)$ does not contain cliques of size c, whereas it does contain such cliques when $n \geq ck$.

2 Hamilton cycles in Kneser graphs

In this work we investigate Hamilton cycles in Kneser graphs, i.e., cycles that visit every vertex exactly once. Kneser graphs have long been conjectured to have a Hamilton cycle, with one notable exception, the Petersen graph $K(5,2)$, which only admits a Hamilton path. As Kneser graphs are vertex-transitive, this is a special case of Lovász' famous conjecture [Lov70], which asserts that every connected vertex-transitive graph admits a Hamilton path. So far, the conjecture for Hamilton cycles in Kneser graphs has been tackled from two angles, namely for sufficiently dense Kneser graphs, and for the sparsest Kneser graphs. From the aforementioned results about the degree and cliques in $K(n,k)$, we see that $K(n,k)$ is relatively dense when n is large w.r.t. k, and relatively sparse otherwise. The sparsest case is when $n = 2k+1$, and the graphs $O_k := K(2k+1,k)$ are also known as *odd graphs*. Intuitively, proving Hamiltonicity should be easier for the dense cases, and harder for the sparse cases.

We first recap the known results for dense Kneser graphs. Heinrich and Wallis [HW78] showed that $K(n,k)$ has a Hamilton cycle if $n \geq 2k + k/(\sqrt[k]{2}-1) = (1+o(1))k^2/\ln 2$. This was improved by B. Chen and Lih [CL87], whose results imply that $K(n,k)$ has a Hamilton cycle if $n \geq (1+o(1))k^2/\log k$; see [CI96]. In another breakthrough, Y. Chen [Che00] showed that $K(n,k)$ is Hamiltonian when $n \geq 3k$. A particularly nice and clean proof for the cases where $n = ck$, $c \in \{3,4,\ldots\}$, was obtained by Y. Chen and Füredi [CF02], later extended by Bellmann and Schülke to any $n \geq 4k$ [BS21]. The asymptotically best result known to date, again due to Y. Chen [Che03], is that $K(n,k)$ has a Hamilton cycle if $n \geq (3k+1+\sqrt{5k^2-2k+1})/2 = (1+o(1))2.618\ldots \cdot k$. With the help of computers, Shields and Savage [SS04] found Hamilton cycles in $K(n,k)$ for all $n \leq 27$ (except for the Petersen graph).

We now briefly summarize the Hamiltonicity story of the sparsest Kneser graphs, namely the odd graphs. Note that $O_k = K(2k+1,k)$ has degree $k+1$, which is only logarithmic in the number of vertices. The conjecture that O_k has a Hamilton cycle for all $k \geq 3$ originated in the 1970s, in papers by Meredith and Lloyd [ML72, ML73] and by Biggs [Big79]. Already Balaban [Bal72] exhibited a Hamilton cycle for the cases $k = 3$

and $k = 4$, and Meredith and Lloyd described one for $k = 5$ and $k = 6$. Later, Mather [Mat76] solved the case $k = 7$. Mütze, Nummenpalo and Walczak [MNW21] finally settled the problem for all odd graphs, proving that O_k has a Hamilton cycle for every $k \geq 3$. Already much earlier, Johnson [Joh11] provided an inductive argument that establishes Hamiltonicity of $K(n, k)$ provided that the existence of Hamilton cycles is known for several smaller Kneser graphs. Combining his result with the unconditional results from [MNW21] yields that $K(2k+2^a, k)$ has a Hamilton cycle for all $k \geq 3$ and $a \geq 0$. These results still leave infinitely many open cases, the sparsest one of which is the family $K(2k + 3, k)$ for $k \geq 1$.

The main contribution of this paper is to settle the conjecture on Hamilton cycles in Kneser graphs affirmatively in full generality.

Theorem 1. *For all $k \geq 1$ and $n \geq 2k + 1$, the Kneser graph $K(n, k)$ has a Hamilton cycle, unless it is the Petersen graph, i.e., $(n, k) = (5, 2)$.*

More generally, our work settles all known instances of Lovász' conjecture for vertex-transitive graphs defined by intersecting set systems. As we shall see, Kneser graphs are the hardest cases among them to prove, i.e., with the help of Theorem 1 the Hamiltonicity of the more general families of graphs can be settled easily.

3 Generalized Johnson graphs

The *generalized Johnson graph* $J(n, k, s)$ has as vertices all k-element subsets of $[n]$, and an edge between any two sets A and B that satisfy $|A \cap B| = s$, i.e., the intersection of A and B has size exactly s. To ensure that the graph is connected, we assume that $s < k$ and $n \geq 2k - s + \mathbf{1}_{[s=0]}$, where $\mathbf{1}_{[s=0]}$ denotes the indicator function that equals 1 if $s = 0$ and 0 otherwise. Generalized Johnson graphs are sometimes called 'uniform subset graphs' in the literature, and they are also vertex-transitive. Furthermore, by taking complements, we see that $J(n, k, s)$ is isomorphic to $J(n, n - k, n - 2k + s)$. Clearly, Kneser graphs are special generalized Johnson graphs obtained for $s = 0$. On the other hand, the graphs obtained for $s = k - 1$ are known as (ordinary) *Johnson graphs* $J(n, k) := J(n, k, k - 1)$.

Chen and Lih [CL87] conjectured that all graphs $J(n, k, s)$ admit a Hamilton cycle except the Petersen graph $J(5, 2, 0) = J(5, 3, 1)$, and this problem was reiterated in Gould's survey [Gou91]. In their original paper, Chen and Lih settled the cases $s \in \{k - 1, k - 2, k - 3\}$. For the Johnson graphs $J(n, k) = J(n, k, k - 1)$, strong Hamiltonicity properties are known [TL73, JR94, Kno94].

We generalize Theorem 1 further, by showing that all connected generalized Johnson graphs admit a Hamilton cycle. This resolves Chen and Lih's conjecture affirmatively in full generality.

Theorem 2. *For all $k \geq 1$, $0 \leq s < k$, and $n \geq 2k - s + \mathbf{1}_{[s=0]}$ the generalized Johnson graph $J(n, k, s)$ has a Hamilton cycle, unless it is the Petersen graph, i.e., $(n, k, s) \in \{(5, 2, 0), (5, 3, 1)\}$.*

4 Bipartite Kneser graphs and the middle levels problem

For integers $k \geq 1$ and $n \geq 2k+1$, the *bipartite Kneser graph* $H(n,k)$ has as vertices all k-element and $(n-k)$-element subsets of $[n]$, and an edge between any two sets A and B that satisfy $A \subseteq B$. It is easy to see that bipartite Kneser graphs are also vertex-transitive. As $H(n,k)$ is the bipartite double cover of $K(n,k)$, Hamiltonicity of $K(n,k)$ is harder than the Hamiltonicity of $H(n,k)$.

Lemma 3. *If $K(n,k)$ admits a Hamilton cycle, then $H(n,k)$ admits a Hamilton cycle or path.*

The sparsest bipartite Kneser graphs $M_k := H(2k+1,k)$ are known as *middle levels graphs*, as they are isomorphic to the subgraph of the $(2k+1)$-dimensional hypercube induced by the middle two levels. The well-known *middle levels conjecture* asserts that M_k has a Hamilton cycle for all $k \geq 1$. This conjecture was raised in the 1980s, settled affirmatively in [Müt16], and a short proof was given in [GMN18]. More generally, all bipartite Kneser graphs $H(n,k)$ were shown to have a Hamilton cycle in [MS17]. Via Lemma 3, our Theorem 1 thus also yields a new alternative proof for the Hamiltonicity of bipartite Kneser graphs. Consequently, our results in this paper settle Lovász' conjecture for all known families of vertex-transitive graphs that are defined by intersecting set systems.

5 Proof ideas

It turns out that Theorem 1 can be used to establish Theorem 2 by a simple inductive construction. Consequently, the main work in this paper is to prove Theorem 1. In this extended abstract, we only sketch the main ideas for this proof, for details see [MMN22].

As mentioned before, Mütze, Nummenpalo and Walczak [MNW21] proved that $K(n,k)$ has a Hamilton cycle for $n = 2k+1$ and all $k \geq 3$. Combining this result with Johnson's construction [Joh11] shows that $K(n,k)$ has a Hamilton cycle for $n = 2k+2^a$ and all $k \geq 3$ and $a \geq 0$, in particular for $n = 2k+2$. The techniques developed in this paper work whenever $n \geq 2k+3$, and thus they settle all remaining cases of Theorem 1. Our proof does not work in the cases $n = 2k+1$ and $n = 2k+2$, so the two earlier constructions do not become obsolete.

We follow a two-step approach to construct a Hamilton cycle in $K(n,k)$ for $n \geq 2k+3$. In the first step, we construct a *cycle factor* in the graph, i.e., a collection of disjoint cycles that together visit all vertices. In the second step, we join the cycles of the factor to a single cycle.

5.1 Cycle factor construction

The starting point is to consider the characteristic vectors of the vertices of $K(n,k)$. For every k-element subset of $[n]$, this is a bitstring of length n with exactly k many 1s at the positions corresponding to the elements of the set. For example, the vertex $\{1,7,9\}$ of $K(9,3)$ is represented by the bitstring 100000101. Clearly, two sets A and B that are vertices of $K(n,k)$ are disjoint if and only if the corresponding bitstrings have no 1s at the same positions.

Our construction of a cycle factor in the Kneser graph $K(n,k)$ uses the following simple rule based on parenthesis matching, a technique pioneered by Greene and Kleitman [GK76]: Given a vertex represented by a bitstring x, we interpret the 1s in x as opening brackets and the 0s as closing brackets, and we match closest pairs of opening and closing brackets in the natural way, which will leave some 0s unmatched. This matching is done *cyclically* across the boundary of x, i.e., x is considered as a cyclic string. We write $f(x)$ for the vertex obtained from x by complementing all matched bits, leaving the unmatched bits unchanged. For example, $x = 100000101$ is interpreted as $x = ()))))()(= ())$`---`$()($, where each `-` denotes an unmatched closing bracket, and then complementing matched bits (the first three and last three in this case) yields the vertex $f(x) = 011000010$. Repeatedly applying f to every vertex partitions the vertices of the Kneser graph into cycles, and we write $C(x) := (x, f(x), f^2(x), \ldots)$ for the cycle containing x. For example, for x from before we obtain $C(x) = (100000101, 011000010, 000110001, 100001100, 010000011, \ldots, 000011010)$. Figure 1 shows several more examples of cycles generated by this parenthesis matching rule.

5.2 Analysis via gliders

The next key step is to understand the structure of the cycles generated by f. We describe the evolution of a bitstring x under repeated applications of f by a kinetic system of multiple gliders that move at different speeds and that interact over time, reminiscent of the gliders in Conway's Game of Life. This physical interpretation and its analysis are one of the main innovations of this paper. Specifically, we view each application of f as one unit of time moving forward. Furthermore, we partition the matched bits of x into groups, and each of these groups is called a *glider*. A glider has a *speed* associated to it, which is given by the number of 1s in its group. For example, in the cycle shown in Figure 1 (a), there is a single matched 1 and the corresponding matched 0, and together these two bits form a glider of speed 1 that moves one step to the right in every time step. Applying f means going down to the next row in the picture, so the time axis points downwards. Similarly, in Figure 1 (b), there are two matched 1s and the corresponding two matched 0s, and together these four bits form a glider of speed 2 that moves two steps to the right in every time step. As we see from these examples, a single glider of speed v simply moves uniformly, following the basic physics law $s(t) = s(0) + v \cdot t$, where t is the time (i.e., the number of applications of f) and $s(t)$ is the position of the glider in the bitstring as a function of time (modulo n). The situation gets more interesting and complicated when gliders of different speeds interact with each other. For example, in Figure 1 (c), there is one glider of speed 2 and one glider of speed 1. As long as these groups of bits are separated, each glider moves uniformly as before. However, when the speed 2 glider catches up with the speed 1 glider, an overtaking occurs. During an overtaking, the faster glider receives a boost, whereas the slower glider is delayed. This can be captured by augmenting the corresponding equations of motion by introducing an additional term that involves a variable counting the number of overtakings, making the equations non-uniform. For more than two gliders, the equations of motion can be generalized accordingly, by introducing such overtaking counters between any pair of gliders. Nevertheless, as the reader may appreciate from Figure 1 (d), in general it is highly nontrivial to recognize from an arbitrary bitstring x which of its matched bits belong to which glider, and consequently which glider is currently overtaking which other glider. Note that in general the gliders will not be nicely separated, but will be involved in simultaneous interactions, so that the groups of bits forming the gliders will be interleaved in complicated ways.

From the aforementioned physics interpretation we obtain that the number of gliders and their speeds are invariant along each cycle. For example, in Figure 1 (d), every bitstring along this cycle has three gliders of speeds 1, 2 and 3. From the equations of motion we also derive another crucial property, namely that no glider stands still forever, but will move eventually. Note that the speed 1 glider in Figure 1 (d) stands still between time

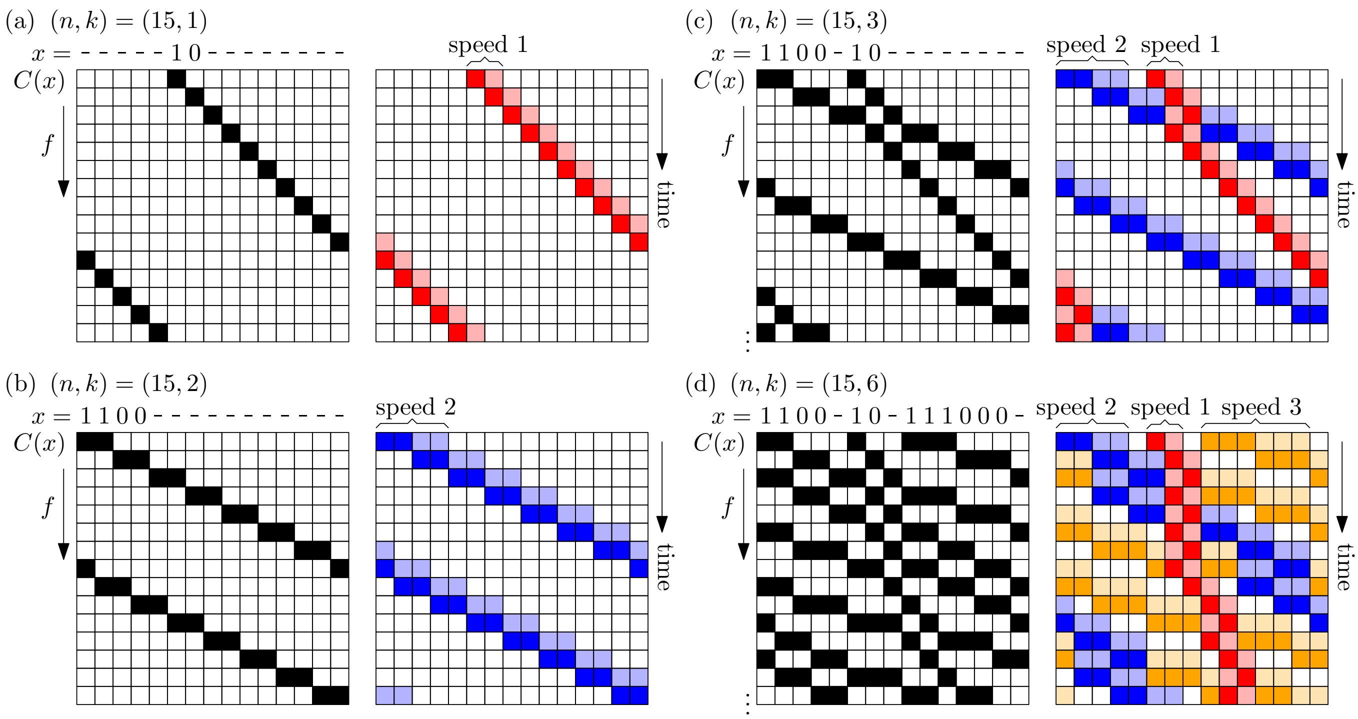


Figure 1: Cycles of our factor in different Kneser graphs $K(n,k)$. The cycles in (a) and (b) are shown completely, whereas in (c) and (d) only the first 15 vertices are shown. Vertices are represented by characteristic vectors, with 1s and 0s shown as black and white squares, resp. In each pair of figures, the right hand side shows the interpretation of certain groups of bits as gliders, and their movement over time. Matched bits belonging to the same glider are colored in the same color, 1-bits filled opaquely, and 0-bits filled transparently. (a) one glider of speed 1; (b) one glider of speed 2; (c) two gliders with speeds 1 and 2 that participate in an overtaking; (d) three gliders of speeds 1, 2 and 3 that participate in multiple overtakings. Animations of these examples are available at [Müt23].

steps 2–8, as during those steps it is overtaken once by the speed 2 glider, and twice by the speed 3 glider (wrapping around the boundary). We establish this fact by linear algebra, by showing that the determinant of the linear systems of equations that governs the gliders' movements is non-singular.

For the reader's entertainment, we programmed an interactive animation of gliders over time, and we encourage experimentation with this code, which can be found at [Müt23].

5.3 Gluing the cycles together

To join the cycles of our factor to a single Hamilton cycle, we consider a 4-cycle D that shares two opposite edges with two cycles C, C' from our factor. Clearly, the symmetric difference of the edge sets $(C \cup C')\Delta D$ yields a single cycle on the same vertex set as $C \cup C'$. We may repeatedly apply such gluing operations until all cycles are joined to a single Hamilton cycle. The two main technical obstacles here are: (a) All of the 4-cycles used for the gluing must be edge-disjoint, so that none of the gluings interfere with each other. (b) The gluings must achieve connectivity, i.e., every cycle must be connected to every other cycle via a sequence of gluings. To control the gluing, we consider the speeds of gliders in a bitstring x in non-increasing order. As the sum of speeds equals k, this sequence forms a number partition of k. To establish (b) we choose gluings that guarantee a lexicographic increase in those number partitions. Specifically, we glue cycles $C(x)$ and $C(y)$ for which the glider speeds in y are obtained from those in x by decreasing the speed of a glider of minimum speed by 1, and by increasing the speed of another glider by 1. This ensures that the number partition of k associated with y is lexicographically larger than that of x. Unfortunately, it is not always possible to use gluings that guarantee such immediate lexicographic improvement. In some cases we have to use gluings where a small lexicographic decrease occurs. We then argue that subsequent gluings compensate for this defect such that the overall effect is again a lexicographic improvement. For example, from a vertex with associated number partition $(4, 4)$, the first gluing may lead to a vertex with number partition $(4, 3, 1)$, and the next gluing may lead to $(5, 3)$. While $(4, 4) \to (4, 3, 1)$ is a lexicographic decrease instead of an increase, overall $(4, 4) \to (4, 3, 1) \to (5, 3)$ is a lexicographic increase.

References

[Bal72] A. T. Balaban. Chemical graphs. XIII. Combinatorial patterns. *Rev. Roumain Math. Pures Appl.*, 17:3–16, 1972.

[Big79] N. Biggs. Some odd graph theory. In *Second International Conference on Combinatorial Mathematics (New York, 1978)*, volume 319 of *Ann. New York Acad. Sci.*, pages 71–81. New York Acad. Sci., New York, 1979.

[BS21] J. Bellmann and B. Schülke. Short proof that Kneser graphs are Hamiltonian for $n \geq 4k$. *Discrete Math.*, 344(7):Paper No. 112430, 2 pp., 2021.

[CF02] Y. Chen and Z. Füredi. Hamiltonian Kneser graphs. *Combinatorica*, 22(1):147–149, 2002.

[Che00] Y. Chen. Kneser graphs are Hamiltonian for $n \geq 3k$. *J. Combin. Theory Ser. B*, 80(1):69–79, 2000.

[Che03] Y. Chen. Triangle-free Hamiltonian Kneser graphs. *J. Combin. Theory Ser. B*, 89(1):1–16, 2003.

[CI96] W. E. Clark and M. E. H. Ismail. Binomial and Q-binomial coefficient inequalities related to the Hamiltonicity of the Kneser graphs and their Q-analogues. *J. Combin. Theory Ser. A*, 76(1):83–98, 1996.

[CL87] B. Chen and K. Lih. Hamiltonian uniform subset graphs. *J. Combin. Theory Ser. B*, 42(3):257–263, 1987.

[EKR61] P. Erdős, C. Ko, and R. Rado. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser. (2)*, 12:313–320, 1961.

[GK76] C. Greene and D. J. Kleitman. Strong versions of Sperner's theorem. *J. Combin. Theory Ser. A*, 20(1):80–88, 1976.

[GMN18] P. Gregor, T. Mütze, and J. Nummenpalo. A short proof of the middle levels theorem. *Discrete Anal.*, pages Paper No. 8, 12 pp., 2018.

[Gou91] R. J. Gould. Updating the Hamiltonian problem—a survey. *J. Graph Theory*, 15(2):121–157, 1991.

[HW78] K. Heinrich and W. D. Wallis. Hamiltonian cycles in certain graphs. *J. Austral. Math. Soc. Ser. A*, 26(1):89–98, 1978.

[Joh11] J. R. Johnson. An inductive construction for Hamilton cycles in Kneser graphs. *Electron. J. Combin.*, 18(1):Paper 189, 12 pp., 2011.

[JR94] M. Jiang and F. Ruskey. Determining the Hamilton-connectedness of certain vertex-transitive graphs. *Discrete Math.*, 133(1-3):159–169, 1994.

[Kno94] M. Knor. Gray codes in graphs. *Math. Slovaca*, 44(4):395–412, 1994.

[Lov70] L. Lovász. Problem 11. In *Combinatorial Structures and Their Applications (Proc. Calgary Internat. Conf., Calgary, AB, 1969)*. Gordon and Breach, New York, 1970.

[Lov78] L. Lovász. Kneser's conjecture, chromatic number, and homotopy. *J. Combin. Theory Ser. A*, 25(3):319–324, 1978.

[Mat76] M. Mather. The Rugby footballers of Croam. *J. Combin. Theory Ser. B*, 20(1):62–63, 1976.

[ML72] G. H. J. Meredith and E. K. Lloyd. The Hamiltonian graphs O_4 to O_7. In *Combinatorics (Proc. Conf. Combinatorial Math., Math. Inst., Oxford, 1972)*, pages 229–236. Inst. Math. Appl., Southend-on-Sea, 1972.

[ML73] G. H. J. Meredith and E. K. Lloyd. The footballers of Croam. *J. Combin. Theory Ser. B*, 15:161–166, 1973.

[MMN22] A. Merino, T. Mütze, and Namrata. Kneser graphs are Hamiltonian. *arXiv:2212.03918.* Full preprint version of the present article, 2022.

[MNW21] T. Mütze, J. Nummenpalo, and B. Walczak. Sparse Kneser graphs are Hamiltonian. *J. Lond. Math. Soc. (2)*, 103(4):1253–1275, 2021.

[MS17] T. Mütze and P. Su. Bipartite Kneser graphs are Hamiltonian. *Combinatorica*, 37(6):1207–1219, 2017.

[Müt16] T. Mütze. Proof of the middle levels conjecture. *Proc. Lond. Math. Soc.*, 112(4):677–713, 2016.

[Müt23] T. Mütze. Gliders in Kneser graphs, 2023. `http://tmuetze.de/gliders.html`.

[SS04] I. Shields and C. D. Savage. A note on Hamilton cycles in Kneser graphs. *Bull. Inst. Combin. Appl.*, 40:13–22, 2004.

[TL73] D. Tang and C. Liu. Distance-2 cyclic chaining of constant-weight codes. *IEEE Trans. Comput.*, C-22:176–180, 1973.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Vizing's Interchange Conjecture holds for simple graphs

(Extended abstract)

Jonathan Narboni*

Abstract

In 1964 Vizing proved that starting from any k-edge-coloring of a graph G one can reach, using only Kempe swaps, a $(\Delta+1)$-edge-coloring of G where Δ is the maximum degree of G. One year later he conjectured that one can also reach a Δ-edge-coloring of G if there exists one. Bonamy *et. al* proved that the conjecture is true for the case of triangle-free graphs. In this paper we prove the conjecture for all simple graphs.

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-102

1 Introduction

In 1964 Vizing [6] proved that the chromatic index of a graph G (*i.e.*the minimum number of colors needed to properly colors the edges of G), denoted by $\chi'(G)$, is at most $\Delta(G)+1$ colors, where $\Delta(G)$ is the maximum degree of G.

Theorem 1. *Any simple graph G satisfy $\chi'(G) \leqslant \Delta(G)+1$.*

The proof heavily relies on the use of *Kempe changes*. Kempe changes were introduced by Kempe in his unsuccessful attempt to prove the 4-color theorem, but it turns out that this concept became one of the most fruitful tool in graph coloring. Throughout this paper, we will mostly focus on simple graphs (with no multiple edges), and thus, except stated explicitly, the graphs we consider are never multigraphs.

Moreover, we only consider proper edge-colorings, and so we will only write colorings to denote proper edge-colorings. Given a graph G and a coloring β, a Kempe chains C is

*The author is supported by National Science Center of Poland grant 2019/34/E/ST6/00443.

a maximal bichromatic component (Kempe chains were invented in the context of vertex-coloring, but the principle remains the same for edge-coloring). Applying a *Kempe swap* (or Kempe change) on C consists in switching the two colors in this component. Since C is maximal, the coloring obtained after the swap is guaranteed to be a proper coloring, and if C is not the unique maximal bichromatic component containing these two colors, the coloring obtained after the swap is a coloring different from β, as the partition of the edges is different.

The Kempe swaps induce an equivalence relation on the set of colorings of a graph G; two colorings β and β' are equivalent if one can find a sequence of Kempe swaps to transform β into β'. In 1964, Vizing actually proved a stronger statement, he proved that any k-coloring of a graph G (with $k > \Delta(G)$) is equivalent to a $(\Delta(G)+1)$-coloring of G.

Theorem 2. *Let G be a graph and β a k-coloring of G (with $k > \Delta(G)$). Then there exists a $(\Delta(G)+1)$-coloring β' equivalent to β.*

Note that some graphs only need Δ colors to be properly colored and thus the existence of an equivalent optimal coloring is not guaranteed by the theorem. One year later, Vizing [7] proved that this result is generalizable to multigraphs, and states the following conjecture (see [5] for more on the history of this conjecture):

Conjecture 3 (Vizing's Interchange Conjecture)**.** *For any multigraph G and for any k-coloring β of G, is there always an optimal coloring equivalent to β ?*

Note that both in Theorem 1 and in Conjecture 3 we do not have the choice in the target coloring. If we can choose a specific target optimal coloring, then the question can be reformulated as a reconfiguration question.

Question 4. For any multigraph G and for any k-coloring β, is any optimal coloring always equivalent to β ?

If true, Question 4 would imply the following conjecture, as it suffices to take the optimal target coloring as an intermediate between the two non-optimal colorings.

Conjecture 5. *Let G be a graph and let $k > \chi'(G)$. Then any two k-colorings are equivalent.*

Mohar proved the case where we have at least two more colors than the optimal [4].

Theorem 6 ([4])**.** *Let G be a graph. Then all $(\chi'(G)+2)$-colorings are equivalent.*

When considering case where we only have more color than the optimal ($k = \chi'(G)+1$), McDonald & al. proved Conjecture 5 for graphs with maximum degree 3 [3], Asratian and Casselgren proved it for graphs with maximum degree 4 [1], and Bonamy & al. proved that the conjecture is true for triangle-free graphs. In this paper, we prove that the conjecture is true for all graphs.

Theorem 7. *Let G be a graph. Then all its $(\chi'(G)+1)$-colorings are Kempe-equivalent.*

Theorem 7 is a direct consequence of the following Lemma which is the main result of this paper.

Lemma 8. *Let G be a graph. Then any $(\chi'(G)+1)$-coloring of G is equivalent to any $\chi'(G)$-coloring of G.*

Note that Lemma 8 together with Theorem 1 directly imply Conjecture 3 for (simple) graphs. Indeed, if G is a graph and β is a k-coloring of G with $k > \chi'(G)$, by Theorem 1, the coloring β is equivalent to a $(\Delta(G)+1)$-coloring of G, and thus it is equivalent to a $(\chi'(G)+1)$-coloring β' of G. Apply Lemma 8 to this coloring β' gives an optimal coloring equivalent to β, and thus proves Conjecture 3.

2 General setting of the proof

The proof inherits the technical setup of [2], in this section, we introduce this setting, and give the general outline of the proof of the main result. The majority of the proofs have been deferred to the appendices.

2.1 Reduction to $\chi'(G)$-regular graphs

The general setting of the proof follows that of [2] which itself follows that of [3] and of [1]. We first show that we can reduce the problem to the class of regular graphs.

Lemma 9. *Let G be a graph. Then there exists a graph $\chi'(G)$-regular graph G' such that:*

- *G is an induced subgraph of G,*
- *any $(\chi'(G)+1)$-coloring of G can be completed into a $(\chi'(G)+1)$-coloring of G', and*
- *if two $(\chi'(G)+1)$-colorings of G' are equivalent, then their restrictions to G are also equivalent.*

Note that colorings in regular graphs are easier to handle due to the following two properties:

- for any $(\Delta(G))$-coloring of a $\chi'(G)$-regular graph G, every vertex v is incident to exactly one edge of each color, and each color class is a perfect matching, and
- for any $(\Delta(G)+1)$-coloring α of a $\chi'(G)$-regular graph G, every vertex v is incident to all but one color, we call this color the *missing color* at v, and denote it by $m_\alpha(v)$ (we often drop the α when the coloring is clear from the context).

From now on, in the rest of the paper, we only consider χ'-regular graphs.

2.2 The good, the bad, and the ugly

The general approach to Lemma 8 is an induction on the chromatic index. Given a graph G, a $\Delta(G)$-coloring α and a $(\Delta(G)+1)$-coloring β, our goal is to find a sequence of Kempe swaps to transform β into α. To do so, we will "allign" α and β using the following lemma.

Lemma 10. *Let G be a regular graph, β be a $(\Delta(G)+1)$-coloring of G, α be a $\Delta(G)$-coloring of G, and let c be a color of α. Then the coloring β is equivalent to a $(\Delta(G)+1)$-coloring β' where for any edge e we have $\beta'(e) = c \Leftrightarrow \alpha(e) = c$.*

For any color class c in α, say $c = 1$, the edges colored 1 induce a perfect matching M in G. So once we obtain the coloring β' using Lemma 10 for the color 1, we can then proceed by induction of $G' = G \setminus M$. Remark that $\chi'(G') = \chi'(G) - 1$, and that the restrictions of α and β' to G' only use $\Delta(G) - 1$, and $\Delta(G)$ colors respectively since the color 1 is not used anymore.

Given a $(\Delta(G)+1)$-coloring β of G and a color, say the color 1. Let M be the perfect matching induced by the edges colored 1. We can partition the edges of G into three sets, an edge e is called:

- *good*, if $e \in M$ and $\beta(e) = 1$,
- *bad*, if $e \in M$ and $\beta(e) \neq 1$, and
- *ugly*, if $e \notin M$ and $\beta(e) = 1$.

A vertex missing the color 1 is called a *free vertex*. Toward contradiction, we assume that β is not equivalent to α, and we consider a $(\Delta(G)+1)$-coloring β' equivalent to β which minimizes the number of ugly edges among the colorings equivalent to β that minimize the number of bad edges, we call such a coloring *minimal*. Observe that in a minimal coloring there exists a bad edge. Thus, if we can find a coloring β'' equivalent to β' where the number of bad edges is strictly lower than in β', or with the same number of bad edges, and strictly fewer ugly edges, we get a contradiction. In a minimal coloring, we first have the following property.

Lemma 11. *In a minimal coloring, there exists a bad edge adjacent to an ugly edge and incident with a free vertex.*

2.3 Fan-like tools

In his proof of Theorem 1, Vizing introduced a technical tool to apply Kempe swaps on a coloring in very controlled way: *Vizing fans*. To define them, we will use an auxiliary digraph. Vizing did not use a digraph to define the fans, but this definition will prove to be suitable for our method of proof. Given a graph G, a $(\Delta(G)+1)$-coloring β of G and a vertex v, the directed graph D_v is defined as follows:

- the vertex set of D_v is the set of edges incident with v, and

- we put an arc between two vertices vv_1 and vv_2 of D_v, if the edge vv_2 is colored with the missing color at v_1.

The *fan around v starting at the edge e*, denoted by $X_v(e)$, is the maximal sequence of vertices of D_v reachable from the edge e. It is sometimes more convenient to speak about the color of the starting edge of a fan: If c is a color, then $X_v(c)$ denotes the fan around v starting at the edge colored c incident with v. Note that since the graph G is $\chi'(G)$-regular, each vertex misses exactly one color, and thus, in the digraph D_v, each vertex has outdegree at most 1. Hence a fan $\mathcal{X}$ is well-defined and we only have three possibilities for the fan $\mathcal{X}$:

- $\mathcal{X}$ is a path,
- $\mathcal{X}$ is a cycle, or
- $\mathcal{X}$ is a *comet* (*i.e.*, a path with an additional arc between the last vertex of the path and an internal vertex of the path).

If $\mathcal{X} = (vv_1, \cdots, vv_k)$ is a fan, v is called the central vertex of the fan, and vv_1 and vv_k are respectively called the first and the last edge of the fan (similarly, v_1 and v_k are the first and last vertex of $\mathcal{X}$ respectively). For any fan $\mathcal{V} = (vv_1, \cdots, vv_k)$ in a coloring β, $V(\mathcal{V})$ denotes the set of vertices $\{v_1, \cdots v_k\}$, and $E(\mathcal{V})$ denotes the set of edges $\{vv_1, \cdots vv_k\}$. We denote by $\beta(\mathcal{V})$ the set of colors involved in $\mathcal{V}$ (*i.e.* $\beta(\mathcal{V}) = \beta(E(\mathcal{V})) \cup m(V(\mathcal{V})) \cup m(v)$); if $\mathcal{V}$ involves the color c, $M(X, c)$ denotes the vertex of $V(\mathcal{V})$ missing the color c if this vertex is unique.

Given a $(\Delta(G)+1)$-coloring β of G, and fan $\mathcal{X} = (vv_1, \cdots, vv_k)$ which is a cycle around a vertex v, where each vertex v_i misses the color i (and so each edge vv_i is colored $(i-1)$), we can define the coloring $\beta' = X^{-1}(\beta)$ as follows:

- for any edge vv_i not in $\mathcal{X}$, $\beta'(vv_i) = \beta(vv_i)$, and
- for any edge vv_i in $\mathcal{X}$, $\beta'(vv_i) = i$ and $m(v_i) = i-1$

The coloring $\mathcal{X}^{-1}(\beta)$ is called the *invert* of $\mathcal{X}$, and we say that X is *invertible* if $\mathcal{X}$ and $\mathcal{X}^{-1}(\beta)$ are equivalent. If the cycle $\mathcal{X}$ is invertible, inverting $\mathcal{X}$ in β means applying a sequence of Kempe swaps to obtain $\mathcal{X}^{-1}(\beta)$ from the coloring β. In this paper, we prove that in any coloring, any cycle is invertible. This is the key ingredient of the proof of Lemma 10.

Lemma 12. *In any $(\chi'(G)+1)$-coloring of a $\chi'(G)$-regular graph G, any cycle is invertible.*

The proof of Lemma 12 is an induction on the size of the cycles. Towards contradiction, assume that there exist non-invertible cycles. A *minimum* cycle $\mathcal{V}$ is a non-invertible cycle of minimum size (*i.e.*, in any coloring, any smaller cycle is invertible).

A cycle of size 2 is clearly invertible as it only consists of a single Kempe chain composed of exactly two edges: to invert the cycle, it suffices to apply a Kempe swap on this component; so the size of a minimum cycle is at least 3. To prove the lemma, we need the two following results.

Lemma 13. *Let $\mathcal{V}$ be a minimum cycle around a vertex v. For any color c different from $m(v)$, the fan $X_v(c)$ is a cycle.*

Lemma 14. *Let $\mathcal{V}$ be a minimum cycle around a vertex v, and $\mathcal{X}$ and $\mathcal{Y}$ be two cycles around v. For any pair of vertices (z, z') in $(\mathcal{V} \cup \mathcal{X} \cup \mathcal{Y})^2$, the fan $\mathcal{Z} = X_z(m(z'))$ is a cycle containing z'.*

Proof of Lemma 12. To prove the Lemma, we prove that the graph G only consists of an even clique where each vertex misses a different color. This is a contradiction since in any $(\Delta(G) + 1)$-coloring of an even clique, for any color c, the number of vertices missing the color c is always even. By Lemma 13, all the fans around v are cycles, so each neighbor of v misses a different color. Moreover, by Lemma 14, there is an edge between each pair of neighbors of v, so $G = N[v] = K_{\Delta(G)+1}$. By construction, G is $\Delta(G)$-colorable, so G is an even clique and each vertex misses a different color as desired. □

The key ingredient of the proof of Lemma 13 is the notion of *entangleness* between two fans. Let $\mathcal{X}$ and $\mathcal{X}'$ be two fans, the fans $\mathcal{X}$ and $\mathcal{X}'$ are called *entangled* if for any color $c \in \beta(\mathcal{X}) \cap \beta(\mathcal{X}')$, we have $M(\mathcal{X}, c) = M(\mathcal{X}', c)$. Thus if two fans that are entangled and share a color, then their central vertices have a common neighbor. Note that if G is a triangle free graph, and uv is an edge of G, then a fan around v cannot be entangled with a fan around u if these two fans share a color. To prove Lemma 13 we need the two following lemmas.

Lemma 15. *Let $\mathcal{V}$ be a minimum cycle in a coloring β and let u and u' be two vertices of $\mathcal{V}$. Then fan $\mathcal{U} = X_u(m(u')) = (uu_1, \cdots, uu_l)$ is a cycle entangled with $\mathcal{V}$.*

Lemma 16. *Let $\mathcal{V}$ be a minimum cycle in a coloring β, u and u' be two vertices of $\mathcal{V}$, and $\mathcal{U} = X_u(m(u')) = (uu_1, \cdots, uu_l)$. Then for any $j \leqslant l$, the fan $X_v(\beta(uu_j))$ is a cycle.*

Note that by Lemma 15, we can directly conclude that $N[v]$ is a clique. The proof of Lemma 16 is pretty involved and technical and consists of finding a sequence of Kempe swaps to invert a minimum cycle $\mathcal{V}$ and thus reaching a contradiction. To do so we will use two meta-operations based on Kempe swaps, namely the inversion of fans that are paths, and the inversion of fans that are cycles smaller that the cycle $\mathcal{V}$. The key ingredient of the proof is to consider a whole equivalence class of colorings where the cycle $\mathcal{V}$ is minimum.

Let $X \subseteq E(G) \cup V(G)$, β a coloring and β' a coloring obtained from β by swapping a component C. The component is called X-stable if :

- for any $v \in X$, $m^{\beta}(v) = m^{\beta'}(v)$, and
- for any $e \in X$, $\beta(e) = \beta'(e)$.

In this case, the coloring β' is called X-*identical* to β.

If $S = (C_1, \cdots, C_k)$ is a sequence of swaps to transform a coloring β into a coloring β' where each C_j is a Kempe swap. The sequence S^{-1} is defined as the sequence of swaps $(C_k, \cdots, C_1)$. Such a sequence is called X-stable if each C_j is X-stable. If a sequence

S is X-stable, then the coloring obtained after apply S to β is called X-*equivalent* to β. Note that the notion of X-equivalence is stronger than the notion of X-identity. Since two colorings β and β' may be X-identical but not X-equivalent if there exists a coloring β'' in the sequence between β and β' that is not X-identical to β.

The following observation gives a relation between X-equivalence and $(G \setminus X)$-identity between colorings.

Observation 17. *Let β be a coloring, $X \subseteq V(G) \cup E(G)$, β_1 a coloring X-equivalent to β, and β_2 a coloring $(G \setminus X)$-identical to β_1. Then, there exists a coloring β_3 equivalent to β_2 that is X-identical to β_2 and $(G \setminus X)$-identical to β.*

If $\mathcal{X}$ is a fan, when two colorings are $(V(\mathcal{X}) \cup E(\mathcal{X}))$-identical (respectively $(V(\mathcal{X}) \cup E(\mathcal{X}))$-equivalent), we simply write that the two colorings are $\mathcal{X}$-identical (respectively $\mathcal{X}$-equivalent). Similarly, if two colorings are $((V(G) \cup E(G)) \setminus X)$-identical (respectively $((V(G) \cup E(G)) \setminus X)$-equivalent), we simply write that the two colorings are $(G \setminus X)$-identical (respectively $(G \setminus X)$-equivalent).

Remark that if $\mathcal{V}$ is a cycle in a coloring β, then the coloring $\mathcal{V}^{-1}(\beta)$ is $(G \setminus \mathcal{V})$-identical to β. So from the previous observation we have the following corollary.

Corollary 18. *Let $\mathcal{V}$ be a cycle in a coloring β. If there exists a coloring β' $\mathcal{V}$-equivalent to β where $\mathcal{V}$ is invertible, then $\mathcal{V}$ is invertible in β.*

From the previous corollary, we derive the following observation.

Observation 19. *Let $\mathcal{V}$ be a minimum cycle in coloring β, and β' a coloring $\mathcal{V}$-equivalent to β. Then in the coloring β', the sequence $\mathcal{V}$ is also a minimum cycle such that for any $e \in E(\mathcal{V})$, $\beta(e) = \beta'(e)$, and for any $v \in V(\mathcal{V})$, $m^{\beta}(v) = m^{\beta'}(v)$.*

We simply say that the cycle $\mathcal{V}$ is the same minimum cycle in the coloring β'. And thus it suffices to find a coloring $\mathcal{V}$-equivalent to β where $\mathcal{V}$ is invertible to reach a contradiction.

Acknowledgements. We thank Marthe Bonamy, František Kardoš and Éric Sopena for helpful discussions, and gratefully thank Penny Haxell for pointing out a mistake in an earlier version.

References

[1] Armen S. Asratian and Carl Johan Casselgren. Solution of Vizing's problem on interchanges for the case of graphs with maximum degree 4 and related results. *Journal of Graph Theory*, 82(4):350–373, 2016.

[2] Marthe Bonamy, Oscar Defrain, Tereza Klimošová, Aurélie Lagoutte, and Jonathan Narboni. On Vizing's edge colouring question. *arXiv preprint arXiv:2107.07900*, 2021.

[3] Jessica McDonald, Bojan Mohar, and Diego Scheide. Kempe equivalence of edge-colorings in subcubic and subquartic graphs. *Journal of Graph Theory*, 70(2):226–239, 2012.

[4] Bojan Mohar. Kempe equivalence of colorings. In *Graph Theory in Paris*, pages 287–297. Springer, 2006.

[5] Michael Stiebitz, Diego Scheide, Bjarne Toft, and Lene M Favrholdt. *Graph edge coloring: Vizing's theorem and Goldberg's conjecture*, volume 75. John Wiley & Sons, 2012.

[6] Vadim G. Vizing. On an estimate of the chromatic class of a p-graph. *Discret Analiz*, 3:25–30, 1964.

[7] Vadim G Vizing. The chromatic class of a multigraph. *Cybernetics*, 1(3):32–41, 1965.

[8] Vadim G. Vizing. Some unsolved problems in graph theory. *Russian Mathematical Surveys*, 23(6):125, 1968.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Product-free sets in the free group

(Extended abstract)

Miquel Ortega Juanjo Rué Oriol Serra *

Abstract

We prove that product-free sets of the free group over a finite alphabet have maximum density 1/2 with respect to the natural measure that assigns total weight one to each set of irreducible words of a given length. This confirms a conjecture of Leader, Letzter, Narayanan and Walters. In more general terms, we actually prove that strongly k-product-free sets have maximum density $1/k$ in terms of the said measure. The bounds are tight.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-103`

1 Introduction

A subset S of a group is said to be *product-free* if there do not exist $x, y, z \in S$ (not necessarily distinct) such that $z = x \cdot y$. Much has been studied about product-free subsets of finite groups, particularly so in the abelian case, where they are usually called *sum-free subsets* (see, for example, the survey by Tao and Vu [8]). The study of product–free subsets in nonabelian groups can be traced back to Babai and Sós [1], see the survey by Kedlaya [3]. Interest on the problem was prompted by the seminal work of Gowers on quasirandom groups [2].

The study of *product-free* sets in discrete infinite structures is more recent. As a first approach to the study of the infinite case, Leader, Letzter, Narayanan, and Walters in [4] proved that product-free subsets of the free *semigroup* on the finite alphabet $\mathcal{A}$ have maximum density 1/2 with respect to the measure that assigns a weight of $|\mathcal{A}|^{-n}$ to every

*Departament de Matemàtiques and Institut de Matemàtiques (IMTech) de la Universitat Politècnica de Catalunya (UPC), and Centre de Recerca Matemàtica (CRM), Barcelona, Spain. E-mail: `miquel.ortega.sanchez-colomer@upc.edu`, `juan.jose.rue@upc.edu`, `oriol.serra@upc.edu`.

word of length n. They conjectured that this is also true for the analogous measure on the free group. The main purpose of the present paper is to provide a proof of this conjecture.

More precisely, let us write $\mathcal{F}$ for the free group over a finite alphabet $\mathcal{A}$. For $n \geq 1$ and $A \subseteq \mathcal{F}$, we write $A(n) = \{w \in A \colon |w| = n\}$ for the set of elements of A whose reduced words have length n, and $A_{\leq n}$ for those that have length smaller or equal than n. We define a measure μ on $\mathcal{F}$ such that $\mu(\{w\}) = 1/|\mathcal{F}(|w|)|$ for all $w \in \mathcal{F}$, so that every layer of words of a given length has the same total weight. Finally, we write $\bar{d}(A) = \limsup_{n\to\infty} \frac{\mu(A_{\leq n})}{\mu(\mathcal{F}_{\leq n})}$ for the upper asymptotic density of A. Our main result can be phrased as follows.

Theorem 1. *Let $S \subseteq \mathcal{F}$ be a product-free subset. Then*

$$\bar{d}(S) \leq \frac{1}{2}. \tag{1}$$

We actually study a generalisation of this result. Following the notion by Łuczak and Schoen [6] we call a subset S of a semigroup *k-product-free* for $k \geq 2$ if there are no $x_1, \ldots, x_k, y \in S$ such that $x_1 \ldots x_k = y$ and, furthermore, we call S *strongly* k-product-free if it is l-product-free for all l with $2 \leq l \leq k$. We are able to prove the following.

Theorem 2. *Let $S \subseteq \mathcal{F}$ be a strongly k-product-free subset for $k \geq 2$. Then*

$$\bar{d}(S) \leq \frac{1}{k}. \tag{2}$$

Fixing an arbitrary $x \in \mathcal{A}$, the natural example of the set $S \subseteq \mathcal{F}$ consisting of words such that the number of x minus the number of x^{-1} in its reduced form is congruent to 1 modulo k, which is strongly k–product–free and has upper asymptotic density $\bar{d}(S) = 1/k$, shows that the upper bound in Theorem 2 is best possible.

We split the proof of Theorem 2 in two steps. The first step consists in reducing the problem to an analogous one over a particular semigroup. The second step consists in proving the theorem over this semigroup, where we may use similar arguments to those of [4]. However, their argument seems to break when considering strongly k-product-free subsets for $k > 3$. We avoid the obstruction we encounter by restricting our analysis to a subsemigroup where S is at least as dense and pseudorandom in a certain weak sense. This is achieved via a density increment argument.

To conclude, we also want to remark that the statement of Theorem 2 also holds in the free semigroup, which actually was the model where we first worked out the proof, and it is a generalisation of the main theorem in [4].

Theorem 3. *For any finite alphabet $\mathcal{A}$, a strongly k-product-free of the free semigroup over $\mathcal{A}$ has upper asymptotic density at most $1/k$.*

Details of the proofs not included here can be found in [7].

2 Reduction to a semigroup

We write $\mathcal{F}^{xy} \subseteq \mathcal{F}$ for the subset of words that begin with x and end in y (with $x, y \in \mathcal{A} \cup \mathcal{A}^{-1}$). The first step of the proof consists in reducing the proof of Theorem 2 to an analogous statement over $\mathcal{F}^{xy}$ with $x \neq y^{-1}$. This ambient space has the advantage of having no cancellation when multiplying, so it is much closer to the case of the free semigroup, and we are then be able to use ideas similar in spirit to those of [4].

For a given family $\mathcal{H} \subseteq \mathcal{F}$ and a subset $A \subseteq \mathcal{H}$, we define the *relative upper density* of A as

$$\bar{d}_{\mathcal{H}}(A) = \limsup_{n\to\infty} \frac{\mu(A_{\leq n})}{\mu(\mathcal{H}_{\leq n})}.$$

The analogous result to Theorem 2 then reads as follows.

Proposition 1. *Let $S \subseteq \mathcal{F}^{xy}$ be a strongly k-product-free with $k \geq 2$ and $x \neq y^{-1}$. Then*

$$\bar{d}_{\mathcal{F}^{xy}}(S) \leq \frac{1}{k}.$$

The proof of Theorem 2 starts by splitting the elements of $S \subset \mathcal{F}$ according to their initial and final letter. If these are not opposite, then we may apply Proposition 1. If they are opposite, the crucial observation is that we are conjugating by a certain letter, and hence the property of being product-free is preserved. We may then split again according to the second and next to last letters. Iterating this argument, the increase in density over $1/k$ must come from words of the form $w\alpha w^{-1}$ for w of large length, which have exponentially low total mass.

3 Proof over a semigroup

Let us give a brief overview of the proof of Proposition 1. Assume fixed $x, y \in \mathcal{A} \cup \mathcal{A}^{-1}$ such that $x \neq y^{-1}$ and write $\mathcal{G} = \mathcal{F}^{xy} \subseteq \mathcal{F}$. Also write $K = \frac{2|\mathcal{A}|}{2|\mathcal{A}|-1}$ for a constant that appears in several arguments, due to the fact that $|\mathcal{F}(i)||\mathcal{F}(j)| = \frac{|\mathcal{F}(i+j)|}{K}$.

We first prove a version of Proposition 1 that depends on the construction of certain subsets of S with appropriate properties. We say $\mathcal{H} \subseteq \mathcal{F}$ is *dense* if $\mu(\mathcal{H}_{\leq n}) > \delta\mu(\mathcal{F}_{\leq n}) > 0$ for a fixed δ and all large enough n. Furthermore, it is a *subsemigroup* if $\mathcal{H} \cdot \mathcal{H} \subseteq \mathcal{H}$, i.e. $\alpha\beta \in \mathcal{H}$ for all $\alpha, \beta \in \mathcal{H}$. Finally, a subset $W \subseteq \mathcal{H}$ has *unique products* in $\mathcal{H}$ if the map

$$\begin{aligned}(W, \mathcal{H}) &\to \mathcal{H}\\ (w, h) &\mapsto w \cdot h\end{aligned}$$

is injective.

For W and $\mathcal{H} \subseteq \mathcal{G}$ satisfying the above properties, we prove a version of Proposition 1 conditional on $\mu(W)$ being close to K.

Lemma 1. *Let $\mathcal{H} \subseteq \mathcal{G}$ be a dense subsemigroup. For any strongly k-product free set $S \subseteq \mathcal{H}$ and finite subset $W \subseteq S$ with unique products it holds that*

$$\bar{d}_{\mathcal{H}}(S) \leq \frac{1}{1 + \frac{\mu(W)}{K} + \cdots + \left(\frac{\mu(W)}{K}\right)^{k-1}}$$

The previous lemma is proved, essentially, through a counting argument. Since S is k-product free, the sets $W^i \cdot S$ will all be disjoint, and hence they cannot be too large. The properties of W allow us to lower bound the size of $W \cdot S$ in terms of the sizes of W and S. We phrase the argument in a probabilistic manner.

In order to prove Proposition 1, we need to build a subset W large enough to apply the previous lemma. We do so in two different ways. We first present a proof for the case $2 \leq k \leq 3$, where we can use a similar argument to the one in [4]. In this case, we may exploit the fact that S is product-free to construct W in a relatively straightforward manner.

The straightforward argument fails for $k > 3$. Thus, we must find another way of building the subset $W \subseteq S$ necessary to apply Lemma 1. To do so, instead of finding such a set directly in $\mathcal{G}$, we find a subset of $\mathcal{G}$ where S is regularly distributed, where the existence of $W \subseteq S$ as we are interested in is much easier to prove and does not depend on S being product-free.

Concretely, we are interested in studying S when we restrict ourselves to words which are divisible in $\mathcal{G}$ by a given factor. For a given $w \in \mathcal{G}$, we write $w\mathcal{G} \subseteq \mathcal{G}$ for the set of words belonging to $\mathcal{G}$ which may be written as $w\alpha$ for $\alpha \in \mathcal{G}$. We then define the following pseudorandomness condition, which measures whether S is evenly distributed when restricted to such sets.

Definition 1. Given $w \in \mathcal{G}$, a subset $S \subset w\mathcal{G}$ is ε-regular in $w\mathcal{G}$ if

$$\bar{d}_{ww'\mathcal{G}}((S \cap (ww'\mathcal{G})) < \bar{d}_{w\mathcal{G}}(S) + \varepsilon \tag{3}$$

for all words $w' \in \mathcal{G}$.

We then prove the analogous statement to Theorem 2 under pseudorandomness assumptions. In particular, the following Lemma implies Theorem 2 when S is ε-regular for all $\varepsilon > 0$.

Lemma 2. *Let $S \subset w\mathcal{G}$ be a strongly k-product free set that is ε-regular in $w\mathcal{G}$, with $w \in \mathcal{G}$, and let $d = \bar{d}_{w\mathcal{G}}(S)$ be its relative upper density. Then*

$$d\left(1 + \frac{d}{d+2\varepsilon} + \cdots + \left(\frac{d}{d+2\varepsilon}\right)^k\right) \leq 1, \tag{4}$$

Finally, we use a density-increment strategy, where failure of pseudorandomness implies an increase in density, to find w such that $S \cap w\mathcal{G}$ is pseudorandom, and apply the previous lemma in this setting.

4 Final remarks

The proof of Theorem 3 is done following the arguments from the previous section, by only replacing the role played by $\mathcal{F}^{xy}$ for $\boldsymbol{F}_{\mathcal{A}}$, and by replacing K by 1. It is also worth noting that the results in [4] concern the upper Banach density of sum-free subsets, which gives slightly stronger results, since the upper Banach density is an upper bound for the upper asymptotic density we consider. For the sake of simplicity, we have not attempted to write down our results for this case, although all arguments go through.

Finally, it would also be interesting to consider the case of k-product-free sets. To state the natural conjecture for this case, define ρ as

$$\rho(l) = \min\left(\{l \in \mathbb{N} \colon l \nmid k-1\}\right).$$

Then we believe the following to be true.

Conjecture 1. *Let $S \subseteq \mathcal{F}$ be a k-product-free subset for $k \geq 2$. Then*

$$\bar{d}(S) \leq \frac{1}{\rho(k)} \tag{5}$$

This is analogous to a result of Łuczak and Schoen [5], which proves the corresponding statement over the integers.

5 Acknowledgments

We acknowledge the support of the grants MTM2017-82166-P, PID2020-113082GB-I00, and the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence in R&D (CEX2020-001084-M). Miquel Ortega also acknowledges the support of the FPI grant PRE2021-099120.

References

[1] László Babai and Vera T. Sós. Sidon sets in groups and induced subgraphs of Cayley graphs. *European J. Combin.*, 6(2):101–114, 1985.

[2] W. T. Gowers. Quasirandom groups. *Combin. Probab. Comput.*, 17(3):363–387, 2008.

[3] Kiran S. Kedlaya. Product-free subsets of groups, then and now. 479:169–177, 2009.

[4] Imre Leader, Shoham Letzter, Bhargav Narayanan, and Mark Walters. Product-free sets in the free semigroup. *European Journal of Combinatorics*, 83:103003, 2020.

[5] Tomasz Łuczak and Tomasz Schoen. On infinite sum-free sets of natural numbers. *J. Number Theory*, 66(2):211–224, 1997.

[6] Tomasz Łuczak and Tomasz Schoen. Sum-free subsets of right cancellative semigroups. *European J. Combin.*, 22(7):999–1002, 2001.

[7] Miquel Ortega, Juanjo Rué, and Oriol Serra. Product-free sets in the free group. arXiv 2302.03748, 2023.

[8] Terence Tao and Van Vu. Sum-free sets in groups: a survey. *J. Comb.*, 8(3):541–552, 2017.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Minimum vertex degree conditions for loose spanning trees in 3-graphs

(Extended abstract)

Yanitsa Pehova* Kalina Petrova†

Abstract

In 1995, Komlós, Sárközy and Szemerédi showed that for large n, every n-vertex graph with minimum degree at least $(1/2+\gamma)n$ contains all spanning trees of bounded degree. We consider a generalization of this result to loose spanning hypertrees, that is, linear hypergraphs obtained by successively appending edges sharing a single vertex with a previous edge, in 3-graphs. We show that for all γ and Δ, and n large, every n-vertex 3-uniform hypergraph of minimum vertex degree $(5/9+\gamma)\binom{n}{2}$ contains every loose spanning tree with maximum vertex degree Δ. This bound is asymptotically tight, since some loose trees contain perfect matchings.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-104`

1 Introduction

A classical result of Komlós, Sárközy and Szemerédi [4] states that for large n, any n-vertex graph with minimum degree $(1/2 + \varepsilon)n$ contains every spanning tree of bounded degree. Since a Hamilton path is a tree of bounded degree, the constant $1/2$ is best possible by any construction showing that Dirac's theorem is best possible.

*Department of Mathematics, London School of Economics, WC2A 2AE London, United Kingdom. E-mail: `y.pehova@lse.ac.uk`. Supported by the Engineering and Physical Sciences Research Council, UK Research and Innovation [grant number EP/V038168/1].

†Department of Computer Science, ETH, 8092 Zürich, Switzerland. E-mail: `kalina.petrova@inf.ethz.ch`. Supported by grant no. CRSII5 173721 of the Swiss National Science Foundation.

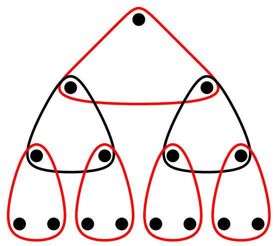


Figure 1: The complete binary loose tree with 4 levels. A perfect matching is shown in red.

We consider the corresponding problem for hypergraphs. We work with a combinatorial definition of a k*-uniform* ℓ*-tree* – a k-uniform hypergraph admitting an edge ordering $e_1, \dots, e_m$ such that each e_i shares ℓ vertices with one previous edge in the ordering[1]. Such orderings we call *valid*, and the edges which can be last in a valid ordering we call *leaves*. We will call 1-trees *loose* (also known in the literature as *linear*). Similarly, a $(k-1)$-tree is also called a *tight tree*. We denote by $\delta_\ell(H)$ the *minimum* ℓ*-degree* of a k-graph H, that is, the minimum number of edges containing a set of ℓ vertices of H. Maximum degree is defined accordingly.

Not much is known about extensions of Komlós, Sárközy and Szemerédi's result to general k-uniform ℓ-trees, apart from a recent result of Pavez-Signé, Sanhueza-Matamala and Stein [5, 6] which shows that minimum $\delta_{k-1}(H) \geqslant (1/2+\gamma)n$ forces the existence of any tight spanning tree T with $\Delta_1(T) \leqslant \Delta$.

Buß, Hàn and Schacht [1] showed that if $\delta_1(H) \geqslant \left(\frac{7}{16}+\varepsilon\right)\binom{n}{2}$, then H contains a loose Hamilton cycle – a cycle whose adjacent edges share exactly one vertex. The constant 7/16 is best possible, and in a later paper Han and Zhao [3] gave the exact threshold.

In light of this, one may conjecture that 3-graphs with minimum vertex degree $\left(\frac{7}{16}+\varepsilon\right)\binom{n}{2}$ also contain every loose tree of bounded degree. However, this is not the case. Consider the *complete binary loose tree* as shown in Figure 1. A complete binary loose tree T_b with an even number of levels contains a perfect matching, so any 3-graph without a perfect matching will also not contain T_b. The asymptotic minimum degree threshold for perfect matchings in 3-graphs was shown to be 5/9 by Hàn, Person and Schacht [2]. Their asymptotic bound was later made exact by Treglown, Kühn and Osthus [8]. This is tight as witnessed by the hypergraph on vertex set $A \cup B$ with $|A| = n/3 - 1$ and $|B| = 2n/3 + 1$ consisting of all edges with at least one vertex in A. Therefore, the minimum vertex degree threshold for the existence of bounded degree loose spanning trees must be at least 5/9. We show that this is in fact the correct threshold.

Theorem 1.1. *For all $\gamma > 0$ and $\Delta \in \mathbb{N}$ there exists $n_0 \in \mathbb{N}$ such that any 3-graph H on $n \geqslant n_0$ vertices with n odd and $\delta_1(H) \geqslant \left(\frac{5}{9}+\gamma\right)\binom{n}{2}$ contains every n-vertex loose tree T with $\Delta_1(T) \leqslant \Delta$.*

[1]More formally, for each $i \geqslant 2$ there exists $j < i$ such that $e_i \cap \bigcup_{j'<i} e_{j'} \subseteq e_j$ and $|e_i \cap e_j| = \ell$.

2 Proof of Theorem 1.1

Our proof employs a classic recipe prescribed by the absorbing method.

Step 1. Find an *absorbing set* $\mathcal{A}$ in our host graph. In Absorbing set lemma we show existence and in Absorbing lemma we show its absorbing properties.

Step 2. Embed a small proportion of our tree T in a way that covers the relevant vertices of $\mathcal{A}$. This is Covering lemma.

Step 3. Use the regularity method to extend this embedding to almost all of T. This is Approximate embedding lemma.

Step 4. Use $\mathcal{A}$ to extend the embedding to all of T.

As an expansion of this sketch we give statements and proof ideas of the four lemmas used above, and show how they imply Theorem 1.1.

The proof of the following lemma is analogous to [6, Proposition 9.4 and Lemma 9.5]. It uses the fact that in a graph with minimum vertex degree $\left(\frac{1}{2}+o(1)\right)\binom{n}{2}$ every triple of vertices (w_1, w_2, w_3) has a positive density of absorbing pairs of Δ-stars (see Figure 2). We denote the set of such star-pairs by $A_\Delta(w_1, w_2, w_3)$. Subsampling these over all triples of vertices with the appropriate probability gives a large absorbing set.

Absorbing set lemma. *Let $1/n \ll \alpha \ll \beta \ll \gamma, 1/\Delta$. Let H be a 3-graph on n vertices with $\delta_1(H) \geqslant \left(\frac{1}{2}+\gamma\right)\binom{n}{2}$. Then there exists a set $\mathcal{A}$ of at most βn vertex-disjoint pairs of Δ-stars such that for every triple (w_1, w_2, w_3) of distinct vertices in H we have $|A_\Delta(w_1, w_2, w_3) \cap \mathcal{A}| \geq \alpha n$.*

The following lemma shows that the set $\mathcal{A}$ in fact absorbs – given a partial embedding of a tree which covers $\mathcal{A}$, we can use $\mathcal{A}$ to find a full embedding. Intuitively, this is possible because given a triple (w_1, w_2, w_3) and one of its absorbing star-pairs (S_{u_2}, S_{u_3}), we can add the edge $\{w_1, u_2, u_3\}$ to the partial embedding by switching u_2 for w_2 and u_3 for w_3 (see Figure 2). Repeating this switch enough times gives a full embedding of T.

Absorbing lemma. *Let $1/n \ll \eta < \alpha < 1/\Delta$. Let T be a loose 3-tree on n vertices of maximum degree Δ with a valid ordering of the edges $e_1, \ldots, e_{(n-1)/2}$ and let $T_0 = \{e_1, \ldots, e_{(n'-1)/2}\}$ be a subtree of T on $n' \geq (1-\eta)n$ vertices. Let H be a 3-graph on n vertices, and ϕ be an embedding $\phi : V(T_0) \to V(H)$. Suppose $\mathcal{A}$ is a family of vertex-disjoint pairs of Δ-stars such that every tuple in $\mathcal{A}$ is covered by ϕ and $|A_\Delta(w_1, w_2, w_3) \cap \mathcal{A}| \geq \alpha n$ for every triple (w_1, w_2, w_3) of distinct vertices of H. Then there is an embedding of T into H.*

The proof of the following covering lemma is analogous to [6, Lemma 9.7].

Covering lemma. *Let $1/n \ll \beta \ll \nu \ll \gamma, 1/\Delta$. Let H be a 3-graph on n vertices with minimum degree $\left(\frac{1}{2}+\gamma\right)\binom{n}{2}$ and let T be a loose tree on νn vertices with maximum degree*

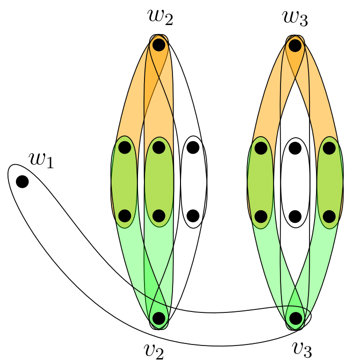


Figure 2: A pair (S_{v_2}, S_{v_3}) of 3-stars which is absorbing for (w_1, w_2, w_3) and covered by an embedding ϕ. Images of edges under ϕ are shown in green. The crucial property of this structure is that the two green stars in ϕ can be switched for the two orange stars plus an extra edge at w_1, thus extending the embedding.

Δ. *Let $\mathcal{A}$ be a set of at most βn pairwise vertex-disjoint absorbing star-pairs in H. Then there is an embedding $\phi : V(T) \to V(H)$ such that every absorbing tuple in $\mathcal{A}$ is covered by ϕ.*

In the following lemma we show that a bounded-degree tree T of size almost n can be embedded in our host graph H. To prove this, we first apply the weak regularity lemma to H to obtain an ε-regular partition of H. The cluster graph inherits the minimum degree of H, and so by the main result in [7] it contains a tight Hamilton cycle $\mathcal{C} = (V_1, ..., V_t)$. The properties of the regular partition give an embedding of T as long as we can produce what we call a *valid assignment* $a : V(T) \to [t]$ of its vertices to the clusters of $\mathcal{C}$. A valid assignment satisfies the following two properties:

- the total number of vertices assigned to each V_j does not exceed $(1 - \eta)|V_j|$, where $\eta \gg \varepsilon$,
- all edges of T are assigned to edges of $\mathcal{C}$.

Our key idea for finding a valid assignment is to break down the almost-spanning tree into linear-sized pieces, assign these pieces to different edges of $\mathcal{C}$, and then 'wrap' around the tight Hamilton cycle $\mathcal{C}$ to connect the pieces to each other. When assigning a piece of our tree to an edge of $\mathcal{C}$, we always make sure to leave approximately the same number of vertices unused in each cluster of that edge of $\mathcal{C}$, so that there is always at least one edge with the capacity to assign an extra piece to it. Since $\mathcal{C}$ has constantly many edges, wrapping around it to connect the pieces only uses up constantly many vertices and so does not interfere with our balance invariant.

Approximate embedding lemma. *Let $1/n \ll \eta \ll \gamma, 1/\Delta$, and let H be a 3-graph on n vertices with $\delta_1(H) \geq \left(\frac{5}{9} + \gamma\right)\binom{n}{2}$. Let T be a loose tree of maximum degree Δ on at least $(1 - \eta)n$ vertices. Then for every $x \in V(T)$ and $z \in V(H)$, there exists an embedding of T into H that maps x to z.*

We are now ready to put these five lemmas together to prove our main result.

Proof of Theorem 1.1. Let $1/n \ll \eta < \alpha \ll \beta \ll \nu \ll \gamma, 1/\Delta$, where n is odd.

We first apply Absorbing set lemma to get a set $\mathcal{A}$ of at most βn pairwise vertex-disjoint pairs of stars, such that for every triple (w_1, w_2, w_3) of vertices in H we have that $|A_\Delta(w_1, w_2, w_3) \cap \mathcal{A}| \geq \alpha n$.

Next, root T arbitrarily at some vertex r and find a subtree $T_x \subset T$ of size $\nu n \leqslant v(T_x) \leqslant 2\Delta\nu n$. This can be done by setting $x := r$ and, until x has a child y whose subtree has at least νn vertices, set $x := y$. At some point this process reaches a vertex x whose subtree T_x has at least νn vertices, but all its children's subtrees have fewer than νn vertices, implying that $v(T_x) \leqslant 2\Delta\nu n$. Let $\nu' := v(T_x)/n$ and apply Covering lemma with $\nu := \nu'$ and $T := T_x$ to find an embedding $\phi_1 : V(T_x) \to V(H)$ such that every pair of stars in $\mathcal{A}$ is covered by ϕ_1. Denote $\phi_1(x) = z$.

Now let $H_1 := (H \setminus \phi_1(T_x)) \cup \{z\}$ and note that $\delta_1(H_1) \geq \left(\frac{5}{9} + \frac{\gamma}{2}\right)\binom{|H_1|}{2}$. Let $T_1 := (T \setminus T_x) \cup \{x\}$ and root T_1 at x. Remove leaf edges from T_1 repeatedly to get T_2 such that $v(T_1) - v(T_2) = \eta|H_1|$. Apply Approximate embedding lemma with $H := H_1$ and $T := T_2$ to find an embedding ϕ_2 of T_2 into H_1 with $\phi_2(x) = z$.

Finally, let $T_3 := T_x \cup T_2$ and note that $v(T_3) = n - \eta|H_1| \geqslant (1-\eta)n$. Combine ϕ_1 and ϕ_2 into an embedding ϕ_3 of T_3 into H, which can be done since $\phi_1(T_x) \cap \phi_2(T_2) = \phi_1(x) = \phi_2(x)$. Then the tree T_3, the embedding ϕ_3, and the set of absorbing tuples $\mathcal{A}$ satisfy the conditions of Absorbing lemma, which we can apply to get an embedding of T in H. □

References

[1] E. Buß, H. Hàn, and M. Schacht. Minimum vertex degree conditions for loose Hamilton cycles in 3-uniform hypergraphs. *J. Combin. Theory Ser. B*, 103(6):658–678, 2013.

[2] H. Hàn, Y. Person, and M. Schacht. On perfect matchings in uniform hypergraphs with large minimum vertex degree. *SIAM J. Discrete Math.*, 23(2):732–748, 2009.

[3] J. Han and Y. Zhao. Minimum vertex degree threshold for loose Hamilton cycles in 3-uniform hypergraphs. *Journal of Combinatorial Theory, Series B*, 114:70–96, 2015.

[4] J. Komlós, G. N. Sárközy, and E. Szemerédi. Proof of a packing conjecture of Bollobás. *Combin. Probab. Comput.*, 4(3):241–255, 1995.

[5] M. Pavez-Signé, N. Sanhueza-Matamala, and M. Stein. Dirac-type conditions for spanning bounded-degree hypertrees. In *Extended Abstracts EuroComb 2021*, pages 586–592. Springer, 2021.

[6] M. Pavez-Signé, N. Sanhueza-Matamala, and M. Stein. Dirac-type conditions for spanning bounded-degree hypertrees, Dec 2021, arXiv:2012.09824.

[7] C. Reiher, V. Rödl, A. Ruciński, M. Schacht, and E. Szemerédi. Minimum vertex degree condition for tight Hamiltonian cycles in 3-uniform hypergraphs. *Proceedings of the London Mathematical Society*, 119(2):409–439, 2019.

[8] A. Treglown, D. Kühn, and D. Osthus. Matchings in 3-uniform hypergraphs of large minimum vertex degree. *Electronic Notes in Discrete Mathematics*, 38:813–818, 2011. The Sixth European Conference on Combinatorics, Graph Theory and Applications, EuroComb 2011.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Strengthening the Directed Brooks' Theorem for oriented graphs and consequences on digraph redicolouring

(Extended abstract)

Lucas Picasarri-Arrieta*

Abstract

Let $D = (V, A)$ be a digraph. We define $\Delta_{\max}(D)$ as the maximum of $\{\max(d^+(v), d^-(v)) \mid v \in V\}$ and $\Delta_{\min}(D)$ as the maximum of $\{\min(d^+(v), d^-(v)) \mid v \in V\}$. It is known that the dichromatic number of D is at most $\Delta_{\min}(D) + 1$. In this work, we prove that every digraph D which has dichromatic number exactly $\Delta_{\min}(D) + 1$ must contain the directed join of $\overleftrightarrow{K_r}$ and $\overleftrightarrow{K_s}$ for some r, s such that $r + s = \Delta_{\min}(D) + 1$, except if $\Delta_{\min}(D) = 2$ in which case D must contain a digon. In particular, every oriented graph $\vec{G}$ with $\Delta_{\min}(\vec{G}) \geq 2$ has dichromatic number at most $\Delta_{\min}(\vec{G})$.

Let $\vec{G}$ be an oriented graph of order n such that $\Delta_{\min}(\vec{G}) \leq 1$. Given two 2-dicolourings of $\vec{G}$, we show that we can transform one into the other in at most n steps, by recolouring one vertex at each step while maintaining a dicolouring at any step. Furthermore, we prove that, for every oriented graph $\vec{G}$ on n vertices, the distance between two k-dicolourings is at most $2\Delta_{\min}(\vec{G})n$ when $k \geq \Delta_{\min}(\vec{G}) + 1$.

We then extend a theorem of Feghali, Johnson and Paulusma to digraphs. We prove that, for every digraph D with $\Delta_{\max}(D) = \Delta \geq 3$ and every $k \geq \Delta + 1$, the k-dicolouring graph of D consists of isolated vertices and at most one further component that has diameter at most $c_\Delta n^2$, where $c_\Delta = O(\Delta^2)$ is a constant depending only on Δ.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-105`

*Université Côte d'Azur, CNRS, I3S, INRIA, Sophia Antipolis, France.
E-mail: `lucas.picasarri-arrieta@inria.fr`. Supported by research grant DIGRAPHS ANR-19-CE48-0013 and by the French government, through the EUR DS4H Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-17-EURE-0004.

1 Introduction

1.1 Graph (re)colouring

Given a graph $G = (V, E)$, a *k-colouring* of G is a function $c : V \to \{1, \ldots, k\}$ such that, for every edge $xy \in E$, we have $c(x) \neq c(y)$. The *chromatic number* of G, denoted by $\chi(G)$, is the smallest k such that G admits a k-colouring. The *maximum degree* of G, denoted by $\Delta(G)$, is the degree of the vertex with the greatest number of edges incident to it. A simple greedy procedure shows that, for any graph G, $\chi(G) \leq \Delta(G) + 1$. The celebrated theorem of Brooks [6] characterizes the graphs for which equality holds.

Theorem 1 ([6])**.** *A connected graph G satisfies $\chi(G) = \Delta(G) + 1$ if and only if G is an odd cycle or a complete graph.*

For any $k \geq \chi(G)$, the *k-colouring graph* of G, denoted by $\mathcal{C}_k(G)$, is the graph whose vertices are the k-colourings of G and in which two k-colourings are adjacent if they differ by the colour of exactly one vertex. A path between two given colourings in $\mathcal{C}_k(G)$ corresponds to a *recolouring sequence*. In the last fifteen years, since the papers of Bonsma, Cereceda, van den Heuvel and Johnson [7, 3, 8, 9], graph recolouring has been studied by many researchers in graph theory. Feghali, Johnson and Paulusma [10] proved the following analogue of Brooks' Theorem for graphs recolouring.

Theorem 2 ([10])**.** *Let $G = (V, E)$ be a connected graph with $\Delta(G) = \Delta \geq 3$, $k \geq \Delta + 1$, and α, β two k-colourings of G. Then at least one of the following holds:*

- *α or β is an isolated vertex in $\mathcal{C}_k(G)$, or*
- *there is a recolouring sequence of length at most $c_\Delta |V|^2$ between α and β, where $c_\Delta = O(\Delta)$ is a constant depending on Δ.*

Considering graphs of bounded maximum degree, Theorem 2 has been very recently improved by Bousquet, Feuilloley, Heinrich and Rabie (see [4]). They prove that there is a recolouring sequence between α and β of length at most $f(\Delta)|V|$ for some computable function f, except if α or β is an isolated vertex in $\mathcal{C}_k(G)$.

1.2 Digraph (re)dicolouring

In this paper, we are looking for extensions of the previous results on graphs colouring and recolouring to digraphs.

Let D be a digraph. A *digon* is a pair of arcs in opposite directions between the same vertices. An *oriented graph* is a digraph with no digon. The *bidirected graph* associated to a graph G, denoted by $\overleftrightarrow{G}$, is the digraph obtained from G, by replacing every edge by a digon. The *underlying graph* of D, denoted by $UG(D)$, is the undirected graph G with vertex set $V(D)$ in which uv is an edge if and only if uv or vu is an arc of D.

Let v be a vertex of a digraph D. The *out-degree* (resp. *in-degree*) of v, denoted by $d^+(v)$ (resp. $d^-(v)$), is the number of arcs leaving (resp. entering) v. We define the *maximum degree* of v as $d_{\max}(v) = \max\{d^+(v), d^-(v)\}$, and the *minimum degree* of v as $d_{\min}(v) = \min\{d^+(v), d^-(v)\}$. We can then define the corresponding maximum degrees of D: $\Delta_{\max}(D) = \max_{v \in V(D)}(d_{\max}(v))$ and $\Delta_{\min}(D) = \max_{v \in V(D)}(d_{\min}(v))$. A digraph D is Δ-*diregular* if, for every vertex $v \in V(D)$, $d^-(v) = d^+(v) = \Delta$. The *directed join* of D_1 and D_2, denoted by $D_1 \Rightarrow D_2$, is the digraph obtained from disjoint copies of D_1 and D_2 by adding all arcs from the copy of D_1 to the copy of D_2.

In 1982, Neumann-Lara [12] introduced the notions of dicolouring and dichromatic number, which generalize the ones of colouring and chromatic number. A k-*dicolouring* of D is a function $c : V(D) \to \{1, \dots, k\}$ such that $c^{-1}(i)$ induces an acyclic subdigraph in D for each $i \in \{1, \dots, k\}$. The *dichromatic number* of D, denoted by $\vec{\chi}(D)$, is the smallest k such that D admits a k-dicolouring. There is a one-to-one correspondence between the k-colourings of a graph G and the k-dicolourings of the associated bidirected graph $\overleftrightarrow{G}$, and in particular $\chi(G) = \vec{\chi}(\overleftrightarrow{G})$. Hence every result on graph colourings can be seen as a result on dicolourings of bidirected graphs, and it is natural to study whether the result can be extended to all digraphs.

The directed version of Brooks' Theorem was first proved by Harutyunyan and Mohar in [11] (see also [1]).

Theorem 3 (Directed Brooks' Theorem)**.** *Let D be a connected digraph. Then $\vec{\chi}(D) \leq \Delta_{\max}(D) + 1$ and equality holds if and only if D is a directed cycle, a bidirected odd cycle or a bidirected complete graph of order at least 4.*

It is easy to prove, by a simple greedy procedure, that every digraph D can be dicoloured with $\Delta_{\min}(D) + 1$ colours. Hence, one can wonder if Brooks' Theorem can be extended to digraphs using $\Delta_{\min}(D)$ instead of $\Delta_{\max}(D)$. Our main result is the following.

Theorem 4. *Let D be a digraph. If $\vec{\chi}(D) = \Delta_{\min}(D) + 1$, then one of the following holds:*

- $\Delta_{\min}(D) \leq 1$, *or*
- $\Delta_{\min}(D) = 2$ *and D contains* $\overleftrightarrow{K}_2$, *or*
- $\Delta_{\min}(D) \geq 3$ *and D contains* $\overleftrightarrow{K}_r \Rightarrow \overleftrightarrow{K}_s$, *for some $r, s \geq 0$ such that* $r + s = \Delta_{\min}(D) + 1$.

In particular, the following is a direct consequence of Theorem 4.

Corollary 5. *Let D be a digraph. If $\vec{\chi}(D) = \Delta_{\min}(D) + 1$, then D contains the complete bidirected graph on $\left\lceil \frac{\Delta_{\min}(D)+1}{2} \right\rceil$ vertices as a subdigraph.*

Corollary 5 is best possible: if we restrict D to not contain the complete bidirected graph on $\left\lceil \frac{\Delta_{\min}(D)+1}{2} \right\rceil + 1$ vertices, then deciding $\vec{\chi}(D) \leq \Delta_{\min}(D)$ is NP-complete (see [13]). Since an oriented graph does not contain any digon, Corollary 5 directly implies the following.

Corollary 6. *Let $\vec{G}$ be an oriented graph. If $\Delta_{\min}(\vec{G}) \geq 2$, then $\vec{\chi}(\vec{G}) \leq \Delta_{\min}(\vec{G})$.*

For any $k \geq \vec{\chi}(D)$, the *k-dicolouring graph* of D, denoted by $\mathcal{D}_k(D)$, is the graph whose vertices are the k-dicolourings of D and in which two k-dicolourings are adjacent if they differ by the colour of exactly one vertex. Observe that $\mathcal{C}_k(G) = \mathcal{D}_k(\overleftrightarrow{G})$ for any bidirected graph $\overleftrightarrow{G}$. A *redicolouring sequence* between two dicolourings is a path between these dicolourings in $\mathcal{D}_k(D)$.

Digraph redicolouring was first introduced in [5], where the authors generalized different results on graph recolouring to digraphs, and proved some specific results when restricted to oriented graphs. In particular, they studied the k-dicolouring graph of digraphs with bounded degeneracy or bounded maximum average degree, and they show that finding a redicolouring sequence between two given k-dicolourings of a digraph is PSPACE-complete for every fixed $k \geq 2$. Dealing with the maximum degree of a digraph, they proved that, given an orientation of a subcubic graph $\vec{G}$ on n vertices, its 2-dicolouring graph $\mathcal{D}_2(\vec{G})$ is connected and has diameter at most $2n$ and they asked if this bound can be improved. We answer this question by proving the following theorem.

Theorem 7. *Let $\vec{G}$ be an oriented graph of order n such that $\Delta_{\min}(\vec{G}) \leq 1$. Then $\mathcal{D}_2(\vec{G})$ is connected and has diameter exactly n.*

In particular, if $\vec{G}$ is an orientation of a subcubic graph, then $\Delta_{\min}(\vec{G}) \leq 1$ (because $d^+(v) + d^-(v) \leq 3$ for every vertex v), and so $\mathcal{D}_2(\vec{G})$ has diameter exactly n. Furthermore, we prove the following as a consequence of Corollary 6 and Theorem 7.

Corollary 8. *Let $\vec{G}$ be an oriented graph of order n with $\Delta_{\min}(\vec{G}) = \Delta \geq 1$, and let $k \geq \Delta + 1$. Then $\mathcal{D}_k(\vec{G})$ is connected and has diameter at most $2\Delta n$.*

Corollary 8 does not hold for digraphs in general: indeed, $\overleftrightarrow{P_n}$, the bidirected path on n vertices, satisfies $\Delta_{\min}(\overleftrightarrow{P_n}) = 2$ and $\mathcal{D}_3(\overleftrightarrow{P_n}) = \mathcal{C}_3(P_n)$ has diameter $\Omega(n^2)$, as proved in [2]. Our last result is the following extension of Theorem 2 to digraphs.

Theorem 9. *Let $D = (V, A)$ be a connected digraph with $\Delta_{\max}(D) = \Delta \geq 3$, $k \geq \Delta + 1$, and α, β two k-dicolourings of D. Then at least one of the following holds:*

- *α or β is an isolated vertex in $\mathcal{D}_k(G)$, or*
- *there is a redicolouring sequence of length at most $c_\Delta |V|^2$ between α and β, where $c_\Delta = O(\Delta^2)$ is a constant depending only on Δ.*

Furthermore we prove that $\mathcal{D}_k(D)$ has an isolated vertex if and only if D is bidirected and its underlying graph has one. Thus, an obstruction in Theorem 9 is exactly the bidirected graph of an obstruction in Theorem 2.

In the next section we prove Theorem 4. The integrality of the proofs of the results in this extended abstract can be found in [13].

2 Proof of Theorem 4

A digraph D is *k-dicritical* if $\vec{\chi}(D) = k$ and for every vertex $v \in V(D)$, $\vec{\chi}(D - v) < k$. Observe that every digraph with dichromatic number at least k contains a k-dicritical subdigraph. Let $\mathcal{F}_2$ be $\{\overleftrightarrow{K_2}\}$, and for each $\Delta \geq 3$, we define $\mathcal{F}_\Delta = \{\overleftrightarrow{K_r} \Rightarrow \overleftrightarrow{K_s} \mid r, s \geq 0 \text{ and } r + s = \Delta + 1\}$. A digraph D is *$\mathcal{F}_\Delta$-free* if it does not contain F as a subdigraph, for any $F \in \mathcal{F}_\Delta$. Theorem 4 can then be reformulated as follows.

Theorem 4. *Let D be a digraph with $\Delta_{\min}(D) = \Delta \geq 2$. If D is $\mathcal{F}_\Delta$-free, then $\vec{\chi}(D) \leq \Delta$.*

Proof. Let D be a digraph such that $\Delta_{\min}(D) = \Delta \geq 2$ and $\vec{\chi}(D) = \Delta + 1$. We will show that D contains some $F \in \mathcal{F}_\Delta$ as a subdigraph.

Let (X, Y) be a partition of $V(D)$ such that for each $x \in X$, $d^+(x) \leq \Delta$, and for each $y \in Y$, $d^-(y) \leq \Delta$. We define the digraph $\tilde{D}$ as follows:

- $V(\tilde{D}) = V(D)$,
- $A(\tilde{D}) = A(D\langle X\rangle) \cup A(D\langle Y\rangle) \cup \{xy, yx \mid xy \in A(D), x \in X, y \in Y\}$.

Let us first prove that $\vec{\chi}(\tilde{D}) \geq \Delta + 1$. Assume for a contradiction that there exists a Δ-dicolouring c of $\tilde{D}$. Then D, coloured with c, must contain a monochromatic directed cycle C. Now C is not contained in X nor Y, for otherwise C would be a monochromatic directed cycle of $D\langle X\rangle$ or $D\langle Y\rangle$ and so a monochromatic directed cycle of $\tilde{D}$. Thus C contains an arc xy from X to Y. But then, $\{xy, yx\}$ is a monochromatic digon in $\tilde{D}$, a contradiction.

Since $\vec{\chi}(\tilde{D}) \geq \Delta + 1$, there is a $(\Delta + 1)$-dicritical subdigraph H of $\tilde{D}$. By dicriticality of H, for every vertex $v \in V(H)$, $d_H^+(v) \geq \Delta$ and $d_H^-(v) \geq \Delta$, for otherwise a Δ-dicolouring of $H - v$ could be extended to H by choosing for v a colour which is not appearing in its out-neighbourhood or in its in-neighbourhood. We define X_H as $X \cap V(H)$ and Y_H as $Y \cap V(H)$. Note that both $H\langle X_H\rangle$ and $H\langle Y_H\rangle$ are subdigraphs of D.

We will now prove that H is Δ-diregular. Let ℓ be the number of digons between X_H and Y_H in H. Observe that, by definition of X and H, for each vertex $x \in X_H$, $d_H^+(x) = \Delta$. Note also that, in H, ℓ is exactly the number of arcs leaving X_H and exactly the number of arcs entering X_H. We get:

$$\Delta|X_H| = \sum_{x \in X_H} d_H^+(x) = \ell + |A(H\langle X_H\rangle)| = \sum_{x \in X_H} d_H^-(x)$$

which implies, since H is dicritical, $d_H^+(x) = d_H^-(x) = \Delta$ for every vertex $x \in X_H$. Using a symmetric argument, we prove that $\Delta|Y_H| = \sum_{y \in Y_H} d_H^+(y)$, implying $d_H^+(y) = d_H^-(y) = \Delta$ for every vertex $y \in Y_H$.

Since H is Δ-diregular, then in particular $\Delta_{\max}(H) = \Delta$. Hence, because $\vec{\chi}(H) = \Delta + 1$, by Theorem 3, either $\Delta = 2$ and H is a bidirected odd cycle, or $\Delta \geq 3$ and H is the bidirected complete graph on $\Delta + 1$ vertices. In both cases, $D\langle V(H)\rangle$ contains a digraph of $\mathcal{F}_\Delta$ as a subdigraph. $\square$

References

[1] Pierre Aboulker and Guillaume Aubian. Four proofs of the directed Brooks' Theorem. *arXiv preprint arXiv:2109.01600*, 2021.

[2] Marthe Bonamy, Matthew Johnson, Ioannis Lignos, Viresh Patel, and Daniël Paulusma. Reconfiguration graphs for vertex colourings of chordal and chordal bipartite graphs. *Journal of Combinatorial Optimization*, 27(1):132–143, 2014.

[3] Paul Bonsma and Luis Cereceda. Finding paths between graph colourings: PSPACE-completeness and superpolynomial distances. *Theoretical Computer Science*, 410(50):5215–5226, 2009. Mathematical Foundations of Computer Science (MFCS 2007).

[4] Nicolas Bousquet, Laurent Feuilloley, Marc Heinrich, and Mikaël Rabie. Short and local transformations between $(\Delta + 1)$-colorings. *arXiv preprint arXiv:2203.08885*, 2022.

[5] Nicolas Bousquet, Frédéric Havet, Nicolas Nisse, Lucas Picasarri-Arrieta, and Amadeus Reinald. Digraph redicolouring. *arXiv preprint arXiv:2301.03417*, 2023.

[6] Rowland Leonard Brooks. On colouring the nodes of a network. *Mathematical Proceedings of the Cambridge Philosophical Society*, 37(2):194–197, 1941.

[7] Luis Cereceda, Jan van den Heuvel, and Matthew Johnson. Connectedness of the graph of vertex-colourings. *Discrete Mathematics*, 308(5):913–919, 2008. Selected Papers from 20th British Combinatorial Conference.

[8] Luis Cereceda, Jan Van den Heuvel, and Matthew Johnson. Mixing 3-colourings in bipartite graphs. *European Journal of Combinatorics*, 30(7):1593–1606, 2009.

[9] Luis Cereceda, Jan van den Heuvel, and Matthew Johnson. Finding paths between 3-colorings. *Journal of Graph Theory*, 67(1):69–82, 2011.

[10] Carl Feghali, Matthew Johnson, and Daniël Paulusma. A reconfigurations analogue of Brooks' Theorem and its consequences. *Journal of Graph Theory*, 83(4):340–358, 2016.

[11] Ararat Harutyunyan and Bojan Mohar. Gallai's theorem for list coloring of digraphs. *SIAM Journal on Discrete Mathematics*, 25(1):170–180, 2011.

[12] Victor Neumann-Lara. The dichromatic number of a digraph. *J. Combin. Theory Ser. B.*, 33:265–270, 1982.

[13] Lucas Picasarri-Arrieta. Strengthening the Directed Brooks' Theorem for oriented graphs and consequences on digraph redicolouring. *arXiv preprint arXiv:2301.04881*, 2023.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

On asymptotic confirmation of the Faudree–Lehel Conjecture on the irregularity strength of graphs

(Extended abstract)

Jakub Przybyło* Fan Wei†

Abstract

We call a multigraph irregular if it has pairwise distinct vertex degrees. No nontrivial (simple) graph is thus irregular. The irregularity strength of a graph G, $s(G)$, is a specific measure of the "level of irregularity" of G. It might be defined as the least k such that one may obtain an irregular multigraph of G by multiplying any selected edges of G, each into at most k its copies. In other words, $s(G)$ is the least k admitting a $\{1, 2, \ldots, k\}$-weighting of the edges of G assuring distinct weighted degrees for all the vertices, where the weighted degree of a vertex is the sum of its incident weights. The most well-known open problem concerning this graph invariant is the conjecture posed in 1987 by Faudree and Lehel that there exists an absolute constant C such that $s(G) \leq \frac{n}{d} + C$ for each d-regular graph G with n vertices and $d \geq 2$, whereas a straightforward counting argument implies that $s(G) \geq \frac{n}{d} + \frac{d-1}{d}$. Until very recently this conjecture had remained widely open. We shall discuss recent results confirming it asymptotically, up to a lower order term. If time permits we shall also mention a few related problems, such as the 1–2–3 Conjecture or the concept of irregular subgraphs, introduced recently by Alon and Wei, and progress in research concerning these.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-106`

*AGH University, Faculty of Applied Mathematics, Al. A. Mickiewicza 30, 30-059 Krakow, Poland. E-mail: `jakubprz@agh.edu.pl`.

†Department of Mathematics, Duke University, Durham, NC 27710, USA. E-mail: `fw97@math.duke.edu`. Supported by NSF Award DMS-1953958.

1 Introduction

One of the most basic facts in graph theory is that every simple graph of order at least two contains a pair of vertices with equal degrees. Thus it cannot be irregular, where by *irregular* we mean: containing pairwise distinct vertex degrees. There are irregular multigraphs, though. In fact any (simple) graph $G = (V, E)$ can be turned into an irregular multigraph throughout multiplying some of its edges, as long as G has no isolated edge and at most one isolated vertex. The least k such that it is feasible using at most k copies of every edge is called the irregularity strength of G and denoted $s(G)$; we set $s(G) = \infty$ if this is not possible at all. Note that equivalently, $s(G)$ may be defined as the least positive integer k for which there is an edge k-weighting, that is a function $\omega : E \to \{1, 2, \dots, k\}$ such that each vertex $v \in V$ is attributed a distinct weighted degree $d_\omega(v) := \sum_{u \in N(v)} \omega(uv)$. This graph invariant was introduced in the 80s by Chartrand et al. [11] in relation to research on the concept of irregular graphs of Chartrand, Erdős and Oellermann [10]. In general it is known that $s(G) \le n-1$ for all graphs with n vertices for which the parameter is finite except K_3 [3, 31], and this upper bound is tight, e.g. for the family of stars. It can however be significantly decreased for graphs without small degree vertices. In particular, it is easy to verify that $s(G) \ge \frac{n}{d} + \frac{d-1}{d}$ for d-regular graphs, while the central open problem of this field is the following conjecture of Faudree and Lehel [17] from 1987 (posed first as a question by Jacobson, see [29]).

Conjecture 1. *There exists a constant C such that $s(G) \le \frac{n}{d} + C$ for every d-regular graph G with $d \ge 2$ and order n.*

This problem "energized the study of the irregularity strength", as Cuckler and Lazebnik stated in [12], and still remains open. A significant step forward towards solving it was achieved in 2002 by Frieze, Gould, Karoński, and Pfender [20], who used the probabilistic method to prove the first linear bound $s(G) \le 48(n/d) + 1$ for $d \le \sqrt{n}$, and a super-linear one $s(G) \le 240(\log n)(n/d) + 1$ in the remaining cases. They also proved similar bounds for general graphs, with d replaced by the minimum degree δ. For example, they showed that $s(G) = O(n/\delta)$ for the maximum degree $\Delta \le n^{1/2}$. The linear bounds in n/δ was further extended to the case when $d \ge 10^{4/3} n^{2/3} \log^{1/3} n$ and $\delta \ge 10 n^{3/4} \log^{1/4} n$, respectively, by Cuckler and Lazebnik [12]. The first general and unified linear bound in n/δ for the full spectrum of (n, δ) was delivered by Przybyło [34, 35], who used a constructive approach to prove that $s(G) \le 16(n/d) + 6$ and $s(G) \le 112(n/\delta) + 28$, respectively. Since then several attempts based on inventive new algorithms have been conducted in pursuit towards improvement of the multiplicative constant in front of n/δ, see e.g. [23, 24, 30]. The best result among these is due to Kalkowski, Karoński, and Pfender [24], who invented a deterministic algorithm implying that in general, $s(G) \le 6\lceil n/\delta \rceil$ for graphs with minimum degree $\delta \ge 1$ and without isolated edges. Conjecture 1 throughout more than 35 years since its formulation was an inspiration for many results, see e.g. [3, 5, 8, 12, 13, 15, 16, 17, 18, 20, 22, 24, 30, 31, 33, 34, 35], and various related problems and concepts, giving rise to a reach and vital branch of graph theory, see [21, 29] for surveys devoted to some of them. Only just recently it was proved by Przybyło [32] that the Faudree-Lehel Conjecture holds asymptotically almost surely for random graphs $G(n,p)$ (which are typically "close to" regular graphs), for any constant p, and holds asymptotically for a wide spectrum of values of d [33].

2 Main Results

Developing research from [33], we managed to confirm asymptotically, up to a lower order term, an extension of Conjecture 1 towards the setting of general graphs.

Theorem 2 ([38])**.** *For every $\varepsilon \in (0, 0.25)$, there are absolute constants C_1, C_2 such that for each graph G with n vertices and minimum degree $\delta > 0$ which does not contain isolated edges, $s(G) \leq \frac{n}{\delta}(1 + \frac{C_1}{\delta^\varepsilon}) + C_2$.*

We also confirmed that the generalization of Faudree-Lehel Conjecture holds, not only asymptotically, for relatively dense graphs.

Theorem 3 ([38])**.** *For every $0.8 < \alpha \leq 1$, there is an absolute constant C such that for each graph G with n vertices and minimum degree $\delta \geq n^\alpha$, $s(G) \leq \frac{n}{\delta} + C$.*

In the case of regular graphs exclusively, we also provided a much shorter argument, implying a more specific result directly related with Conjecture 1.

Theorem 4 ([39])**.** *Given any $\varepsilon \in (0, 0.25)$, for every d-regular graph G with n vertices, if d is sufficiently large in terms of ε, $s(G) < \frac{n}{d}(1 + \frac{14}{d^\varepsilon}) + 28$.*

Theorem 5 ([39])**.** *Given any $0.8 < \alpha \leq 1$, for every d-regular graph G on n vertices with $d \geq n^\alpha$, if d is sufficiently large in terms of α, then $s(G) < \frac{n}{d} + 28$.*

3 Main Ideas

3.1 General Graphs

A very vague general idea behind our construction yielding Theorems 2 and 3 is to randomly partition V into a big set B and a small set S, where $|S| = (n/\delta) \cdot o(\delta)$, in a special and controlled manner. We then first randomly modify the edge weights so that almost all vertices in B have distinct weighted degrees. Finally, we locally adjust weighted degrees of the rest of the vertices in order to differentiate them in entire G.

Our approach can be divided into three main steps.

Step 1 relies on a random construction assuring relatively sparse distribution of weighted degrees of the vertices in B, i.e. without too many vertex weights in any of the predefined intervals partitioning positive integers. A general, yet still imprecise idea here is to assign to every vertex v a random variable $X_v \sim U[0,1]$, and then attribute an edge uv a small weight if $X_u + X_v$ is small, and a large weight, otherwise. This way a small value of X_v

pulls the weighted degree of v downwards, while a large value of X_v pushes its weighted degree up.

Step 2 concentrates around modifications of weights of the edges between B and S, resulting in relatively small weights' shifts, attributing pairwise distinct weighted degrees to all but a small set of "bad vertices" in B. Note that in order to be able to achieve our goal, we must assure that the (randomly chosen) set S is large enough to guarantee sufficiently many edges between B and S.

In **Step 3** we modify mainly weights of the edges within S (and a small fraction of the edges outside S) in order to differentiate weighted degrees in S mostly. For this purpose we associate to these vertices special weighted degrees, which were earlier deliberately not used within step 2. While distinguishing weighted degrees in S we in particular benefit from the fact that S is small in comparison to B, and thus vertices in S have on average large fraction of their incident edges in $E(S, B)$ (statistically much larger than the fraction of edges in S). This allows taking on essential preparatory measures prior to step 3 (in step 1) assuring sparse weighted degrees' distribution within S and facilitating the mentioned final cleanup in this set. Throughout the construction we moreover specify several types of "bad vertices", which do not fulfill one of a list of certain conditions and cannot be distinguished according to major procedures. The aggregated set of these is however small enough to be taken care of in a special manner within step 3.

3.2 Regular Graphs

In order to provide much shorter proof of more specific results in the case of regular graphs, i.e. Theorems 4 and 5 (directly referring to Conjecture 1), we use in a way similar general 3-step approach, exploiting in particular random variables $X_v \sim U[0, 1]$ associated with vertices. We however phrase our construction differently, using quantization and the Lovász Local Lemma, which was redundant in the construction above. This time we may guarantee that weighted degrees of the vertices in the big set B are arranged very tightly, in fact these form a sequence of $|B|$ consecutive integers. We moreover again benefit from S being small compared to B, this time by assigning heavy weights between S and B, thus guaranteeing that weighted degrees of vertices in (the small set) S are all larger than those in B (as random choice of S and B results, with positive probability, in many edges joining vertices in S with those in B). Still particular preparatory measures need to be undertaken within our special initial random vertex and edge partitions, in order to facilitate later final weighted degrees distinction within S. We refer the reader to [38, 39] for more details of our randomized constructions.

4 Related Concepts

One of the most well known variants of the irregularity strength is its local correspondent, within which one confines to requiring distinct weighted degrees only for adjacent vertices. This concept was introduced in 2004 by Karoński, Łuczak and Thomason [26] together with an intriguing conjecture that just weights 1, 2 and 3 are sufficient for every graphs without isolated edges within such a setting. This so-called 1–2–3 Conjecture swiftly became yet another central problem of this field, and gained considerable attention, comparable to the Conjecture of Faudree and Lehel, cf. in particular [1, 2, 6, 7, 9, 14, 25, 26, 27, 28, 36, 37, 40, 41, 42, 43, 44]. In 2021 the 1–2–3 Conjecture was proven to hold for regular graphs with large enough degrees [36], while in 2022 also for general graphs with minimum degree $\delta = \Omega(\log \Delta)$ [37]. Lately Keusch [28] proved that actually weights $1, 2, 3, 4$ always suffice, whereas very recently the same author announced [27] to finally resolve the conjecture in the affirmative.

Also recently yet another related concept was proposed by Alon and Wei [4]. Roughly speaking they posed a conjecture that every graph contains a spanning subgraph which is (globally) almost as irregular as possible. More precisely they asked if any d-regular graph on n vertices contains a spanning subgraph in which the number of vertices of each degree between 0 and d deviates from $\frac{n}{d+1}$ by at most 2, and similarly, if every graph on n vertices, not necessarily regular, with minimum degree δ contains a spanning subgraph in which the number of vertices of each degree does not exceed $\frac{n}{\delta+1} + 2$. They also supported the conjectures by showing in particular that if $d^3 \log n \leq o(n)$ then every d-regular graph with n vertices contains a spanning subgraph in which the number of vertices of each degree between 0 and d is $(1 + o(1))\frac{n}{d+1}$, and a similar result for general graphs. Some of these results were also later significantly strengthened by Fox, Luo and Pham [19].

The mentioned problems are just the tip of the iceberg of related concepts. An extensive list of other related issues can in particular be found in Gallian's survey [21].

References

[1] L. Addario-Berry, K. Dalal, C. McDiarmid, B.A. Reed, A. Thomason, *Vertex-Colouring Edge-Weightings*, Combinatorica 27(1) (2007) 1–12.

[2] L. Addario-Berry, K. Dalal, B.A. Reed, *Degree Constrained Subgraphs*, Discrete Appl. Math. 156(7) (2008) 1168–1174.

[3] M. Aigner, E. Triesch, *Irregular assignments of trees and forests*, SIAM J. Discrete Math. 3(4) (1990) 439–449.

[4] N. Alon, F. Wei, *Irregular subgraphs*, Combinatorics, Probability and Computing 32(2) (2023) 269–283. doi:10.1017/S0963548322000220

[5] D. Amar, *Irregularity strength of regular graphs of large degree*, Discrete Math. 114 (1993) 9–17.

[6] T. Bartnicki, J. Grytczuk, S. Niwczyk, *Weight Choosability of Graphs*, J. Graph Theory 60(3) (2009) 242–256.

[7] J. Bensmail, *A 1-2-3-4 result for the 1-2-3 Conjecture in 5-regular graphs*, Discrete Appl. Math. 257 (2019) 31–39.

[8] T. Bohman, D. Kravitz, *On the irregularity strength of trees*, J. Graph Theory 45 (2004) 241–254.

[9] L. Cao, *Total weight choosability of graphs: Towards the 1–2–3-conjecture*, J. Combin. Theory Ser. B 149(1-2) 109–146

[10] G. Chartrand, P. Erdős, O.R. Oellermann, *How to Define an Irregular Graph*, College Math. J. 19(1) (1988) 36–42.

[11] G. Chartrand, M.S. Jacobson, J. Lehel, O.R. Oellermann, S. Ruiz, F. Saba, *Irregular networks*, Congr. Numer. 64 (1988) 197–210.

[12] B. Cuckler, F. Lazebnik, *Irregularity Strength of Dense Graphs*, J. Graph Theory 58(4) (2008) 299–313.

[13] J.H. Dinitz, D.K. Garnick, A. Gyárfás, *On the irregularity strength of the $m \times n$ grid*, J. Graph Theory 16 (1992) 355–374.

[14] A. Dudek, D. Wajc, *On the complexity of vertex-coloring edge-weightings*, Discrete Math. Theor. Comput. Sci. 13(3) (2011) 45–50.

[15] G. Ebert, J. Hemmeter, F. Lazebnik, A.J. Woldar, *On the irregularity strength of some graphs*, Congr. Numer. 71 (1990) 39–52.

[16] R.J. Faudree, M.S. Jacobson, J. Lehel, R. Schelp, *Irregular networks, regular graphs and integer matrices with distinct row and column sums*, Discrete Math. 76 (1989) 223–240.

[17] R.J. Faudree, J. Lehel, *Bound on the irregularity strength of regular graphs*, Colloq Math Soc Jańos Bolyai, 52, Combinatorics, Eger North Holland, Amsterdam, (1987), 247–256.

[18] M. Ferrara, R.J. Gould, M. Karoński, F. Pfender, *An iterative approach to graph irregularity strength*, Discrete Appl. Math. 158 (2010) 1189–1194.

[19] J. Fox, S. Luo, H.T. Pham, *On random irregular subgraphs*, arXiv:2207.13651.

[20] A. Frieze, R.J. Gould, M. Karoński, F. Pfender, *On Graph Irregularity Strength*, J. Graph Theory 41(2) (2002) 120–137.

[21] J.A. Gallian, *Graph Labeling*, Electron. J. Combin. (2019) 1–535, Dynamic survey DS6.

[22] A. Gyárfás, *The irregularity strength of $K_{m,m}$ is 4 for odd m*, Discrete Math. 71 (1998) 273–274.

[23] M. Kalkowski, *A note on 1,2-Conjecture*, in Ph.D. Thesis, Poznań, 2009.

[24] M. Kalkowski, M. Karoński, F. Pfender, *A new upper bound for the irregularity strength of graphs*, SIAM J. Discrete Math. 25 (2011) 1319–1321.

[25] M. Kalkowski, M. Karoński, F. Pfender, *Vertex-coloring edge-weightings: Towards the 1-2-3 conjecture*, J. Combin. Theory Ser. B 100 (2010) 347–349.

[26] M. Karoński, T. Łuczak, A. Thomason, *Edge weights and vertex colours*, J. Combin. Theory Ser. B 91 (2004) 151–157.

[27] R. Keusch, *A Solution to the 1–2–3 Conjecture*, arXiv:2303.02611.

[28] R. Keusch, *Vertex-coloring graphs with 4-edge-weightings*, Combinatorica, to appear.

[29] J. Lehel, *Facts and quests on degree irregular assignments*, Graph Theory, Combinatorics and Applications, Willey, New York, 1991, 765–782.

[30] P. Majerski, J. Przybyło, *On the irregularity strength of dense graphs*, SIAM J. Discrete Math. 28(1) (2014) 197–205.

[31] T. Nierhoff, *A tight bound on the irregularity strength of graphs*, SIAM J. Discrete Math. 13(3) (2000) 313–323.

[32] J. Przybyło, *A generalization of Faudree-Lehel Conjecture holds almost surely for random graphs*, Random Struct Alg. 61 (2022) 383–396.

[33] J. Przybyło, *Asymptotic confirmation of the Faudree-Lehel Conjecture on irregularity strength for all but extreme degrees*, J. Graph Theory 100 (2022) 189–204.

[34] J. Przybyło, *Irregularity strength of regular graphs*, Electron. J. Combin. 15(1) (2008) ♯R82.

[35] J. Przybyło, *Linear bound on the irregularity strength and the total vertex irregularity strength of graphs*, SIAM J. Discrete Math. 23(1) (2009) 511–516.

[36] J. Przybyło, *The 1–2–3 Conjecture almost holds for regular graphs*, J. Combin. Theory Ser. B 147 (2021) 183–200.

[37] J. Przybyło, *The 1–2–3 Conjecture holds for graphs with large enough minimum degree*, Combinatorica 42 (2022) 1487–1512.

[38] J. Przybyło, F. Wei, *On the asymptotic confirmation of the Faudree-Lehel Conjecture for general graphs*, Combinatorica (2023), https://doi.org/10.1007/s00493-023-00036-5.

[39] J. Przybyło, F. Wei, *Short proof of the asymptotic confirmation of the Faudree-Lehel Conjecture*, arXiv:2109.13095.

[40] C. Thomassen, Y. Wu, C.Q. Zhang, *The 3-flow conjecture, factors modulo k, and the 1-2-3 conjecture*, J. Combin. Theory Ser. B 121 (2016) 308–325.

[41] T. Wang, Q. Yu, *On vertex-coloring 13-edge-weighting*, Front. Math. China 3(4) (2008) 581–587.

[42] T. Wong, X. Zhu, *Every graph is (2,3)-choosable*, Combinatorica 36(1) (2016) 121–127.

[43] L. Zhong, *The 1–2–3–conjecture holds for dense graphs*, J. Graph Theory 90 (2019) 561–564.

[44] X. Zhu, *Every nice graph is (1,5)-choosable*, J. Combin. Theory Ser. B 157 (2022) 524–551.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

On Hypergraph Supports

(Extended abstract)

Rajiv Raman* Karamjeet Singh†

Abstract

Let $\mathcal{H} = (X, \mathcal{E})$ be a hypergraph. A support is a graph Q on X such that for each $E \in \mathcal{E}$, the subgraph of Q on the elements in E is connected. We consider hypergraphs defined on a host graph. Given a graph $G = (V, E)$, with $c : V \to \{\mathbf{r}, \mathbf{b}\}$ and a collection of connected subgraphs $\mathcal{H}$ of G, a primal support is a graph Q on $\mathbf{b}(V)$ such that for each $H \in \mathcal{H}$, the subgraph $Q[\mathbf{b}(H)]$ on vertices $\mathbf{b}(H) = H \cap c^{-1}(\mathbf{b})$ is connected. A *dual support* is a graph Q^* on $\mathcal{H}$ s.t. for each $v \in X$, the subgraph $Q^*[\mathcal{H}_v]$ is connected, where $\mathcal{H}_v = \{H \in \mathcal{H} : v \in H\}$. We present sufficient conditions on the host graph and hyperedges so that the resulting support comes from a restricted family.

We primarily study two classes of graphs: (1) If the host graph has genus g and the hypergraphs satisfy a topological condition of being *cross-free*, then there is a primal and a dual support of genus at most g. (2) If the host graph has treewidth t and the hyperedges satisfy a combinatorial condition of being *non-piercing*, then there exist primal and dual supports of treewidth $O(2^t)$. We show that this exponential blow-up is sometimes necessary. As an intermediate case, we also study the case when the host graph is outerplanar. Finally, we show applications of our results to packing and covering, and coloring problems on geometric hypergraphs.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-107`

*IIIT-Delhi, 110020 New Delhi, India. E-mail: `rajiv@iiitd.ac.in` Part of this work was done when the first author was at LIMOS, Université Clermont Auvergne, and was partially supported by the French government research program "Investissements d'Avenir" through the IDEX-ISITE initiative 16-IDEX-0001 (CAP 20-25).

†IIIT-Delhi, 110020 New Delhi, India. E-mail: `karamjeets@iiitd.ac.in`

1 Introduction

A hypergraph $(X, \mathcal{E})$ is defined by a set X of elements and a collection $\mathcal{E}$ of subsets of X. A *support* is a graph Q on X s.t. $\forall\ E \in \mathcal{E}$, the subgraph $Q[E]$ induced by the vertices of E is connected. The notion of a support was introduced by Voloshina and Feinberg [28] in the context of VLSI circuits. Since then, this notion has found wide applicability in several areas, such as visualizing hypergraphs [7, 8, 9, 10, 11, 16, 18], in the design of networks [2, 4, 5, 12, 17, 22, 25], and similar notions have been used in the analysis of local search algorithms for geometric problems [3, 6, 13, 23, 24, 27].

Any hypergraph clearly has a support: a complete graph on X is a support. The problem becomes interesting if we introduce a global constraint on the graph that is in *tension* with the *local* connectivity requirement for each hyperedge. In particular, we are interested in restrictions on the hypergraph that guarantees the existence of a support from a *sparse* family of graphs, namely a family with sub-linear sized separators[1]. A support from a family $\mathcal{G}$ of graphs is called a $\mathcal{G}$ support.

Our motivation to study the existence of such supports comes primarily from the analysis of *local search* algorithms for several packing and covering problems defined by geometric objects in the Euclidean plane. With the aim of extending the analysis techniques from the plane to other surfaces, we study hypergraphs defined on a sparse *host graph*.

A *geometric hypergraph* is defined by a set P of points in $\mathbb{R}^2$, and a set $\mathcal{D}$ of *regions*, or subsets of $\mathbb{R}^2$, where the hyperedges are defined by $D \cap P$ for each $D \in \mathcal{D}$. We call this hypergraph a *primal hypergraph* to distinguish it from other hypergraphs we will define shortly.

If $\mathcal{D}$ is a collection of *pseudodisks*,[2] Pyrga and Ray [26] showed that the hypergraph $(P, \mathcal{D})$ admits a planar support. Raman and Ray in [27] generalized this result to show that if the regions in a geometric hypergraph are *non-piercing*[3], then the hypergraph $(P, \mathcal{D})$ admits a planar support. The authors also show that the *dual hypergraph* $(\mathcal{D}, \{\mathcal{D}_p\}_{p\in P})$, where for each $p \in P$, $\mathcal{D}_p = \{D \in \mathcal{D} : D \ni p\}$ admits a planar support.

For an arrangement $\mathcal{D}$ of pseudodisks in the plane in general position[4], let G denote the *dual arrangement graph*, whose vertices are the cells in the arrangement, and two cells are adjacent if they share an arc of the boundary of a pseudodisk. It is easy to see that the dual arrangement graph G is a plane graph, and each pseudodisk D corresponds to a connected subgraph H_D of G. Further, for any pair of subgraphs $H_D, H_{D'}$ corresponding to

[1] A graph has a sublinear sized separator if there is some constant $\epsilon > 0$, and $c > 0$ such that there is a set S of size $O(|V|^{1-\epsilon})$ such that $G \setminus S$ can be partitioned into two subgraphs A and B s.t. there is no edge in $G \setminus S$ between a vertex in A and a vertex in B, and s.t. $|V(A)|, |V(B)| \leq c|V(G)|$.

[2] A collection of simple Jordan curves define a set of pseudocircles if each pair intersects either 0 or twice. A collection of bounded regions whose boundaries are a collection pseudocircles are a collection of pseudodisks.

[3] a collection of regions $\mathcal{H}$, where each $H \in \mathcal{H}$ is a path-connected region bounded by a simple Jordan curve (possibly with holes) is non-piercing if both $H \setminus H'$ and $H' \setminus H$ are connected.

[4] An arrangement $\mathcal{D}$ of pseudodisks in the plane is in general position if there are no three pseudodisks whose boundaries pass through a common point, and at each intersection point, the boundaries of the pair of pseudodisks defining the intersection point properly cross.

pseudodisks D and D', respectively, the graphs $G[H_D \setminus H_{D'}]$ and $G[H_{D'} \setminus H_D]$ are connected, i.e., $\mathcal{H} = \{H_D : D \in \mathcal{D}\}$ is a collection of *non-piercing* subgraphs of G. This motivates the following definition:

Definition 1 (Non-piercing)**.** *For a graph G, a collection of connected subgraphs $\mathcal{H}$ of G is non-piercing if for any two subgraphs $H, H' \in \mathcal{H}$, both $G[H \setminus H']$ and $G[H' \setminus H]$ are connected.*

For each cell c in the arrangement $\mathcal{D}$, let v_c denote the vertex in the dual arrangement graph G corresponding to this cell. If $c \cap P \neq \emptyset$, set $color(v_c) = \mathbf{b}$, and if $c \cap P = \emptyset$, set $color(v_c) = \mathbf{r}$. Let $\mathbf{b}(V)$ and $\mathbf{r}(V)$ denote respectively, the subsets of vertices in $color^{-1}(\mathbf{b})$ and $color^{-1}(\mathbf{r})$. The result of Pyrga and Ray [26] translated in this context says that for a plane graph G with $color : V \to \{\mathbf{r}, \mathbf{b}\}$, and a collection of connected subgraphs $\mathcal{H}$ s.t. $\mathcal{H}$ is non-piercing, there is a planar support on $color^{-1}(\mathbf{b})$, i.e., a planar graph Q on the vertices $color^{-1}(\mathbf{b})$ s.t. the subgraph $Q[H]$ is connected for each $H \in \mathcal{H}$.

Our goal is to generalize the results above to the setting where the host graph comes from a non-trivial graph class. If the host graph G has bounded genus and $\mathcal{H}$ is a collection of connected non-piercing subgraphs of G, we may expect, as in the case of pseudodisk hypergraphs above, that there exist bounded genus supports for the primal and dual hypergraphs. While the statement may be true for primal hypergraphs. However, for the dual hypergraph, this is not true: Let G be the torus grid graph $T_{n,n} = C_n \square C_n$[29]. The subgraphs are the n non-contractible cycles perpendicular to the hole, and the n non-contractible cycles parallel to the hole. Each vertex of the graph belongs to exactly two subgraphs forcing them to be adjacent in the dual support, and thus the dual support is $K_{n,n}$ which is not embeddable on the torus for large enough n.

For bounded genus graphs, we show that if the subgraphs satisfy a condition of being *cross-free*, then there exists bounded genus supports for the primal as well as the dual hypergraphs. In the plane, the cross-free condition is weaker than the non-piercing condition, but these two conditions are incomparable on higher genus surfaces.

If we restrict attention to host graphs of bounded treewidth, we show that if the subgraphs are non-piercing then both the primal and dual supports have bounded treewidth. However, the treewidth of the support could be exponentially larger than the treewidth of the host graph. Along the way, we also consider outerplanar graphs, which have treewidth 2. Here, we show a distinction between the primal and dual settings. For the primal setting, the cross-free condition on the hypergraphs is sufficient to obtain an outerplanar support, while in the dual setting, we show that restricting the subgraphs to be non-piercing is a sufficient condition to obtain a dual outerplanar support.

2 Preliminaries

Let $\mathcal{H}$ be a collection of connected subgraphs of a graph $G = (V, E)$. This defines a hypergraph $(V, \mathcal{H})$. We call the pair $(G, \mathcal{H})$ a graph system. If G comes from a class $\mathcal{G}$ of graphs, and $\mathcal{H}$ satisfies property P we say that $(G, \mathcal{H})$ is a P-$\mathcal{G}$ system. Further, if G has genus g, we say that $(G, \mathcal{H})$ has genus g. In particular, if G is planar, we say that $(G, \mathcal{H})$ is a planar system. Let $c : V \to \{\mathbf{r}, \mathbf{b}\}$ be a coloring of V with two colors. Let $\mathbf{b}(V)$ and $\mathbf{r}(V)$ denote respectively $c^{-1}(\mathbf{b})$ and $c^{-1}(\mathbf{r})$.

For a graph system $(G, \mathcal{H})$, a *primal support* is a graph Q on $\mathbf{b}(V)$ s.t. $\forall\ H \in \mathcal{H}$, $Q[\mathbf{b}(H)]$ is connected[5], i.e., a support for the *primal hypergraph* $(V, \mathcal{H})$. A *dual support* is a graph Q^* on $\mathcal{H}$ s.t. $\forall\ v \in V$, $Q^*[\mathcal{H}_v]$ is connected[6], where $\mathcal{H}_v = \{H \in \mathcal{H} : H \ni v\}$, i.e., a support for the *dual hypergraph* $(\mathcal{H}, \{\mathcal{H}_v\}_{v \in V(G)})$. For a graph G and two families of connected subgraphs $\mathcal{H}$ and $\mathcal{K}$ of G, let the 3-tuple $(G, \mathcal{H}, \mathcal{K})$ denote an *intersection system*. An *intersection support* is a graph $\tilde{Q}$ that is a support for the *intersection hypergraph* $(\mathcal{H}, \{\mathcal{H}_K\}_{K \in \mathcal{K}})$, where $\mathcal{H}_K = \{H \in \mathcal{H} : K \cap H \neq \emptyset\}$. The notion of an intersection hypergraph generalizes both the primal and dual hypergraphs defined above.

3 Bounded genus graphs

Let $(G, \mathcal{H})$ be a graph system of genus g. Consider a cellular[7] embedding of G in an oriented surface of genus g. For a pair of subgraphs $H, H' \in \mathcal{H}$, we define the notion of a *reduced graph* that is required for the definition of a *cross-free* system.

Definition 2 (Reduced graph)**.** *Let $(G, \mathcal{H})$ be a graph system with G cellularly embedded in an oriented surface. For any two subgraphs $H, H' \in \mathcal{H}$, the* reduced graph $R(H, H')$ *is the embedded graph obtained from G by contracting all edges, both of whose end-points lie in $H \cap H'$, where multi-edges and self-loops are retained.*

Note that if G can be embedded in a surface Σ, then so can be $R(H, H')$.

Definition 3 (Cross-free at v)**.** *A graph system $(G, \mathcal{H})$ with G cellularly embedded in an oriented surface, is cross-free at a vertex $v \in V(G)$ if for any two subgraphs $H, H' \in \mathcal{H}_v$, the following holds: Let $\tilde{v}$ be the image of v in the reduced graph $R(H, H')$. Then, there are no 4 edges $e_i = \{\tilde{v}, v_i\}$ in $R(H, H')$, $i = 1, \ldots, 4$ incident to $\tilde{v}$ in cyclic order around $\tilde{v}$, s.t. $v_1, v_3 \in H \setminus H'$, and $v_2, v_4 \in H' \setminus H$.*

If there is an embedding of G s.t. $(G, \mathcal{H})$ is cross-free at every vertex of G, we say that $(G, \mathcal{H})$ is cross-free.

By the Jordan curve theorem, it follows that if $(G, \mathcal{H})$ is a non-piercing planar system, then the graph system $(G, \mathcal{H})$ is cross-free. It is easy to construct examples to show that the reverse direction does not hold in the plane.

[5]Note that we cannot simply project each H on $\mathbf{b}(V)$ as the resulting subgraphs may not be connected in G.

[6]To make the definition symmetric, we could have considered a coloring $c : \mathcal{H} \to \{\mathbf{r}, \mathbf{b}\}$, and required that $Q^*[\mathcal{H}^{\mathbf{b}_v}]$ be connected for each $v \in V$, where $\mathcal{H}^{\mathbf{b}_v} = \{H \in \mathcal{H} : H \ni v \text{ and } c(H) = \mathbf{b}\}$. However, this problem reduces to constructing a dual support restricted to the hypergraphs $\mathcal{H}^{\mathbf{b}} = \{H \in \mathcal{H} : c(H) = \mathbf{b}\}$. Therefore, in the dual setting, it is sufficient to study the uncolored version of the problem.

[7]A cellular, or 2-cell embedding of a graph G on a surface is an embedding where the edges are non-crossing, and each face is homeomorphic to a disk.

Theorem 4. *Let $(G,\mathcal{H})$ be a cross-free system of genus g, with $c : V \to \{\mathbf{r},\mathbf{b}\}$. Then, there is a support Q of genus at most g on $\mathbf{b}(V)$ i.e., $Q[\mathbf{b}(H)]$ is connected for each $H \in \mathcal{H}$.*

Theorem 5. *Let $(G,\mathcal{H})$ be a cross-free system of genus g, then, there is a support Q^* on $\mathcal{H}$ of genus at most g i.e., $Q^*[\mathcal{H}_v]$ is connected for each $v \in V$.*

Theorem 6. *Let $(G,\mathcal{H},\mathcal{K})$ be a cross-free intersection system of genus g. Then, there exists an intersection support $\tilde{Q}$ on $\mathcal{H}$ of genus at most g.*

In all the results above, we use the notion of *Vertex Bypassing* defined below:

Definition 7. *Let G be embedded in an oriented surface Σ. Let $N(v) = (v_1,\dots,v_k,v_1)$ be the cyclic order of vertices around v. The Vertex Bypassing of v is defined as follows:*

1. *Subdivide each edge $\{v,v_i\}$ by a vertex u_i. Construct a cycle $C = (u_1,\dots,u_k,u_1)$ by joining consecutive vertices u_i,u_{i+1} (with indices taken $\mod k$) by a simple arc not intersecting the edges of G s.t. the resulting graph remains embedded in Σ. Remove the vertex v. Let G'' denote the resulting graph.*

2. *$\forall\ H \in \mathcal{H}_v$, s.t. $\{v,v_i\} \in H$, let H'' denote the subgraph of G'' on $(H \setminus \{v\}) \cup \{\cup_{v_i \in H''} u_i\}$ Let $\mathcal{H}''_v = \{H'' : H \in \mathcal{H}_v\}$. Let $\mathcal{H}'' = (\mathcal{H} \setminus \mathcal{H}_v) \cup \mathcal{H}''_v$ (Note that the subgraphs in $\mathcal{H}'_v$ may not be connected).*

3. *Add a set D of non-intersecting chords[8] in C so that $\forall\ H \in \mathcal{H}''$, H induces a connected subgraph in $C \cup D$, and the resulting subgraphs remain cross-free.*

Let $(G',\mathcal{H}')$ be the resulting system.

The heart of the proof is in showing that Step 3 can be done, i.e., there exists a set of non-intersecting chords that we can add in C so that the resulting subgraphs are connected, and the system remains cross-free. Assuming we can apply vertex bypassing, the proof of Theorem 5 follows by repeatedly applying vertex bypassing to a vertex of *maximum depth* in G, i.e., to a vertex v in G maximizing $|\{H \in \mathcal{H} : H \ni v\}|$, until each vertex of the graph is in at most one subgraph. We can then obtain a support by contracting the edges in each subgraph. The proof of Theorem 6 follows by using Theorem 5 and techniques from the proof of Theorem 4.

4 Bounded Treewidth graphs

We show that if $(G,\mathcal{H})$ is a graph system, and $\mathcal{H}$ is a collection of non-piercing subgraphs then both the primal and dual supports have treewidth $O(2^{tw(G)})$ and this exponential blow-up in the treewidth of the support is sometimes necessary.

[8] We use the term non-intersecting to mean internally non-intersecting.

Theorem 8. *Let $(G, \mathcal{H})$ be a non-piercing graph system. Let $c : V(G) \to \{\mathbf{r}, \mathbf{b}\}$ be a 2-coloring of the vertices $V(G)$ of G. Then, there is a support Q on $\mathbf{b}(V)$ s.t. $tw(Q) \leq 3 \cdot 2^{tw(G)}$. Further, Q can be computed in time polynomial in $|G|, |H|$ if G has bounded treewidth. There exist non-piercing graph systems $(G, \mathcal{H})$ where any support has size $\Omega(2^{tw(G)})$.*

Theorem 9. *Let $(G, \mathcal{H})$ be a non-piercing graph system. There is a dual support Q^* on $\mathcal{H}$ s.t. $tw(Q^*) \leq 4 \cdot 2^{tw(G)}$. Further, Q^* can be computed in time polynomial in $|G|, |\mathcal{H}|$ if G has bounded treewidth. There exist non-piercing graph systems $(G, \mathcal{H})$ where any dual support has size $\Omega(2^{tw(G)})$.*

5 Outerplanar Graphs

Let $(G, \mathcal{H})$ be an outerplanar graph system. In the setting of outerplanar graphs, there is a difference between the primal and dual settings. In the primal setting, if the subgraphs are cross-free, then there is a primal support that is also outerplanar. In the dual setting, the cross-free condition is not sufficient. We show an example below. However, restricting the subgraphs to be non-piercing is sufficient for the system to admit a dual outerplanar support.

Consider a triangle drawn in the plane (with straight-line segments) with vertices $\{1, 2, 3\}$. Subdivide the segments $\{1, 2\}, \{2, 3\}$ and $\{1, 3\}$ by points 4, 5, and 6 respectively. Add a triangle on the points $4, 5$ and 6. This defines an embedding of an asteroid-triple G. The subgraphs $\mathcal{H}$ are those by the points $\{1, 4, 2\}, \{2, 5, 3\}, \{1, 6, 3\}$ and $\{4, 5, 6\}$. It is easy to check that $(G, \mathcal{H})$ is cross free, and the support for the dual is K_4 which is not outerplanar.

Theorem 10. *Let $(G, \mathcal{H})$ be an outerplanar cross-free system, with $c : V(G) \to \{\mathbf{r}, \mathbf{b}\}$, a 2-coloring of the vertices $V(G)$ of G. Then, there is an outerplanar support Q on $\mathbf{b}(V)$ i.e., $Q[\mathbf{b}(H)]$ is connected for each $H \in \mathcal{H}$.*

Theorem 11. *Let $(G, \mathcal{H})$ be a non-piercing outerplanar system. Then, there is an outerplanar dual support Q^* on $\mathcal{H}$.*

6 Applications

In this section, we describe some applications of the existence of supports. Raman and Ray [27] showed that for an intersection hypergraph defined on a set of non-piercing regions in the plane, there is a planar support (See [27] for precise definitions), which implies a support for both the primal and dual settings for the hypergraphs defined by points and non-piercing regions in the plane.

Since graphs of genus g admit separators of size $O(\sqrt{gn})$ [15], all the algorithmic consequence of [27] generalize to cross-free systems on bounded genus graphs. Instead of describing a long sequence of results that follow from the existence of supports, we highlight just three results that follow as a consequence of Theorem 6.

Theorem 12. *Let $(G, \mathcal{H})$ be a cross-free system of genus g, then there exists*

1. *a PTAS for the Dominating Set problem, i.e., find $\mathcal{H}' \subseteq \mathcal{H}$ of minimum cardinality s.t. for each $H \in \mathcal{H}$, either $H \in \mathcal{H}'$ or $H \cap H' \neq \emptyset$ for some $H' \in \mathcal{H}'$.*

2. *a PTAS for the problem of packing points when each subgraph $H \in \mathcal{H}$ has capacity D_H bounded by a constant, i.e., find $V' \subseteq V$ of maximum cardinality s.t. $|H \cap V'| \leq D_H$ for each $H \in \mathcal{H}$.*

3. *a PTAS for the problem of packing subgraphs when each vertex $v \in V$ has capacity D_v bounded by a constant, i.e., find $\mathcal{H}' \subseteq \mathcal{H}$ of maximum cardinality s.t. $|\{H \in \mathcal{H}' : H \ni v\}| \leq D_v$ for each $v \in V$.*

Keller and Smorodinsky [19] showed that the intersection hypergraph of disks in the plane can be colored with 4 colors, and this was generalized by Keszegh [20] for pseudodisks, which was further generalized in [27] to show that the intersection hypergraph of non-piercing regions is 4-colorable. As a consequence of Theorem 6, we obtain the following.

Theorem 13. *Let $(G, \mathcal{H}, \mathcal{K})$ be a cross-free intersection system of genus g. Then, $\mathcal{H}$ can be colored with at most $\frac{7+\sqrt{1+24g}}{2}$ colors such that no hyperedge $\mathcal{H}_K$ is monochromatic.*

Proof. By Theorem 6, $(G, \mathcal{H}, \mathcal{K})$ has a support $\tilde{Q}$ of genus at most g. Now, $\chi(\tilde{Q}) \leq \frac{7+\sqrt{1+24g}}{2}$ [14]. Since $\tilde{Q}$ is a support, for each $K \in \mathcal{K}$, there is an edge between some two subgraphs $H, H' \in \mathcal{H}_K$. Therefore, no hyperedge $\mathcal{H}_K$ is monochromatic. $\square$

Keszegh and Pàlvölgyi [21] introduced the notion of $ABAB$-free hypergraphs. Ackerman et al., [1] show that these are equivalent to hypergraphs with a *stabbed pseudo-disk representation*, i.e., each $S \in \mathcal{S}$ is mapped to a closed and bounded region D_S containing the origin whose boundary is a simple Jordan curve, each $x \in X$ is mapped to a point p_x in $\mathbb{R}^2$ s.t. $p_x \in D_S$ iff $x \in S$. The regions $\mathcal{D} = \{D_S : S \in \mathcal{S}\}$ form a stabbed pseudodisk arrangement. Let $(P, \mathcal{D})$ denote the embedding of the hypergraph where $P = \{p_x : x \in X\}$.

The authors show that to any stabbed pseudodisk arrangement $\mathcal{D}$ and a set P of points, we can add additional pseudodisks $\mathcal{D}'$ s.t. (i) each $D' \in \mathcal{D}'$ contains exactly 2 points of P, (ii) $\mathcal{D} \cup \mathcal{D}'$ is a pseudodisk arrangement, and (iii) Each $D \in \mathcal{D}$ s.t. $|D \cap P| \geq 3$ contains a pseudodisk $D' \in \mathcal{D}'$. The graph on P whose edges are defined by $\mathcal{D}'$ is called the *delaunay graph* of the arrangement. They show that $ABAB$-free hypergaphs are 3 colorable by showing that the delaunay graph is outerplanar. This result follows from Theorem 10 since a support for cross-free outerplanar graph system satisfies the properties of delaunay graph above.

References

[1] Eyal Ackerman, Balázs Keszegh, and Dömötör Pálvölgyi. Coloring hypergraphs defined by stabbed pseudo-disks and abab-free hypergraphs. *SIAM Journal on Discrete Mathematics*, 34(4):2250–2269, 2020.

[2] Emmanuelle Anceaume, Maria Gradinariu, Ajoy Kumar Datta, Gwendal Simon, and Antonino Virgillito. A semantic overlay for self-peer-to-peer publish/subscribe. In *26th IEEE International Conference on Distributed Computing Systems (ICDCS'06)*, pages 22–22. IEEE, 2006.

[3] Daniel Antunes, Claire Mathieu, and Nabil H. Mustafa. Combinatorics of local search: An optimal 4-local hall's theorem for planar graphs. In *25th Annual European Symposium on Algorithms, ESA 2017, September 4-6, 2017, Vienna, Austria*, pages 8:1–8:13, 2017.

[4] Roberto Baldoni, Roberto Beraldi, Vivien Quema, Leonardo Querzoni, and Sara Tucci-Piergiovanni. Tera: topic-based event routing for peer-to-peer architectures. In *Proceedings of the 2007 inaugural international conference on Distributed event-based systems*, pages 2–13, 2007.

[5] Roberto Baldoni, Roberto Beraldi, Leonardo Querzoni, and Antonino Virgillito. Efficient publish/subscribe through a self-organizing broker overlay and its application to siena. *The Computer Journal*, 50(4):444–459, 2007.

[6] Aniket Basu Roy, Sathish Govindarajan, Rajiv Raman, and Saurabh Ray. Packing and covering with non-piercing regions. *Discrete & Computational Geometry*, 2018.

[7] Sergey Bereg, Krzysztof Fleszar, Philipp Kindermann, Sergey Pupyrev, Joachim Spoerhase, and Alexander Wolff. Colored non-crossing euclidean steiner forest. In *Algorithms and Computation: 26th International Symposium, ISAAC 2015, Nagoya, Japan, December 9-11, 2015, Proceedings*, pages 429–441. Springer, 2015.

[8] Sergey Bereg, Minghui Jiang, Boting Yang, and Binhai Zhu. On the red/blue spanning tree problem. *Theoretical computer science*, 412(23):2459–2467, 2011.

[9] Ulrik Brandes, Sabine Cornelsen, Barbara Pampel, and Arnaud Sallaberry. Blocks of hypergraphs: applied to hypergraphs and outerplanarity. In *Combinatorial Algorithms: 21st International Workshop, IWOCA 2010, London, UK, July 26-28, 2010, Revised Selected Papers 21*, pages 201–211. Springer, 2011.

[10] Ulrik Brandes, Sabine Cornelsen, Barbara Pampel, and Arnaud Sallaberry. Path-based supports for hypergraphs. *Journal of Discrete Algorithms*, 14:248–261, 2012.

[11] Kevin Buchin, Marc J van Kreveld, Henk Meijer, Bettina Speckmann, and KAB Verbeek. On planar supports for hypergraphs. *Journal of Graph Algorithms and Applications*, 15(4):533–549, 2011.

[12] Raphaël Chand and Pascal Felber. Semantic peer-to-peer overlays for publish/subscribe networks. In *Euro-Par 2005 Parallel Processing: 11th International Euro-Par Conference, Lisbon, Portugal, August 30-September 2, 2005. Proceedings 11*, pages 1194–1204. Springer, 2005.

[13] Vincent Cohen-Addad and Claire Mathieu. Effectiveness of local search for geometric optimization. In *Proceedings of the Thirty-first International Symposium on Computational Geometry*, SoCG '15, pages 329–343, Dagstuhl, Germany, 2015. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[14] Reinhard Diestel. Graph theory 3rd ed. *Graduate texts in mathematics*, 173(33):12, 2005.

[15] John R Gilbert, Joan P Hutchinson, and Robert Endre Tarjan. A separator theorem for graphs of bounded genus. *Journal of Algorithms*, 5(3):391–407, 1984.

[16] Frédéric Havet, Dorian Mazauric, Viet-Ha Nguyen, and Rémi Watrigant. Overlaying a hypergraph with a graph with bounded maximum degree. *Discrete Applied Mathematics*, 319:394–406, 2022.

[17] Jun Hosoda, Juraj Hromkovič, Taisuke Izumi, Hirotaka Ono, Monika Steinová, and Koichi Wada. On the approximability and hardness of minimum topic connected overlay and its special instances. *Theoretical Computer Science*, 429:144–154, 2012.

[18] Ferran Hurtado, Matias Korman, Marc van Kreveld, Maarten Löffler, Vera Sacristán, Akiyoshi Shioura, Rodrigo I Silveira, Bettina Speckmann, and Takeshi Tokuyama. Colored spanning graphs for set visualization. *Computational Geometry*, 68:262–276, 2018.

[19] Chaya Keller and Shakhar Smorodinsky. Conflict-free coloring of intersection graphs of geometric objects. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 2397–2411, 2018.

[20] Balázs Keszegh. Coloring intersection hypergraphs of pseudo-disks. *Discret. Comput. Geom.*, 64(3):942–964, 2020.

[21] Balázs Keszegh and Dömötör Pálvölgyi. An abstract approach to polychromatic coloring: shallow hitting sets in aba-free hypergraphs and pseudohalfplanes. *Journal of Computational Geometry*, 10(1):1–26, 2019.

[22] Ephraim Korach and Michal Stern. The clustering matroid and the optimal clustering tree. *Mathematical Programming*, 98:385–414, 2003.

[23] Erik Krohn, Matt Gibson, Gaurav Kanade, and Kasturi Varadarajan. Guarding terrains via local search. *Journal of Computational Geometry*, 5(1):168–178, 2014.

[24] Nabil H Mustafa and Saurabh Ray. Improved results on geometric hitting set problems. *Discrete & Computational Geometry*, 44(4):883–895, 2010.

[25] Melih Onus and Andréa W Richa. Minimum maximum-degree publish–subscribe overlay network design. *IEEE/ACM Transactions on Networking*, 19(5):1331–1343, 2011.

[26] Evangelia Pyrga and Saurabh Ray. New existence proofs for ϵ-nets. In *Proceedings of the Twenty-fourth Annual Symposium on Computational Geometry*, SoCG '08, pages 199–207, New York, NY, USA, 2008. ACM.

[27] Rajiv Raman and Saurabh Ray. Constructing planar support for non-piercing regions. *Discrete & Computational Geometry*, 64(3):1098–1122, 2020.

[28] AA Voloshina and VZ Feinberg. Planarity of hypergraphs. In *DOKLADY AKADEMII NAUK BELARUSI*, volume 28, pages 309–311. ACADEMII NAUK BELARUSI F SCORINA PR 66, ROOM 403, MINSK, BYELARUS 220072, 1984.

[29] Eric W Weisstein. Torus grid graph. *https://mathworld. wolfram. com/*, 2016.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

The Rado Multiplicity Problem in Vector Spaces over Finite Fields

(Extended abstract)

Juanjo Rué* Christoph Spiegel†

Abstract

We study an analogue of the Ramsey multiplicity problem for additive structures, establishing the minimum number of monochromatic 3-APs in 3-colorings of $\mathbb{F}_3^n$ and obtaining the first non-trivial lower bound for the minimum number of monochromatic 4-APs in 2-colorings of $\mathbb{F}_5^n$. The former parallels results by Cumings et al. [3] in extremal graph theory and the latter improves upon results of Saad and Wolf [13]. Lower bounds are notably obtained by extending the flag algebra calculus of Razborov [11].

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-108

1 Introduction

In 1959 Goodman [7] proved that asymptotically at least a quarter of all vertex triples in any graph must either form a clique or an independent set. This lead to the study of the *Ramsey multiplicity problem*, where one would like to determine the minimum number of monochromatic cliques of prescribed size over any edge-coloring of the complete graph [5,

*Departament de Matemàtiques and Institut de Matemàtiques (IMTech) de la Universitat Politècnica de Catalunya (UPC), and Centre de Recerca Matemàtica (CRM), Barcelona, Spain. Spain. E-mail: juan.jose.rue@upc.edu This work was partially funded by the grants MTM2017-82166-P, PID2020-113082GBI00, and the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence in R&D (CEX2020-001084-M).

†Zuse Institute Berlin, Department AIS2T, Germany. E-mail: spiegel@zib.de. This work was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – The Berlin Mathematics Research Center MATH+ (EXC-2046/1, project ID: 390685689).

15, 2]. Recently there has been an increased interest in studying the arithmetic analogue of this type of question, originally initiated when Graham, Rödl, and Ruczinsky [8] gave an asymptotic lower bound for the minimum number of monochromatic Schur triples in 2-colorings of the first n integers in 1996, see also [12, 14, 4, 1].

In this extended abstract, we focus on the analogue of the Ramsey multiplicity problem for specific additive structures in vector spaces over finite fields of small order. Let $q \in \mathbb{N}$ be a fixed prime power throughout and write $\mathbb{F}_q$ for the finite field with q elements. Given a subset $T \subseteq \mathbb{F}_q^n$ and a linear map L defined by some matrix $A \in \mathcal{M}^{r \times m}(\mathbb{Z})$ with integer entries co-prime to q, we are interested in studying the set $\mathcal{S}_L(T) = \{\mathbf{s} = (s_1, \ldots, s_m) \in T^m : L(\mathbf{s}) = \mathbf{0}$ and $s_i \neq s_j$ for $i \neq j\}$ of solutions with all-distinct entries in T. Throughout, we will assume that A is of full rank and that $\mathcal{S}_L(\mathbb{F}_q^n) \neq \emptyset$. We will also write $s_L(T) = |\mathcal{S}_L(T)|/|\mathcal{S}_L(\mathbb{F}_q^n)|$. Writing $[c] = \{1, \ldots, c\}$ for some given number of colors $c \in \mathbb{N}$, we call $\gamma : \mathbb{F}_q^n \to [c]$ a *c-coloring* of dimension $\dim(\gamma) = n$ and let $\gamma^{(i)}$ denote the set of elements colored with color $1 \leq i \leq c$ as well as $\Gamma_c(q, n)$ for the set of all c-colorings of $\mathbb{F}_q^n$. The *Rado multiplicity problem* is concerned with determining

$$m_c(L, q) = \lim_{n \to \infty} \min_{\gamma \in \Gamma_c(q,n)} s_L(\gamma^{(1)}) + \ldots + s_L(\gamma^{(c)}). \tag{1}$$

The limit exists by monotonicity and we have $0 \leq m_c(L, q) \leq 1$ by definition. Rado's theorem establishes that $m_c(L, q) > 0$ and we say that L is *c-common* for q if $m_c(L, q) = c^{1-m}$, that is if the minimum number of monochromatic solutions is attained in expectation by a uniform random coloring. For $r = 1$ a result of Cameron, Cilleruelo, and Serra [1] establishes that any L is 2-common if m is odd. When m is even, Saad and Wolf [13] showed that any L where the coefficients can be partitioned into pairs, with each pair summing to zero, is 2-common. Fox, Pham, and Zhao [6] showed that this sufficient condition is in fact also necessary. The case when $r > 1$ is much less understood, with Kamčev, Liebenau, and Morrison [9] recently characterizing a large family of non-common linear maps by showing that any L that 'induces' some smaller 2×4 linear map is uncommon. Focusing on specific values of q, Král, Lamaison, and Pach [10] also recently characterized the 2-common L for $q = 2$ when $r = 2$ and m is assumed to be odd. When $q = 5$, the most relevant additive structures to study is that of 4-APs. Saad and Wolf [13] showed that they are not 2-common by establishing an upper bound of $1/8 - 7 \cdot 2^{10} \cdot 5^{-2} \approx 0.1247 < 2^{-4}$. We establish the first non-trivial lower bound for this problem and an improved upper bound.

Proposition 1.1. *We have* $1/10 < m(L_{4\text{-}AP}, 5) \leq 13/126 = 0.1\overline{031746}$.

Going beyond 4-APs, we can also show that $m(L_{5\text{-AP}}, 5) \leq 1/26 < 2^{-4}$, establishing that 5-APs are likewise not 2-common in $\mathbb{F}_5$, but in this case did not obtain any meaningful lower bound. The study of monochromatic structures in colorings with more than two colors has also proven relevant in extremal graph theory. Most notably, Cummings et al. [3] extended the results of Goodman [7] by establishing the exact Ramsey multiplicity of triangles in 3-colorings and showing that they are not 3-common despite being 2-common. We consider a similar question and establish the exact multiplicity of 3-APs in 3-colorings of $\mathbb{F}_3^n$.

Theorem 1.2. *We have* $m_3(L_{3\text{-}AP}, 3) = 1/27$.

We can also show that $0.04486 \leq m_3(L_{\text{Schur}}, 2) \leq 1/16$ as well as $m_3(L_{\text{Schur}}, 3) \leq 7/81$, establishing that Schur triples are also not 3-common for $q = 2$ and $q = 3$. Upper bounds of all results are obtained through explicit blowup-type constructions. Lower bounds in the graph theoretic setting have recently been obtained through a computational approach relying on flag algebras due to Razborov [11]. This approach has been extended to different contexts, but so far seems to not have been explored in the arithmetic setting.

2 The correct notion of isomorphism

Let us omit q and c from notation, so in particular we write $\Gamma(n) = \Gamma_c(q, n)$ for the set of all c-colorings of dimension n as well as $\Gamma = \bigcup_{n=0}^{\infty} \Gamma(n)$. The 0-dimensional vector space consist of a single point, that is $\mathbb{F}_q^0 = \{0\}$, and we write e_j for the j-th canonical unit basis vector of $\mathbb{F}_q^n$ for $1 \leq j \leq n$ as well as e_0 for the zero vector.

Definition 2.1. *We refer to an affine linear map* $\varphi : \mathbb{F}_q^k \to \mathbb{F}_q^n$ *as a* morphism *and say that it is* t-fixed *for some* $t \geq 0$ *if* $\varphi(e_j) = e_j$ *for all* $0 \leq j \leq t$. *A morphism is a* monomorphism *whenever it is injective and a monomorphism is an* isomorphism *whenever* $n = k$.

Out of notational convenience, we extend the range of t to -1 in order to include *unfixed* morphisms and will always use t^+ to denote $\max\{t, 0\}$. For a given $t \geq -1$ and $n \geq k \geq t^+$, we let $\mathrm{M}_t(k; n)$ denote the set of t-fixed morphisms from $\mathbb{F}_q^k$ to $\mathbb{F}_q^n$ up to t-fixed isomorphism of $\mathbb{F}_q^k$. We likewise write $\mathrm{Mon}_t(k; n)$ for the set of monomorphisms with the same properties. Given $k_1, \ldots, k_m \geq t^+$ and $n \geq k_1 + \ldots + k_m - (m-1)\, t^+$, we let $\mathrm{Mon}_t(k_1, \ldots, k_m; n)$ denote the set of all tuples of monomorphisms $(\varphi_1, \ldots, \varphi_m) \in \mathrm{Mon}_t(k_1; n) \times \ldots \times \mathrm{Mon}_t(k_m; n)$ overlapping only in the t-fixed subspace.

Using these notions, we say two colorings $\gamma_1, \gamma_2 \in \Gamma(n)$ are *t-fixed isomorphic* for some $t \geq -1$, denoted by $\gamma_1 \cong_t \gamma_2$, if there exists a t-fixed isomorphism $\varphi : \mathbb{F}_q^n \to \mathbb{F}_q^n$ satisfying $\gamma_1 \equiv \gamma_2 \circ \varphi$. We let $\Gamma_t(n) = \Gamma(n)/\cong_t$ denote the set of all c-colorings of $\mathbb{F}_q^n$ up to t-fixed isomorphism and also write $\Gamma_t = \bigcup_{n \geq t^+} \Gamma_t(n)$. Given $k_1, \ldots, k_m \geq t^+$ and $n \geq k_1 + \ldots + k_m - (m-1)\, t^+$, the *density* $p_t(\delta_1, \ldots, \delta_m; \gamma)$ of some colorings $\delta_1 \in \Gamma_t(k_1)$, ..., $\delta_m \in \Gamma_t(k_m)$ in $\gamma \in \Gamma_t(n)$ is defined as the probability that a a tuple of t-fixed monomorphism chosen uniformly at random from $\mathrm{Mon}_t(k_1, \ldots, k_m; n)$ induces copies of $\delta_1, \ldots, \delta_m$ in γ. For $n \geq k \geq t^+$, we also let the *degenerate density* $p_t^d(\delta; \gamma)$ of some $\delta \in \Gamma_t(k)$ in γ denote the probability that a not-necessarily-injective t-fixed morphism does the same.

3 The correct notion of solution

In order to develop the flag algebra approach, the density of solutions needs to be representable as the weighted density of particular colorings, motivating the following definition.

Definition 3.1. *For any $t \geq -1$ and $n \geq t^+$, the* t-fixed dimension $\dim_t(s)$ *of* $\mathbf{s} \in \mathcal{S}_L(\mathbb{F}_q^n)$ *is the smallest $k \geq t^+$ for which there exists a t-fixed k-dim. subspace of $\mathbb{F}_q^n$ containing* $\mathbf{s}$.

We will only need the unfixed and 0-fixed dimension and denote by $\dim_t(L)$ the largest t-fixed dimension of any solution to a given linear map L. In general, $\dim_t(L) = m - r + t$ for any linear map L when $t \geq 0$ as well as $\dim_{-1}(L) = m - r - 1$ when L is *invariant*, that is if for any solution $\mathbf{s} = (x_1, \ldots, x_m) \in \mathcal{S}_L(\mathbb{F}_q^n)$ and element $a \in \mathbb{F}_q^n$ we have $a + \mathbf{s} = (a + x_1, \ldots, a + x_m) \in \mathcal{S}'_L(\mathbb{F}_q^n)$. We say that L is *admissible* if $t \geq 0$ or if $t = -1$ and L is invariant. A solution $\mathbf{s} \in \mathcal{S}_L(\mathbb{F}_q^n)$ for some admissible L is *t-fixed fully dimensional* if $\dim_t(s)$ attains the respective upper bound. For a given set $T \subseteq \mathbb{F}_q^n$, we denote the set of fully dimensional solutions to some admissible L by $\mathcal{S}_L^t(T)$ and write $s_L^t(T) = |\mathcal{S}_L^t(T)| \,/\, |\mathcal{S}_L^t(\mathbb{F}_q^n)|$.

The important property that we make use of is that each fully-dimensional solution defines a unique $\dim(L)$-dimensional t-fixed subspace in which it lies and that for any $t \geq -1$, admissible L, and $n \geq t \geq 0$, the number of solutions in a subset of $\mathbb{F}_q^n$ is invariant under t-fixed isomorphism. The same would not hold for $t = -1$ if L was not invariant.

4 The flag algebras for additive structures

For any $t \geq 0$, we refer to elements of $\Gamma_t(t) = \Gamma(t)$ as *types* of dimension t. We also introduce a unique *empty type*, denoted by $\varnothing$, of dimension $t = -1$. For a given type τ of dimension t, we refer to a coloring $F \in \Gamma_t(n)$ satisfying $F \circ \mathrm{id}_{t,n} \equiv \tau$ as a *flag of type* τ, where $\mathrm{id}_{t,n}$ denotes the unique t-fixed isomorphism from $\mathbb{F}_q^t$ to $\mathbb{F}_q^n$ and the requirement is vacantly true for $t = -1$. We will write $\mathcal{F}_n^\tau$ for the set of all flags of given type τ and dimension n as well as $\mathcal{F}^\tau = \bigcup_n \mathcal{F}_n^\tau$.

Definition 4.1. *The* flag algebra $\mathcal{A}^\tau$ *of type τ is given by equipping $\mathbb{R}\mathcal{F}^\tau / \mathcal{K}^\tau$, where $\mathcal{K}^\tau = \{F - \sum_{F' \in \mathcal{F}_n^\tau} p_t(F; F')\, F' : F \in \mathcal{F}^\tau, n \geq \dim(F)\}$, with the product given by the the bilinear extension of $F_1 \cdot F_2 = \sum_{H \in \mathcal{F}_n^\tau} p_t(F_1, F_2; H)\, H + \mathcal{K}^\tau$ defined for any two flags $F_1, F_2 \in \mathcal{F}^\tau$ and arbitrary $n \geq \dim(F_1) + \dim(F_2) - \dim(\tau)$.*

Assume we are given a parameter $\lambda : \Gamma \to \mathbb{R}$ that is invariant under t_λ-fixed isomorphisms for some $t_\lambda \geq -1$ and that satisfies $\lambda(\gamma) = \sum_{\beta \in \Gamma_{t_\lambda}(n)} \lambda(\beta)\, p_{t_\lambda}(\beta, \gamma)$ for some $n_\lambda \in \mathbb{N}$ and all $\gamma \in \Gamma_{t_\lambda}$, where $n_\lambda \leq n \leq \dim(\gamma)$. Monochromatic fully-dimensional solutions to a given linear map L define such a parameter with $t_\lambda = 0$ for general L and $t_\lambda = -1$ for invariant ones, where in either case $n_\lambda = \dim_{t_\lambda}(L)$. We are interested in determining

$$\lambda^\star = \lim_{n \to \infty} \min_{\gamma \in \Gamma_{t_\lambda}(n)} \lambda(\gamma). \tag{2}$$

Writing $C_\lambda^\tau = \sum_{\beta \in \mathcal{F}_{n_\lambda}^\tau} \lambda(\beta)\, \beta$ for any type τ of dimension t_λ, our problem of determining $\lambda^\star$ can be restated through the conic optimization problem

$$\lambda^\star = \max\{\lambda' \in \mathbb{R} : C_\lambda^\tau \geq \lambda' \text{ for all types } \tau \text{ of dimension } t_\lambda\}, \tag{3}$$

where we write $\mathrm{Hom}^+(\mathcal{A}^\tau, \mathbb{R})$ for the set of positive homomorphisms, that is algebra homomorphisms $\phi \in \mathrm{Hom}(\mathcal{A}^\tau, \mathbb{R})$ satisfying $\phi(F) \geq 0$ for any $F \in \mathcal{F}^\tau$, and $\mathcal{S}^\tau = \{f \in \mathcal{A}^\tau : \phi(f) \geq 0 \text{ for all } \phi \in \mathrm{Hom}^+(\mathcal{A}^\tau, \mathbb{R})\}$ for the *semantic cone* of type τ. Noting that we can define a linear *downward operator* $[\![\cdot]\!]_{t_\lambda}\colon \mathcal{A}^\tau \to \mathcal{A}^{\tau_\lambda}$ for any type τ of dimension $t \geq t_\lambda$ that satisfies $[\![\mathcal{S}^\tau]\!]_{t_\lambda} \subseteq \mathcal{S}^{\tau_\lambda}$, we can derive a lower bound by defining a set of types $\mathcal{T}$ as well as sets of algebra elements $\mathcal{B}_{\tau'} \subset \mathcal{A}^{\tau'}$ and establishing that

$$C_\lambda^\tau \geq \lambda' + \sum_{\tau' \in \mathcal{T}} \sum_{f \in \mathcal{B}_{\tau'}} [\![f^2]\!]_{t_\lambda}. \tag{4}$$

References

[1] P. J. Cameron, J. Cilleruelo, and O. Serra. On monochromatic solutions of equations in groups. *Revista Matemática Iberoamericana*, 23(1):385–395, 2007.

[2] D. Conlon. On the Ramsey multiplicity of complete graphs. *arXiv preprint arXiv:0711.4999*, 2007.

[3] J. Cummings, D. Král', F. Pfender, K. Sperfeld, A. Treglown, and M. Young. Monochromatic triangles in three-coloured graphs. *Journal of Combinatorial Theory, Series B*, 103(4):489–503, 2013.

[4] B. A. Datskovsky. On the number of monochromatic Schur triples. *Advances in Applied Mathematics*, 31(1):193–198, 2003.

[5] P. Erdős. On the number of complete subgraphs contained in certain graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl*, 7(3):459–464, 1962.

[6] J. Fox, H. T. Pham, and Y. Zhao. Common and Sidorenko linear equations. *The Quarterly Journal of Mathematics*, 72(4):1223–1234, 2021.

[7] A. W. Goodman. On sets of acquaintances and strangers at any party. *The American Mathematical Monthly*, 66(9):778–783, 1959.

[8] R. Graham, V. Rödl, and A. Ruciński. On Schur properties of random subsets of integers. *Journal of Number Theory*, 61(2):388–408, 1996.

[9] N. Kamčev, A. Liebenau, and N. Morrison. Towards a characterisation of Sidorenko systems. *arXiv preprint arXiv:2107.14413*, 2021.

[10] A. Lamaison, P. P. Pach, et al. Common systems of two equations over the binary field. In *Discrete Mathematics Days 2022*, pages 169–173. Editorial de la Universidad de Cantabria, 2022.

[11] A. A. Razborov. Flag algebras. *The Journal of Symbolic Logic*, 72(4):1239–1282, 2007.

[12] A. Robertson and D. Zeilberger. A 2-coloring of $[1, n]$ can have $(1/22)n^2 + o(n)$ monochromatic Schur triples, but not less! *The Electronic Journal of Combinatorics*, 5(1):R19, 1998.

[13] A. Saad and J. Wolf. Ramsey multiplicity of linear patterns in certain finite abelian groups. *The Quarterly Journal of Mathematics*, 68(1):125–140, 2017.

[14] T. Schoen. The number of monochromatic Schur triples. *European Journal of Combinatorics*, 20(8):855–866, 1999.

[15] A. Thomason. A disproof of a conjecture of Erdős in Ramsey theory. *Journal of the London Mathematical Society*, 2(2):246–255, 1989.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Unified study of the phase transition for block-weighted random planar maps

(Extended abstract)

Zéphyr Salvy*

Abstract

In [Fleurat, Salvy 2023], we introduced a model of block-weighted random maps that undergoes a phase transition as the density of separating elements changes. The purpose of this note is to demonstrate that the methodology we developed can be extended to many other families of maps. We prove that a phase transition exists and provide detailed information about the size of the largest blocks in each regime.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-109`

1 Introduction

A *planar map* $\mathfrak{m}$ is the proper embedding into the two-dimensional sphere of a connected planar finite multigraph, considered up to homeomorphisms. Maps exhibit very rich combinatorial and probabilistic properties, which have been the focus of an extensive literature. Many families of planar maps have very nice counting formulas [Tut63]. A key aspect of planar maps is that they can be decomposed, typically into components of higher connectivity degree. Such decompositions typically relate one family of planar maps to another and gives an equation between their generating series.

Theses types of decompositions were initially introduced by Tutte [Tut63] to obtain some enumerative results about planar maps. But they also play a major role in the enumerative study of planar graphs [GN09]. They allow to study certain models of discrete metric spaces in theoretical physics [Bon16]. In view of applications to random generation

*Univ Gustave Eiffel, CNRS, LIGM, F-77454 Marne-la-Vallée, France. E-mail: `zephyr.salvy@univ-eiffel.fr`.

[Sch99], the decomposition of planar maps has been systematised in [BFSS01], where a uniform treatment via analytic combinatorics is developed. A probabilistic approach was later derived using encoding of maps via enriched trees [Stu18, AB19].

Planar map models exhibit *universality*, meaning that many natural classes of random maps show similar behavior as their size grows to infinity. When taking an object of size n uniformly at random among all objects in a class and appropriately rescaling its distance, the sequence of random objects converges to a certain metric space. This was first proved for uniform quadrangulations by Miermont [Mie13] and Le Gall [LG13], and since then, results have been extended to other families of maps, including uniform triangulations and uniform $2q$-angulations ($q \geqslant 2$) [LG13], uniform simple triangulations and uniform simple quadrangulations [ABA17], bipartite planar maps with a prescribed face-degree sequence [Mar18], $(2q+1)$-angulations [ABA21] and Eulerian triangulations [Car21].

In a previous article [FS23], together with Fleurat, we studied a model of random maps, depending on a parameter u which controls the density of separating elements. We proved that this model exhibits a phase transition as u varies, and that it interpolates between the Brownian sphere and the Brownian tree of Aldous [Ald91]. This approach — which we detailed in [FS23] for general maps and their 2-connected cores and for general quadrangulations and their simple cores — can be applied to other decompositions, such as those in [BFSS01, Table 3] (which is partially reproduced in Table 1), and this is the focus of this note. We restrict our study to decomposition schemes without "coreless" maps (the decompositions involving coreless maps, such as 2-connected maps into 3-connected components, bring further difficulties, which we expect to handle with some more work).

Let us give some formalism for decompositions. A map is said to be *loopless* if it does not contain any loop; 2-*connected* if it does not contain any *cut vertex* (*i.e.* a vertex whose removal deconnects the map) and *simple* if it has neither loops nor multiple edges. Planar maps can be decomposed into loopless (or 2-connected, or simple, or 2-connected simple...) components, which are the so-called "blocks". It is also the case for *bipartite* maps, whose vertices can be properly bicolored in black and white; and for *triangulations*, whose faces all have degree 3. The latter can be decomposed into *irreducible* components, in which every 3-cycle defines a face. We consider eight models here (see Table 1):

1. Loopless maps decomposed into simple blocks;
2. General maps decomposed into 2-connected blocks;
3. 2-connected maps decomposed into 2-connected simple blocks;
4. Bipartite maps decomposed into bipartite simple blocks;
5. Bipartite maps decomposed into bipartite 2-connected blocks;
6. Bipartite 2-connected maps decomposed into bipartite 2-connected simple blocks;
7. Loopless triangulations decomposed into triangular simple blocks;
8. Simple triangulations decomposed into triangular irreducible blocks.

In general, the *size* $|\mathfrak{m}|$ of a planar map $\mathfrak{m}$ is its number of edges. In a decomposition scheme, we let $M(z) = \sum_{n\in\mathbb{Z}_{\geqslant 0}} m_n z^n$ be the generating series of the class of maps to be decomposed and similarly $B(z) = \sum_{n\in\mathbb{Z}_{\geqslant 0}} b_n z^n$ be the generating series of the class of "blocks" into which the maps are decomposed.

Setting $M(z,u) = \sum_{\mathfrak{m}\in\mathcal{M}} z^{|\mathfrak{m}|}u^{b(\mathfrak{m})}$, where $b(\mathfrak{m})$ is the number of blocks of positive size in $\mathfrak{m}$, all the models we consider — which are listed in Table 1 — satisfy

$$M(z,u) = uB(H(z, M(z,u))), \tag{1}$$

or, for the last one,

$$M(z,u) = (1 + M(z,u)) \times uB(H(z, M(z,u))). \tag{2}$$

For example, for the decomposition of general maps into 2-connected ones (which is the case studied in [FS23]), one has

$$M(z,u) = uB(z(1 + M(z,u))^2).$$

For $u > 0$, denote by $\rho(u)$ the radius of convergence of $z \mapsto M(z,u)$. In view of the form of Equations (1) and (2) and in particular that they are non-linear, it holds that $M(\rho(u),u) < \infty$.

In the following, $\mathbf{M}$ and $\mathbf{M}_n$ are random variables drawn according to the following probability distributions. For $u \in \mathbb{R}_{>0}$, $n \in \mathbb{Z}_{>0}$ and $\mathfrak{m} \in \mathcal{M}$, we set

$$\mathbb{P}_u(\mathfrak{m}) = \frac{\rho(u)^{|\mathfrak{m}|}u^{b(\mathfrak{m})}}{M(\rho(u),u)} \qquad \text{and} \qquad \mathbb{P}_{n,u}(\mathfrak{m}) = \frac{u^{b(\mathfrak{m})}}{[z^n]M(z,u)}\mathbb{1}_{|\mathfrak{m}|=n}.$$

Regarding enumeration in our setting, we show the following by analytic methods (details are omitted in this short note).

Theorem 1. *For any model described in Table 1, where maps are decomposed into blocks weighted with a weight $u > 0$, there exists a critical value u_C at which the model undergoes a phase transition. As u varies, there exists $c(u) > 0$ such that*

$$[z^n]M(z,u) \sim \begin{cases} c(u)n^{-5/2}\rho(u)^{-n} & \text{if } u < u_C \\ c(u_C)n^{-5/3}\rho(u_C)^{-n} & \text{if } u = u_C \\ c(u)n^{-3/2}\rho(u)^{-n} & \text{if } u > u_C \end{cases}.$$

All the constants involved in Theorem 1 are explicit. Table 2 gives the expressions for u_C, $\rho(u)$ and $M(\rho(u),u)$ when $u \leqslant u_C$.

The polynomial correction for $u < u_C$ (subcritical case) is the same than for planar maps, whereas when $u > u_C$ (supercritical case) it is the same than for plane trees. Moreover, at $u = u_C$, a new asymptotic behaviour appears with a polynomial correction in $n^{-5/3}$.

In this note, we also focus on another aspect of the phase transition, namely the size of the largest blocks. We show that if $u < u_C$, a condensation phenomenon occurs and the

Scheme	maps, $M(z)$	blocks, $B(z)$	submaps, $H(z, M)$
1	loopless, $M_2(z)$	simple, $M_3(z)$	$z(1+M)$
2	all, $M_1(z)$	2-connected, $M_4(z)$	$z(1+M)^2$
3	2-connected $M_4(z) - z$	2-connected simple, $M_5(z)$	$z(1+M)$
4	bipartite, $B_1(z)$	bipartite simple, $B_2(z)$	$z(1+M)$
5	bipartite, $B_1(z)$	bipartite 2-connected, $B_4(z)$	$z(1+M)^2$
6	bipartite 2-connected, $B_4(z)$	bipartite 2-connected simple $B_5(z)$	$z(1+M)$
7	loopless triangulations, $T_1(z)$	simple triangulations, $z + zT_2(z)$	$z(1+M)^3$
8	simple triangulations, $T_2(z)$	irreducible triangulations, $T_3(z)$	$z(1+M)^2$

Table 1: Partial reproduction of [BFSS01, Table 3], which describes composition schemas of the form $\mathcal{M} = \mathcal{B} \circ \mathcal{H}$ except the last one where $\mathcal{M} = (1 + \mathcal{M}) \times \mathcal{B} \circ \mathcal{H}$. The terminology and notation were slightly changed. For all i, $[z^n]M_i(z)$ and $[z^n]B_i(z)$ is the number of such maps with n edges. $[z^n]T_1(z)$ (resp. $[z^n]T_2(z)$ and $[z^n]T_3(z)$) is the number of loopless (resp. simple or irreducible) triangulations with $n+2$ (resp. $n+3$) vertices.

largest block is of size $\Theta(n)$; when $u > u_C$, the largest block is of size $\Theta(\log(n))$; for $u = u_C$, the largest block is of size $\Theta(n^{2/3})$ (Theorem 3). For the subcritical case, as in [FS23], we follow the probabilistic approach of [AB19] (whereas [BFSS01] gives an analytic approach).

These results further support that the scaling limits should be the Brownian sphere when $u < u_C$, the Brownian tree when $u > u_C$ and the stable tree of parameter $3/2$ when $u = u_C$. This was proved for the decomposition of quadrangulations into simple components [FS23], and we expect this phenomenon to be generic. For model 2, the critical scaling limit was established in [FS23] and the supercritcal one in [Stu20a]. For model 5, the supercritical case was also established [Stu20a].

2 Tree structure

We explain here how an underlying tree structure can be associated to each of the models of Table 1. As a first step, we rewrite the decomposition equations in the standard Lagrangian form $M(z) = z \times \Phi(M(z))$ for some function Φ, taking the weight u into account. (Beware that Equations (1) and (2) are not of this form as the products by z are inside H.)

Proposition 1. *For all models listed in Table 1, there exists a generating function Φ with nonnegative coefficients such that*

$$M(z,u) = z \times \Phi(M(z,u), u). \tag{3}$$

Proof sketch. We discuss how the rewriting is done for two cases: general maps into 2-connected components, and simple triangulations into irreducible components.

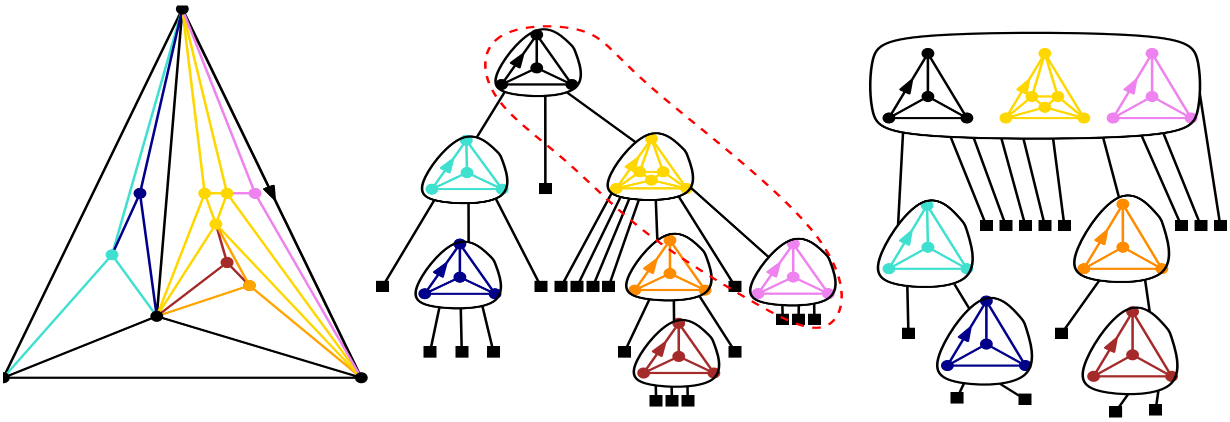


Figure 1: A simple triangulation, its classical tree of (irreducible) blocks, the adapted tree where some blocks are grouped into sequences.

We start by maps decomposed into 2-connected components. In order to do the rewriting, we need to change the size parameter, which we take as the number of half-edges plus one. Accordingly, we set $\hat{M}(z,u) = z(1+M(z^2,u))$. The equation then becomes

$$\hat{M}(z,u) = z\left(1 + uB(\hat{M}(z,u)^2)\right),$$

which is of the desired form.

The second case we discuss is the decomposition of simple triangulations into irreducible components. We also need to change the size parameter, taking here the number of inner faces instead of the number of vertices. A further ingredient compared to the first case is that we need to group the components into sequences to obtain an equation in Lagrangian form. Let $\hat{z}$ count internal faces. The generating series of simple (resp. irreducible) triangulations counted by internal faces $\hat{T}_2(\hat{z},u)$ (resp. $\widetilde{T}_3(\hat{z})$) is closely related to $T_2(z,u)$ (resp. $T_3(z)$) since a triangulation with $n+3$ vertices has $2n+2$ faces so $2n+1$ internal faces. Then, denoting $\hat{T}_3(\hat{z}) = \widetilde{T}_3(\hat{z})/\hat{z}$, It holds that

$$\hat{T}_2(\hat{z},u) = \hat{z} + u\hat{T}_2(\hat{z},u)\hat{T}_3(\hat{T}_2(\hat{z},u)), \qquad \text{so} \qquad \hat{T}_2(\hat{z},u) = \frac{\hat{z}}{1-u\hat{T}_3(\hat{T}_2(\hat{z},u))}.$$

Therefore, we set $\Phi(M,u) = \frac{1}{1-u\hat{T}_3(M)}$. This corresponds to the vertices of the tree encoding a sequence of irreducible triangulations, which is represented on Fig. 1[1]. $\square$

The *block tree* $T_{\mathfrak{m}}$ of a map $\mathfrak{m}$ is the tree associated to the the decomposition of $\mathfrak{m}$ as expressed in Proposition 1. Each node of the tree correspond to an object φ counted by Φ. The subtrees hanging at a node (corresponding to some φ) are the trees of the components substituted into φ.

[1]With additional work, one can do the same for simple quadrangulations decomposed into irreducible ones.

Scheme	u_C	$\rho(u)$	$M(\rho(u),u)$	$E(u)$	$1-E(1)$
1	$\frac{81}{17}$	$\frac{27}{8(5u+27)}$	$\frac{5u}{27}$	$\frac{32u}{3(5u+27)}$	$\frac{2}{3}$
2	$\frac{9}{5}$	$\frac{4}{3(u^2+6u+9)}$	$\frac{u}{3}$	$\frac{8u}{3(u+3)}$	$\frac{1}{3}$
3	$\frac{135}{7}$	$\frac{128}{27(5u+27)}$	$\frac{25u^2+135u+128}{27(5u+27)}$	$\frac{32u}{5(5u+27)}$	$\frac{4}{5}$
4	$\frac{36}{11}$	$\frac{5}{8(u+4)}$	$\frac{u}{4}$	$\frac{20u}{9(u+4)}$	$\frac{5}{9}$
5	$\frac{52}{27}$	$\frac{25}{8(u^2+8u+16)}$	$\frac{u}{4}$	$\frac{40u}{13(u+4)}$	$\frac{5}{13}$
6	$\frac{68}{3}$	$\frac{125}{128(u+4)}$	$\frac{u}{4}$	$\frac{20u}{17(u+4)}$	$\frac{13}{17}$
7	$\frac{16}{7}$	$\frac{54}{u^3+24u^2+192u+512}$	$\frac{u}{8}$	$\frac{9u}{2(u+8)}$	$\frac{1}{2}$
8	$\frac{64}{37}$	$\frac{25}{6912}u^2-\frac{5}{108}u+\frac{4}{27}$	$\frac{5u}{32-5u}$	$\frac{27u}{2(32-5u)}$	$\frac{1}{2}$

Table 2: Values of u_C, $\rho(u)$, $M(\rho(u),u)$ and $E(u)$ when $u \leqslant u_C$ for all the decomposition schemes of Table 1.

As a consequence of (3), if we set a probability measure μ^u such that, for $k \in \mathbb{Z}_{\geqslant 0}$,

$$\mu^u(k) = \frac{[X^k]\Phi(X,u)y(u)^k}{\Phi(y(u),u)}$$

for $y(u) = M(\rho(u),u)$ (using the definition of M given in (3)); then, using the fact that $T_{\mathfrak{m}}$ and the decoration of its vertices are bijectively linked to $\mathfrak{m}$, we have the following result (again stated in terms of the M of (3)):

Theorem 2. *For all $u > 0$, $T_{\mathbf{M}}$ follows the law of a Galton-Watson tree of reproduction law μ^u. Moreover, $T_{\mathbf{M}_n}$ follows the law of a Galton-Watson tree of reproduction law μ^u conditioned to have n vertices.*

The decomposition tree of a random map of "Lagrangian size" n is a Galton-Watson tree of reproduction law μ^u conditioned to have n vertices. For instance, for maps decomposed into 2-connected components, the tree of a random map with n edges is a Galton-Watson tree conditioned to have $2n+1$ vertices. This enables to put into light a phase transition on the tree structure, using the usual phase transition for Galton-Watson trees [Nev86].

Proposition 2. *The expectation $E(u)$ of μ^u is written down in Table 2.*

3 Results on the size of the largest blocks

Starting from Equations (1) and (2) and Theorem 2, we use the same techniques as in [FS23] to obtain results on decomposition schemes. However, simple triangulations decomposed into irreducible blocks present a challenge, as the vertices of the block trees are not decorated with a single block (or none) but with a sequence of blocks. Hence, the size of the blocks cannot be immediately read from the degrees in the block tree. However, an extreme condensation phenomenon occurs, concentrating mass in only one element of the sequence (as in [Gou98, Th1]), resulting in a similar behaviour.

Denote by $L_{n,j}$ the size of the j-th largest block of $\mathbf{M}_n$. By Theorem 2, the same arguments as in [FS23] apply and the following holds.

Theorem 3. *Models described in Table 1, where maps are decomposed into blocks weighted with a weight $u > 0$, satisfy the following.*

Subcritical case *For all $u < u_C$, we have*

$$L_{n,1} = (1 - E(u))n + O_{\mathbb{P}}(n^{2/3}) \quad \textit{and} \quad L_{n,2} = O_{\mathbb{P}}(n^{2/3}).$$

Moreover, there exists an explicit constant $\tilde{c}(u) > 0$ such that the following joint convergence holds:

$$\left(\frac{1}{n\tilde{c}(u)}\right)^{2/3} ((1 - E(u))n - L_{n,1}, (L_{n,j}, j \geqslant 2)) \xrightarrow[n\to\infty]{(d)} \left(L_1, \left(\Delta L_{(j-1)}, j \geqslant 2\right)\right) \quad (4)$$

where $(L_t)_{t\in[0,1]}$ is a Stable process of parameter $3/2$ such that $\mathbb{E}\left[e^{-sL_1}\right] = e^{\Gamma(-3/2)s^{3/2}}$ and $\Delta L_{(1)} \geqslant \Delta L_{(2)} \geqslant \dots$ is the ranked sequence of its jumps.

Supercritical case *For all $u > u_C$, there exist explicit values $F(u), G(u) > 0$ such that, for all fixed $j \geqslant 1$,*

$$L_{n,j} = F(u)\ln(n) - G(u)\ln(\ln(n)) + O_{\mathbb{P}}(1).$$

Critical case *If $u = u_C$, then*

$$\left(\frac{L_{n,j}}{n^{2/3}}, j \geqslant 1\right) \xrightarrow[n\to\infty]{(d)} \left(E_{(j)}, j \geqslant 1\right),$$

where the $\left(E_{(j)}\right)$ are the ordered atoms of an explicite Point Process, specified in [Jan12, Ex19.27, Rk19.28].

As mentioned in [FS23], for $u = 1$, we retrieve by a probabilistic method the results of [BFSS01, Table 4], established by analytic techniques: indeed, our $1 - E(1)$ corresponds to their α_0. The probabilistic approach we follow was first developed by [AB19] and has the advantage that we obtain a joint limit law for the largest block and the subsequent ones. It can also yield local limit theorems for the size of the largest block, as is discussed by Stufler in [Stu20b].

References

[AB19] Louigi Addario-Berry. A probabilistic approach to block sizes in random maps. *ALEA - Latin American Journal of Probability and Mathematical Statistics*, XVI:1–13, 2019.

[ABA17] Louigi Addario-Berry and Marie Albenque. The scaling limit of random simple triangulations and random simple quadrangulations. *The Annals of Probability*, 45(5):2767 – 2825, 2017.

[ABA21] Louigi Addario-Berry and Marie Albenque. Convergence of odd-angulations via symmetrization of labeled trees. *Annales Henri Lebesgue*, 4:653–683, 2021.

[Ald91] David Aldous. The continuum random tree. II. an overview. *Stochastic analysis*, 167:23–70, 1991.

[BFSS01] Cyril Banderier, Philippe Flajolet, Gilles Schaeffer, and Michèle Soria. Random Maps, Coalescing Saddles, Singularity Analysis, and Airy Phenomena. *Random Struct. Algorithms*, 19(3–4):194–246, oct 2001.

[Bon16] Valentin Bonzom. Large N limits in tensor models: Towards more universality classes of colored triangulations in dimension $d \geqslant 2$. *Symmetry, Integrability and Geometry: Methods and Applications*, 12(073):39, 2016.

[Car21] Ariane Carrance. Convergence of Eulerian triangulations. *Electronic Journal of Probability*, (26):1–48, 2021.

[FS23] William Fleurat and Zéphyr Salvy. A phase transition in block-weighted random maps. 2023.

[GN09] Omer Giménez and Marc Noy. Asymptotic enumeration and limit laws of planar graphs. *Journal of the American Mathematical Society*, 22(2):309–329, 2009.

[Gou98] Xavier Gourdon. Largest component in random combinatorial structures. *Discrete Mathematics*, 180(1):185–209, 1998.

[Jan12] Svante Janson. Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation. *Probability Surveys*, 9(none):103 – 252, 2012.

[LG13] Jean-François Le Gall. Uniqueness and universality of the Brownian map. *The Annals of Probability*, 41(4):2880 – 2960, 2013.

[Mar18] Cyril Marzouk. Scaling limits of random bipartite planar maps with a prescribed degree sequence. *Random Struct. Algorithms*, 53(3):448–503, 2018.

[Mie13] Grégory Miermont. The Brownian map is the scaling limit of uniform random plane quadrangulations. *Acta Mathematica*, 210(2):319–401, 2013.

[Nev86] Jacques Neveu. Arbres et processus de Galton-Watson. *Annales de l'I.H.P. Probabilités et statistiques*, 22(2):199–207, 1986.

[Sch99] Gilles Schaeffer. Random sampling of large planar maps and convex polyhedra. In *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*, STOC '99, pages 760–769, New York, NY, USA, 1999. Association for Computing Machinery.

[Stu18] Benedikt Stufler. Random enriched trees with applications to random graphs. *Electronic Journal of Combinatorics*, 25(3), 2018.

[Stu20a] Benedikt Stufler. Limits of random tree-like discrete structures. *Probability Surveys*, 17(none):318 – 477, 2020.

[Stu20b] Benedikt Stufler. On the maximal offspring in a subcritical branching process. *Electronic Journal of Probability*, 25(none):1 – 62, 2020.

[Tut63] W. T. Tutte. A census of planar maps. *Canadian Journal of Mathematics*, 15:249–271, 1963.

Acknowledgements

The author would like to thank their supervisors Marie Albenque and Éric Fusy for their guidance and support.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

The Turán Number of Surfaces

(Extended abstract)

Maya Sankar*

Abstract

We show that there is a constant c such that any 3-uniform hypergraph $\mathcal{H}$ with n vertices and at least $cn^{5/2}$ edges contains a triangulation of the real projective plane as a sub-hypergraph. This resolves a conjecture of Kupavskii, Polyanskii, Tomon, and Zakharov. Furthermore, our work, combined with prior results, asymptotically determines the Turán number of all surfaces.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-110`

1 Introduction

Turán-type questions are fundamental in the study of extremal combinatorics. Given a fixed r-uniform hypergraph $\mathcal{F}$, its *Turán number* $\mathrm{ex}(n, \mathcal{F})$ is the maximum number of edges in an r-uniform hypergraph $\mathcal{H}$ on n vertices which does not contain $\mathcal{F}$ as a sub-hypergraph. Estimating Turán numbers for hypergraphs remains a largely open problem; we refer the reader to the surveys [1, 3, 9] for a general overview.

In this paper, we investigate a topological variant of this problem. Any r-uniform hypergraph $\mathcal{H}$ may be viewed as an $(r-1)$-dimensional simplicial complex whose facets are the edges of $\mathcal{H}$. Similarly, one may ask if any sub-hypergraph of $\mathcal{H}$ is homeomorphic to a given $(r-1)$-dimensional simplicial complex X. This topological perspective yields many natural generalizations of graph properties to higher dimensions. For example, one analogue of Hamiltonian cycles in 3-uniform hypergraphs that has received some attention (see [2, 8]) is a spanning sub-hypergraph homeomorphic to the 2-sphere. Additionally,

*Department of Mathematics, Stanford University. Email: `mayars@stanford.edu`. Supported by a Fannie and John Hertz Foundation Fellowship and NSF Graduate Research Fellowship DGE-1656518.

one is naturally interested in the following extremal quantity. Let X be a closed $(r-1)$-dimensional manifold. Denote by $\mathrm{ex}_{\mathrm{hom}}(n, X)$ the maximum number of edges in a r-uniform hypergraph $\mathcal{H}$ on n vertices such that no sub-hypergraph of $\mathcal{H}$ is homeomorphic (as a simplicial complex) to X. This is the *Turán number* of the topological space X.

As part of his program in high-dimensional combinatorics, Linial [7] asked for the asymptotics of $\mathrm{ex}_{\mathrm{hom}}(n, X)$ when $r \geq 3$. Linial's question was partially motivated by the work of Sós, Erdős, and Brown [11] some decades prior, which showed that $\mathrm{ex}_{\mathrm{hom}}(n, X) = \Theta(n^{5/2})$ when X is the 2-sphere $\mathbb{S}^2$. Linial [7] sketched a new proof of the lower bound $\mathrm{ex}_{\mathrm{hom}}(n, \mathbb{S}^2) = \Omega(n^{5/2})$ which generalized to all closed, connected 2-manifolds X; this proof is given rigorously in [5, §2]. We call such a 2-manifold a *surface*.

All surfaces fall into one of three categories: the sphere $\mathbb{S}^2$, the connected sum of $g \geq 1$ tori, or the connected sum of $k \geq 1$ real projective planes. Until recently, it was unknown if the lower bound of $n^{5/2}$ was asymptotically tight for the latter two classes. Indeed, Linial [6, 7] repeatedly conjectured a matching upper bound for the torus $\mathbb{T}^2$, i.e. that $\mathrm{ex}_{\mathrm{hom}}(n, \mathbb{T}^2) = O(n^{5/2})$. Kupavskii, Polyanskii, Tomon, and Zakharov [5] proved Linial's conjecture in 2020. Additionally, they showed that if two surfaces X_1, X_2 satisfy $\mathrm{ex}_{\mathrm{hom}}(n, X_i) = O(n^{5/2})$, their connected sum $X_1 \# X_2$ also satisfies $\mathrm{ex}_{\mathrm{hom}}(n, X_1 \# X_2) = O(n^{5/2})$, thereby extending the upper bound $\mathrm{ex}_{\mathrm{hom}}(n, X) = O(n^{5/2})$ to orientable surfaces of the form $X = \mathbb{T}^2 \# \cdots \# \mathbb{T}^2$. They were unable to derive the corresponding result for any non-orientable surfaces, but conjectured that the same bound applies to all surfaces.

Our main result is the resolution of this conjecture. We show that Linial's lower bound is asymptotically tight for the real projective plane $\mathbb{RP}^2$.

Theorem 1.1. *We have* $\mathrm{ex}_{\mathrm{hom}}(n, \mathbb{RP}^2) = O(n^{5/2})$.

By Kupavskii, Polyanskii, Tomon, and Zakharov's result on connected sums, this bound generalizes to all non-orientable surfaces $X = \mathbb{RP}^2 \# \cdots \# \mathbb{RP}^2$. Combining our work with the results of Sós, Erdős, and Brown [11] for the sphere and Kupavskii, Polyanskii, Tomon, and Zakharov [5] for all other orientable surfaces, we completely determine the asymptotics of $\mathrm{ex}_{\mathrm{hom}}(n, X)$ for any surface X.

Theorem 1.2. *Let X be any surface. Then* $\mathrm{ex}_{\mathrm{hom}}(n, X) = \Theta(n^{5/2})$*, where the constant coefficients may depend on the surface X.*

In the remaining two sections, we sketch the proof of Theorem 1.1. We first describe how to build a hypergraph homeomorphic to $\mathbb{RP}^2$ out of smaller substructures. Then, we give an overview of the probabilistic techniques required to locate these substructures.

2 Deconstructing $\mathbb{RP}^2$

Our proof of Theorem 1.1 begins by identifying conditions under which a 3-uniform hypergraph $\mathcal{H}$ contains a sub-hypergraph homeomorphic to $\mathbb{RP}^2$.

We decompose $\mathbb{RP}^2$ as two copies of $\mathbb{D}^2$ attached to $\mathbb{S}^1 \vee \mathbb{S}^1$. Consider the standard representation of $\mathbb{RP}^2$ as a disk with boundary glued to itself antipodally — this is pictured

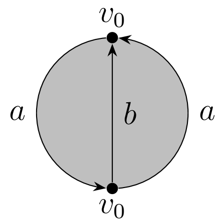


Figure 1: Two loops a and b in $\mathbb{RP}^2$ based at the point v_0. Here, $\mathbb{RP}^2$ is depicted as a disk with boundary points identified antipodally.

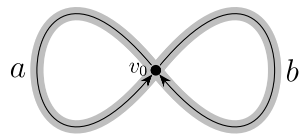


Figure 2: Two loops a and b in $\mathbb{S}^1 \vee \mathbb{S}^1$ sharing the same basepoint v_0.

in Fig. 1. Let a and b be the loops in $\mathbb{RP}^2$ depicted, with a traversing half the boundary of the disk and b a diameter of the disk. The union of a and b, shown in Fig. 2, is a subspace of $\mathbb{RP}^2$ homeomorphic to $\mathbb{S}^1 \vee \mathbb{S}^1$. Moreover, $\mathbb{RP}^2$ can be recovered from this subspace by attaching two copies of $\mathbb{D}^2$ — one corresponding to each semicircular region of Fig. 1 — to the concatenated loops ab and $a^{-1}b$. This is summarized in the following proposition.

Proposition 2.1. *Let a and b be the two loops in $\mathbb{S}^1 \vee \mathbb{S}^1$ shown in Fig. 2. Form a CW complex from $\mathbb{S}^1 \vee \mathbb{S}^1$ by attaching one disk to the loop ab and another disk to the loop $a^{-1}b$. The resulting topological space is homeomorphic to $\mathbb{RP}^2$.*

Let $\mathbb{D}^{2-}$ be the quotient of $\mathbb{D}^2$ obtained by gluing together two points x, y on the boundary of $\mathbb{D}^2$. Proposition 2.1 decomposes $\mathbb{RP}^2$ as a union of two copies of $\mathbb{D}^{2-}$ intersecting on their shared boundary, a subspace of $\mathbb{RP}^2$ homeomorphic to $\mathbb{S}^1 \vee \mathbb{S}^1$.

Now, suppose $\mathcal{H}$ is a 3-uniform hypergraph. For a vertex $u \in V(\mathcal{H})$, we denote by $\mathcal{H}_u$ its *link graph*, the graph on $V(\mathcal{H}) \setminus \{u\}$ whose edges vw correspond to 3-edges $uvw \in E(\mathcal{H})$. For distinct vertices u and u' of $\mathcal{H}$, we write $\mathcal{H}_{u,u'} = \mathcal{H}_u \cap \mathcal{H}_{u'}$; that is, $\mathcal{H}_{u,u'}$ is the graph on $V(\mathcal{H}) \setminus \{u, u'\}$ with edge set $E(\mathcal{H}_u) \cap E(\mathcal{H}_{u'})$.

One might attempt to build $\mathbb{RP}^2$ using the following naïve approach. Choose vertices u, u' and cycles $C, C' \subseteq \mathcal{H}_{u,u'}$ so that C and C' intersect in a single vertex v_0, implying that $C \cup C'$ is homeomorphic to $\mathbb{S}^1 \vee \mathbb{S}^1$. Let $\mathcal{A}, \mathcal{A}' \subseteq \mathcal{H}$ be sub-hypergraphs induced by the edge sets $E(\mathcal{A}) = \{ue : e \in E(C) \cup E(C')\}$ and $E(\mathcal{A}') = \{u'e : e \in E(C) \cup E(C')\}$. One hopes that $\mathcal{A}$ and $\mathcal{A}'$ are copies of $\mathbb{D}^{2-}$ whose union is homeomorphic to $\mathbb{RP}^2$, and indeed this is almost true. However, the 1-simplex uv_0 (resp. $u'v_0$) is contained in four different edges of $\mathcal{A}$ (resp. $\mathcal{A}'$), so neither $\mathcal{A}$ nor $\mathcal{A}'$ is homeomorphic to $\mathbb{D}^{2-}$.

To obtain a homeomorphic copy of $\mathbb{RP}^2$, we alter the hypergraphs $\mathcal{A}$ and $\mathcal{A}'$ to avoid these four-way intersections. The resulting construction is pictured in Fig. 3. Let v_1, v_2

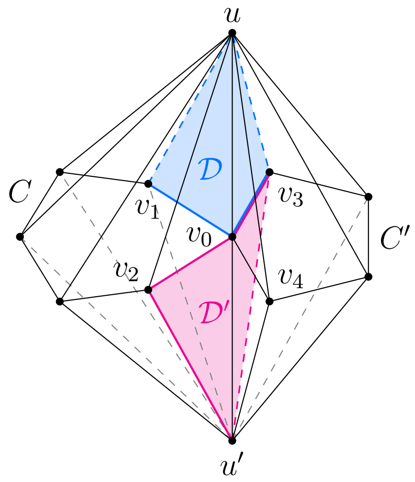


Figure 3: Building $\mathbb{RP}^2$ from cycles $C, C' \subseteq \mathcal{H}_{u,u'}$ and disks $\mathcal{D}, \mathcal{D}' \subseteq \mathcal{H}$.

be the two neighbors of v_0 in C, and let v_3, v_4 be the two neighbors of v_0 in C'. Consider the edge subsets $D = \{uv_0v_1, uv_0v_3\} \subseteq E(\mathcal{A})$ and $D' = \{u'v_0v_2, u'v_0v_3\} \subseteq E(\mathcal{A}')$, which correspond to disks with boundaries $v_0v_1uv_3v_0$ and $v_0v_2u'v_3v_0$ respectively. We locate alternate sub-hypergraphs $\mathcal{D}, \mathcal{D}' \subseteq \mathcal{H}$ homeomorphic to disks with the same boundaries, and replace D and D' with them. If $\mathcal{D}$ and $\mathcal{D}'$ are chosen appropriately, the altered hypergraphs $(\mathcal{A} \setminus D) \cup \mathcal{D}$ and $(\mathcal{A}' \setminus D') \cup \mathcal{D}'$ are homeomorphic to $\mathbb{D}^{2-}$ with shared boundary $C \cup C'$. In fact, they are created by attaching disks to $C \cup C'$ along the loops $v_0v_2 \overset{C}{\cdots} v_1v_0v_3 \overset{C'}{\cdots} v_4v_0$ and $v_0v_1 \overset{C}{\cdots} v_2v_0v_3 \overset{C'}{\cdots} v_4v_0$, respectively. Using Proposition 2.1, one can show that their union is homeomorphic to $\mathbb{RP}^2$.

3 Probabilistic Techniques

We have reduced Theorem 1.1 to finding substructures $C, C', \mathcal{D}, \mathcal{D}'$ of a 3-uniform hypergraph $\mathcal{H}$ arranged as in Fig. 3. We locate these substructures via a probabilistic approach, analyzing the likelihood that a randomly chosen subset of $V(\mathcal{H})$ will contain each of these substructures. To quantify these probabilities, we require some new definitions. Write $U \subseteq_p V$ to indicate that U is a randomly chosen subset of V, containing each vertex independently with probability p.

Definition 3.1. Fix $p, \epsilon \in (0, 1]$. Let $\mathcal{H}$ be a 3-uniform hypergraph and let $x_1, \ldots, x_4$ be four distinct vertices of $\mathcal{H}$. Sampling $U \subseteq_p V(\mathcal{H})$, let $A_{x_1x_2x_3x_4}$ be the event that there is some sub-hypergraph $\mathcal{D} \subseteq \mathcal{H}[\{x_1, \ldots, x_4\} \cup U]$ which is homeomorphic to a disk bounded by the 4-cycle $x_1x_2x_3x_4$, and which contains neither 1-simplex x_1x_3 or x_2x_4. We say the 4-cycle $x_1 \cdots x_4$ is (p, ϵ)-*disk-coverable* if $\Pr[A_{x_1\cdots x_4}] \geq 1 - \epsilon$.

Kupavskii, Polyanskii, Tomon, and Zakharov implicitly studied disk-coverability when upper-bounding the Turán number of the torus in [5]. They introduced the following related notion.

Definition 3.2. Fix $p, \epsilon \in (0,1]$. Let G be a graph and $e = xy$ an edge of G. Sample $U \subseteq_p V(G)$ and let A_e be the event that there is a cycle containing xy in $G[U \cup \{x,y\}]$. We say the edge e is (p,ϵ)-*admissible* if $\Pr[A_e] \geq 1 - \epsilon$.

The concept of admissibility is useful due to the following observation: if an edge vw in a link graph $\mathcal{H}_{u,u'}$ is (p,ϵ)-admissible, then the 4-cycle $uvu'w$ is (p,ϵ)-disk-coverable. This is because any cycle $v_0 \cdots v_\ell$ with $v = v_0$ and $w = w_\ell$ gives rise to a sub-hypergraph with edge set

$$\bigcup_{i=0}^{\ell-1} \{v_i v_{i+1} u, v_i v_{i+1} u'\},$$

which is homeomorphic to a disk with boundary $uvu'w$.

At this point, we may sketch the proof of Theorem 1.1. Given a hypergraph $\mathcal{H}$, we locate vertices $u, u', v_0, \cdots, v_4$, cycles $C, C' \subseteq \mathcal{H}_{u,u'}$, and disks $\mathcal{D}, \mathcal{D}' \subseteq \mathcal{H}$ as pictured in Fig. 3.

Proof Overview of Theorem 1.1. Let $p = 1/4$ and fix ϵ (to be determined later). Let $\mathcal{H}$ be a 3-uniform hypergraph with at least $cn^{5/2}$ edges. If c is sufficiently large in terms of p and ϵ then, using techniques from [5], we may pass to a sub-hypergraph $\mathcal{H}' \subseteq \mathcal{H}$ with at least $\frac{c}{2}n^{5/2}$ edges such that for any neighboring edges $xyz, x'yz \in \mathcal{H}'$, the 4-cycle $xyx'z$ is (p,ϵ)-disk-coverable in $\mathcal{H}$.

We locate vertices u, u', a graph $G \subseteq \mathcal{H}'_{u,u'}$, and incident edges $v_0v_1, v_0v_3 \in E(G)$ which are both (p,δ)-admissible in G, with $\delta = 1/3$. Additionally, we show that $\deg_G(v_0)$ is at most some fixed constant d, which is computed in terms of the admissibility parameters (p,δ). The details of this step may be found in [10].

Choose $v_2 \in N_G(v_0)$ uniformly at random and partition $V(G) = U_1 \cup \cdots \cup U_4$ by placing each vertex in a given set U_i independently with probability $p = 1/4$. Consider the following three events.

(A1) There are cycles C, C' satisfying the inclusions $v_0v_1 \subset C \subset G[U_1 \cup \{v_0, v_1\}]$ and $v_0v_3 \subset C' \subset G[U_2 \cup \{v_0, v_3\}]$.

(A2) The cycle C described in (A1) contains $v_1v_0v_2$ as a subpath.

(A3) There are $\mathcal{D}, \mathcal{D}' \subseteq \mathcal{H}$ homeomorphic to disks with boundaries $uv_1v_0v_3$ and $u'v_2v_0v_3$ whose non-boundary vertices are contained in U_3 and U_4, respectively. Moreover, $\mathcal{D}$ does not contain the 1-simplex uv_0, and $\mathcal{D}'$ does not contain the 1-simplex $u'v_0$.

If all three events hold simultaneously, then the structures $C, C', \mathcal{D}, \mathcal{D}'$ do not intersect except at the vertices u, u', v_0, v_1, v_3. To obtain a homeomorphic copy of $\mathbb{RP}^2$, we must additionally ensure that the structures do not contain any of these five vertices unless mentioned in (A1) and (A3). This is summarized in the following two conditions.

(B1) We have $v_3 \notin C$ and $v_1 \notin C'$.

(B2) We have $u' \notin \mathcal{D}$ and $u \notin \mathcal{D}'$.

It remains to show that the five events (A1), (A2), (A3), (B1), (B2) occur simultaneously with positive probability.

Because the edges v_0v_1 and v_0v_3 are (p,δ)-admissible in G, the event (A1) occurs with probability at least $1-2\delta = 1/3$. To additionally show that $v_3 \notin C$ and $v_1 \notin C'$, as in (B1), we check that the edges v_0v_1 and v_0v_3 are admissible (with suitable parameters) in $G - v_3$ and $G - v_1$, respectively. If $xy \in E(G)$ is a (p,δ)-admissible edge in G and $G' = G - z$ is a subgraph created by deleting a third vertex z from G, then

$$\begin{aligned} &\Pr_{U'\subseteq_p V(G')}[\nexists \text{ cycle in } G'[U' \cup \{x,y\}] \text{ containing } xy] \\ &= \Pr_{U\subseteq_p V(G)}[\nexists \text{ cycle in } G[U \cup \{x,y\}] \text{ containing } xy \mid z \notin U] \\ &\leq \frac{\Pr_{U\subseteq_p V(G)}[\nexists \text{ cycle in } G[U \cup \{x,y\}] \text{ containing } xy]}{\Pr_{U\subseteq V(G)}[z \notin U]} \leq \frac{\delta}{1-p}. \end{aligned}$$

It follows that the edges v_0v_1 and v_0v_3 are $(p, \frac{\delta}{1-p})$-admissible in $G - v_3$ and $G - v_1$, respectively. Thus, with probability at least $1 - \frac{2\delta}{1-p} = 1/9$, there are cycles C, C' satisfying (A1) and (B1).

Notice that (A2) is independent of (A1) and (B1) — the latter two events depend only on the choice of $U_1, \dots, U_4$, while (A2) depends on the choice of v_2. It follows that (A1), (B1), (A2) simultaneously hold with probability at least $\frac{1}{9}\Pr[(\text{A2})] = 1/9d$.

Lastly, we consider (A3) and (B2). Observe that uv_1v_0 and uv_3v_0 are neighboring edges of $\mathcal{H}'$, so $uv_1v_0v_3$ is (p,ϵ)-disk-coverable in $\mathcal{H}$. Similarly, $u'v_2v_0v_3$ is also (p,ϵ)-disk-coverable in $\mathcal{H}$. It follows that (A3) holds with probability at least $1 - 2\epsilon$. To additionally include (B2), we note that $uv_1v_0v_3$ and $u'v_2v_0v_3$ are $(p, \frac{\epsilon}{1-p})$-disk-coverable in $\mathcal{H} - u'$ and $\mathcal{H} - u$, respectively, by a calculation analogous to that for (B1) above. Thus, (A3) and (B2) hold simultaneously with probability at least $1 - \frac{2\epsilon}{1-p} \geq 1 - 3\epsilon$.

By a union bound, the five events (A1), (B1), (A2), (A3), (B2) hold simultaneously with probability at least $1/9d - 3\epsilon$. Thus, assuming that ϵ was chosen to satisfy $\epsilon < 1/27d$, there is a sub-hypergraph of $\mathcal{H}$ homeomorphic to $\mathbb{RP}^2$. □

References

[1] Z. Füredi and M. Simonovits, The history of degenerate (bipartite) extremal graph problems, in *Erdős centennial*, *Bolyai Soc. Math. Stud.*, vol. 25, János Bolyai Math. Soc., Budapest, 2013, pp. 169–264.

[2] A. Georgakopoulos, J. Haslegrave, R. Montgomery, and B. Narayanan, Spanning surfaces in 3-graphs, *J. Eur. Math. Soc.* **24** (2022), 303–339.

[3] P. Keevash, Hypergraph Turán problems, in *Surveys in combinatorics 2011*, *London Math. Soc. Lecture Note Ser.*, vol. 392, Cambridge Univ. Press, Cambridge, 2011, pp. 83–139.

[4] P. Keevash, J. Long, B. Narayanan, and A. Scott, A universal exponent for homeomorphs, *Israel J. Math.* **243** (2021), 141–154.

[5] A. Kupavskii, A. Polyanskii, I. Tomon, and D. Zakharov, The extremal number of surfaces, *Int. Math. Res. Not.* **2022** 13246–13271.

[6] N. Linial, What is high-dimensional combinatorics?, Random-Approx, 2008.

[7] N. Linial, Challenges of high-dimensional combinatorics, Lovász's Seventieth Birthday Conference, 2018.

[8] Z. Luria and R. J. Tessler, A sharp threshold for spanning 2-spheres in random 2-complexes, *Proc. Lond. Math. Soc.* **119** (2019), 733–780.

[9] D. Mubayi, O. Pikhurko, and B. Sudakov, Hypergraph Turán problem: some open questions, in *AIM workshop problem lists, manuscript*, 2011.

[10] M. Sankar, The Turán number of surfaces, 2022. Preprint available at arXiv:2210.11041.

[11] V. T. Sós, P. Erdős, and W. G. Brown, On the existence of triangulated spheres in 3-graphs, and related problems, *Period. Math. Hungar.* **3** (1973), 221–228.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

On universal singular exponents in equations with one catalytic parameter of order one

(Extended abstract)

Gilles Schaeffer*

Abstract

Equations with one catalytic variable and one univariate unkown, also known as *discrete difference equations of order one*, form a familly of combinatorially relevant functional equations first discussed in full generality by Bousquet-Mélou and Jehanne (2006) who proved that their power serie solutions are algebraic. Drmota, Noy and Yu (2022) recently showed that in the non linear case the singular expansions of these series have a universal dominant term of order 3/2, as opposed to the dominant square root term of generic ℕ-algebraic series. Their direct analysis of the cancellation underlying this behavior is a tour de force of singular analysis. We show that the result can instead be given a straightforward explanation by showing that the derivative of the solution series conforms to the standard square root singular behavior. Consequences also include an atypical, but generic in this situation, $n^{5/4}$ asymptotic behavior for the cumulated values of the underlying catalytic parameter.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-111`

Context and known results. Our interest is in families of bigraded combinatorial structures $\mathscr{F} = (\mathscr{F}_{n,k})_{n,k\geqslant 0}$ whose bivariate generating series satisfiy a so-called *equation with one catalytic variable and one univariate unknown* [3], or *discrete difference equation of order one* [1]. Many examples of such combinatorial structures have surfaced over the

*LIX, CNRS, Ecole polytechnique, Institut Polytechnique de Paris, Palaiseau, France. E-mail: `gilles.schaeffer@lix.polytechnique.fr`. Partially supported by the projects ANR-16-CE40-0009-01 (GATO), ANR-20-CE48-0018 (3DMaps), and ANR-21-CE48-0017 (LambdaComb).

last fifty years in the combinatorial literature, among which proeminent topics of recent interest like *lattice paths*, *rooted planar maps*, *two-stack-sortable permutations*, *normal linear lambda terms*, *intervals of the Tamari lattice*, *figting fish* or *tree parking functions* (see *e.g.* references in [1, 3, 7, 8, 5]).

Catalytic equations as a generic class of functional equations were notably studied by Bousquet-Mélou and Jehanne [3] who proved that the power series solutions of catalytic equations with one catalytic variable (and arbitrary order) are in general algebraic functions. Further explicit universal results were more recently obtained for order one catalytic equations: complexity issues were discussed by Bostan *et al* [1], while Drmota, Noy and Yu [7] exibited a universal critical exponant governing the polynomial correction in the asymptotic of the coefficients of the solution series (see also Chen [5] for more detailed universality results concerning a slightly more restricted class of functions, and Chapuy [4] for earlier partial results along the same lines as [7]).

Here we concentrate on the generic order one catalytic equation with one variable, as studied by Drmota, Noy and Yu [7]:

$$F(u) = Q\left(F(u), \tfrac{1}{u}(F(u)-f), u, t\right) \tag{1}$$

where $F(u) \equiv F(u,t) = \sum_{n,k\geqslant 0} |\mathcal{F}_{n,k}| t^n u^k$ and $f \equiv f(t) = F(0,t)$ are respectively the ordinary generating functions of elements of $\mathcal{F}$ and of $\not{f} = \cup_{n\geqslant 0}\mathcal{F}_{n,0}$, with t marking the size n and u marking the secondary parameter k (refered to as the *catalytic parameter*), and where $Q(v,w,u,t) = \sum_{i,j,k,\ell\geqslant 0} q_{i,j,k,\ell} v^i w^j u^k t^\ell$, is assumed[1] to take the form $Q(v,w,u,t) = Q_0(u)+tQ_+(v,w,u,t)$ with Q_0 and Q_+ polynomials with non negative coefficients such that $Q''_{\mathrm{vv}} + Q''_{\mathrm{vw}} + Q''_{\mathrm{ww}} \neq 0$ (non linearity condition), and $Q'_w + Q''_{\mathrm{vu}} \neq 0$ (catalytic condition).

Following Bousquet-Mélou and Jehanne [3], we consider the derivative of Equation (1) with respect to u, as given formally by the standard chain rule for derivation:

$$\begin{aligned}\frac{\partial F}{\partial u}(u) \;=\; & \left(\frac{\partial Q}{\partial v}\left(F(u), \tfrac{1}{u}(F(u)-f), u, t\right) + \frac{\partial Q}{\partial w}\left(F(u), \tfrac{1}{u}(F(u)-f), u, t\right)\tfrac{1}{u}\right)\frac{\partial F}{\partial u}(u) \\ & + \left(\frac{\partial Q}{\partial u}\left(F(u), \tfrac{1}{u}(F(u)-f), u, t\right) - \frac{\partial Q}{\partial w}\left(F(u), \tfrac{1}{u}(F(u)-f), u, t\right)\tfrac{1}{u^2}(F(u)-f)\right).\end{aligned} \tag{2}$$

Under our assumptions on Q, upon extracting coefficients of successives powers of t, the equation

$$U = U\frac{\partial Q}{\partial v}\left(F(U), \tfrac{1}{U}(F(U)-f), U, t\right) + \frac{\partial Q}{\partial w}\left(F(U), \tfrac{1}{U}(F(U)-f), U, t\right)$$

extracted from the first line of Equation (2) is seen to have a unique power series solution $U \equiv U(t)$ in $t \cdot \mathbb{Q}[t]$, and $\frac{\partial F}{\partial u}(U(t))$ is a well defined power series in t. The substitution

[1] Actually our analysis applies in a slighltly more general analytic setting, provided Q has non negative coefficients and the equation is non degenerate, non linear and catalytic, but we stick for simplicity with the polynomiality assumptions used in [3, 7] which covers most known examples.

$u = U$ then cancels the first line of Equation (2), and therefore also its second line. This implies that the series U, $V \equiv V(t) = F(U)$, and $W \equiv W(t) = \frac{F(U)-f}{U}$ satisfy the systems

$$\left\{\begin{array}{rcl} V &=& Q(V,W,U,t), \\ U &=& U\frac{\partial Q}{\partial v}(V,W,U,t) + \frac{\partial Q}{\partial w}(V,W,U,t), \\ 0 &=& \frac{\partial Q}{\partial u}(V,W,U,t) - \frac{\partial Q}{\partial w}(V,W,U,t)\frac{W}{U}, \\ f &=& V - UW, \end{array}\right. \text{ and } \left\{\begin{array}{rcl} V &=& Q(V,W,U), \\ U &=& U\frac{\partial Q}{\partial v}(V,W,U) + \frac{\partial Q}{\partial w}(V,W,U), \\ W &=& W\frac{\partial Q}{\partial v}(V,W,U) + \frac{\partial Q}{\partial u}(V,W,U), \\ f &=& V - UW. \end{array}\right. \tag{3}$$

where the second system is obtained, following Drmota, Noy and Yu [7], by replacing Line 3 of the first system by the linear combination $\frac{W}{U}(\text{Line}\,2) + (\text{Line}\,3)$.

This system was used by Drmota, Noy and Yu [7] to derive the asymptotic behavior of the coefficients of the series f under the aforementioned assumption that $Q(v,w,u,t) = Q_0(u) + tQ_+(v,w,u,t)$ where Q_0 and Q_+ are polynomials of $\mathbb{Q}[v,w,u,t]$ with non negative coefficients. Apart in the linear case and in a few other simple degenerate situations discussed in [7], these assumptions imply that the three series V, U, and W are the unique power series solutions of a system of three polynomial equations with non negative coefficients whose dependancy graph is strongly connected: the celebrated Drmota-Lalley-Woods theorem [10, Thm VII.6, p. 489] then immediately yields that these series have a common dominant singular behavior of the square root type:

$$\begin{array}{rcl} V(t) &=& a_V - b_V\sqrt{1-t/\rho} + c_V(1-t/\rho) + d_V(1-t/\rho)^{3/2} + O((1-t/\rho)^2), \\ W(t) &=& a_W - b_W\sqrt{1-t/\rho} + c_W(1-t/\rho) + d_W(1-t/\rho)^{3/2} + O((1-t/\rho)^2), \\ U(t) &=& a_U - b_U\sqrt{1-t/\rho} + c_U(1-t/\rho) + d_U(1-t/\rho)^{3/2} + O((1-t/\rho)^2), \end{array} \tag{4}$$

near their common radius of convergence $\rho > 0$, for positive constants a_V, b_V, a_W, b_W, a_U, and b_U and constants c_V, d_V, c_W, d_W, c_U, and d_U that can be explicitely expressed in terms of ρ, Q and its derivatives.

A first computation with these explicit expressions shows that a systematic cancellation of the square root terms occurs when these singular expansions are pluged in the expression $f(t) = V(t) - W(t)U(t)$, so that f is generically expected to admit a singular expansion with the next possible higher order $3/2$:

$$f = \alpha + \beta(1-t/\rho) - \gamma(1-t/\rho)^{3/2} + O((1-t/\rho)^2).$$

Using higher order expansions given by Drmota-Lalley-Wood theorem, the constant γ can be in turn expressed in terms of higher derivatives of Q. However showing that γ is positive is non trivial (as opposed to the easy statement that it is non negative), and Dmota, Noy and Yu [7] develop a quite delicate analysis to obtain this result, showing that the exponent $3/2$ is indeed universal.

Under standard technical aperiodicity conditions (see the detailed discussion in [7]), classical transfer theorems [10, Thm VI.3, p390] then imply Drmota, Noy and Yu's main result [7, Thm 2] that the coefficients of f behave as $[t^n]f(t) \sim \text{cte}\cdot\rho^{-n}\cdot n^{-5/2}$. This is a beautiful achievement to be compared for instance to the standard universal $\text{cte}\cdot\rho^{-n}\cdot n^{-3/2}$ asymptotic behavior of the coefficients of generating series of irreducible context free structures amenable to the Drmota-Lalley-Wood theorem [10, Thm VII.5, p483].

Our short derivation. The purpose of this note is to make an observation that allows to circumvent the delicate analysis of the cancellation at the heart of Drmota, Noy and Yu's approach, and to give a direct explanation of the universal asymptotic behavior of the coefficients of f. Consider the derivation of Equation (1) with respect to t :

$$\begin{array}{rcl}\frac{\partial F}{\partial t}(u) & = & \left(\frac{\partial Q}{\partial v}\left(F(u),\frac{1}{u}(F(u)-f),u,t\right)+\frac{\partial Q}{\partial w}\left(F(u),\frac{1}{u}(F(u)-f),u,t\right)\frac{1}{u}\right)\frac{\partial F}{\partial t}(u)\\ & + & \left(\frac{\partial Q}{\partial w}\left(F(u),\frac{1}{u}(F(u)-f),u,t\right)\right)\frac{(-1)}{u}\frac{\partial f}{\partial t}+\frac{\partial Q}{\partial t}\left(F(u),\frac{1}{u}(F(u)-f),u,t\right).\end{array} \tag{5}$$

In view of the chain rule for derivation, the first line of Equation (5) has the exact same form as the first line of Equation (2) so that it also cancels upon substituting $u = U$. As a consequence, the second line of Equation (5) yields the identity

$$\left(\frac{\partial Q}{\partial w}(V,W,U,t)\right)\frac{(-1)}{U}\frac{\partial f}{\partial t}+\frac{\partial Q}{\partial t}(V,W,U,t)=0. \tag{6}$$

Using again a linear combination, $\frac{1}{U}\frac{\partial f}{\partial t}$ (Line 2 of System (3)) + (Equation (6)), we obtain:

$$\frac{\partial f}{\partial t}=\frac{\partial f}{\partial t}\frac{\partial Q}{\partial v}(V,W,U,t)+\frac{\partial Q}{\partial t}(V,W,U,t).$$

Upon letting $S \equiv S(t)$ denote the unique formal power series in $\mathbb{Q}[t]$ solution of the equation $S = 1 + S\frac{\partial Q}{\partial v}(V,W,U,t)$, we obtain the larger but completely non negative system

$$\left\{\begin{array}{rcl} V & = & Q(V,W,U,t),\\ S & = & 1+S\frac{\partial Q}{\partial v}(V,W,U,t),\\ U & = & S\frac{\partial Q}{\partial w}(V,W,U,t),\\ W & = & S\frac{\partial Q}{\partial u}(V,W,U,t),\\ \frac{\partial f}{\partial t} & = & S\frac{\partial Q}{\partial t}(V,W,U,t).\end{array}\right. \tag{7}$$

Observe in particular that apart in a few degenerate cases (which are the same already listed in [7]) the dependancy graph of the four first unknowns $\{F,S,U,W\}$ in the four first equations of System (7) is strongly connected. Hence with the same assumptions as above, the hypotheses of the classical Drmota-Lalley-Woods theorem are satisfied again. As a consequence these four series have a singular expansion of the form $a_x - b_x\sqrt{1-t/\rho} + O(1-t/\rho)$ near their common dominant singularity $\rho > 0$, with computable positive constants a_x and b_x specific to each series $x \in \{V,S,W,U\}$.

Our main observation is then that, since singular expansions of the square root type are preserved via finite products and sums, the last equation of our system immediately provides a singular expansion of the square root type for the derivative of f:

$$\frac{\partial f}{\partial t}(t)=S(t)\frac{\partial Q}{\partial t}(V(t),W(t),U(t),t)=\alpha'-\beta'\sqrt{1-t/\rho}+O(1-t/\rho).$$

In particular the positivity of β' immediately follows from the positivity of the coefficients of Q and of the various constants a_* and b_*.

Under the usual aperiodicity conditions, this immediately yields that

$$[t^n]\frac{\partial f}{\partial t}(t)\sim \text{cte}\cdot\rho^{-n}\cdot n^{-3/2}\quad\text{and}\quad [t^n]f(t)=\frac{1}{n}[t^{n-1}]\frac{\partial f}{\partial t}(t)\sim \text{cte}\cdot\rho^{-n}\cdot n^{-5/2}.$$

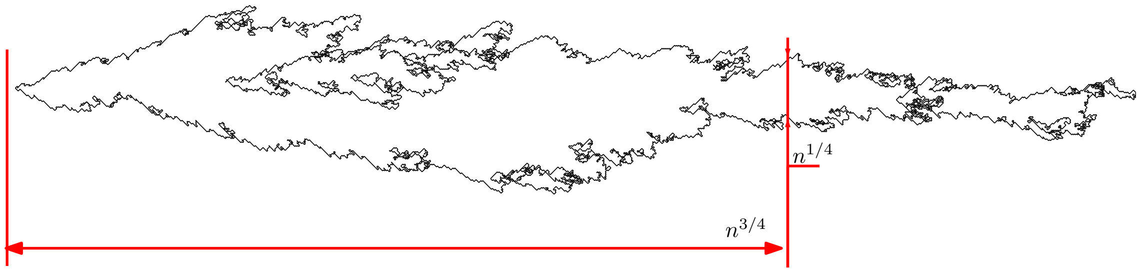


Figure 1: A fighting fish and its average parameters

Cumulated catalytic parameter. Equation like (1) typically arise from a combinatorial recursive decomposition using a so-called catalytic parameter, an auxiliary parameter which characterizes the substructures involved in the decomposition. Refinements of Equations (7) then allow to derive the asymptotic behavior of a number of interesting combinatorial parameters that are closely related to the catalytic parameter: typically the average depth of a node in the decomposition trees with root parameter 0 is of order $n^{3/4}$, and the average value of a random substructure is of order $n^{1/4}$. In this context the basic quantity that governs these behavior is the cumulated value of the catalytic parameter over all decreasing substructures of the structures of size n. This quantity is captured by the series $Z = U \cdot \frac{\partial F}{\partial u}(U)$, whose singular behavior can be derived upon derivating Equation (2) a second time with respect to u and using Equations (7): the series Z satisfies a quadratic equation with coefficients that depends on derivatives of Q evaluated at V, W and U, whose discriminant cancels at first order at the singularity, leading to a dominant term $(1-t/\rho)^{1/4}$ in the singular expansion.

A concrete example is that of fighting fish, where the cumulated label corresponds to a variant of the area (namely the area in the narrowing columns of the fish), in terms of which the width, as well as the depth of a random point of the boundary can be computed, leading to the results illustrated by Figure 1. The random instance displayed in Figure 1 has size 10000 and was generated using a random sampling algorithm for non separable maps [11] combined with the recent bijection of Duchi and Henriet [8].

Concluding remarks. From a combinatorial point of view, and in accordance with Schützenberger methodology [2], the fact that the derivative of the solution $f(t)$ of an order one equations with one catalytic variable is the solution of a system of polynomial equations with non negative coefficients suggests that if a combinatorial family admits a *first order recursive specification with one catalytic variable*, its derivative family should enjoy a *context free specification* in the sense of [10, Chapter VII.6]. This is the topic of a forthcoming article of Duchi and the author [9].

From an analytic point of view a natural question is to understand if the results can be extended to higher order catalytic equations with one catalytic variable, that are expected to share the same universal asymptotic behavior. A first remarkable achievment in this direction was recently obtained by Drmota and Hainzl for second order equations [6].

References

[1] A. Bostan, H. Notarantonio, and M. S. E. Din, *Algorithms for discrete differential equations of order 1*, in ISSAC, Lille, France, Jul 2022.

[2] M. Bousquet-Mélou, *Rational and algebraic series in combinatorial enumeration*, in Proceedings of the Internatinal Congress of Mathematicians, 2006.

[3] M. Bousquet-Mélou and A. Jehanne, *Polynomial equations with one catalytic variable, algebraic series and map enumeration*, J. Comb. Theory, Ser. B, 96 (2006), pp. 623–672.

[4] G. Chapuy, *Combinatoire d'objets planaires via leurs fonctions génératrices.* Mémoire de M2, 2006.

[5] L. Chen, *Enumeration of fully parked trees.* 2021.

[6] M. Drmota and E.-M. Hainzl, *Universal Properties of Catalytic Variable Equations*, in AofA, vol. 225 of LIPIcs, 2022, pp. 7–1.

[7] M. Drmota, M. Noy, and G.-R. Yu, *Universal singular exponents in catalytic variable equations*, Journal of Combinatorial Theory, Series A, 185 (2022), p. 105522.

[8] E. Duchi and C. Henriet, *Combinatorics of fighting fish, planar maps and tamari intervals*, in FPSAC, Bengalore. Séminaire Lotharingien de Combinatoire, no. 86B.83, 2022.

[9] E. Duchi and G. Schaeffer, *Manuscript in preparation.*

[10] P. Flajolet and R. Sedgewick, *Analytic combinatorics*, Cambridge University Press, 2009.

[11] G. Schaeffer, *Random sampling of large planar maps and convex polyhedra*, in Proceedings of the ACM Symposium on Theory of Computing (Atlanta, GA, 1999), ACM, New York, 1999, pp. 760–769.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

A new approach for the Brown-Erdős-Sós problem

(Extended abstract)

Asaf Shapira* Mykhaylo Tyomkyn†

Abstract

The celebrated Brown-Erdős-Sós conjecture states that for every fixed e, every 3-uniform hypergraph with $\Omega(n^2)$ edges contains e edges spanned by $e+3$ vertices. Up to this date all the approaches towards resolving this problem relied on highly involved applications of the hypergraph regularity method, and yet they supplied only approximate versions of the conjecture, producing e edges spanned by $e+O(\log e/\log\log e)$ vertices. We describe a completely different approach, which reduces the problem to a variant of another well-known conjecture in extremal graph theory. A resolution of the latter would resolve the Brown-Erdős-Sós conjecture up to an absolute additive constant.

DOI: https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-112

1 Introduction

1.1 Background and previous results

Some of the most well studied problems in extremal combinatorics are those asking which objects are guaranteed to appear in "dense" objects. Among notable examples are Roth's Theorem [18] on 3-term arithmetic progressions in dense sets of integers, and the Kővári-Sós-Turán Theorem [16] on bipartite subgraphs of dense graphs. In this paper we consider

*School of Mathematics, Tel Aviv University, Tel Aviv 69978, Israel. E-mail: asafico@tau.ac.il.
Supported in part by ERC Consolidator Grant 863438 and NSF-BSF Grant 20196.
†Department of Applied Mathematics, Charles University. Email: tyomkyn@kam.mff.cuni.cz.
Supported in part by ERC Synergy Grant DYNASNET 810115 and GAČR Grant 22-19073S.

a question raised by Brown, Erdős and Sós in 1973 [3, 2], which is one of the most famous open problems of this type.

Given an integer $e \geq 3$, one would expect a dense 3-uniform hypergraph (3-graph for short) to contain e edges spanned by a small number of vertices. To quantify this, let (v, e)-*configuration* denote a set of e edges spanned by at most v vertices. The Brown–Erdős–Sós Conjecture (BESC) states that for every fixed $e \geq 3$ and all large enough n, every 3-graph with $\Omega(n^2)$ edges contains an $(e+3, e)$-configuration. Despite a lot of effort over the past 50 years, the BESC is only known to hold for $e = 3$, due to a result of Ruzsa and Szemerédi [21].

Since even the $e = 4$ case of the BESC seems hopeless, it is natural to try to prove approximate versions of the conjecture, namely that 3-graphs with $\Omega(n^2)$ edges contain $(e + f(e), e)$-configurations, for some slowly growing function f. The first result of the above type was obtained by Sárközy and Selkow [22] who showed that every 3-graph with $\Omega(n^2)$ edges contains for every fixed e an $(e + 2 + \lfloor \log_2 e \rfloor, e)$-configuration. This was improved by Solymosi and Solymosi [23] for the special case $e = 10$ from 15 to 14 vertices. A general asymptotic improvement of the result of [23] was obtained recently by Conlon, Gishboliner, Levanzov and Shapira [8], who proved the existence of $(e + O(\log e/ \log\log e), e)$-configurations.

Besides its intrinsic interest, the BESC turned out to be one of the most influential problems in extremal combinatorics. For example, the proof of the case $e = 3$ [21] was one of the first applications of Szemerédi's regularity lemma [24], and further introduced the famous graph removal lemma. One of the main motivations for the development of the celebrated hypergraph regularity method [11, 17, 19, 20, 26] was the hope that it will lead to a resolution of BESC. While this did not materialize, the hypergraph regularity method was instrumental in the latest works [8, 23]. However, although the above proofs rely on highly involved applications of the hypergraph regularity method, it appears that the following natural approximate version of the BESC is beyond their reach.

Conjecture 1.1 (Constant deficiency BESC)**.** *There is an absolute constant d so that for every e and every large enough n, every 3-graph with $\Omega(n^2)$ edges contains an $(e + d, e)$-configuration.*

1.2 A new approach for Conjecture 1.1

Our aim in this paper is to reduce Conjecture 1.1 to a problem involving graphs. Let us denote by $\mathrm{ex}(n, H)$ the maximum number of edges in an n vertex graph not containing a copy of H as a subgraph. The Kővári-Sós-Turán Theorem [16] which we mentioned above, states that for every fixed $t \leq s$, we have $\mathrm{ex}(n, K_{s,t}) = O(n^{2-1/t})$ where $K_{s,t}$ is the complete bipartite graph with parts of size t and s. This bound is known to be tight for large s, see [4] for recent progress and references. One of the main research directions in extremal graph theory is to obtain better bounds for sparser bipartite graphs. One such problem was raised by Erdős [9], who conjectured that if H is a t-degenerate bipartite graph then $\mathrm{ex}(n, H) = O(n^{2-1/t})$. While there are some approximate results towards this conjecture [1, 10, 13, 15], the question is open even for $t = 2$. Note that in general, the conjectured bound $O(n^{2-1/t})$ for t-degenerate bipartite graphs cannot be improved since the aforementioned $K_{s,t}$ is t-degenerate. In particular, the bound is tight for every t-degenerate H which contains a copy of $K_{s,t}$. In light of this, Conlon [5] conjectured that if we assume that a t-degenerate bipartite graph H has no $K_{t,t}$ then we have $\mathrm{ex}(n, H) = O(n^{2-1/t-\delta})$ for some $\delta = \delta(H) > 0$. Lending plausibility to this conjecture, Sudakov and Tomon [25] showed that if all vertices in one of the parts of H have degree at most t but H has no $K_{t,t}$ then $\mathrm{ex}(n, H) = o(n^{2-1/t})$. For $t = 2$ Conlon's conjecture can be stated as:

Conjecture 1.2 (Conlon [5])**.** *For every 2-degenerate C_4-free bipartite graph H there exists a constant $\delta = \delta(H) > 0$ such that*

$$ex(n, H) = O(n^{3/2-\delta}) .$$

There are several results supporting Conjecture 1.2. For example, Conlon and Lee [7] proved that if H is a bipartite graph so that each vertex in one of H's sides has maximum degree 2 (such a graph is clearly 2-degenerate) and H is C_4-free then $\mathrm{ex}(n, H) = O(n^{3/2-\delta})$ for some $\delta = \delta(H) > 0$. Further results in this direction were obtained in [6, 14].

Let $\mathcal{H}_{k,t}$ be the family of 2-degenerate graphs on k vertices and $2k - t$ edges. We raise the following weaker version of Conjecture 1.2.

Conjecture 1.3. *There are absolute constants t, k_0 such that for every $k \geq k_0$ and large enough n, every graph with $\Omega(n^{3/2})$ edges contains a copy of* **some** *$H \in \mathcal{H}_{k,t}$.*

Let us briefly explain why Conjecture 1.3 is indeed weaker than Conjecture 1.2. It is not hard to see that for every t and large enough k, the family $\mathcal{H}_{k,t}$ contains C_4-free graphs (see Claim 3.1). Conjecture 1.2 then states that if G has $\Omega(n^{3/2})$ edges then G should contain a copy of **every** $H \in \mathcal{H}_{k,t}$ which is C_4-free, while Conjecture 1.3 only asks G to contain a copy of some $H \in \mathcal{H}_{k,t}$. Note also that Conjecture 1.3 is weaker than the statement that for every $k \geq k_0$ we have $\mathrm{ex}(n, H) = o(n^{3/2})$ for some $H \in \mathcal{H}_{k,t}$, which is itself weaker than Conjecture 1.2.

Our main result in this paper is the following alternative approach for resolving Conjecture 1.1.

Theorem 1.4. *Conjecture 1.3 implies Conjecture 1.1.*

Before turning to the proof of Theorem 1.4, we mention that it might very well be the case that in Conjecture 1.3 we can replace the lower bound $\Omega(n^{3/2})$ by $\Omega(n^{3/2-\delta})$ for some $\delta = \delta(k) > 0$. Indeed, this bound is implied by Conjecture 1.2. It is not hard to see that in this case the proof of Theorem 1.4 would give that for some absolute constant d and for every e there is $\varepsilon = \varepsilon(e) > 0$ so that one can find $(e + d, e)$-configurations in every 3-graph with $n^{2-\varepsilon}$ edges. Such a result would be an approximate version of a conjecture suggested by Gowers and Long [12], stating that 3-graphs with $n^{2-\varepsilon}$ edges contain $(e+4, e)$-configurations.

2 Proof of Theorem 1.4

To avoid confusion, we will refer to edges of a 3-graph as hyperedges. Fix $e \geq 3$ and let $\mathcal{G}$ be a 3-graph with n vertices and $\Omega(n^2)$ hyperedges. We will rely on the well known observation that in the context of the BESC one can assume that $\mathcal{G}$ is linear and 3-partite on vertex sets $(\mathcal{A}, \mathcal{B}, \mathcal{C})$. We now apply a variant of the construction of Solymosi and Solymosi [23]. Given $\mathcal{G}$, define an auxiliary bipartite multigraph G' as follows. Set $V(G') = (A, B)$ where $A = \binom{\mathcal{A}}{2}$ and $B = \binom{\mathcal{B}}{2}$. For two vertices $\{a_1, a_2\} \in A$ and $\{b_1, b_2\} \in B$ put an edge between them if there is a $c \in \mathcal{C}$ so that a_1b_1c and a_2b_2c are hyperedges of $\mathcal{G}$, and (independently) put an edge between them if there is a $c' \in \mathcal{C}$ such that a_1b_2c' and a_2b_1c' are hyperedges of $\mathcal{G}$. Since $\mathcal{G}$ is linear, each pair of vertices in G' are connected by at most 2 edges. If we let $d(c)$ denote the degree of a vertex $c \in \mathcal{C}$ in $\mathcal{G}$ then

$$|E(G')| = \sum_{c \in \mathcal{C}} \binom{d(c)}{2} \geq |\mathcal{C}| \binom{\frac{1}{|\mathcal{C}|} \sum_{c \in \mathcal{C}} d(c)}{2} = |\mathcal{C}| \binom{|E(\mathcal{G})|/|\mathcal{C}|}{2} \geq \frac{|E(\mathcal{G})|^2}{4|\mathcal{C}|}.$$

Since $e(\mathcal{G}) = \Omega(n^2)$, $|\mathcal{C}| \leq n$, and $|V(G')| \leq n^2$, we obtain $|E(G')| = \Omega(|V(G')|^{3/2})$. Since, as noted above, each pair of vertices in G' are connected by at most 2 edges, G' has a simple subgraph G which also contains $\Omega(|V(G)|^{3/2})$ edges. Therefore, if k_0 and t are the constants from Conjecture 1.3 and n is large enough, then we may assume the following.

Observation 2.1. *For every $k_0 \leq k \leq e$, the graph G contains a 2-degenerate bipartite graph F on k vertices with at least $2k - t$ edges.*

We would now like to understand what kind of (v, e)-configuration in $\mathcal{G}$ we get by "unpacking" each of the graphs F in Observation 2.1. Optimistically, if $v_1, \ldots, v_k$ is the ordering of $V(F)$ certifying its 2-degeneracy, then every time we add a vertex v_i to $v_1, \ldots, v_{i-1}$ of degree 2 to the previous vertices, we expect to get 4 new vertices in $\mathcal{G}$; these are c_1, c_2 and either a_1, a_2 (if $v_i \in A$) or b_1, b_2 (if $v_i \in B$). We also expect to get 4 new hyperedges in $\mathcal{G}$; these are the 4 hyperedges that correspond to the 2 new edges in G that connect v_i to 2 of the vertices $v_1, \ldots, v_{i-1}$. If this holds for all but a bounded number of F's vertices, then we will get a $(4k, 4k - O_k(1))$ configuration, hence taking $k \approx e/4$ would finish the proof. Unfortunately, we do not know how to prove such a statement, since in certain cases (see below) some of the 4 vertices/hyperedges might have already appeared when adding one of the previous vertices v_j. Instead, the main idea in Lemma 2.2 below is to show that F gives rise to a $(e' + d, e')$-configuration, so that if e' is not very close to $4k$ (as in the optimistic analysis above) then we have $d \leq 0$. It is then easy to show how repeated applications of Lemma 2.2 give Theorem 1.4. In what follows G and $\mathcal{G}$ are those we discussed above.

Lemma 2.2. *Let $k \geq t \geq 4$ be integers, and suppose F is a 2-degenerate subgraph of G with k vertices and $2k - t$ edges. Then $\mathcal{G}$ contains a subgraph $\mathcal{F}$ such that*

(1) $|V(\mathcal{F})| - 4t \leq |E(\mathcal{F})| \leq 4k$, and

(2) Either $|E(\mathcal{F})| \geq 4k - 10^4 t^3$ or $|E(\mathcal{F})| \geq |V(\mathcal{F})| > 0$.

For the proof of Lemma 1.4 we refer the reader to the full version of the paper. We will now show how to derive Theorem 1.4 from Lemma 2.2. Assuming Conjecture 1.3 holds with constants t, k_0 we show that Conjecture 1.1 holds with $d = \max\{24k_0, 3(4t + 10^4t^3)\}$. Indeed, we claim that for every $0 \le e' \le e$ we can find e' hyperedges in $\mathcal{G}$ spanned by at most $e'+d$ vertices. If $e' \le \max\{8k_0, 4t+10^4t^3\}$, we just take e' arbitrary hyperedges from $\mathcal{G}$. For larger e' we apply Lemma 2.2 with the above t and with $k = \lfloor e'/4 \rfloor \ge k_0$ (by Observation 2.1 we know that G contains an F with these parameters). If the lemma returns a configuration $\mathcal{F}'$ whose number of edges satisfies $e' - 10^4t^3 - 4 \le |E(\mathcal{F}')| \le e'$ (and is on at most $e' + 4t$ vertices), we just add to $\mathcal{F}'$ arbitrarily chosen $e' - |E(\mathcal{F}')| \le 10^4t^3 + 4$ hyperedges to get a set of e' edges on at most $e' + d$ vertices. Otherwise, we have $|E(\mathcal{F}')| \ge |V(\mathcal{F}')| > 0$ so we can remove $\mathcal{F}'$ from $\mathcal{G}$ and then restart the process with $e'' = e' - |E(\mathcal{F}')|$ (the 3-graph $\mathcal{G} \setminus \mathcal{F}$ still has $\Omega(n^2)$ hyperedges assuming n is large). We will obtain a set $\mathcal{F}''$ of e'' hyperedges on at most $e'' + d$ vertices, and can then return $\mathcal{F}'' \cup \mathcal{F}'$ as the set of e' hyperedges on at most $e' + d$ vertices.

3 C_4-free graphs in $\mathcal{H}_{k,t}$

We say that a graph is *exactly-*$(2,t)$*-degenerate* if it can be obtained from a set of t isolated vertices by repeatedly adding new vertices of degree exactly 2. Note that every exactly-$(2,t)$-degenerate graph belongs to $\mathcal{H}_{k,t}$. The following claim shows that $\mathcal{H}_{k,t}$ contains not only C_4-free graphs, but in fact graphs of arbitrary large girth.

Claim 3.1. *For every g there is $t = t(g)$ so that for every $k \ge t$, there is a k-vertex exactly-$(2,t)$-degenerate bipartite graph of girth at least g.*

Proof. We claim that starting with an independent set of size $t = t(g)$ (to be chosen later), we can repeatedly add vertices so that each k-vertex graph in the sequence is exactly-$(2,t)$-degenerate, bipartite, of girth at least g, and in addition satisfies the following two conditions: (i) it has maximum degree at most 8 and (ii) it has a bipartition into two set of sizes $\lceil k/2 \rceil$ and $\lfloor k/2 \rfloor$. The initial independent set under a balanced bipartition clearly satisfies these two conditions, so let us show how to add a vertex and maintain them. Suppose the graph has $k-1$ vertices and bipartition into sets A, B satisfying $|A| \le |B|$. Since it has maximum degree at most 8, it contains $O(k)$ pairs of vertices connected by a path of length at most $g-2$. Since the average degree of the vertices in B is less than 4, at least half the vertices have degree at most 7. Hence, at least $\binom{(k-1)/4}{2} \ge \frac{k^2}{50}$ of the pairs of vertices in B both have degree at most 7. Assuming t is large enough so that $k \ge t$ satisfies $\frac{k^2}{50} - O(k) > 1$, we thus have a pair of vertices $u, v \in B$ so that both of them have degree at most 7 and there is no path of length at most $g-2$ connecting them. Hence, we can add a new vertex to A and connect it to u and v. □

Acknowledgement: We would like to thank David Conlon for useful discussions.

References

[1] N. Alon, M. Krivelevich, and B. Sudakov. Turán numbers of bipartite graphs and related Ramsey-type questions, Comb. Probab. Comput 12 (2003), 477–494.

[2] W. G. Brown, P. Erdős and V.T. Sós, Some extremal problems on r-graphs, New Directions in the Theory of Graphs, Proc. 3rd Ann Arbor Conference on Graph Theorey, Academic Press, New York, 1973, 55-63.

[3] W. Brown, P. Erdős and V. Sós. On the existence of triangulated spheres in 3-graphs, and related problems, Periodica Mathematica Hungarica, 3(3–4) (1973), 221-228.

[4] B. Bukh, Extremal graphs without exponentially-small bicliques, manuscipt 2022.

[5] D. Conlon, private commuication, 2022.

[6] D. Conlon, O. Janzer and J. Lee, More on the extremal number of subdivisions, Combinatorica 41 (2021), 465-494.

[7] D. Conlon and J. Lee, On the extremal number of subdivisions, Int. Math. Res. Not. 2021, 9122-9145.

[8] D. Conlon, L. Gishboliner, Y. Levanzov and A. Shapira, A new bound for the Brown–Erdős–Sós problem, J. Combin. Theory Ser. B. 158 (2023), 1-35.

[9] P. Erdős, Some recent results on extremal problems in graph theory. Results, Theory of Graphs (Internat. Sympos., Rome, 1966), pages 117–123, 1967.

[10] Z. Füredi, On a Turán type problem of Erdős, Combinatorica 11 (1991), 75–79.

[11] W. T. Gowers, Hypergraph regularity and the multidimensional Szemerédi theorem, Ann. of Math. 166 (2007), 897–946.

[12] W. T. Gowers and J. Long, The length of an s-increasing sequence of r-tuples, Combin. Probab. Comput. 30 (2021), 686–721.

[13] A. Grzesik, O. Janzer and Z. L. Nagy, The Turán number of blow-ups of trees, J. Combin. Theory Ser. B. 156 (2022), 299-309.

[14] O. Janzer, The extremal number of the subdivisions of the complete bipartite graph, SIAM J. Discrete Math. 34 (2020), 241-250.

[15] O. Janzer, Disproof of a conjecture of Erdős and Simonovits on the Turán number of graphs with minimum degree 3, Int. Math. Res. Not., to appear.

[16] T. Kővári, V. T. Sós and P. Turán, On a problem of K. Zarankiewicz, Colloquium Math. 3 (1954), 50-57.

[17] B. Nagle, V. Rödl and M. Schacht, The counting lemma for regular k-uniform hypergraphs, Random Structures Algorithms 28 (2006), 113–179.

[18] K.F. Roth, On certain sets of integers, J. London Math. Soc. 28 (1953), 104-109.

[19] V. Rödl and J. Skokan, Regularity lemma for k-uniform hypergraphs, Random Structures Algorithms 25 (2004), 1–42.

[20] V. Rödl and J. Skokan, Applications of the regularity lemma for uniform hypergraphs, Random Structures Algorithms 28 (2006), 180–194.

[21] I. Ruzsa and E. Szemerédi, Triple systems with no six points carrying three triangles, in Combinatorics (Keszthely, 1976), Coll. Math. Soc. J. Bolyai 18, Volume II, 939-945.

[22] G. N. Sárközy and S. Selkow, An extension of the Ruzsa-Szemerédi theorem, Combinatorica 25 (2004), 77-84.

[23] D. Solymosi and J. Solymosi, Small cores in 3-uniform hypergraphs, J. Combin. Theory Ser. B. 122 (2017), 897-910.

[24] E. Szemerédi, Regular partitions of graphs, In: *Proc. Colloque Inter. CNRS*, 1978, 399-401.

[25] B. Sudakov and I. Tomon, Turán number of bipartite graphs with no $K_{t,t}$, Proc. Amer. Math. Soc. 148 (2020), 2811-2818.

[26] T. Tao, A variant of the hypergraph removal lemma, J. Combin. Theory Ser. A 113 (2006), 1257–1280.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Semidegree, edge density and antidirected subgraphs

(Extended abstract)

Maya Stein* Camila Zárate-Guerén†

Abstract

An oriented graph is called *antidirected* if it has no directed path with 2 edges. We prove that asymptotically, any oriented graph D of minimum semidegree greater than $\frac{k}{2}$ contains every balanced antidirected tree of bounded degree and with k edges, and D also contains every antidirected subdivision H of a sufficiently small complete graph K_h, with a mild restriction on the lengths of the antidirected paths in H replacing the edges of K_h, and with H having a total of k edges.
Further, we address a conjecture of Addario-Berry, Havet, Linhares Sales, Reed and Thomassé stating that every digraph on n vertices and with more than $(k-1)n$ edges contains all antidirected trees with k edges. We show that their conjecture is asymptotically true in oriented graphs for balanced antidirected trees of bounded degree and size linear in n.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-113`

1 Introduction

In extremal graph theory, a typical task is to determine conditions on the minimum or the average degree of a graph G (the 'host graph') which guarantee that G contains some

*University of Chile, Chile. E-mail: `mstein@dim.uchile.cl`. Supported by ANID Fondecyt Regular Grant 1221905, by FAPESP-ANID Investigación Conjunta grant 2019/13364-7, by RandNET (RISE project H2020-EU.1.3.3) and by ANID PIA CMM FB210005.
†University of Birmingham, UK. E-mail: `ciz230@student.bham.ac.uk`. Supported by the University of Birmingham, by RandNET (RISE project H2020-EU.1.3.3) and by ANID PIA CMM FB210005.

specific subgraph. We study this type of question for oriented host graphs and restricting ourselves to finding antidirected subgraphs.

We present three theorems. The first of these relates high semidegree to the existence of balanced antidirected trees, where we call an oriented tree *balanced* if its bipartition classes have the same size.

Theorem 1.1. *For all $\eta \in (0,1)$, $c \in \mathbb{N}$ there is n_0 such that for all $n \geq n_0$ and $k \geq \eta n$, every oriented graph D on n vertices with $\delta^0(D) > (1+\eta)\frac{k}{2}$ contains every balanced antidirected tree T with k edges and with $\Delta(T) \leq (\log(n))^c$.*

The second theorem relates high semidegree to the existence of antidirected subdivisions of complete graphs. For $h, k \in \mathbb{N}$, consider any subdivision H of K_h where each edge $e \in E(K_h)$ is substituted by a path with $g(e)$ edges, with $\sum_{e \in E(K_h)} g(e) = k$, and such that the edges of K_h with $g(e) < 3$ induce a forest in K_h. If H has antidirected orientations, then call any antidirected orientation of H a *long k-edge antidirected subdivision of K_h*.

Theorem 1.2. *For all $\eta \in (0,1)$ there are $n_0 \in \mathbb{N}$, $\gamma > 0$ such that for each $n \geq n_0$, each $k \geq \eta n$ and each $h \leq \gamma\sqrt{n}$ the following holds. Every oriented graph D on n vertices with $\delta^0(D) > (1+\eta)\frac{k}{2}$ contains each long k-edge antidirected subdivision of K_h.*

The third theorem related high edge density to the existence of balanced antidirected trees.

Theorem 1.3. *For all $\eta \in (0,1)$, $c \in \mathbb{N}$, there is $n_0 \in \mathbb{N}$ such that for every $n \geq n_0$ and every $k \geq \eta n$, every oriented graph D on n vertices with more than $(1+\eta)(k-1)n$ edges contains each balanced antidirected tree T with k edges and $\Delta(T) \leq (\log(n))^c$.*

Each of our theorems will be motivated and discussed in one of the sections below (the sections follow the same order we chose for stating the theorems here). We provide a discussion of the context of the results and include a sketch of the proof of each of the results. We refer to [14] for more discussion and full proofs.

2 Paths and trees

Dirac (see [5]) observed that if an undirected connected graph G on at least $k+1$ vertices satisfies $\delta(G) \geq \frac{k}{2}$, then G contains a k-edge path (here and later, k is any natural number, independent of the order of the host graph). Trying to translate this result to oriented graphs, a natural possibility would be to replace the minimum degree by the minimum semidegree $\delta^0(D)$, which is defined as the minimum over all the in- and all the out-degrees of the vertices in D, and to ask for certain oriented paths in D.

Jackson [7] showed that every oriented graph D with $\delta^0(D) > \frac{k}{2}$ contains the directed path on k edges. The first author conjectured [13] that in this result, the directed path can be replaced with any oriented path of the same length. This conjecture is best possible for directed paths [7] and also for antidirected paths: observe that in an ℓ-blow-up of the directed triangle (where each vertex is replaced with ℓ independent vertices), any antidirected path covers at most 2ℓ vertices. We show that the conjecture from [13] is asymptotically true:

Corollary 2.1. *For all $\eta \in (0,1)$ there is n_0 such that for all $n \geq n_0$ and $k \geq \eta n$ every oriented graph D on n vertices with $\delta^0(D) > (1+\eta)\frac{k}{2}$ contains every antidirected path with k edges.*

To see that Corollary 2.1 follows from Theorem 1.1, observe that any path either is balanced, or can be extended by one to become balanced. (In the latter case we apply Theorem 1.1 with a sufficiently smaller η.)

Note that the class of trees considered in Theorem 1.1 is very similar to antidirected paths, not only because of the balancedness, but also because of the bounded degree. In graphs, there is a very well-known result for finding bounded degree trees by Komlós, Sárközy and Szemerédi [10]. It states that asymptotically, every graph of minimum degree larger than $\frac{n}{2}$ contains every spanning tree of maximum degree at most $O(\frac{n}{\log n})$. Recently, this result was extended to digraphs by Kathapurkar and Montgomery [8]. Theorem 1.1 can be considered as a version for smaller antidirected trees of Kathapurkar and Montgomery's result, in oriented graphs.

We actually prove a stronger version of Theorem 1.1, namely Theorem 2.2 below, which allows us to choose where the root of the antidirected tree goes. This more general result will be useful in the proof of Theorem 1.3.

Theorem 2.2. *For all $\eta \in (0,1)$, $c \in \mathbb{N}$ there is n_0 such that for all $n \geq n_0$ and $k \geq \eta n$ the following holds for every oriented graph D on n vertices with $\delta^0(D) > (1+\eta)\frac{k}{2}$, and every balanced antidirected tree T with k edges and $\Delta(T) \leq (\log(n))^c$. For each set $V^* \subseteq V(D)$ with $|V^*| \geq \eta n$ and for each $x \in V(T)$, there is an embedding of T in D with x mapped to V^*.*

Sketch of the proof of Theorem 2.2

Given an oriented graph D and an antidirected tree T fulfilling the conditions of the theorem, we apply the digraph regularity lemma to D to find a partition into a bounded number of clusters C_i. The reduced oriented graph R will have a minimum semidegree similar to the one of D (proportionally). Let $x \in V(T)$ and $V^* \subseteq V(D)$ be given, with $|V^*| \geq \eta n$. Note that at least one cluster C_i contains $\eta|C_i|$ vertices from V^*. Let C^* be one such cluster.

Next, we need the concept of a *connected antimatching*: this is a set M of disjoint edges in D such that every pair of edges in M is connected by an *antidirected walk* or simply *antiwalk*, which is a sequence of edges that alternate direction. The *length* of an antidirected walk is its number of edges, where we count repeated edges once for each time they appear.

We show that the minimum semidegree in the reduced oriented graph R suffices to ensure that R contains a large connected antimatching M. Further, the antiwalks connecting the edges of M have bounded length:

Lemma 2.3 (Lemma 4.8 in [14]). *Let $t \in \mathbb{N}^+$, let D be an oriented graph with $\delta^0(D) \geq t$, and let $w \in V(D)$. Then D has a connected antimatching $M = \{a_ib_i\}_{1\leq i\leq t}$ of size t, with $w = a_1$, and such that, for every $1 \leq i \leq t$, there is an antiwalk of length at most $8t$ containing a_ib_i and a_1b_1.*

Now we turn to our antidirected tree T. We decompose T into a family $\mathcal{T}$ of small subtrees, connected by a constant number of vertices. This type of decomposition has been widely used lately, appearing for the first time in [3]. We prove that it is possible to assign the trees in $\mathcal{T}$ to edges of M in a way that they will fit comfortably into the corresponding clusters, while respecting the orientations. We let P_i denote the set of trees in $\mathcal{T}$ that are assigned to the clusters associated to the edge $a_ib_i \in M$.

We now embed T as follows. In each step, we embed one small tree $S \in \mathcal{T}$. When we choose a new small tree to embed, we make sure that we keep the embedded part connected in the underlying tree. We embed the first d levels of S into the clusters of an antiwalk W_S in R that starts in the cluster containing the image of the parent of the root of S and ends in a_ib_i, if $S \in P_i$. The remaining levels of S are embedded into the clusters corresponding to a_i and b_i.

Since T has bounded maximum degree, the union of the first d levels of the trees in $\mathcal{T}$ is very small, and therefore it is not a problem that the first d levels of each $S \in \mathcal{T}$ are embedded in the connecting antidirected walk W_S. After going through all $S \in \mathcal{T}$, we have embedded all of T. For the full proof see [14]. $\square$

3 Subdivisions and cycles

Mader [11] proved that there is a function $g(h)$ such that every (undirected) graph of minimum degree at least $g(h)$ contains a subdivision of the complete graph K_h. Thomassen [15] showed that a direct translation of this result to digraphs is not true. Mader [12] suggested to replace the subdivision of the complete digraph with the *transitive tournament*, i.e. the tournament without directed cycles:

Conjecture 3.1 (Mader [12]). *There is a function $f(h)$ such that every digraph of minimum outdegree at least $f(h)$ contains a subdivision of the transitive tournament of order h.*

This conjecture is open even for $h = 5$. Aboulker, Cohen, Havet, Lochet, Moura and Thomassé [1] observed that in Conjecture 3.1, the minimum outdegree can be replaced with the minimum semidegree, and the resulting conjecture is equivalent to Conjecture 3.1. Our Theorem 1.2 can be seen as a version of Conjecture 3.1 for oriented graphs and antidirected subdivisions of K_h.

For $h = 3$, the objects found in Theorem 1.2 are antidirected cycles. In the existing literature, there are already a number of results on finding oriented cycles with conditions on the minimum semidegree. We will quickly discuss those related to antidirected cycles.

For an oriented cycle C, the *cycle type* of C is defined as the number of forward edges minus the number of backwards edges of C. Note that antidirected cycles have cycle type

0. Kelly, Kühn and Osthus [9] showed that for each $k \geq 3$ and $\eta > 0$ every large enough n-vertex oriented graph of minimum semidegree at least ηn contains all oriented cycles of length at most k and cycle type 0. Further, $\delta^0(D) \geq \frac{3n}{8} + o(n)$ is enough to find a copy of any oriented cycle of length between 3 and n in an oriented graph D [9]. Both results give (quite different) bounds on the semidegree for antidirected cycles. While in the first result, the cycle is small compared to n, in the second result there are antidirected cycles of any even length. Theorem 1.2 provides us with an intermediate semidegree bound for finding an antidirected cycle of medium length:

Corollary 3.2. *For all $\eta \in (0,1)$ there is n_0 such that for all $n \geq n_0$ and $k \geq \eta n$, every oriented graph D on n vertices with $\delta^0(D) > (1+\eta)\frac{k}{2}$ contains any antidirected cycle of length at most k.*

Indeed, this corollary follows from Theorem 1.2 since any antidirected cycle with more than four edges can be expressed as a long antisubdivision of K_3, while antidirected C_4 is guaranteed by the results of [9].

Sketch of the proof of Theorem 1.2

Let D be an oriented graph satisfying the conditions of Theorem 1.2. Let a long k-antisubdivision of K_h be given and remove two consecutive inner vertices (along with all adjacent edges) from one of the long antidirected paths of this antisubdivision. Keep removing two vertices from other long antidirected paths until we are left with an antidirected tree T. Denote by $\mathcal{P}$ the set of long antidirected paths of which we removed vertices.

As in the proof of Theorem 2.2, we find a connected antimatching M in the reduced graph R of D. We embed the branch vertices of the antisubdivision into a pair of clusters B, C, such that BC is some fixed edge of M. We start embedding the long antipaths in the clusters corresponding to edges of M, using the antiwalks given by Lemma 2.3 to move between the matching edges.

The only vertices left are the ones removed at the beginning. Since their neighbours are already embedded in $B \cup C$, they can be embedded in $B \cup C$ by regularity. For all details see [14].

4 Edge density

In 1970, Graham [6] confirmed a conjecture he attributes to Erdős: for every antidirected tree T there is a constant c_T such that every sufficiently large directed graph D on n vertices and with at least $c_T n$ edges contains T. A similar statement is false for other oriented trees [2, 4]. In 1982, Burr [4] gave an improvement of Graham's result: Every n-vertex digraph D with more than $4kn$ edges contains each antidirected tree T on k edges, and provides an example where $(k-1)n$ edges are not sufficient. In 2013, Addario-Berry, Havet, Linhares Sales, Reed and Thomassé [2] formulate the following conjecture.

Conjecture 4.1 (Addario-Berry et al. [2]). *Every n-vertex digraph D with more than $(k-1)n$ edges contains each antidirected tree on k edges.*

Theorem 1.3 implies that Conjecture 4.1 is approximately true in oriented graphs for all balanced antidirected trees of bounded maximum degree.

Sketch of the proof of Theorem 1.3

Given the antidirected tree T and the oriented graph D as in the theorem, we start by finding a non-empty oriented subgraph D' of D where each vertex has either out-degree at least $\frac{k}{2}$ or out-degree 0, and either in-degree at least $\frac{k}{2}$ or in-degree 0 (see Lemma 7.1 in [14]). We construct a new oriented graph D'' consisting of four copies of D', two of them with all edges reversed. Because of the way we put those copies together, D'' will have minimum semidegree greater than $\frac{k}{2}$.

Using Theorem 2.2, we embed T into D'', with the root v of T embedded in one of the copies of D' with the original orientations. Taking a little more care, we can ensure that an edge at v is also embedded in this copy. It is then easy to deduce that all of T is embedded into the same copy. Since $D' \subseteq D$, we proved the statement. For the full proof see [14]. $\square$

References

[1] Pierre Aboulker, Nathann Cohen, Frédéric Havet, William Lochet, Phablo Moura, and Stéphan Thomassé. Subdivisions in digraphs of large out-degree or large dichromatic number. *The Electronic Journal of Combinatorics*, 26:3.19, 2019.

[2] Louigi Addario-Berry, Frédéric Havet, Cláudia Linhares Sales, Bruce Reed, and Stéphan Thomassé. Oriented trees in digraphs. *Discrete Mathematics*, 313(8):967–974, 2013.

[3] Miklós Ajtai, János Komlós, and Endre Szemerédi. On a conjecture of Loebl. *Proc. of the 7th International Conference on Graph theory, combinatorics, and algorithms*, 1(2):1135–1146, 1995.

[4] Stefan Burr. Subtrees of directed graphs and hypergraphs. *Proc. of the 11th Southeastern Conference on Combinatorics, Graph Theory and Combinatorics (Florida Atlantic Univ., Boca Raton, Fla.)*, 28(I):227–239, 1980.

[5] Paul Erdős and Tibor Gallai. On maximal paths and circuits of graphs. *Acta Mathematica Academiae Scientiarum Hungarica*, 10:337–356, 1959.

[6] Ronald L. Graham. On subtrees of directed graphs with no path of length exceeding one. *Canad. Math. Bull.*, 13(3):329–332, 1970.

[7] Bill Jackson. Long paths and cycles in oriented graphs. *Journal of Graph Theory*, 5(2):145–157, 1981.

[8] Amarja Kathapurkar and Richard Montgomery. Spanning trees in dense directed graphs. *J. Combinatorial Theory Series B*, 156:223–249, 2022.

[9] Luke Kelly, Daniela Kühn, and Deryk Osthus. Cycles of given length in oriented graphs. *J. Combinatorial Theory Series B*, 100(3):251–264, 2010.

[10] János Komlós, Gábor N. Sárközy, and Endre Szemerédi. Proof of a packing conjecture of Bollobás. *Combinatorics, Probability and Computing*, 4(3):241–255, 1995.

[11] Wolfgang Mader. Homomorphieeigenschaften und mittlere kantendichte von graphen. *Mathematische Annalen*, 174:265–268, 1967.

[12] Wolfgang Mader. Degree and local connectivity in digraphs. *Combinatorica*, 5:161–165, 1985.

[13] Maya Stein. Tree containment and degree conditions. In Raigorodskii, A.M., Rassias, M.T. (editors). *Discrete Mathematics and Applications*, 165:459–486, 2020. Springer.

[14] Maya Stein and Camila Zárate-Guerén. Antidirected subgraphs of oriented graphs. Preprint 2022, arXiv 2212.00769.

[15] Carster Thomassen. Even cycles in directed graphs. *European Journal of Combinatorics*, 6(1):85–89, 1985.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

The structure of Sidon set systems

(Extended abstract)

Maximilian Wötzel*

Abstract

A family $\mathcal{F} \subset 2^G$ of subsets of an abelian group G is a *Sidon system* if the sumsets $A+B$ with $A, B \in \mathcal{F}$ are pairwise distinct. Cilleruelo, Serra and the author previously proved that the maximum size $F_k(n)$ of a Sidon system consisting of k-subsets of the first n positive integers satisfies $C_k n^{k-1} \leq F_k(n) \leq \binom{n-1}{k-1} + n - k$ for some constant C_k only depending on k. We close the gap by proving an essentially tight structural result that in particular implies $F_k(n) \geq (1-o(1))\binom{n}{k-1}$. We also use this to establish a result about the size of the largest Sidon system in the binomial random family $\binom{[n]}{k}_p$. Extensions to h-fold sumsets for any fixed $h \geq 3$ are also obtained.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-114`

Introduction and main results

A subset A of an abelian group G is a *Sidon set* if the twofold sums of elements in A are pairwise distinct. The study of Sidon sets in the integers is a classical topic in additive number theory, see for instance the survey of O'Bryant [8]. A topic of particular interest is to determine the maximum size of a Sidon set contained in the first n positive integers. Seminal results of Erdős and Turán [5] concerning the upper bound, as well as Ruzsa [9], Bose [2] and Singer [11] for the lower bound established the following result.

Theorem 1 ([5, 11, 2, 9]). *A maximum size Sidon set $A \subset [n]$ satisfies $|A| = (1 \pm o(1))\sqrt{n}$.*

*Korteweg-de Vries Institute for Mathematics, Universiteit van Amsterdam, 1098 XG Amsterdam, The Netherlands. E-mail: `m.wotzel@uva.nl`. Supported by the Dutch Science Council (NWO) under grant number OCENW.M20.009.

Interestingly, it is still an open problem to establish the lower order behavior of this cardinality. The main lower order term in the upper bound stood at $n^{1/4}$ since 1969, due to Lindström [7], but the leading constant has recently been pushed below 1 due to Balogh, Füredi and Souktik [1]. The main question, whether the lower order term diverges or not, is still wide open.

One can naturally extend the notion of a Sidon set to set systems. Recall that the *sumset* (or *Minkowski sum*) of two sets A and B is defined as

$$A + B = \{a + b : a \in A, b \in B\}.$$

For an integer $h \geq 2$, we will often write hA as shorthand for the sumset $(h-1)A + A$. In [3], Cilleruelo, Serra and the author defined the following notion of a *Sidon system*.

Definition 2. Let $\mathcal{F} \subset 2^G$ be a family of subsets of an abelian group G. Then $\mathcal{F}$ is a *Sidon system* if for any $A, B, C, D \in \mathcal{F}$ it holds that

$$A + B = C + D \iff \{A, B\} = \{C, D\}.$$

So a Sidon set is just a Sidon system composed entirely of singleton sets. Another way to interpret Sidon systems in an abelian group G is as Sidon sets in the abelian monoid of subsets of G together with the sumset operation.

The size and structure of large Sidon systems of k-sets in $[n]$. In [3] the authors established the following analogue to Theorem 1. We write $F_k(n)$ for the maximum cardinality of a Sidon system composed entirely of k-element subsets of $[n]$.

Theorem 3 ([3])**.** *Let $n > k \geq 2$ be positive integers. Then there exists a constant C_k only depending on k such that*

$$C_k n^{k-1} \leq F_k(n) \leq \binom{n-1}{k-1} + n - k.$$

The major problem left open in [3] was to conclude whether the upper bound in Theorem 3 is asymptotically correct. In fact, a case analysis in the specific case of $k = 3$ did establish this fact. Actually, the authors formulated a stronger conjecture on the structure of Sidon systems, motivated by the proof of the upper bound in Theorem 3.

For integers $n > k \geq 2$, define the set system

$$\binom{[n]}{k}_0 = \{A \subset \{0, 1, \ldots, n\} : |A| = k,\, 0 \in A\}.$$

Then the following conjecture was posed implicitly in [3].

Conjecture 4 ([3])**.** *Let $n > k \geq 3$, and suppose $\mathcal{F} \subset \binom{[n]}{k}$ is any family of k-subsets of the first n integers such that for every $A \in \binom{[n]}{k}_0$ it holds that*

$$|\{x \in \mathbb{Z} : A + x \in \mathcal{F}\}| \leq 1. \tag{1}$$

Then one can remove $o(n^{k-1})$ sets from $\mathcal{F}$ to make it a Sidon system. In particular, by starting with any family that satisfies Eq. (1) *with equality,*

$$F_k(n) \sim \frac{n^{k-1}}{(k-1)!}.$$

As mentioned above, a motivation for Conjecture 4 is the following observation which is one of the main ideas going into proving the upper bound in Theorem 3. For distinct $A, B \in \binom{[n]}{k}_0$ such that $x+A, y+A, u+B, v+B$ are pairwise distinct sets in a Sidon system $\mathcal{F}$, we must have $|x-y| \neq |u-v|$. Since the minimum element of any set in $\binom{[n]}{k}$ can be at most $n-k+1$, their positive differences must lie in $[n-k]$. Hence, after starting with a set system $\mathcal{F}$ as described in Conjecture 4, one can only add at most $n-k$ additional sets to it before it necessarily contains a violation to the Sidon condition. If $k \geq 3$, we see that n is negligible when compared to n^{k-1}, and so here these additional sets can be ignored. The same is not true for $k = 2$, and in fact, while the first part of Conjecture 4 holds here, the second does not: In [3] the authors showed that the family

$$\mathcal{F} = \{\{1, n-i\} + \{0, i\} : i = 1, \dots, n-1\}$$

is a Sidon system. We see that every set in $\binom{[n-1]}{2}_0$ that has a translation in $\mathcal{F}$ except for $\{0, n-1\}$ in fact has two of them. It is also not difficult to check that the size of this family matches the upper bound given by Theorem 3.

As our first result, we resolve Conjecture 4 in the affirmative. Recall that for an integer $h \geq 2$, a subset $A \subset G$ of an abelian group G is called a *B_h-set* if for any $a_1, \dots, a_h, b_1, \dots, b_h \in A$ it holds that

$$a_1 + \cdots + a_h = b_1 + \cdots + b_h \iff \{a_1, \dots, a_h\} = \{b_1, \dots, b_h\} \text{ as multisets.}$$

This generalizes the notion of Sidon sets by observing that Sidon sets are B_2-sets. We prove the following result.

Theorem 5. *For any positive integer k, there exists an integer $\ell(k) = \ell$ such that the following holds. Let $A, B, C, D \subset \mathbb{R}$ be B_ℓ-sets of cardinality k all having the same minimal element. Then*

$$A + B = C + D \iff \{A, B\} = \{C, D\}.$$

Note that Theorem 5 indeed implies Conjecture 4 by the following argument. Any set in $\binom{[n]}{k}_0$ that is not a B_ℓ-set for some ℓ corresponds to a solution to a system of linear equations

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ \lambda_1 & \lambda_2 & \dots & \lambda_k \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} = \mathbf{0},$$

with $\lambda_i \in \mathbb{Z}$, $\sum |\lambda_i| \leq 2\ell$ and such that there are at least two indices $0 < i < j \leq k$ with $\lambda_i, \lambda_j \neq 0$. In particular, the matrix on the left-hand side has rank 2, and so there are at most n^{k-2} solutions to this system of linear equations in $[n]$. Since there are clearly at most $(2\ell)^k = O_k(1)$ such matrices, we see that $\binom{[n]}{k}_0$ contains $O_k(n^{k-2})$ non-B_ℓ-sets for any ℓ only depending on k, so one can remove their representatives to obtain a Sidon system.

In fact, the following stronger version of Theorem 5 is proved.

Proposition 6. *For any positive integer k and h, there exists an integer $\ell(k,h) = \ell$ such that the following holds. Let G be an abelian group, and let $A_1, \ldots, A_h, B_1, \ldots, B_h \subset G$ be B_ℓ-subsets of cardinality k all sharing an element. If there exist indices $i, j \in [h]$ such that $|A_i \cap B_j| \geq 2$, then*

$$A_1 + \cdots + A_h = B_1 + \cdots + B_h \iff \{A_1, \ldots, A_h\} = \{B_1, \ldots, B_h\} \text{ as multisets.}$$

This implies Theorem 5 for $h = 2$, since one can show that in any linearly ordered group, this minimum intersection requirement is satisfied, even without assuming the sets to be B_ℓ. The key tool in proving Proposition 6 is following simple statement, which holds in arbitrary abelian groups.

Lemma 7. *Let $A, B, C \subset G$ be subsets of an abelian group G such that A is a Sidon set. Then for any set $X \subset A$ satisfying $|X| > |C|$, it holds that*

$$X + B \subset A + C \implies B \subset C.$$

It would be interesting to find out whether an intersection size of size 1 in Proposition 6 is actually possible, and we prove some partial results regarding this.

The largest Sidon system in $\binom{[n]}{k}_p$ and δ-additive families. Recall that the binomial random family $\binom{[n]}{k}_p$ is defined such that every k-set $A \subset [n]$ is contained in $\binom{[n]}{k}_p$ independently with probability p. We write $[n]_p$ for $\binom{[n]}{1}_p$. An interesting question is to study a *sparse random* analogue of determining bounds on $F_k(n)$. That is, instead of investigating the size of the largest Sidon system in $\binom{[n]}{k}$, what happens if we do this in $\binom{[n]}{k}_p$? The Sidon set equivalent of this question was answered by Kohayakawa, Lee, Rödl and Samotij in [6] and they discovered an interesting phase transition. Essentially, as long as $p = o(n^{-1/3})$, the expected number of quadruples violating the Sidon set condition is negligible when compared to the expected size of the random set, and hence standard concentration bounds tell us that the size of the largest Sidon subset will be the same as the size of the random set. For p in the range between $n^{-1/3}$ and constant, the situation is similar to that in $[n]$, that is, the size of the largest Sidon subset is approximately the square root of np, the size of the random set. This range can be seen as an example of the *transference principle* (cf. [4, 10]) that says that results in the dense setting can be moved to the sparse random one in appropriate contexts. Since the problem is clearly monotone in nature, the situation when $n^{-2/3} \leq p \leq n^{-1/3}$ is that the largest Sidon subset must stay constant in the exponent at approximately $n^{1/3}$. Let us summarize.

Theorem 8 ([6]). *Let $0 \le a \le 1$ be a fixed constant. Suppose $p = p(n) = (1+o(1))n^{-a}$. There exists a constant $b = b(a)$ such that almost surely the largest Sidon subset of $[n]_p$ has size $n^{b+o(1)}$. Furthermore,*

$$b(a) = \begin{cases} 1-a, & \text{if } 2/3 \le a \le 1, \\ 1/3, & \text{if } 1/3 \le a \le 2/3, \\ (1-a)/2, & \text{if } 0 \le a \le 1/3. \end{cases}$$

Our second main result establishes a somewhat less nuanced analogue of Theorem 8. It will be helpful to change the language from the absence to the appearance of additive structures.

Definition 9. Let G be an abelian group and suppose $A, B, C, D \subset G$ are subsets. We say that (A, B, C, D) forms an *additive quadruple* if $A + B = C + D$, and furthermore, it is called *nontrivial* if $\{A, B\} \ne \{C, D\}$.

Hence, a Sidon system is a family that contains no nontrivial additive quadruples. We can now define a relative version of this concept.

Definition 10. Let G be an abelian group and $\delta > 0$. Then a finite family of subsets $\mathcal{F} \subset 2^G$ is called δ*-additive* if every subfamily $\mathcal{G} \subseteq \mathcal{F}$ with $|\mathcal{G}| \ge \delta|\mathcal{F}|$ contains a nontrivial additive quadruple.

Using Theorem 5, we are able to determine the threshold probability for when $\binom{[n]}{k}_p$ is δ-additive.

Theorem 11. *Let $k \ge 2$ be a fixed integer and $\delta \in (0,1)$. Then there exist constants C, c that only depend on k, δ such that*

$$\lim_{n\to\infty} \mathbf{Pr}\left(\binom{[n]}{k}_p \text{ is } \delta\text{-additive}\right) = \begin{cases} 1, & \text{if } p \ge c/n \\ 0, & \text{if } p \le C/n \end{cases}.$$

Recalling that $F_k(n) \le O_k(n^{k-1})$ by Theorem 3, this immediately gives us the following analogue of Theorem 8.

Corollary 12. *Let $k \ge 2$ be a fixed integer. Then there exist constants C, c that only depend on k such that asymptotically almost surely, the largest Sidon system $\mathcal{F} \subset \binom{[n]}{k}_p$ has size*

$$|\mathcal{F}| = \begin{cases} \Theta(n^{k-1}), & \text{if } p \ge C/n \\ \Theta(n^k p), & \text{if } p \le c/n \end{cases}.$$

In other words, we are essentially always in the regime that one can remove a negligible number of k-subsets in order to transform the random family into a Sidon system comparable to the $p = o(n^{-2/3})$ case for Sidon sets.

References

[1] József Balogh, Zoltán Füredi, and Souktik Roy. An upper bound on the size of Sidon sets. *Am. Math. Mon.*, 130(5):437–445, 2023.

[2] R. C. Bose. An affine analogue of Singer's theorem. *J. Indian Math. Soc., New Ser.*, 6:1–15, 1942.

[3] Javier Cilleruelo, Oriol Serra, and Maximilian Wötzel. Sidon set systems. *Rev. Mat. Iberoam.*, 36(5):1527–1548, 2020.

[4] D. Conlon and W. T. Gowers. Combinatorial theorems in sparse random sets. *Ann. Math. (2)*, 184(2):367–454, 2016.

[5] Pál Erdős and Pál Turán. On a problem of Sidon in additive number theory, and on some related problems. *J. Lond. Math. Soc.*, 16:212–215, 1941.

[6] Yoshiharu Kohayakawa, Sang June Lee, Vojtěch Rödl, and Wojciech Samotij. The number of Sidon sets and the maximum size of Sidon sets contained in a sparse random set of integers. *Random Struct. Algorithms*, 46(1):1–25, 2015.

[7] Bernt Lindström. An inequality for B_2-sequences. *J. Comb. Theory*, 6:211–212, 1969.

[8] Kevin O'Bryant. A complete annotated bibliography of work related to Sidon sequences. *Electron. J. Comb.*, DS11:39, 2004.

[9] Imre Z. Ruzsa. Solving a linear equation in a set of integers. I. *Acta Arith.*, 65(3):259–282, 1993.

[10] Mathias Schacht. Extremal results for random discrete structures. *Ann. Math. (2)*, 184(2):333–365, 2016.

[11] James Singer. A theorem in finite projective geometry and some applications to number theory. *Trans. Am. Math. Soc.*, 43:377–385, 1938.

Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Minimum non-chromatic-λ-choosable graphs

(Extended abstract)

Jialu Zhu* Xuding Zhu†

Abstract

For a multi-set $\lambda = \{k_1, k_2, \ldots, k_q\}$ of positive integers, let $k_\lambda = \sum_{i=1}^q k_i$. A λ-list assignment of G is a list assignment L of G such that the colour set $\bigcup_{v \in V(G)} L(v)$ can be partitioned into the disjoint union $C_1 \cup C_2 \cup \ldots \cup C_q$ of q sets so that for each i and each vertex v of G, $|L(v) \cap C_i| \geq k_i$. We say G is λ-choosable if G is L-colourable for any λ-list assignment L of G. The concept of λ-choosability puts k-colourability and k-choosability in the same framework: If $\lambda = \{k\}$, then λ-choosability is equivalent to k-choosability; if λ consists of k copies of 1, then λ-choosability is equivalent to k-colourability. If G is λ-choosable, then G is k_λ-colourable. On the other hand, there are k_λ-colourable graphs that are not λ-choosable, provided that λ contains an integer larger than 1. Let $\phi(\lambda)$ be the minimum number of vertices in a k_λ-colourable non-λ-choosable graph. This paper determines the value of $\phi(\lambda)$ for all λ.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-115`

1 Introduction

A *proper colouring* of a graph G is a mapping $f : V(G) \to \mathbb{N}$ such that $f(u) \neq f(v)$ for any edge uv of $E(G)$. The *chromatic number* $\chi(G)$ of G is the minimum positive integer k such that G is k-colourable, i.e., there is a proper colouring f of G using colours from $\{1, 2, \ldots, k\}$. The *choice number* $ch(G)$ of G is the minimum positive integer k such that G is k-choosable, i.e., if L is a list assignment which assigns to each vertex v a set $L(v) \subseteq \mathbb{N}$

*Department of Mathematics, Zhejiang Normal University, China. E-mail: `jialuzhu@zjnu.edu.cn`.
†Department of Mathematics, Zhejiang Normal University, China. E-mail: `xdzhu@zjnu.edu.cn`. Supported by NSFC 11971438, U20A2068.

of at least k integers as permissible colours, then there is a proper colouring f of G such that $f(v) \in L(v)$ for each vertex v.

It follows from the definitions that $\chi(G) \le ch(G)$ for any graph G, and it was shown in [5] that bipartite graphs can have arbitrarily large choice number. An interesting problem is for which graphs G, $\chi(G) = ch(G)$. Such graphs are called *chromatic-choosable*. Chromatic-choosable graphs have been studied extensively in the literature. There are a few challenging conjectures that assert certain families of graphs are chromatic-choosable. The most famous problem concerning this concept is perhaps the *list colouring conjecture*, which asserts that line graphs are chromatic-choosable [1]. Another problem concerning chromatic-choosable graphs that has attracted a lot of attention is the minimum order of a non-chromatic-choosable graph with given chromatic number. For a positive integer k, let

$$\phi(k) = \min\{n : \text{ there exists a non-}k\text{-choosable } k\text{-chromatic } n\text{-vertex graph}\}.$$

Ohba [20] conjectured that $\phi(k) \ge 2k+2$. In other words, k-colourable graphs on at most $2k+1$ vertices are k-choosable. This conjecture was studied in many papers [14, 16, 18–22, 24, 25], and was finally confirmed by Noel, Reed and Wu [18]. This lower bound is tight if k is even, i.e., $\phi(k) = 2k+2$ when k is even. Noel [17] further conjectured that if k is odd, then k-colourable graphs on at most $2k+2$ vertices are also k-choosable. Recently, the authors of this paper confirmed Noel's conjecture [28], and determined the value of $\phi(k)$ for all k.

Theorem 1. *[28] For $k \ge 2$,*

$$\phi(k) = \begin{cases} 2k+2, & \text{if } k \text{ is even,} \\ 2k+3, & \text{if } k \text{ is odd.} \end{cases}$$

The concept of λ-choosability is a refinement of choosability introduced in [32]. Assume that $\lambda = \{k_1, k_2, \ldots, k_q\}$ is a multi-set of positive integers. Let $k_\lambda = \sum_{i=1}^q k_i$ and $|\lambda| = q$. A *λ-list assignment* of G is a list assignment L such that the colour set $\bigcup_{v \in V(G)} L(v)$ can be partitioned into the disjoint union $C_1 \cup C_2 \cup \ldots \cup C_q$ of q sets so that for each i and each vertex v of G, $|L(v) \cap C_i| \ge k_i$. Note that for each vertex v, $|L(v)| \ge \sum_{i=1}^q k_i = k_\lambda$. So a λ-list assignment L is a k_λ-list assignment with some restrictions on the set of possible lists. We say G is *λ-choosable* if G is L-colourable for any λ-list assignment L of G.

For a positive integer a, let $m_\lambda(a)$ be the multiplicity of a in λ. If $m_\lambda(a) = m$, then instead of writing m times the integer a, we may write $a \star m$. For example, $\lambda = \{1 \star k_1, 2 \star k_2, 3\}$ means that λ is a multi-set consisting of k_1 copies of 1, k_2 copies of 2 and one copy of 3. If $\lambda = \{k\}$, then λ-choosability is the same as k-choosability; if $\lambda = \{1 \star k\}$, then λ-choosability is equivalent to k-colourability [32]. So the concept of λ-choosability puts k-choosability and k-colourability in the same framework.

Assume that $\lambda = \{k_1, k_2, \ldots, k_q\}$ and $\lambda' = \{k'_1, k'_2, \ldots, k'_p\}$. We say λ' is a *refinement* of λ if $p \ge q$ and there is a partition $I_1 \cup I_2 \cup \ldots \cup I_q$ of $\{1, 2, \ldots, p\}$ such that $\sum_{j \in I_t} k'_j = k_t$ for $t = 1, 2, \ldots, q$. We say λ' is obtained from λ by increasing some parts if $p = q$ and

$k_t \le k'_t$ for $t = 1, 2, \ldots, q$. We write $\lambda \le \lambda'$ if λ' is a refinement of λ'', and λ'' is obtained from λ by increasing some parts. It follows from the definitions that if $\lambda \le \lambda'$, then every λ-choosable graph is λ'-choosable. Conversely, it was proved in [32] that if $\lambda \nleq \lambda'$, then there is a λ-choosable graph which is not λ'-choosable. In particular, λ-choosability implies k_λ-colourability, and if $\lambda \ne \{1 \star k_\lambda\}$, then there are k_λ-colourable graphs that are not λ-choosable.

All the partitions λ of a positive integer k are sandwiched between $\{k\}$ and $\{1 \star k\}$ in the above order. As observed above, $\{k\}$-choosability is the same as k-choosability, and $\{1 \star k\}$-choosability is equivalent to k-colourability. For other partitions λ of k, λ-choosability reveals a complex hierarchy of colourability of graphs sandwiched between k-colourability and k-choosability. The framework of λ-choosability provides room to explore generalizations of colourability and choosability results or problems (see [8, 10, 32])

2 Preliminaries

In this paper, we are interested in Ohba type question for λ-choobility. Similar to the definition of $\phi(k)$, for a multi-set λ of positive integers, we define $\phi(\lambda)$ as follows:

Definition 1. *Assume λ is a multi-set of positive integers. Let*

$$\phi(\lambda) = \min\{n : \text{ there exists a non-}\lambda\text{-choosable } k_\lambda\text{-chromatic } n\text{-vertex graph}\}.$$

If $\lambda = \{1 \star k\}$, then λ-choosable is equivalent to k-colourable. In this case, we set $\phi(\lambda) = \infty$. We call such a multi-set λ trivial. In the following, we only consider non-trivial multi-sets of positive integers.

If $\lambda = \{k\}$, then $\phi(\lambda) = \phi(k)$. The value of $\phi(k)$ is determined in Theorem 1. For general multiset λ of positive integers, the function $\phi(\lambda)$ was first studied in [30]. Let $m_\lambda(\text{odd})$ be the number of odd integers in λ. The following result was proved in [30].

Theorem 2. *For any non-trivial multi-set λ of positive integers,*

$$2k_\lambda + m_\lambda(1) + 2 \leqslant \phi(\lambda) \leqslant \min\{2k_\lambda + m_\lambda(\text{odd}) + 2, 2k_\lambda + 5m_\lambda(1) + 3\}.$$

If $m_\lambda(1) = m_\lambda(\text{odd}) = t$, then it follows from Theorem 2 that $\phi(\lambda) = 2k_\lambda + t + 2$. However, when $m_\lambda(1)$ and $m_\lambda(\text{odd}) - m_\lambda(1)$ are both large, then the gap between the upper and lower bounds for $\phi(\lambda)$ in Theorem 2 becomes large.

3 Main result

This paper proves Theorem 3 below, which strengthens Theorem 1 and Theorem 2 and determines the value of $\phi(\lambda)$ for all λ.

Theorem 3. *Assume λ is a non-trivial multi-set of positive integers. Then*

$$\phi(\lambda) = \min\{2k_\lambda + m_\lambda(\text{odd}) + 2, 2k_\lambda + 3m_\lambda(1) + 3\}.$$

Below is a sketch of the proof of Theorem 3.

By Theorem 2, to prove Theorem 3, it suffices to consider the case that $m_\lambda(\text{odd}) > m_\lambda(1)$.

First we consider the case that $m_\lambda(1) = 0$ and $m_\lambda(\text{odd}) > 0$. In this case, we need to show that $\phi(\lambda) = 2k_\lambda + 3$.

Let $k_\lambda = k$. By Theorem 2, $2k+2 \le \phi(\lambda) \le 2k+3$. So it suffices to show that $\phi(\lambda) \ne 2k+2$, i.e., any graph G with $\chi(G) \le k$ and $|V(G)| \le 2k+2$ is λ-choosable. We only need to consider the case that G is a complete k-partite graph. The following result was proved in [29].

Theorem 4. *Assume G is a complete k-partite graph with $|V(G)| \le 2k+2$. Then G is k-choosable, unless k is even and $G = K_{4,2\star(k-1)}$ or $G = K_{3\star(k/2+1),1\star(k/2-1)}$.*

Thus we may assume that k is even and $G = K_{4,2\star(k-1)}$ or $G = K_{3\star(k/2+1),1\star(k/2-1)}$. We say a k-list assignment L of G is *bad* if G is not L-colourable. All bad assignments for $K_{4,2\star(k-1)}$ and $K_{3\star(k/2+1),1\star(k/2-1)}$ are characterized in [4] and [29], respectively and we can verify that such bad list assignments is not λ-list assignment (using the assumption $m_\lambda(\text{odd}) > 0$). This implies that all graphs $K_{4,2\star(k-1)}$ and $K_{3\star(k/2+1),1\star(k/2-1)}$ are λ-choosable. This completes the proof for the case $m_\lambda(1) = 0$.

Next we consider the case that $m_\lambda(1) = a \ge 1$ and $m_\lambda(\text{odd}) - m_\lambda(1) = c \ge 1$. We need to show that $\phi(\lambda) = \min\{2k+a+c+2, 2k+3a+3\}$. First, we prove the upper bound, i.e.,

$$\phi(\lambda) \le \min\{2k+a+c+2, 2k+3a+3\}.$$

By Theorem 2, $\phi(\lambda) \le 2k+a+c+2$. It remains to show that $\phi(\lambda) \le 2k+3a+3$. Observe that $k_\lambda = k$, $m_\lambda(1) = a$ and $m_\lambda(\text{odd}) = a+c$ implies that $\{1\star a, 2\star(k-a-3c)/2, 3\star c\}$ is a refinement of λ. Hence it suffices to prove the following lemma.

Lemma 5. *Assume $\lambda = \{1\star a, 2\star b, 3\star c\}$ and $k = a+2b+3c$ (and hence $m_\lambda(1) = a$, $m_\lambda(\text{odd}) = a+c$ and $k_\lambda = k$). Then there exists a k-chromatic graph G with $|V(G)| = 2k+3a+3$ which is not λ-choosable.*

Let $G = K_{5\star(a+1),2\star(k-a-1)}$ be the complete k-partite graph with partite sets $U_i = \{u_{i,1}, u_{i,2}, u_{i,3}, u_{i,4}, u_{i,5}\}$ where $i = 1,2,\dots,a+1$, and $V_j = \{v_{j,1}, v_{j,2}\}$ where $j = 1,2,\dots,k-a-1$.

Let $S_i = \{s_{i,1}, s_{i,2}, \dots, s_{i,6}\}$ be pairwise disjoint sets of size 6 where $i = 1,2,\dots,c$ and let $T_i = \{t_{i,1}, t_{i,2}, t_{i,3}, t_{i,4}\}$ be pairwise disjoint sets of size 4 where $i = 1,2,\dots,b$. Let E be a set of a colours, and the sets E, S_i, T_i are pairwise disjoint and let

$$A_1 = \bigcup_{i=1}^{c}\{s_{i,1}, s_{i,3}, s_{i,5}\},\ A_2 = \bigcup_{i=1}^{c}\{s_{i,1}, s_{i,3}, s_{i,6}\},\ A_3 = \bigcup_{i=1}^{c}\{s_{i,1}, s_{i,2}, s_{i,4}\},\ A_4 = \bigcup_{i=1}^{c}\{s_{i,2}, s_{i,3}, s_{i,4}\},$$
$$A_5 = \bigcup_{i=1}^{c}\{s_{i,2}, s_{i,5}, s_{i,6}\},\ A_6 = \bigcup_{i=1}^{c}\{s_{i,1}, s_{i,2}, s_{i,3}\},\ A_7 = \bigcup_{i=1}^{c}\{s_{i,4}, s_{i,5}, s_{i,6}\},$$
$$B_1 = \bigcup_{i=1}^{b}\{t_{i,2}, t_{i,3}\},\ B_2 = \bigcup_{i=1}^{b}\{t_{i,2}, t_{i,4}\},\ B_3 = \bigcup_{i=1}^{b}\{t_{i,1}, t_{i,2}\},\ B_4 = \bigcup_{i=1}^{b}\{t_{i,1}, t_{i,3}\},$$
$$B_5 = \bigcup_{i=1}^{b}\{t_{i,1}, t_{i,4}\},\ B_6 = \bigcup_{i=1}^{b}\{t_{i,1}, t_{i,2}\},\ B_7 = \bigcup_{i=1}^{b}\{t_{i,3}, t_{i,4}\}.$$

Let L be the λ-list assignment of G defined as follows:

$$L(v) = \begin{cases} A_j \cup B_j \cup E, & \text{if } v = u_{i,j}, 1 \leq i \leq a+1, 1 \leq j \leq 5, \\ A_{j+5} \cup B_{j+5} \cup E, & \text{if } v = v_{i,j}, 1 \leq i \leq k-a-1, 1 \leq j \leq 2, \end{cases}$$

It can be proved that L is λ-list assignment and G is not L-colourable. The proof is a little complicated, and the details are omitted.

It remains to prove the lower bound that $\phi(\lambda) \geqslant \min\{2k+3a+3, 2k+a+c+2\}$.

Assume to the contrary that $\phi(\lambda) < \min\{2k+a+c+2, 2k+3a+3\}$ for some λ. We choose such a multi-set $\lambda = \{k_1, k_2, \dots, k_q\}$ with $|\lambda| = q$ minimum. Assume that $k_1 = k_2 = \dots = k_a = 1$ and $3 \leq k_{a+1} \leq k_{a+2} \leq \dots \leq k_{a+c}$ are the odd integers in λ.

Let $n = \min\{2k+a+c+2, 2k+3a+3\}$. Then there is a k-chromatic graph G with $|V(G)| \leq n-1$ which is not λ-choosable. We may assume that G is a complete k-partite graph with $|V(G)| = n-1$ and with partite sets $P_1, P_2, \dots, P_k$ such that $|P_1| \geq |P_2| \geq \dots \geq |P_k|$. For a positive integer i, let

$$I_i = \{j : |P_j| = i\}.$$

Note that $|P_1| \geq 3$ (as $|V(G)| > 2k$). Using the assumption $m_\lambda(1) \geq 1$ and the minimality of $|\lambda|$, we can conclude that $|P_1| \leq 4$, and if $c \leq 2a+1$, then $|P_1| \leq c-2a+3$. Since $a \geq 1$, we know that $c \geq 2a \geq 2$, and if $c = 2$, then $a = 1$ and $|P_1| = 3$.

Definition 2. *A 4-tuple (a_1, a_2, a_3, a_4) of integers is* reducible *if*

$$0 \leq a_i \leq |I_i|, \ \sum_{i=1}^{4} a_i = k_{a+1} \ \textit{and} \ 2k_{a+1}+1 \leq \sum_{i=1}^{4} i a_i \leq 2k_{a+1}+2.$$

Combining with Theorem 4 and the minimality of $|\lambda|$, we conclude that

Claim 6. *There is no reducible 4-tuple.*

It follows from Claim 6 that $|I_2| \leq k_{a+1}-2$ and if $c \geq 3$, then $|I_1| \geq \frac{2}{3}k_{a+1}$ and if $c = 2$, then $|I_1| \geq (k_{a+1}-1)/2$. Recall that $3 \leq |P_1| \leq 4$. By Claim 6, we can conclude that if $|P_1| = 4$, then $|I_3| < \lfloor \frac{k_{a+1}-|I_2|-1}{2} \rfloor$, $|I_4| < \lceil \frac{k_{a+1}-|I_2|-2|I_3|-1}{3} \rceil + 1$ and if $|P_1| = 3$, then $|I_3| < \lceil \frac{k_{a+1}-|I_2|-1}{2} \rceil + 1$. This contradicts to $|V(G)| = n-1 \geq 2k+1$. This completes the proof of Theorem 3.

References

[1] Béla Bollobás and Andrew Harris. List-colourings of graphs. *Graphs Combin.*, 1(2):115–127, 1985.

[2] Hojin Choi and Young Soo Kwon. On t-common list-colorings. *Electron. J. Combin.*, 24(3):Paper 3.32, 10, 2017.

[3] Lech Duraj, Grzegorz Gutowski, and Jakub Kozik. Chip games and paintability. *Electron. J. Combin.*, 23(3):Paper 3.3, 12, 2016.

[4] Hikoe Enomoto, Kyoji Ohba, Katsuhiro Ota, and Junko Sakamoto. Choice number of some complete multi-partite graphs. *Discrete Math.*, 244(1-3):55–66, 2002.

[5] Paul Erdős, Arthur L. Rubin, and Herbert Taylor. Choosability in graphs. In *Proceedings of the West Coast Conference on Combinatorics, Graph Theory and Computing (Humboldt State Univ., Arcata, Calif., 1979)*, Congress. Numer., XXVI, pages 125–157. Utilitas Math., Winnipeg, Man., 1980.

[6] Fred Galvin. The list chromatic index of a bipartite multigraph. *J. Combin. Theory Ser. B*, 63(1):153–158, 1995.

[7] Sylvain Gravier and $Frdric$ Maffray. Choice number of 3-colorable elementary graphs. *Discrete Math.*, 165/166:353–358, 1997. Graphs and combinatorics (Marseille, 1995).

[8] Yangyan Gu, Yiting Jiang, David Wood, and Xuding Zhu. Refined list version of Hadwiger's conjecture. arXiv:2209.07013.

[9] Po-Yi Huang, Tsai-Lien Wong, and Xuding Zhu. Application of polynomial method to on-line list colouring of graphs. *European J. Combin.*, 33(5):872–883, 2012.

[10] Arnfried Kemnitz and Margit Voigt. A note on non-4-list colorable planar graphs. *Electron. J. Combin.*, 25(2):Paper 2.46, 5, 2018.

[11] Seog-Jin Kim, Young Soo Kwon, Daphne Der-Fen Liu, and Xuding Zhu. On-line list colouring of complete multipartite graphs. *Electron. J. Combin.*, 19(1):Paper 41, 13, 2012.

[12] Seog-Jin Kim and Boram Park. Bipartite graphs whose squares are not chromatic-choosable. *Electron. J. Combin.*, 22(1):Paper 1.46, 12, 2015.

[13] Seog-Jin Kim and Boram Park. Counterexamples to the list square coloring conjecture. *J. Graph Theory*, 78(4):239–247, 2015.

[14] Alexandr V. Kostochka, Michael Stiebitz, and Douglas R. Woodall. Ohba's conjecture for graphs with independence number five. *Discrete Math.*, 311(12):996–1005, 2011.

[15] Alexandr V. Kostochka and Douglas R. Woodall. Choosability conjectures and multicircuits. *Discrete Math.*, 240(1-3):123–143, 2001.

[16] Jakub Kozik, Piotr Micek, and Xuding Zhu. Towards an on-line version of $Ohba's$ conjecture. *European J. Combin.*, 36:110–121, 2014.

[17] Jonathan A. Noel. Choosability of graphs with bounded order: Ohba's conjecture and beyond. Master's thesis, McGill University, 2013.

[18] Jonathan A. Noel, Bruce A. Reed, and Hehui Wu. A proof of a conjecture of Ohba. *J. Graph Theory*, 79(2):86–102, 2015.

[19] Jonathan A. Noel, Douglas B. West, Hehui Wu, and Xuding Zhu. Beyond Ohba's conjecture: a bound on the choice number of k-chromatic graphs with n vertices. *European J. Combin.*, 43:295–305, 2015.

[20] Kyoji Ohba. On chromatic-choosable graphs. *J. Graph Theory*, 40(2):130–135, 2002.

[21] Kyoji Ohba. Choice number of complete multipartite graphs with part size at most three. *Ars Combin.*, 72:133–139, 2004.

[22] Bruce Reed and Benny Sudakov. List colouring when the chromatic number is close to the order of the graph. *Combinatorica*, 25(1):117–123, 2005.

[23] Uwe Schauz. Mr. Paint and Mrs. Correct. *Electron. J. Combin.*, 16(1):Research Paper 77, 18, 2009.

[24] Yufa Shen, Wenjie He, Guoping Zheng, and Yanpo Li. Ohba's conjecture is true for graphs with independence number at most three. *Appl. Math. Lett.*, 22(6):938–942, 2009.

[25] Yufa Shen, Wenjie He, Guoping Zheng, Yanning Wang, and Lingmin Zhang. On choosability of some complete multipartite graphs and Ohba's conjecture. *Discrete Math.*, 308(1):136–143, 2008.

[26] M. Voigt. A non-3-choosable planar graph without cycles of length 4 and 5. *Discrete Math.*, 307(7-8):1013–1015, 2007.

[27] Margit Voigt. List colourings of planar graphs. *Discrete Math.*, 120(1-3):215–219, 1993.

[28] Jialu Zhu and Xuding Zhu. Bad list assignments for non-k-choosable k-chromatic graphs with $2k+2$-vertices. arXiv:2202.09756.

[29] Jialu Zhu and Xuding Zhu. Minimum non-chromatic-choosable graphs with given chromatic number. arXiv:2201.02060.

[30] Jialu Zhu and Xuding Zhu. Chromatic λ-choosable and λ-paintable graphs. *J. Graph Theory*, 98(4):642–652, 2021.

[31] Xuding Zhu. On-line list colouring of graphs. *Electron. J. Combin.*, 16(1):Research Paper 127, 16, 2009.

[32] Xuding Zhu. A refinement of choosability of graphs. *J. Combin. Theory Ser. B*, 141:143–164, 2020.