Mathematisches Forschungsinstitut Oberwolfach

Report No. 1/2023

DOI: 10.4171/OWR/2023/1

Combinatorics

Organized by
Peter Keevash, Oxford
Wojciech Samotij, Tel Aviv
Benny Sudakov, Zürich

1 January – 7 January 2023

Abstract. Combinatorics is an area of mathematics primarily concerned with counting and studying properties of discrete objects such as graphs, set systems, partial orders, polyhedra, etc. Combinatorial problems naturally arise in many areas of mathematics, such as algebra, geometry, probability theory, and topology, and in theoretical computer science. Historically, such questions were often studied using ad hoc arguments. However, over the last few decades, the development of general and powerful methods have elevated combinatorics to a thriving branch of mathematics with many connections to other subjects. The workshop brought together the established leading experts and the brightest young talents from different parts of this very broad area in order to discuss the most exciting recent developments, current themes and trends, and the most promising new directions for future research.

Mathematics Subject Classification (2020): 05-XX.

Introduction by the Organizers

The Combinatorics workshop, co-organised by Peter Keevash (Oxford), Wojciech Samotij (Tel Aviv), and Benny Sudakov (Zürich), commenced during the first week of January of 2023. Among the 51 participants that started the New Year at Oberwolfach, there were mathematicians from Canada, several European countries, Israel, the United Kingdom, and the United States. Four additional researchers could participate with the help of the Zoom video conferencing software. The program of the workshop comprised eleven 50-minute-long plenary lectures and fifteen shorter, 25-minute talks. In order to provide a platform for younger researchers, we have accommodated all of them who wanted to speak at the meeting. A lively, one-hour-long problem session, led by Nati Linial (Jerusalem) was held

on Tuesday evening. Last but not least, the daily, individual discussions between various sets of attendees were commonplace and often stretched well into the night. It was a real pleasure to see how well the community has already recovered from the COVID pandemic.

This report contains extended abstracts of all the lectures given at the meeting as well as a summary of the problems presented at the aforementioned open problem session. The lectures covered a diverse range of topics that not only spanned the broad field of combinatorics, but also ventured into several adjacent fields such as algebra, theoretical computer science, probability theory, and complex analysis. The quality of the lectures was exceptionally high and this brief introduction, as well as the extended abstracts that follow, can hardly give justice to the strength of the results presented by our colleagues. To support our overwhelmingly postive evaluation of the scientific part of the workshop, we now briefly highlight three of the most spectacular developments that were communicated during the workshop.

The first highlight is the work of Richard Montgomery on the Ryser–Brualdi–Stein conjecture. A Latin square of order n is an n-by-n grid filled with n symbols in such a way that every symbol appears exactly once in each column and each row. A transversal of a Latin square is a collection of cells that share no column, no row, and are filled with distinct symbols. Interesting examples of Latin squares are multiplication tables of finite groups. This narrow class of examples already shows that one cannot expect every order n Latin square to admit a transversal of size n when n is even. More than fifty years ago, Ryser conjectured that every Latin square of odd order n does admit a transversal of size n, while Brualdi and Stein independently conjectured that each Latin square of even order n admits a transversal of size $n-1$.

Numerous researchers attempted to solve the conjectures of Ryser, Brualdi, and Stein, obtaining stronger and stronger lower bounds on the guaranteed largest size of a transversal in Latin squares of large order. Prior to the work of Montgomery, the state-of-the-art was the result of Keevash, Pokrovskiy, Sudakov, and Yepremyan, who showed that every Latin square of order n has a transversal of size $n - O(\log n/\log\log n)$; this improved on the lower bound of $n - O(\log^2 n)$ proved by Shor almost forty years earlier.

In his lecture, Montgomery presented a sketch of the proof of his resolution of the Brualdi–Stein conjecture for all sufficiently large n.

Theorem 1 (Montgomery)**.** *There is some n_0 such that every Latin square of order $n \geqslant n_0$ contains a transversal of size $n-1$.*

One of the key insights of Montgomery that allows him to prove this beautiful result is separate treatment of Latin squares that are, in some well-defined sense, similar to multiplication tables of finite groups.

Our second highlight is the work of Oleg Pikhurko (joint with Łukasz Grabowski, András Máthé, and Jonathan Noel) on the (nearly) one-hundred-years-old Tarski's Circle Squaring Problem. Tarski's problem asks whether a disc and a square of the same area in $\mathbb{R}^2$ are equidecomposable, that is, whether they can be partitioned into the same finite number of pieces so that the pieces are pairwise isometric.

This question was answered only in 1990 by Laczkovich, who showed that it is possible to partition a disc of unit area into a finite number of pieces and use them to assemble a unit square using translations only. A somewhat unfortunate feature of Laczkovich's proof is that it crucially relies on the Axiom of Choice; consequently, the pieces of his circle squaring cannot be even guaranteed to be (Lebesgue) measurable.

Several years ago, a 'graph-theoretical' approach to the Circle Squaring Problem was pioneered by Grabowski, Máthé, and Pikhurko, who proved that there exists a circle squaring with measurable pieces. Then, Marks and Unger proved that the pieces can be even made Borel. A recent work of Noel, Máthé, and Pikhurko decreased the Borel complexity of the pieces (each piece is a Boolean combination of countable unions of closed sets) and ensured that the boundary of each piece is 'small' in the sense that its upper Minkowski dimension is strictly less than two. The gripping and lucid lecture of Pikhurko gave a high-level outline of the proof of this result.

Our third highlight is the work of István Tomon, joint with János Nagy and Péter Pál Pach, on special coverings of $\mathbb{F}_p^n$ by hyperplanes and their surprising connections with several longstanding open problems in algebra.

A covering of the finite vector space $\mathbb{F}_p^n$ by a collection $\mathcal{H}$ of hyperplanes is called irredundant if no proper subcollection $\mathcal{H}' \subsetneq \mathcal{H}$ covers the whole space. Tomon and his coauthors are interested in bounding the smallest size of an irredundant covering of $\mathbb{F}_p^n$ by a collection $\mathcal{H}$ of hyperplanes whose normal vectors span the whole n-dimensional space, which we will henceforth denote by $f_p(n)$. Whereas it is easy to find, for every prime p and all positive integers n, an irredundant covering of $\mathbb{F}_p^n$ by p hyperplanes (which is clearly the smallest size of any covering), the additional condition makes the problem of $f_p(n)$ decidedly more difficult and interesting.

An elementary dimension argument yields the lower bound $f_p(n) \geqslant n+1$ and an easy blow-up type construction provides the upper bound $f_p(n) \leqslant \lceil pn/2 \rceil + 1$; no other bounds were previously known. The main result of the work of Tomon and his co-authors is the stronger lower bound $f_p(n) \geqslant c \log p / \log\log p \cdot n$ for some (explicit) positive constant c.

It turns out, very surprisingly, that improving the lower bound on $f_p(n)$ by a small multiplicative factor would have several surprising consequences, one of which is a resolution of the Along–Jaeger–Tarsi conjecture from the late 1980s. This conjecture states that every invertible n-by-n matrix A over $\mathbb{F}_p$ admits a vector $v \in \mathbb{F}_p^n$ such that neither v nor Av has a zero coordinate. In his lecture, Tomon presented a short proof of the fact that $f_p(n) > 2n$ implies the conjecture and thus one of the corollaries of the main result of his work is that the conjecture holds for all $p \geqslant 67$, with the exception of $p = 79$.

Acknowledgement: First and foremost, the organisers thank the wonderful staff of the Mathematisches Forschungsinstitut Oberwolfach for all the help before, during, and after the workshop. Additionally, the MFO and the workshop organisers would

like to thank the Simons Foundation for supporting József Balogh in the ‘Simons Visiting Professors’ program at the MFO.

Workshop: Combinatorics

Table of Contents

Abstracts

Brambles, stack number and topological overlap

Sergey Norin

(joint work with David Eppstein, Robert Hickingbotham, Laura Merker, Michał T. Seweryn and David R. Wood)

A *(strict) bramble* $\mathcal{B}$ in a graph G is a collection of connected subgraphs of G closed under taking union. The *order* of a bramble $\mathcal{B}$ is the smallest size of a set of vertices that intersect the vertex set of each of its elements. Brambles have long been part of the graph minor theory toolkit, in particular, because a bramble of high order is an obstruction to existence of a low width tree decomposition [1].

We discuss high dimensional analogues of brambles which extend some of the properties of graph (1-dimensional) brambles. In particular, we show that a d-dimensional bramble of high order in a d-dimensional simplicial complex X is an obstruction to existence of a low multiplicity continuous map from X to $\mathbb{R}^d$ (and more generally to any d-dimensional contractible complex). This can be seen as a qualitative variant of Gromov's topological overlap theorem [2].

Let us now be more precise with the definitions. Let X be an abstract d-dimensional simplicial complex. A collection $\mathcal{B}$ of subcomplexes of X is a *(d-dimensional) bramble* if

- $\mathcal{B}$ is closed under taking unions,
- $H^i(B)$ is trivial for every $0 \leqslant i \leqslant d-1$ and $0 \leqslant i \leqslant d-1$.

In particular, 1-dimensional brambles are exactly strict brambles in graphs, and 2-dimensional brambles are union-closed collections of simply connected subcomplexes of a 2-dimensional complex.

The following discrete theorem about simplicial maps implies the above mentioned continuous result.

Theorem 1. *Let $f : X \to Y$ be a simplicial map between d-dimensional simplicial complexes such that $H^d(Y)$ is trivial, and let $\mathcal{B}$ be a d-dimensional bramble in X. Then*

$$\cap_{B\in\mathcal{B}} f(B) \neq \varnothing.$$

We modify the definition of an *order* $\|\mathcal{B}\|$ of a bramble $\mathcal{B}$ to be equal to the smallest size of a set d-dimensional faces that intersect all elements of the bramble. A refinement of Theorem 1 implies that under the same conditions some point of Y belongs to an image of at least $\|\mathcal{B}\|$, d-dimensional faces of $\mathcal{B}$.

Theorem 1 can be applied to lower bound stack number of graphs, which we now define. For a graph G and ordering $(v_1, \dots, v_n)$ of $V(G)$, two edges $v_iv_j, v_kv_\ell \in E(G)$ *cross* with respect to $(v_1, \dots, v_n)$ if $i < k < j < \ell$. An *s-stack layout* of G consists of an ordering $(v_1, \dots, v_n)$ of $V(G)$ together with a function $\phi : E(G) \to \{1, \dots, s\}$ such that for each $a \in \{1, \dots, s\}$ no two edges in $\phi^{-1}(a)$ cross with respect to $(v_1, \dots, v_n)$. Each set $\phi^{-1}(a)$ is called a *stack*. The *stack-number* $\mathrm{sn}(G)$ of a graph G is the minimum s for which there exists an s-stack layout of G.

In [3] we use a variant of Theorem 1 derived from Gromov's topological overlap theorem [2] to show the following.

Theorem 2. $\mathrm{sn}(P_n \boxtimes P_n \boxtimes P_n) = \Theta(n^{1/3})$, *where* P_n *denotes the* n*-vertex path and* $\boxtimes$ *is the strong product.*

Theorem 2 provides the first explicit example of a graph family with bounded maximum degree and unbounded stack-number. It is also the first example of a bounded degree graph with bounded queue-number and unbounded stack-number. The proof of the lower bound proceeds by considering the natural two-dimensional complex associated with $P_n \boxtimes P_n \boxtimes P_n$ with a 2-face corresponding to each triangle. This complex contains a bramble of order n and so Theorem 1 implies that a map from the complex to the plain associated to an s-stack layout will contain a point coverted by n triangles. A short pigeonhole argument then implies that $s \geqslant n^{1/3}$.

Theorem 1 further implies the asymptotically tight bound of Gromov on overlap number of complete 2-dimensional complex, i.e. that for every map from a complete 2-dimensional complex Δ_n^2 to the plane there is a point that belongs to the image of $\frac{2}{9}\binom{n}{3} + o(n^3)$ triangles via the following lemma.

Lemma 3. *Let* $\mathcal{B}$ *be a set of subcomplexes of* Δ_n^2 *which can be obtained by taking unions of cones of the graphs* $G \subseteq K_n$ *satisfying the following*

- $|V(G)| > 2/3n$,
- *for every* $X \subseteq V(G)$ *with* $|X| \leqslant \frac{1}{2}|V(G)|$, *there are more than* $\frac{1}{2}|X|(n - |X|)$ *edges in* G *from* X *to* $|V(G)| - X$.

Then $\mathcal{B}$ *is a bramble and* $\|\mathcal{B}\| = \frac{2}{9}\binom{n}{3} + o(n^3)$

The questions of obtaining brambles of order matching Gromov's bounds on overlap numbers of complete complexes in higher dimensions, and obtaining further high-dimensional extensions of structural results about brambles remain open.

References

[1] P. D. Seymour and R. Thomas. *Call routing and the ratcatcher*, Combinatorica **14** (1994): 217–241.

[2] M. Gromov, *Singularities, expanders and topology of maps. Part 2: From combinatorics to topology via algebraic isoperimetry*, Geometric and Functional Analysis **20.2** (2010): 416–526.

[3] D. Eppstein, L. Merker, S Norin, M. T. Seweryn, D. R. Wood. *Three-dimensional graph products with unbounded stack-number*, Discrete and Computational Geometry, to appear. (2022) arXiv preprint arXiv:2202.05327.

Common graphs with arbitrary chromatic number

Daniel Král'

(joint work with Jan Volec and Fan Wei)

Ramsey's Theorem [10] started a a significant amount of research on the presence of well-behaved substructures in large structures. In one of its simplest forms, the Ramsey's Theorem asserts that for every complete graph K_n, there exists

an integer N such that any 2-edge-coloring of the complete graph K_N contains a monochromatic copy of K_n. We are concerned with the quantitative version of this problem asking how many monochromatic copies of a graph H necessarily exist in any 2-edge-coloring of the complete graph K_N vertices and when the bound coming from the random construction is optimal.

Goodman's Theorem [5] states that the number of monochromatic copies of the triangle K_3 is asymptotically minimized by the random 2-edge-coloring, i.e., when each edge of a complete graph is colored randomly with one of two colors with probability 1/2. Inspired by this result, we say that a graph H is *common* if the number of monochromatic copies of H is asymptotically minimized by the random 2-edge-coloring of a complete graph. In particular, K_3 is common and more generally every cycle is common as proven in [13].

In 1962, Erdős [3] conjectured that every complete graph is common, and later Burr and Rosta [1] conjectured that every graph is common. However, both of these conjectures turned out to be false. In the late 1980s, Sidorenko [12, 13] showed that a triangle with a pendant edge is not common, and Thomason [14] showed that K_4 is not common. More generally, any graph containing K_4 is not common [8].

A characterization of the class of common graphs is an intriguing open problem and there is even no conjectured description of the class. This problem is closely related to the famous conjecture of Sidorenko [11] and of Erdős and Simonovits [4], which asserts every bipartite graph H has the Sidorenko property, i.e., the number of copies H in any graph is asymptotically at least the number of its copies in the random graph with the same density. Since every graph H with the Sidorenko property is common, as the number of copies of H in each color class is at least the expected number of its copies in the random edge-coloring, the conjecture, if true, would imply that all bipartite are common. Hence, families of bipartite graphs known to have the Sidorenko property provide examples of bipartite graphs that are common.

Common graphs that are not bipartite, i.e., their chromatic number is larger than two, are rare. In particular, Jagger, Šťovíček and Thomason asked whether there exists a common graph with chromatic number at least four. While odd cycles and even wheels [8, 13] are examples of 3-chromatic common graphs, also see [6], the existence of a common graph with chromatic number at least four was open until 2012 when the 5-wheel was proven to be common [7]. The question whether there exist common graphs with arbitrarily large chromatic number has been reiterated in [7], and also by Conlon, Fox and Sudakov in their survey paper [2, Problem 2.28].

Problem 1. *Do there exist common graphs of all chromatic numbers?*

We solve this problem by establishing the following.

Theorem 2. *For every $\ell \in \mathbb{N}$, there exists a connected common graph with chromatic number ℓ.*

Our techniques would however not yield the existence of a 3-connected high-chromatic common graph. So, it is natural to ask whether there exists an ℓ-chromatic k-connected common graph for all $\ell \geqslant 2$ and every $k \geqslant 3$. Ko and Lee [9] answered this in the affirmative by combining our construction and the book product of graphs.

References

[1] S. A. Burr and V. Rosta: *On the Ramsey multiplicities of graphs—problems and recent results*, J. Graph Theory **4** (1980), 347–361.

[2] D. Conlon, J. Fox and B. Sudakov: *Recent developments in graph Ramsey theory*, in: Surveys in combinatorics 2015, *London Math. Soc. Lecture Note Ser.*, volume 424 (2015), 49–118.

[3] P. Erdős: *On the number of complete subgraphs contained in certain graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **7** (1962), 459–464.

[4] P. Erdős and M. Simonovits: *Supersaturated graphs and hypergraphs*, Combinatorica **3** (1983), 181–192.

[5] A. W. Goodman: *On sets of acquaintances and strangers at any party*, Amer. Math. Monthly **66** (1959), 778–783.

[6] A. Grzesik, J. Lee, B. Lidický and J. Volec: *On tripartite common graphs*, to appear in Combin. Probab. Comput.

[7] H. Hatami, J. Hladký, D. Kráľ, S. Norine and A. Razborov: *Non-three-colourable common graphs exist*, Combin. Probab. Comput. **21** (2012), 734–742.

[8] C. Jagger, P. Šťovíček and A. Thomason: *Multiplicities of subgraphs*, Combinatorica **16** (1996), 123–141.

[9] S. Ko and J. Lee: *Common graphs with arbitrary connectivity and chromatic number*, preprint arXiv:2207.09427.

[10] F. P. Ramsey: *On a problem of formal logic*, Proceedings of the London Mathematical Society **2** (1930), 264–286.

[11] A. Sidorenko: *A correlation inequality for bipartite graphs*, Graphs Combin. **9** (1993), 201–204.

[12] A. F. Sidorenko: *Extremal problems in graph theory and functional-analytic inequalities*, Proceedings of the All-Union seminar on discrete mathematics and its applications (Russian) (1986), 99–105.

[13] A. F. Sidorenko: *Cycles in graphs and functional inequalities*, Mat. Zametki **46** (1989), 72–79, 104.

[14] A. Thomason: *A disproof of a conjecture of Erdős in Ramsey theory*, J. London Math. Soc. (2) **39** (1989), 246–255.

Geodesic Geometry on Graphs

Nati Linial

(joint work with Daniel Cizma)

The idea of viewing graphs from a geometric perspective has been immensely fruitful. Most of the existing connections between graph theory and differential geometry concern the eigenvalues of graphs. Here we study graphs from the perspective of *geodesic geometry*. Our main discovery is that for the vast majority of graphs the geodesic theory is way richer than the metric one.

Here is the main object that we study. In a graph $G = (V, E)$ we consider a system of paths $\mathcal{P} = \{P_{u,v} | u, v \in V\}$ where $P_{u,v}$ connects vertices u and v. This

system is said to be *consistent* if it is closed under taking subpaths. Namely, $\mathcal{P}$ has the property that whenever vertices y, z are in $P_{u,v}$, the sub-path of $P_{u,v}$ between them coincides with $P_{y,z}$. A map $w : E \to (0, \infty)$ is said to *induce* $\mathcal{P}$ if for every $u, v \in V$ the path $P_{u,v}$ is w-geodesic. The map w *strictly induces* $\mathcal{P}$ if in addition, for every $u, v \in V$, the path $P_{u,v}$ is the unique w-shortest path between them. We say that G is *metrizable* (resp. strictly metrizable) if every consistent path system is induced by some such w.

Here are our main findings:

- Metrizability is rare: E.g., (i) Every large 2-connected metrizable graph is planar, (ii) No large 3-connected graph is metrizable.
- However, arbitrarily large 2-connected metrizable graphs do exist: E.g., every outerplanar graph is metrizable.
- We reveal some of the structural underpinnings of metrizability. The class of metrizable graphs is closed under the topological minor relation and is characterized by finitely many forbidden topological minors.
- On the computational side, metrizability can be decided in polynomial time.

The role of computers in this work: All of the results that we present can be verified by hand, although this research would not be carried out without our use of the computer. Initially, we proved by hand that Petersen's graph is non-metrizable, but it quickly transpired that we needed a larger supply of such graphs. To this end we wrote a brute-force search program that found eleven such graphs and gave certificates that they are indeed non-metrizable. These certificates are easily verifiable by hand. Our proofs make substantial use of these graphs.

Some Examples. Not all path systems are metrizable. Figure 1 exhibits a non-metrizable path system in the Petersen graph Π.

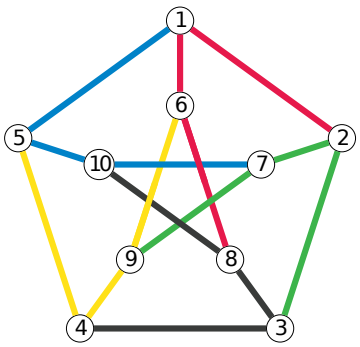

FIGURE 1. Non-metrizable path system in the Petersen Graph

If $uv \in E(\Pi)$, then the path P_{uv} is comprised of the single edge uv. Between any two nonadjacent vertices $x, y \in V(\Pi)$ there is a unique path of length 2. For most such pairs this is taken to be $P_{x,y}$. There are 5 exceptional pairs of nonadjacent

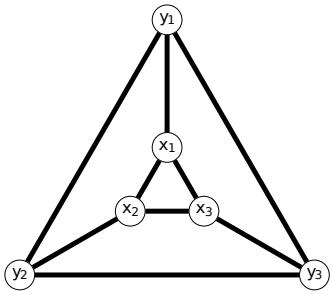

FIGURE 2. A graph which has path system which is a metrizable but not strictly metrizable

vertices, those which are connected by a colored path in Figure 1. For example, $P_{1,7} = 1, 5, 10, 7$. It is easily verified that this path system is consistent, and as we show next, this path system is nonmetrizable. If w is a weight function that induces it, then by considering the colored paths, the following inequalities must hold:

$$\begin{aligned} w_{1,2} + w_{1,6} + w_{6,8} &\leqslant w_{2,3} + w_{3,8} \\ w_{2,3} + w_{2,7} + w_{7,9} &\leqslant w_{3,4} + w_{4,9} \\ w_{3,4} + w_{3,8} + w_{8,10} &\leqslant w_{4,5} + w_{5,10} \\ w_{4,5} + w_{4,9} + w_{6,9} &\leqslant w_{1,5} + w_{1,6} \\ w_{1,5} + w_{5,10} + w_{7,10} &\leqslant w_{1,2} + w_{2,7} \end{aligned}$$

which implies

$$w_{6,8} + w_{7,9} + w_{8,10} + w_{6,9} + w_{7,10} \leqslant 0,$$

showing a weight function inducing these paths cannot be strictly positive.

Figure 2 shows a metrizable path system which is not strictly metrizable. Namely, every edge is the chosen path between its two vertices. For $i = 1, 2, 3$, let $P_{x_i,y_{i+1}} = x_i y_i y_{i+1}$ and $P_{y_i,x_{i+1}} = y_i x_i x_{i+1}$, with incides taken mod 3. It is easy to see that the constant weight function induces this path system. If a weight function w strictly induces this system, then for $i = 1, 2, 3$ the following inequalities must hold:

$$w(x_i y_i) + w(y_i y_{i+1}) < w(x_i x_{i+1}) + w(x_{i+1i} y_{i+1})$$

and

$$w(y_i x_i) + w(x_i x_{i+1}) < w(y_i y_{i+1}) + w(y_{i+1i} x_{i+1})$$

Summing the first inequality for $i = 1, 2, 3$ and canceling identical terms yields

$$\sum_{i=1}^{3} w(y_i y_{i+1}) < \sum_{i=1}^{3} w(x_i x_{i+1}).$$

Similarly, adding up the second inequality gives

$$\sum_{i=1}^{3} w(x_i x_{i+1}) < \sum_{i=1}^{3} w(y_i y_{i+1}),$$

a contradiction.

Combinatorics of Circle Squaring

Oleg Pikhurko
(joint work with András Máthé and Jonathan A. Noel)

1. Introduction

Tarski's Circle Squaring Problem [12] from 1925 asks if a circle (i.e., a circular disk) and a square of the same area in $\mathbb{R}^2$ are *equidecomposable*, that is, whether we can partition the circle into finitely many pieces and apply some isometry to each piece to get a partition of the square. This question was answered affirmatively some 65 years later by Laczkovich who showed in a deep and groundbreaking paper [5] that, in fact, it is possible to square a circle using translations only.

The Axiom of Choice plays a crucial role in his proof and, consequently, the pieces of his circle squaring could not be guaranteed to have any discernible regularity properties. A notable problem (mentioned by e.g., Wagon [13, Appendix C] or Laczkovich [5, Section 10]) has been to determine whether there exist circle squarings using "better structured" pieces. Recently, Grabowski, Máthé and Pikhurko [1] proved that the pieces of a circle squaring can simultaneously be Lebesgue measurable and have the property of Baire. Then, Marks and Unger [9] proved that the pieces can be made Borel. (Let us assume in this paper that the disk and the square are closed and thus Borel sets.) In fact, Marks and Unger [9, Section 7] showed that the pieces of a circle squaring can be chosen to be in $\boldsymbol{\mathcal{B}}(\boldsymbol{\Sigma}^0_4)$, where $\boldsymbol{\Sigma}^0_i$ is the i-th additive class of the standard Borel hierarchy (see e.g., [4, Section 11.B]) and $\boldsymbol{\mathcal{B}}(\mathcal{F})$ denotes the algebra generated by $\mathcal{F}$ (that is, the family of all Boolean combinations of elements from $\mathcal{F}$).

Shortly after his circle-squaring paper, Laczkovich [7, 6] proved a far-reaching generalisation. Before stating it, let us set up some notation. Fix $k \geqslant 1$. Let λ denote the Lebesgue measure on $\mathbb{R}^k$ and let ∂X denote the (topological) boundary of $X \subseteq \mathbb{R}^k$. Recall that the *upper Minkowski dimension*, sometimes called *box* or *grid dimension*, of $X \subseteq \mathbb{R}^k$ is

$$\dim_{\square}(X) := \limsup_{\delta \to 0^+} \frac{\log(N_\delta(X))}{\log(\delta^{-1})}, \tag{1}$$

where $N_\delta(X)$ is the number of boxes from the regular grid in $\mathbb{R}^k$ of side-length δ that intersect X.

Theorem 1 (Laczkovich [7, 6])**.** *If $k \geqslant 1$ and $A, B \subseteq \mathbb{R}^k$ are bounded sets such that $\lambda(A) = \lambda(B) > 0$, $\dim_{\square}(\partial A) < k$ and $\dim_{\square}(\partial B) < k$, then A and B are equidecomposable by translations.*

The subsequent papers [1, 9] on circle squaring in fact prove appropriate "constructive" versions of Theorem 1 and derive the corresponding circle squaring results as special cases. Very recently, Noel, Máthé and Pikhurko [11] decreased the Borel complexity of the pieces by two hierarchy levels and ensured that the boundary of each piece is "small" in a strong sense:

Theorem 2. *In $\mathbb{R}^2$, a closed disk and square of the same area can be equidecomposed using translations so that every piece has boundary of upper Minkowski dimension at most* 1.987*, belongs to $\boldsymbol{\mathcal{B}}(\boldsymbol{\Sigma}^0_2)$ (i.e., is a Boolean combination of F_σ sets), and has positive Lebesgue measure.*

Recall that a subset $X \subseteq \mathbb{R}^k$ is *Jordan measurable* if its indicator function is Riemann integrable. An equivalent definition is that X is bounded and $\lambda(\partial X) = 0$. It easily follows that any bounded set $X \subseteq \mathbb{R}^k$ with $\dim_\square(\partial X) < k$ is Jordan measurable. Therefore, Theorem 2 implies that circle squaring is possible with Jordan measurable pieces, which addresses questions by Laczkovich [8] and Máthé [10, Question 6.2]. An advantage of a Jordan measurable circle squaring is that an arbitrarily large portion of it can be described in an error-free way with finitely many bits of information. Namely, for every $\varepsilon > 0$, if n is large enough, then at most εn^2 boxes of the regular $n \times n$ grid on the equidecomposed unit square can intersect more than one piece and thus $O(n^2)$ bits are enough to describe our equidecompostion up to a set of measure at most ε. (Furthermore, the dimension estimate of Theorem 2 shows that ε, as a function of $n \to \infty$, can be taken to be $n^{-0.013+o(1)}$.)

Theorem 2 is obtained as a special case of the following general result. For a set $A \subseteq \mathbb{R}^k$, let $\mathcal{T}_A := \{A + \boldsymbol{t} : \boldsymbol{t} \in \mathbb{R}^k\}$ consist of all its translations. For a family $\mathcal{F}$ of sets, let $\boldsymbol{\Sigma}(\mathcal{F})$ be the collection of all countable unions of sets in $\mathcal{F}$. Also, recall that $\boldsymbol{\Sigma}^0_1$ stands for the collection of open sets in $\mathbb{R}^k$.

Theorem 3. *If $k \geqslant 1$ and $A, B \subseteq \mathbb{R}^k$ are bounded sets such that $\lambda(A) = \lambda(B) > 0$, $\dim_\square(\partial A) < k$ and $\dim_\square(\partial B) < k$, then A and B are equidecomposable by translations so that all the following statements hold simultaneously:*

(a) *for some explicit $\zeta = \zeta(k, \dim_\square(\partial A), \dim_\square(\partial B)) > 0$ the topological boundary of each piece has upper Minkowski dimension at most $k - \zeta$,*
(b) *each piece belongs to $\boldsymbol{\mathcal{B}}(\boldsymbol{\Sigma}(\boldsymbol{\mathcal{B}}(\boldsymbol{\Sigma}^0_1 \cup \mathcal{T}_A \cup \mathcal{T}_B)))$,*
(c) *if*

$$\lambda\left(\left\{\boldsymbol{t} \in \mathbb{R}^k : (A + \boldsymbol{t}) \cap B \neq \varnothing \text{ and } \lambda\left((A + \boldsymbol{t}) \cap B\right) = 0\right\}\right) = 0 \tag{2}$$

(that is, the set of vectors $\boldsymbol{t} \in \mathbb{R}^k$ such that $(A + \boldsymbol{t}) \cap B$ is non-empty and Lebesgue-null has measure 0), then each piece has positive Lebesgue measure.

If the sets A and B in Theorem 3 are Borel, then, by Part (b), all pieces of the equidecomposition can be taken to be Borel with a good control over their Borel complexity.

On the other hand, if $k \geqslant 3$ and one allows all orientation-preserving isometries of $\mathbb{R}^k$, then the obvious necessary conditions for a set to be equidecomposable to a

cube using Lebesgue (resp. Baire) measurable pieces turn out to be sufficient, see Grabowski, Máthé and Pikhurko [2, Corollary 1.10]. However, nothing like this is known for Borel and Jordan measurable equidecompositions.

2. Some Proof Ideas

Let us give a very high-level outline of the proof of Part (a) of Theorem 3.

Like in the previous work, we assume that A and B are subsets of the torus $\mathbb{T}^k := \mathbb{R}^k/\mathbb{Z}^k$ and do all translations modulo 1. We pick a suitable (somewhat large) integer d and vectors $\boldsymbol{x}_1, \dots, \boldsymbol{x}_d \in \mathbb{T}^k$ satisfying certain conditions (that are satisfied with positive probability by random vectors). Let G_d be the graph on $\mathbb{T}^k$ where we connect $\boldsymbol{u}$ to $\boldsymbol{u} + \sum_{i=1}^d n_i \boldsymbol{x}_i$ for each non-zero $(n_1, \dots, n_d) \in \{-1, 0, 1\}^d$. Assuming that $\boldsymbol{x}_1, \dots, \boldsymbol{x}_d$ do not satisfy any linear dependencies with rational coefficients, each component of G_d is a $(3^d - 1)$-regular graph on a copy of $\mathbb{Z}^d$.

Our aim is to "construct" a bijection from A to B such that, for some constant r, each element of A is moved by the bijection by distance at most r within the graph G_d. Such a bijection naturally gives an equidecomposition between A and B that uses at most $(2r+1)^d$ pieces. As was observed by Marks and Unger [9], the problem of finding such a bijection can be reduced to finding a uniformly bounded integer-valued flow within the graph G_d, where the demand is 1 on A, -1 on B and 0 elsewhere.

As one of the first steps of their proof, Marks and Unger [9] constructed a real-valued (i.e. not necessarily integer-valued) flow f_∞ which satisfies these demands. The flow f_∞ is defined to be a pointwise limit of a sequence of flows f_m that are locally constructed from A and B. Since the collection of bounded subsets of $\mathbb{R}^k$ with boundary of upper Minkowski dimension at most $k-\zeta$ is not a σ-algebra, we should not use the values of f_∞ if we want to produce pieces with this structure. Instead, we work with the locally defined approximations f_m.

For flow rounding (that is, making all flow values integer), we construct Jordan measurable subsets $J_1, J_2, \dots$ of $\mathbb{T}^k$ such that their union $\bigcup_{i=1}^\infty J_i$ is co-null in $\mathbb{T}^k$ and $(J_i)_{i=1}^\infty$ is a *toast sequence*, roughly meaning that each J_i induces only finite (in fact, uniformly bounded) components in G_d and the graph boundaries of all components arising this way are well separated from each other. In fact, each set J_i is a finite union of "strips," i.e., sets of the form $[a, b) \times [0, 1)^{k-1}$; in particular, it is Borel and has a $(k-1)$-dimensional boundary. The idea of using toast sequences to construct satisfying assignments was previously applied to many problems in descriptive combinatorics (with the exact definition of "toast sequence" often being problem-specific). For a systematic treatment of this idea for general actions of $\mathbb{Z}^d$, we refer the reader to Grebík and Rozhoň [3].

We can view the toast sequence $(J_1, J_2, \dots)$ as a process where, at time i, vertices of the set J_i arrive and our algorithm has to decide the value of the final integer flow f on every edge with at least one vertex in this set. We are not allowed to look into the future nor modify any already defined values of the flow f. We prove that, if all things are set up carefully, then this is indeed possible to do and, in fact, there are some constants m_i and R_i such that the value of f on any

edge $xy \in E(G_d)$ intersecting J_i can be computed only from the current picture in the R_i-neighbourhood of $\{x, y\}$ in G_d and the values of the approximation f_{m_i} of f_∞ there. Here, a key challenge is that, when we round the flow on J_i, we have only incomplete information (namely, the flow f_{m_i} which meets the demands only within some small error). The idea that allows us to overcome this difficulty is that, if the cumulative error of f_{m_i} on each component of J_i is small, then whenever our algorithm encounters some inconsistency, it can round it to the nearest integer and produce values that are in fact perfectly compatible with all past and future choices of the algorithm.

The proof coming from the above arguments, with a careful choice of how the size of the components of J_i can grow with i, produces a partial equidecomposition between A and B so that the topological boundaries of the pieces as well as the unmatched part of A and of B have upper Minkowski dimension less than k. Thus if we can extend this equidecomposition to all of A and B (even by using the Axiom of Choice), then we can achieve Part (a) of Theorem 3.

References

[1] Ł. Grabowski, A. Máthé, and O. Pikhurko. Measurable circle squaring. *Annals of Math. (2)*, 185:671–710, 2017.

[2] Ł. Grabowski, A. Máthé, and O. Pikhurko. Measurable equidecompositions for group actions with an expansion property. *J. Europ. Math. Soc*, 24:4277–4326, 2022.

[3] J. Grebík and V. Rozhoň. Local problems on grids from the perspective of distributed algorithms, finitary factors, and descriptive combinatorics. E-print arxiv:2103.08394, 2021.

[4] A. S. Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.

[5] M. Laczkovich. Equidecomposability and discrepancy; a solution of Tarski's circle-squaring problem. *J. Reine Angew. Math.*, 404:77–117, 1990.

[6] M. Laczkovich. Decomposition of sets with small boundary. *J. Lond. Math. Soc.*, 46:58–64, 1992.

[7] M. Laczkovich. Uniformly spread discrete sets in $\mathbf{R}^d$. *J. Lond. Math. Soc.*, 46:39–57, 1992.

[8] M. Laczkovich. The story of squaring the circle, July 2017. Workshop "Geometric measure theory", Warwick, 10–14 July 2017. Talk slides: `http://homepages.warwick.ac.uk/~masfay/workshop_2017/Slides/Laczkovich.pdf`.

[9] A. S. Marks and S. Unger. Borel circle squaring. *Annals of Math. (2)*, 186:581–605, 2017.

[10] A. Máthé. Measurable equidecompositions. In *Proceedings of the International Congress of Mathematicians. Volume 2*, pages 1709–1728. World Scientific, 2018.

[11] A. Máthé, J. A. Noel, and O. Pikhurko. Circle squaring with pieces of small boundary and low borel complexity. E-print arxiv:2202.01412, 2022.

[12] A. Tarski. Problème 38. *Fund. Math.*, 7:381, 1925.

[13] S. Wagon. Circle-squaring in the twentieth century. *Math. Intelligencer*, 3:176–181, 1980/81.

Anticoncentration in Ramsey graphs and a proof of the Erdős–McKay conjecture

Mehtaab Sawhney

(joint work with Matthew Kwan, Ashwin Sah, Lisa Sauermann)

We call a graph G on n-vertices, C-Ramsey if G contains no clique or independent set (e.g. homogenous set) of size $C\log_2(n)$. For the purposes of this abstract, we will imagine C as a fixed constant therefore we are considering graphs with "nearly-optimal" Ramsey behavior. In particular, the fundamental bound of Erdős and Szekeres [8] on diagonal Ramsey numbers implies that no graph is 1/2-Ramsey and the famous probablistic construction of Erdős [7] implies that for $n \geqslant 3$ their exist graphs G which are 2-Ramsey.

Despite significant effort (see for example [2, 16] and references therein), there are no known non-probabilistic constructions of graphs with comparably small homogeneous sets. Therefore Erdős and collaborators posed a series of questions which probe the extent to which C-Ramsey graphs are "random"-like, in particular by asking the extent to which the induced subgraphs of a C-Ramsey graph G are diverse. The first result in this direction, due to Erdős and Szemerédi [9], proved there is $\varepsilon_C > 0$ such that for every C-Ramsey graph G on n vertices we have $e(G) \geqslant \varepsilon_C\binom{n}{2} \geqslant \varepsilon_C n^2/4$. Subsequently Prömel and Rödl [20] (answering a conjecture of Erdős and Hajnal) proved that G is $\delta_C \log n$ universal, Shelah [21] (answering a conjecture of Erdős and Rényi) proved that G contains $2^{\delta_C n}$ non-isomorphic induced subgraphs, Kwan and Sudakov [14] (answering a conjecture of Erdős, Faudree, and Sós) proved that G contains $\delta_C n^{5/2}$ subgraphs that can be distinguished by looking at their edge and vertex numbers; and by Jenssen, Keevash, Long, and Yepremyan [13] proved (improving on a conjecture of Erdős, Faudree, and Sós proved by Bukh and Sudakov [3]) that G contains an induced subgraph with $\delta_C n^{2/3}$ distinct degrees (all for some $\delta_C > 0$ depending on C).

Only one of Erdős' conjectures (on properties of C-Ramsey graphs) from this period has remained open: Erdős and McKay (see [10]) ambitiously conjectured that there is $\delta_C > 0$ such that for any C-Ramsey graph G with n vertices and any integer $0 \leqslant x \leqslant \delta_C n^2$, there is an induced subgraph of G with exactly x edges. Erdős reiterated this problem in several collections of his favorite open problems in combinatorics [10, 11] (also in [12]), and offered one of his notorious monetary prizes ($100) for its solution (see [11, 6, 5]).

Progress on the Erdős–McKay conjecture has come from four different directions. First, Calkin, Frieze and McKay [4] (answering questions raised by Erdős and McKay) proved that for any constants $p \in (0,1)$ and $\eta > 0$, a random graph $\mathbb{G}(n,p)$ typically contains induced subgraphs with all numbers of edges up to $(1-\eta)p\binom{n}{2}$. Second, improving on initial bounds of Erdős and McKay [10], it was proved by Alon, Krivelevich, and Sudakov [1] that there is $\alpha_C > 0$ such that in a C-Ramsey graph on n vertices, one can always find an induced subgraph with any given number of edges up to n^{α_C}. Third, improving on a result of Narayanan, Sahasrabudhe, and Tomon [19], Kwan and Sudakov [15] proved that there is $\delta_C > 0$

such that in any C-Ramsey graph on n vertices contains induced subgraphs with $\delta_C n^2$ different numbers of edges (without making any guarantee on what those numbers of edges are). Finally, Long and Ploscaru [17] recently proved a *bipartite* analog of the Erdős–McKay conjecture.

As our first result, we prove a substantial strengthening of the Erdős–McKay conjecture.

Theorem 1. *Fix $C > 0$ and $\eta > 0$, and let G be a C-Ramsey graph on n vertices, where n is sufficiently large with respect to C and η. Then for any integer x with $0 \leqslant x \leqslant (1-\eta)e(G)$, there is a subset $U \subseteq V(G)$ inducing exactly x edges.*

Given prior results due to Alon, Krivelevich and Sudakov [1], Theorem 1 is actually a simple corollary of the following deeper result on edge-statistics in Ramsey graphs.

Theorem 2. *Fix $C, \lambda > 0$, let G be a C-Ramsey graph on n vertices and let $\lambda \leqslant p \leqslant 1-\lambda$. Then if U is a random subset of $V(G)$ obtained by independently including each vertex with probability p, we have*

$$\sup_{x \in \mathbb{Z}} \Pr[e(G[U]) = x] \leqslant K_{C,\lambda} n^{-3/2}$$

for some $K_{C,\lambda} > 0$ depending only on C, λ. Furthermore, for every fixed $A > 0$, we have

$$\inf_{\substack{x \in \mathbb{Z} \\ |x - p^2 e(G)| \leqslant A n^{3/2}}} \Pr[e(G[U]) = x] \geqslant \kappa_{C,A,\lambda} n^{-3/2}$$

for some $\kappa_{C,A,\lambda} > 0$ depending only on C, A, λ, if n is sufficiently large in terms of C, λ and A.

The proof of Theorem 2 involves a range of tools from Fourier analysis, random matrix theory, the theory of Boolean functions, probabilistic combinatorics, and low-rank approximation. To give a sense of the starting point of our analysis, let us consider $e(G[U])$ as a random variable and let A denote the adjacency matrix of G. Let v denote the indicator vector of the set U and $x_u = 2v_u - 1$ for all $u \in V(G)$. We then see

$$e(G[U]) \stackrel{d.}{=} \frac{x^T A x}{2} \stackrel{d.}{=} \frac{e(G)}{4} + \frac{1}{4} \sum_{v \in V(G)} \deg_G(v) x_v + \frac{1}{4} \sum_{uv \in E(G)} x_u x_v.$$

Given this rewriting of the random variable (which is essentially the Fourier Walsh expansion) we have that $e(G[U])$ breaks into understanding a linear contribution $\sum_{v \in V(G)} \deg_G(v) x_v$ and a quadratic contribution $\frac{1}{4} \sum_{uv \in E(G)} x_u x_v$. A straightforward computation proves that

$$\mathrm{Var}\Big(\sum_{v \in V(G)} \deg_G(v) x_v\Big) \asymp n^3, \ \mathrm{Var}\Big(\sum_{uv \in E(G)} x_u x_v\Big) \asymp n^2.$$

Therefore a standard quantitative central limit theorem implies that $e(G[U])$ satisfies a bulk-central limit theorem; the first obstacle in our proof is that this does not necessarily hold down to the finest scale.

To see this, consider a random $(n/2)$-regular graph G (which with high probability is 100-Ramsey). It follows that $\sum_{v\in V(G)} \deg_G(v)x_v \in (n/2)\mathbb{Z}$; this unfortunately is problematic as the remaining randomness in $\sum_{uv\in E(G)} x_u x_v$ can be seen to not provide a sufficient smoothing to even have a "local" central limit theorem. (In particular the distribution of $e(G[U])$ exhibits spikes within its distribution.) This phenomenon is quantified via the regularized least common denominator (RLCD) introduced by Vershynin [22] in the context of random matrix theory. In the case when the RLCD is suitably large, one can essentially prove a local central limit theorem and when the RLCD is sufficiently small the degrees of the graph G can be grouped into a small number of buckets such that the degree are essentially equal within a bucket. In this case it turns out to be fruitful to consider the distribution when the number of vertices chosen in a bucket is fixed and for the sake of simplicity when G is regular it suffices to consider a single bucket.

At this point the key issue is to understand the distribution of $\sum_{uv\in E(G)} x_u x_v$ conditional on $\sum_{u\in V(G)} x_u$. For the sake of simplicity, when x_u are independent (although the conditional problem poses additional technical difficulties). The crucial property we would require for the quadratic part to "smooth" out the distribution of the linear part would be that $\sum_{uv\in E(G)} x_u x_v$ is both anti-concentrated and has an unbounded tail in at least one direction. At this point, both Boolean functions and low-rank approximation enter as $\sum_{uv\in E(G)} x_u x_v = \frac{x^T A x}{2}$ has the same bulk distribution as when x is replaced by $z \sim \mathcal{N}(0,1)^{\otimes n}$ due to *Gaussian invariance principle* of Mossel, O'Donnell, and Oleszkiewicz [18]. Given this, and rotational invariance of the Gaussian ensemble, the bulk distribution is determined by the eigenvalues values of A. In particular, we prove under the assumption that A is not well approximated by low-rank matrices in Frobenious norm, the quadratic Gaussian polynomial associated to A satisfies the necessary anticoncentration and tail behavior under the assumption and that the adjacency matrices of Ramsey graphs are not well-approximated by low-rank forms. (In fact for our proof we require the stronger property that the matrices are not well approximated by block low-rank forms which is substantially more involved.)

Even at this point the remainder of the proof is still rather involved; in particular this summary omits how one proves via decoupling that the distribution is smooth at all other scales, how one integrates the above distribution for a fixed size sample to obtain Theorem 2 among a host of other technical issues.

References

[1] Noga Alon, Michael Krivelevich, and Benny Sudakov, *Induced subgraphs of prescribed size*, J. Graph Theory **43** (2003), 239–251.

[2] Boaz Barak, Anup Rao, Ronen Shaltiel, and Avi Wigderson, *2-source dispersers for $n^{o(1)}$ entropy, and Ramsey graphs beating the Frankl-Wilson construction*, Ann. of Math. (2) **176** (2012), 1483–1543.

[3] Boris Bukh and Benny Sudakov, *Induced subgraphs of Ramsey graphs with many distinct degrees*, J. Combin. Theory Ser. B **97** (2007), 612–619.

[4] Neil Calkin, Alan Frieze, and Brendan D. McKay, *On subgraph sizes in random graphs*, Combin. Probab. Comput. **1** (1992), 123–134.

[5] F. R. K. Chung, *Open problems of Paul Erdős in graph theory*, J. Graph Theory **25** (1997), 3–36.
[6] Fan Chung and Ron Graham, *Erdős on graphs*, A K Peters, Ltd., Wellesley, MA, 1998, His legacy of unsolved problems.
[7] P. Erdős, *Some remarks on the theory of graphs*, Bull. Amer. Math. Soc. **53** (1947), 292–294.
[8] P. Erdős and G. Szekeres, *A combinatorial problem in geometry*, Compositio Math. **2** (1935), 463–470.
[9] P. Erdős and A. Szemerédi, *On a Ramsey type theorem*, Period. Math. Hungar. **2** (1972), 295–299.
[10] Paul Erdős, *Some of my favourite problems in various branches of combinatorics.* Combinatorics 92 (Catania, 1992), Matematiche (Catania) **47** (1992), 231–240 (1993).
[11] Paul Erdős, *Some of my favourite problems in number theory, combinatorics, and geometry*, Resenhas **2** (1995), 165–186, Combinatorics Week (Portuguese) (São Paulo, 1994).
[12] Paul Erdős, *Some recent problems and results in graph theory.* The Second Krakow Conference on Graph Theory (Zgorzelisko, 1994), Discrete Math. **164** (1997), 81–85.
[13] Matthew Jenssen, Peter Keevash, Eoin Long, and Liana Yepremyan, *Distinct degrees in induced subgraphs*, Proc. Amer. Math. Soc. **148** (2020), 3835–3846.
[14] Matthew Kwan and Benny Sudakov, *Proof of a conjecture on induced subgraphs of Ramsey graphs*, Trans. Amer. Math. Soc. **372** (2019), 5571–5594.
[15] Matthew Kwan and Benny Sudakov, *Ramsey graphs induce subgraphs of quadratically many sizes*, Int. Math. Res. Not. IMRN (2020), 1621–1638.
[16] Xin Li, *Non-malleable extractors and non-malleable codes: partially optimal constructions*, 34th Computational Complexity Conference, LIPIcs. Leibniz Int. Proc. Inform., vol. 137, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2019, pp. Art. No. 28, 49.
[17] Eoin Long and Laurențiu Ploscaru, *A bipartite version of the Erdős–McKay conjecture*, arXiv:2207.12874.
[18] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz, *Noise stability of functions with low influences: invariance and optimality*, Ann. of Math. (2) **171** (2010), 295–341.
[19] Bhargav Narayanan, Julian Sahasrabudhe, and István Tomon, *Ramsey graphs induce subgraphs of many different sizes*, Combinatorica **39** (2019), 215–237.
[20] Hans Jürgen Prömel and Vojtěech Rödl, *Non-Ramsey graphs are $c \log n$-universal*, J. Combin. Theory Ser. A **88** (1999), 379–384.
[21] Saharon Shelah, *Erdős and Rényi conjecture*, J. Combin. Theory Ser. A **82** (1998), 179–185.
[22] Roman Vershynin, *Invertibility of symmetric random matrices*, Random Structures Algorithms **44** (2014), 135–182.

Product free sets in groups

Noam Lifshitz
(joint work with Peter Keevash and Dor Minzer)

1. Erdős–Ko–Rado type theorems

A family $\mathcal{F} \subseteq \binom{[n]}{k}$ is said to be *intersecting* if $A \cap B \neq \varnothing$ for all $A, B \in \mathcal{F}$. One example of an intersecting family is the *dictator*

$$\left\{ A \in \binom{[n]}{k} : i \in A \right\}.$$

The Erdős–Ko–Rado theorem [5] says the following.

Theorem 1. *Let $k < \frac{n}{2}$. The extremal intersecting family in $\binom{[n]}{k}$ is the dictator.*

The Erdős–Ko–Rado theorem opened a whole sub-field of extremal combinatorics concerning extremal problems whose solution is a dictator. The terminology stems from the theory of Boolean functions, where the dictators are given by the functions $f(x) = x_i$. The dictators are special case of juntas. A function $f\colon \{0,1\}^n \to \{0,1\}$ is said to be a j-junta if it depends on j variables. The corresponding notion for a family of sets is given by setting $\mathcal{F} \subseteq \binom{[n]}{k}$ to be a j-junta if there exists a set J of size j, such that the question whether a set A belongs to $\mathcal{F}$ or not depends only on $A \cap J$. Dinur and Friedgut showed that every intersecting family is essentially contained in an intersecting junta. Keller and Lifshitz then showed that the solution to various long standing Erdős–Ko–Rado type theorems is indeed a dictator by proving a similar approximation by junta theorem and then upgrading the rough junta structure into an exact extremal result. In this talk we concern a similar phenomenon where a rough junta structure is upgraded into an exact extremal solution by a dictator, but this time the problem comes from group theory and additive combinatorics.

2. Product free sets in groups

Let G be a group. A subset $A \subseteq G$ is said to be product free if for all a, b in A the group element ab is not in A. On 1985 Babai and Sós [1] posed the following problem.

Problem 2. *Let $A \subseteq A_n$ be a produduct free subset of the alternating group. What's the largest possible value of $|A|$?*

This problem will turn out to be an Erdős–Ko–Rado type problem, where the extremal solution is a dictator of sorts. What are dictators in A_n? By analogy from Boolean functions and set systems it makes sense to define $A \subseteq A_n$ to be a dictator if there exists $x \in [n]$, such that the question whether σ belongs to A or not depends only on $\sigma(x)$. In other words, the dictators are the sets of the form

$$D_{x,I} = \{\sigma \in A_n : \sigma(x) \in I\}.$$

The only problem with this definition is that the dictators are no longer product free. This led Kedlaya to give the following modification of the dictator

$$K_{x,I} = \{\sigma \in D_{x,I} : \sigma(I) \subseteq I^c\}$$

which is easily seen to be product free. The optimal value of I is of order $\Theta(\sqrt{n})$ and for such size of I the condition that $\sigma(I) \subseteq I^c$ occurs with a constant probability. Babai and Sós conjectured that there exists a much larger product free subset $A \subseteq A_n$ of density $\Omega(1)$. This was refuted by Gowers [7] who showed in 2008 that if $A \subseteq A_n$ is product free, then $\frac{|A|}{|A_n|} \leqslant (n-1)^{\frac{1}{3}}$. This was later improved by Eberhard [4] to

$$\frac{|A|}{|A_n|} \leqslant n^{-1/2} O\left(\log^{\frac{7}{2}} n\right).$$

We completely solve this problem when n is sufficiently large.

Theorem 3 (Keevash, L. and Minzer)**.** *There exists n_0, such that for all $n \geqslant n_0$ the extremal product free set in A_n is of the form $K_{x,I}$ for some x, I.*

3. Symmetry and pseudorandomness

The problem of determining the largest product free sets is closely related to the study of pseudorandomness in Cayley graphs. Indeed, a set A is product free if and only if in the Cayley graph $\mathrm{Cay}(G, A)$ the set A is independent when viewed as a set of vertices. Such a large independent set shouldn't exist in a random graph. Now the symmetries of the Cayley graph $\mathrm{Cay}(G, A)$ contain the right action of the group G on the set of vertices. Indeed, this follows from the fact that right multiplication commutes with multiplication from the left. We are used to think of symmetry and randomness as contradictory properties. However, the paradoxical idea in Gowers proof, (which originated in the work of Sarnak and Xue) is that symmetry can sometimes imply pseudorandomness.

Definition 4. *We say that a graph $G = (V, E)$ is ϵ-mixing if for each two set A, B of densitities $\alpha := \frac{|A|}{|V|}, \beta := \frac{|B|}{|V|} \geqslant \epsilon$ we have*

$$\frac{\Pr_{\{v,u\}\sim E}[v \in A, u \in B]}{\alpha\beta} \in (0.99, 1.01).$$

We say that a group G is ϵ-mixing if each Cayley graph $\mathrm{Cay}(G, A)$ of density $\frac{|A|}{|G|} \geqslant \epsilon$ is ϵ-mixing.

Gowers then coined the notion of a quasirandom group. A group G is said to be D-quasirandom if the minimal dimension of an irreducible representation of G is $\geqslant d$. Let us write $\epsilon(G)$ for the minimal ϵ for which G is ϵ-mixing and $D(G)$ for the minimal dimension of an irreducible representation of G. Gowers Showed that if G is D-quasirandom, then it is $O\left(D^{-\frac{1}{3}}\right)$-mixing. Nikolov and Pyber used the classification of finite simple groups to show that some converse of this holds and actually $D(G)$ and $\frac{1}{\epsilon(G)}$ are polynomially related for every finite group G. It is a remarkable fact that the combinatorial quantity $\frac{1}{\epsilon(G)}$ and the representation theoretic quantity are actually polynomially equivalent.

4. Analysis of Boolean functions

Let $f\colon \{-1,1\}^n \to \mathbb{R}$. The Fourier expansion of f is given by

$$f = \sum_{S\subseteq[n]} \widehat{f}(S)\chi_S,$$

where $\chi_S = \prod_{i\in S} x_i$. Given $x \in \{-1,1\}^n$ the ρ-noisy distribution is given by choosing $y \in \{-1,1\}^n$ according to the following process. Independently for each i we remember x_i with probability ρ and set $y_i = x_i$, and with probability $1-\rho$ we forget x_i and choose $y_i \sim \{-1,1\}$ uniformly at random.

The noise operator T_ρ on the Boolean cube is given by

$$\mathrm{T}_\rho f(x) = \mathbb{E}_{y\sim N_\rho}[f(x)].$$

The noise operator has the following Fourier formula

$$\mathrm{T}_\rho f = \sum_{S\subseteq[n]} \rho^{|S|}\hat{f}(S)\,\chi_S.$$

One of the central results in the analysis of the Boolean cube is the Bonami–Gross–Beckner hypercontractive inequality [3, 8, 2]

$$\|\mathrm{T}_\rho f\|_q \leqslant \|f\|_p$$

for all $\rho \leqslant \sqrt{\frac{p-1}{q-1}}$ and $q > p$. It is called hypercontractivity since the L_q-unit ball is always contained in the L_p-unit ball. The hypercontractive inequality then says that the noise operator squeezes the large L_p ball into the small L_q-ball. Hypercontractivity immediately implies that if f is a homogeneous polynomial of degree d, then $\|f\|_4 \leqslant \sqrt{3}^d\|f\|_2$. On the other hand when $f = 1_A$ and $\mathbb{E}[f] = \alpha$ we have $\|f\|_4 = \alpha^{-\frac{1}{4}}\|f\|_2$. This shows that indicators of small sets and low degree functions behave very differently. In fact, one can use the hypercontractive inequality to deduce that small sets are essentially orthogonal to all low degree functions. Here the degree of a function is meant to be its degree as a multilinear polynomial.

5. Going beyond quasirandomness via hypercontractivity

When studying product free sets our idea is to incorporate the above tools from the Boolean world into the symmetric group. It turns out that this plan fits perfectly with the representation theory of the alternating group. In $L^2(A_n)$ we also have a degree decomposition. We write $V_{\leqslant d}$ for the linear space spanned by products of d dictators. We then write $V_{=d} = V_{\leqslant d} \cap V_{\leqslant d-1}^{\perp}$. The space $V_{=d}$ is an A_n-bimodule. Moreover, it turns out that all its sub-representations have dimension $\geqslant \left(\frac{cn}{d}\right)^d$. This shows that the higher the degree is the higher of a minimal dimension of an irreducible representation is. Our idea is now to say that our product free sets A satisfies that 1_A is essentially orthogonal to all the low degree representations and effectively our group should behave like a group with a much better quasirandomness parameter $D(G)$. The problem with this approach is that hypercontractivity no longer holds in the symmetric group. Instead Filmus, Kindler, Lifshitz, and Minzer [6] proved a refined notion called hypercontractivity for global functions in the symmetric group. Their bound lies in the base of our approach. The work of Filmus et al is based on an earlier works of Keevash, Lifshitz, Long, and Minzer [9].

References

[1] L. Babai, V. Sós, Sidon sets in groups and induced subgraphs of Cayley graphs, European J. Combin. 6 (1985), 101–114.

[2] William Beckner. Inequalities in Fourier analysis. Annals of Mathematics, pages 159–182, 1975

[3] Aline Bonami. Étude des coefficients de Fourier des fonctions de $L^p(G)$. In Annales de l'institut Fourier, volume 20(2), pages 335–402, 1970.

[4] Sean Eberhard. Product mixing in the alternating group. Discrete Analysis, 2:18pp, 2016.

[5] Paul Erdős, Chao Ko, and Richard Rado. Intersection theorems for systems of finite sets. The Oxford Quarterly Journal of Mathematics, 12(1):313–320, 1961.
[6] Yuval Filmus, Guy Kindler, Noam Lifshitz, and Dor Minzer. Hypercontractivity on the symmetric group. arXiv preprint arXiv:2009.05503, 2020.
[7] W. Tim Gowers. Quasirandom groups. Combinatorics, Probability and Computing, 17(3):363–387, 2008.
[8] L. Gross. Logarithmic Sobolev inequalities. Amer. J. Math., 97:1061–1083, 1975.
[9] Peter Keevash, Noam Lifshitz, Eoin Long, and Dor Minzer. Global hypercontractivity and its applications. arXiv preprint arXiv:2103.04604, 2021.

The space of triangulations of a planar point set

Uli Wagner
(joint work with Anna Lubiw, Zuzana Masárová, and Emo Welzl)

Let $P \subset \mathbf{R}^2$ be a set of n points in general position (no collinear triples). A *triangulation* of P is a maximal straight-edge embedded plane graph with vertex set P. If P is fixed, we may identify a triangulation with its set of edges, which is an inclusion-maximal subset T of segments spanned by pairs of points in P such that any two segments in T are *non-crossing* (they are either disjoint or share a common endpoint). Two triangulations T and T' of P differ by a *flip* (also called *edge flip* or *diagonal flip*) if their symmetric difference $T \oplus T'$ consists of two segments $e \in T$ and $e' \in T'$ that form the diagonals of a convex quadrilateral whose interior does not contain any points of P.

The *flip graph* of P is the abstract graph whose vertices are the triangulations of P, and two triangulations are connected by an edge in the flip graph if they differ by a flip. It has been well-known since the work of Lawson [4] that the flip graph is connected. In joint work with Welzl [7], we strengthen this and determine the vertex connectivity of the flip graph. A trivial upper bound for the vertex connectivity of any graph is its minimal degree. Hurtado, Noy, and Urrutia [3] showed that the minimum degree $\delta(P)$ of the flip graph of a set P of n points in general position is at least $\lceil \frac{n-4}{2} \rceil$ (i.e., any triangulation of P contains at least that many flippable edges), and this is tight in the worst-case. We show that this upper bound is sharp (provided n is sufficiently large).

Theorem 1 (W.–Welzl)**.** *There exists an integer n_0 such that the flip graph of any set P of $n \geqslant n_0$ points in general position in $\mathbf{R}^2$ is $\delta(P)$-vertex connected, where $\delta(P) \geqslant \lceil \frac{n-4}{2} \rceil$ of the flip graph. Moreover, the flip graph is always at least $\lceil \frac{n-4}{2} \rceil$-vertex connected (without any assumption on n).*

A flip from a triangulation T to a triangulation T' that exchanges diagonal $e \in T$ and $e' \in T'$ naturally induces a bijection $T \to T'$ between the two sets of edges, by mapping e to e' and mapping every other edge to itself. Thus, every walk $T \rightsquigarrow T'$ in the flip graph starting at some triangulation T and ending at a triangulation T' induces a bijection $T \to T'$, by composing the bijections corresponding to individual flips. Which bijections are *realizable by flips* in this way? In joint work with Masárová and Lubiw[5] we answer this question. Affirming a conjecture of

Bose, Lubiw, Pathak, and Verdonschot [1], we show that the following necessary condition is also sufficient (which also yields a polynomial-time algorithm for testing whether a given bijection is realizable by flips). Given a bijection $\pi\colon T \to T'$ between the edge sets of two triangulations of a point set P, let us say that π is *locally realizable edge by edge* if, for every edge $e \in T$ and its image $e' = \pi(e) \in T'$, there is a walk $T \rightsquigarrow T'$ in the flip graph such that the induced bijection maps e to e' (but may differ from π on other edges).

Theorem 2 (Lubiw-Masárová-W.)**.** *Given two triangulations T and T' of a finite point set $P \subseteq \mathbf{R}^2$ and a bijection $\pi\colon T \to T'$, the bijection π is realizable by flips if and only if π is locally realizable edge by edge.*

The following fact plays a central role in proofs of both Theorems 1 and 2. In the special case where P is a set of n points in convex position, it is well-known that the flip graph of P is the 1-skeleton of an $(n-3)$-dimensional polytope, the *associahedron* $\mathcal{A}_{n-3}$ (see, e.g., the book [2] for more background and further references). For general point sets, the flip graph is no longer the graph of a convex polytope, but it can be seen as 1-skeleton of a higher-dimensional polytopal complex $\mathcal{F}(P)$, called the *flip complex*, which is contractible. This follows from results of Orden and Santos [6] (and, unaware of this, we rediscovered this and gave a different proof in [5]). Moreover, the faces of the flip complex have a very simple description: We say that a set S of pairwise non-crossing segments spanned by points in P is a *subdivision* of P if S contains all convex hull edges the geometric graph (P, S) is connected, and every bounded region of $\mathbf{R}^2 \backslash (\bigcup S)$ is convex. Every subdivision S naturally corresponds to a product of associahedra, one for each convex region determined by S, and these products of associahedra are the faces of the flip complex.

References

[1] Bose, P., Lubiw, A., Pathak, V., and Verdonschot, S. Flipping edge-labelled triangulations. *Comput. Geom. 68* (2018), 309–326.

[2] De Loera, J. A., Rambau, J., and Santos, F. *Triangulations: Structures for Algorithms and Applications*. Springer, 2010.

[3] Hurtado, F., Noy, M., and Urrutia, J. Flipping edges in triangulations. *Discrete Comput. Geom.* 22 (1999), no. 3, 333–346.

[4] Lawson, C. L. Transforming triangulations. *Discrete Math.* 3 (1972), no. 4, 365–372.

[5] Lubiw, A., Masárová, Z., and Wagner, U. A proof of the orbit conjecture for flipping edge-labelled triangulations. *Discrete Comput. Geom.* 61 (2019), no. 4, 880–898.

[6] Orden, D., and Santos, F. The polytope of non-crossing graphs on a planar point set. *Discrete Comput. Geom.* 33 (2005), no. 2, 275–305.

[7] Wagner, U. and Welzl, E. Connectivity of triangulation flip graphs in the plane. *Discrete Comput. Geom.* 68 (2022), no. 4, 1227–1284

Decomposing cubic graphs into linear forests

Shoham Letzter
(joint work with Gal Kronenberg, Alexey Pokrovskiy, and Liana Yepremyan)

1. Introduction

A well-known conjecture, due to Akiyama, Exoo, and Harary [2] (1980), dubbed the *Linear arboricity conjecture*, asserts that every graph G can be decomposed into at most $\frac{\Delta(G)+1}{2}$ linear forests.

This conjecture is open in general, but is known when $\Delta(G) \in \{2, 3, 4, 5, 6, 8, 10\}$, and for various families of graph G, such as complete and complete bipartite graphs, trees, planar graphs, and graphs with high girth. It is also known to hold, with high probability, for random graphs and random regular graphs. Finally, it is known asymptotically: Lang and Postle (2020) showed [9] that $\frac{\Delta}{2} + O(\Delta^{1/2}\text{polylog}\Delta)$ linear forests suffice for decomposing a graph with maximum degree Δ, improving previous asymptotic results by Alon [4] (1988) and Ferber, Fox and Jain [6] (2020).

While the conjecture appears to be hard in general, it actually has quite an easy proof when $\Delta(G) = 3$ (due to Akiyama and Chvátal [1] 1981). In particular, it shows that every cubic graph can be decomposed into two linear forests (in fact, this statement is equivalent to the linear arboricity conjecture for $\Delta(G) = 3$). It is thus interesting to try and impose more structure on the two linear forests.

One such strengthening is due to Thomassen [10] (1999), who showed that every cubic graph can be decomposed into two linear forests with components of length at most 5. This proved a conjecture of Bermond, Fouquet, Habib, and Péroche [5] (1984) and improved on results due to Jackson and Wormald [7] (1996) and Aldred and Wormald [3] (1998).

A different direction was considered by Wormald (1987), who conjectured [11] that every cubic graph on $4n$ vertices can be decomposed into two isomorphic linear forests. Prior to our work, the conjecture was known to hold only for some specific families of cubic graphs. We prove [8] Wormald's conjecture for large connected cubic grpahs.

Theorem 1. *Every connected cubic graph on $4n$ vertices, where n is large, can be decomposed into two isomorphic linear forests.*

2. Proof sketch

We now give an overview of our proof. The main idea is to first colour a small part of the graph in a very structured way, so that it can later be used to make small fixes to the full colouring, and then colour the rest of the graph in a random way, while guaranteeing that the monochromatic components are (not too long) paths. Using the randomness, we show that the two colour classes are almost isomorphic. We then use the pre-coloured graph to fix the imbalance and make the colour classes isomorphic, thus completing the proof.

2.1. **The approximate result.** While this is not the first step in the process, we now describe an approximate solution of Wormald's conjecture, and later explain how to obtain an appropriate partial pre-colouring. For the purpose of this explanation, our task is to red-blue colour a (large) given cubic graph such that the colour classes are "almost" isomorphic, that is, the difference between the number of red and blue components isomorphic to a path of length t is small, for all t. For this, we wish to colour the graph randomly, while maintaining certain properties (such as the monochromatic components being paths).

Our random colouring will consist of three steps. For the first step, we use Thomassen's result about the existence of a 2-colouring where each monochromatic component is a path of length at most 5; we denote the two colours here by purple and green. The first random step colours each purple or green component by one of the two possible alternating red-blue colourings, chosen uniformly at random and independently. Notice that this random red-blue colouring of G has no vertices incident with three edges of the same colour. Moreover, the symmetry between the colours and the bound on the lengths of purple and green paths would allow us to show that the colours are, in some sense, close to being isomorphic. However, there is nothing preventing the appearance of cycles, and we could not rule out the existence of very long monochromatic paths, a complication for concentration and for the final rebalancing.

This brings us to the second random step, which will be broken into two parts, and whose purpose is to eliminate monochromatic cycles. Here, we first do something very intuitive: we simply flip the colour of one edge e_C of each monochromatic cycle C, choosing the edge uniformly at random and independently. Unsurprisingly, while this breaks all monochromatic cycles that existed before the first step, new monochromatic cycles can appear. Luckily, a small fix saves us and eliminates all monochromatic cycles. The fix essentially consists of reswapping the colour of e_C for some originally monochromatic cycles C, and swapping the colour of a neighbouring edge of e_C in C, while making choices randomly and independently.

The next and final random step is designed to break long paths. Here the idea is less intuitive. We let each monochromatic path P choose one of its boundary edges (namely, edges of the opposite colour that touch an end of P) uniformly at random and independently. Then, for each edge e that was chosen by two paths, we flip the colour of e. This somewhat strange process has several benefits: first, with high probability, it swaps an edge of each monochromatic path of length at least $1000 \log n$; second, it creates no monochromatic cycles; and, third, it does not allow more than two monochromatic paths to join up (more precisely, monochromatic paths in the new colouring have at most one edge whose colour was swapped).

Finally, we analyse the resulting random colouring, and show that its colour classes are almost isomorphic. We accomplish this goal via McDiarmid's inequality, using the independence of the various random decisions, as well as the fact that each decision has a small impact on the resulting graph.

2.2. **Gadgets.** Next, we wish to balance the number of red and blue components isomorphic to P_t, for every t, so as make the two colour classes isomorphic. For this, the main idea is to pre-colour a small part of the graph, thereby creating many *gadgets* that can later be used for balancing the number of paths in each length.

We define a blue *ℓ-gadget* in a cubic graph G to be a red-blue subgraph $H \subseteq G$, which has a recolouring H', so that for every red-blue colouring of G that extends H (and whose monochromatic components are paths), if we recolour H according to H', the monochromatic component counts change as follows (where $b(P_t)$ and $r(P_t)$ are the numbers of, respectively, blue and red components that are paths of length t): $b(P_\ell)$ decreases by exactly 1; $b(P_t)$ with $t < \ell$ changes only slightly; $b(P_t)$ with $t > \ell$ and $r(P_t)$ with any t do not change. Such a gadget (and its counterpart with roles of colours reversed) will be used to equalise $r(P_t)$ and $b(P_t)$, and we balance the paths from the longest to the shortest.

Finding gadgets is the most technical and lengthy part of our proof, which we do not elaborate on here. Very roughly speaking, we start from a long enough geodesic (that is why our proof only works for large connected cubic graphs), and find a gadget within a small radius of the geodesic.

2.3. **Exact result.** Our work on gadgets shows that in large enough connected cubic graphs we can find many gadgets of length up to $\varepsilon \log n$ (for some small $\varepsilon > 0$). We incorporate those in the random colouring described above, to obtain a red-blue colouring where the two components are almost isomorphic linear forests with components of length at most $10^4 \log n$, and which contains many gadgets of each length up to $m := 10^4\sqrt{\log n}$. We use a different argument to balance the number of gadgets of length longer than m, and then use the gadgets to balance the remaining lengths.

References

[1] J. Akiyama and V. Chvátal, *A short proof of the linear arboricity for cubic graphs*, Bull. Liber. Arts Sci. NMS **2** (1981), 1–3.

[2] J. Akiyama, G. Exoo, and F. Harary, *Covering and packing in graphs. III: Cyclic and acyclic invariants*, Mathematica Slovaca **30** (1980), no. 4, 405–417.

[3] R. E. L. Aldred and N. C. Wormald, *More on the linear k-arboricity of regular graphs*, Australas. J. Combin. **18** (1998), 97–104.

[4] N. Alon, *The linear arboricity of graphs*, Israel J. Math. **62** (1988), no. 3, 311–325.

[5] J.-C. Bermond, J.-L. Fouquet, M. Habib, and B. Peroche, *On linear k-arboricity*, Discr. Math. **52** (1984), no. 2-3, 123–132.

[6] A. Ferber, J. Fox, and V. Jain, *Towards the linear arboricity conjecture*, J. Combin. Theory Ser. B **142** (2020), 56–79.

[7] B. Jackson and N. C. Wormald, *On the linear k-arboricity of cubic graphs*, Discr. Math. **162** (1996), no. 1-3, 293–297.

[8] G. Kronenberg, S. Letzter, A. Pokrovskiy, and L. Yepremyan, *Decomposing cubic graphs into isomorphic linear forests*, arXiv:2210.11458 (2022).

[9] R. Lang and L. Postle, *An improved bound for the linear arboricity conjecture*, arXiv:2008.04251 (2020).

[10] C. Thomassen, *Two-coloring the edges of a cubic graph such that each monochromatic component is a path of length at most 5*, J. Combin. Theory Ser. B **75** (1999), no. 1, 100–109.
[11] N. Wormald, *Problem 13*, Ars Combinatoria **23** (1987), 332–334.

Hamilton cycles in pseudorandom graphs

David Munhá Correia
(joint work with Stefan Glock and Benny Sudakov)

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. 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 [11], 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 [8] 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 [20] proved that the Erdős-Renyi random graph model $G(n,p)$ with $p \geqslant C \log n/n$ 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. 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 $3 \leqslant d \leqslant n-1$ (for this result, the reader is referred to the papers of Cooper, Frieze and Reed [9] and Krivelevich, Sudakov, Vu and Wormald [14] 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 whose second largest eigenvalue of G in absolute value, $\lambda(G)$, is such that $\lambda(G) \leqslant \lambda$. 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 [16] for detailed survey concerning pseudorandom graphs.

In a rather influential paper, Krivelevich and Sudakov [15] 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

$$d/\lambda \geqslant \frac{1000 \log n (\log\log\log n)}{(\log\log n)^2} \tag{1}$$

guarantees that any (n, d, λ)-graph contains a Hamilton cycle. This result 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 multiplicative ratio of d/λ is already sufficient to imply Hamiltonicity. Krivelevich and Sudakov [15] conjectured that it should suffice that d/λ is only a large enough constant.

Conjecture 1. *There exists an absolute constant $K > 0$ such that any (n, d, λ)-graph with $d/\lambda \geqslant K$ contains a Hamilton cycle.*

Despite the plethora of incentives, there has been no improvement until now on the Krivelevich and Sudakov bound. We make significant progress towards Conjecture 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. *There exists a constant $C > 0$ such that any (n, d, λ)-graph with $d/\lambda \geqslant C(\log n)^{1/3}$ contains a Hamilton cycle.*

The proof of the above result relies on Pósa's rotation-extension method with various new twists. Namely, we needed to develop some techniques in order to use this method in a *robust* manner.

Secondly, we confirm Conjecture 1 in full when the degree is polynomial in the order of the graph. More generally, we prove the following result.

Theorem 3. *There exists a constant $C > 0$ such that any (n, d, λ)-graph with $d \geqslant (\log n)^5$ and $d/\lambda \geqslant C \log_d n$ contains a Hamilton cycle.*

In particular, this implies Conjecture 1 when d is polynomial in n.

Corollary 4. *For every $\alpha > 0$ there exists $K > 0$ such that any (n, d, λ)-graph with $d \geqslant n^\alpha$ and $d/\lambda \geqslant K$ contains a Hamilton cycle.*

In fact, Theorem 3 is a corollary of a more general statement that we prove which informally states that (n, d, λ)-graphs with many vertex-disjoint cycles are Hamiltonian.

2. Applications and related problems

Both Theorem 2 and Corollary 4 immediately yield improvements in several applications which made use of the Krivelevich and Sudakov result. One application is an important special case of a famous open question of Lovász [17] from 1969 concerning the Hamiltonicity of a certain class of well-behaved graphs (see e.g., [10] and its references for more background on the problem).

Conjecture 5. *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 [21].

Conjecture 6. *Every connected Cayley graph is Hamiltonian.*

For these conjectures, a proof is currently out of sight. Given this, it is natural to consider the "random" version of Conjecture 6. Indeed, Alon and Roichman [2] 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 6 is to show that $\Gamma(G, S)$ is almost surely Hamiltonian, which itself is also a conjecture of Pak and Radoičić [19]. In fact, this relates directly to Conjecture 1 since it can be shown, generalizing the result of Alon and Roichman, that if $|S| \geqslant C \log |G|$ for some large constant C, then $\Gamma(G, S)$ is almost surely an (n, d, λ)-graph with $d/\lambda \geqslant K$ for some large constant K. Hence, Conjecture 1 would imply the Hamiltonicity of $\Gamma(G, S)$. Improving on several earlier results [6, 15, 18], we use Theorem 2 to prove that if $|S| \geqslant \log^{5/3} n$, then $\Gamma(G, S)$ is almost surely Hamiltonian.

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 [1] studied the emergence of more complex additive structures in such subgroups. In particular, they proved that when A as above has size $|A| \geqslant q^{3/4}(\log q)^{1/2-o(1)}$, 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 Corollary 4, 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 do there exist $x, y \in A$ such that $x + y \in A$ but also much more complex structures.

Finally, we also use the methods in the proof of Theorem 3 for a problem related to Conjecture 6 concerning the existence of Hamiltonian Cayley graphs $\Gamma(G, S)$ with a small set of generators S. Motivated by this conjecture, Pak and Radoičić [19] 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 their result. Using the methods we developed for the proof of Theorem 3 we give a classification-free proof that there is always such a set S with $|S| = O(\log n)$.

References

[1] N. Alon and J. Bourgain. Additive patterns in multiplicative subgroups. *Geometric and Functional Analysis*, 24(3):721–739, 2014.

[2] N. Alon and Y. Roichman. Random Cayley graphs and expander. *Random Structures Algorithms*, 5(2):271–284, 1994.

[3] S. Akbari, O. Etesami, H. Mahini and M. Mahmoody. On rainbow cycles in edge colored complete graphs. *Australasian Journal of Combinatorics*, 37:33–42, 2007.

[4] L. Babai. Long cycles in vertex-transitive graphs. *Journal of Graph Theory*, 3(3):301–304, 1979.

[5] B. Bollobás. The evolution of sparse graphs. *Graph Theory and Combinatorics*, Academic Press, London 35–57, 1984.

[6] D. Christofides and K. Markstrom. Random Latin square graphs. *Random Structures Algorithms*, 41(1):47–65, 2012.

[7] D. Christofides, J. Hladký and A. Máthé. Hamilton cycles in dense vertex-transitive graphs. *Journal of Combinatorial Theory, Series B*, 109:34–72, 2014.

[8] V. Chvátal and P. Erdős. A note on Hamiltonian circuits. *Discrete Mathematics*, 2(2):111–113, 1972.

[9] C. Cooper, A. Frieze and B. Reed. Random regular graphs of non-constant degree: connectivity and Hamiltonicity. *Combinatorics, Probability and Computing*, 11(3):249–261, 2002.

[10] S. Curran and J. Gallian. Hamiltonian cycles and paths in Cayley graphs and digraphs - a survey. *Discrete Mathematics*, 156(1-3):1–18, 1996.

[11] G. A. Dirac. Some theorems on abstract graphs. *Proceedings of the London Mathematical Society*, 3(1):69–81, 1952.

[12] J. Komlós and E. Szemerédi. Limit distribution for the existence of Hamiltonian cycles in a random graph. *Discrete Mathematics*, 43(1):55–63, 1983.

[13] A. Korshunov. Solution of a problem of Erdős and Rényi on Hamilton cycles in non-oriented graphs. *Soviet Math. Dokl*, 17:760–764, 1976.

[14] M. Krivelevich, B. Sudakov, V. Vu and N. Wormald. Random regular graphs of high degree. *Random Structures Algorithms*, 18(4):346–363, 2001.

[15] M. Krivelevich and B. Sudakov. Sparse pseudo-random graphs are Hamiltonian. *Journal of Graph Theory*, 42(1):17–33, 2003.

[16] M. Krivelevich and B. Sudakov. Pseudo-random graphs. *More sets, graphs and numbers*, Springer, 199–262, 2006.

[17] L.Lovász. Combinatorial structures and their applications. *Proc. Calgary Internat. Conf., Calgary, Alberta*, 1970:243–246, 1969.

[18] J. Meng and H. Qiongxiang. Almost all Cayley graphs are Hamiltonian. *Acta Mathematica Sinica*, 12(2):151–155, 1996.

[19] I. Pak and ,R. Radoičić. Hamiltonian paths in Cayley graphs. *Discrete Mathematics*, 309(17):5501–5508, 2009.

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

[21] E. Rapaport-Strasser. Cayley color groups and Hamilton lines. *Scripta Mathematica*, 24:51–58, 1959.

The Exact Rank of Sparse Random Graphs

Matthew Kwan
(joint work with Margalit Glasgow, Ashwin Sah, and Mehtaab Sawhney)

A foundational theorem in combinatorial random matrix theory, due to Komlós [16, 17], says that discrete random matrices with i.i.d. entries are typically nonsingular (over $\mathbb{R}$). In particular, let B be an $n \times n$ random matrix with i.i.d. Bernoulli(p) entries (meaning that each entry B_{ij} satisfies $\Pr[B_{ij} = 1] = p$ and $\Pr[B_{ij} = 0] = 1 - p$). For constant $p \in (0, 1)$ and $n \to \infty$, such a random matrix is nonsingular *with high probability* ("whp" for short): that is, $\lim_{n\to\infty} \Pr[B \text{ is singular}] = 0$.

A huge number of strengthenings and variations of Komlós' theorem have been considered over the years. Two particular highlights include a result of Tikhomirov [19] that for constant $0 < p \leqslant 1/2$, the singularity probability is $(1 - p + o(1))^n$, and a result of Costello, Tao, and Vu [9] that *symmetric* discrete random matrices are also nonsingular whp. A symmetric binary matrix can be interpreted as the adjacency matrix of a graph, so the Costello–Tao–Vu theorem has an interpretation in terms of random graphs: for constant $p \in (0, 1)$, an Erdős–Rényi random graph $G \sim \mathbb{G}(n, p)$ has nonsingular adjacency matrix whp[1]. Actually, Komlós' theorem can be interpreted in graph-theoretic terms as well: the random matrix B described above can be interpreted as the *biadjacency* matrix of a *bipartite* Erdős–Rényi random graph $G \sim \mathbb{G}(n, n, p)$ (where one of the parts corresponds to the rows of the matrix, and the other part corresponds to the columns).

If p decays too rapidly with n (in particular, if $p \leqslant (1 - \varepsilon) \log n/n$ for some constant $\varepsilon > 0$), then for reasons related to the coupon collector problem, a typical outcome of $G \sim \mathbb{G}(n, p)$ (respectively, $G \sim \mathbb{G}(n, n, p)$) has isolated vertices, meaning that its adjacency matrix (respectively, biadjacency matrix) has all-zero rows and is therefore singular. In fact, $\log n/n$ is a *sharp threshold* for singularity, in the sense that if $p \geqslant (1 + \varepsilon) \log n/n$ (and p is bounded away from 1) then a typical $G \sim \mathbb{G}(n, p)$ (respectively, $G \sim \mathbb{G}(n, n, p)$) has nonsingular adjacency matrix (respectively, nonsingular biadjacency matrix). This seems to have been first case observed by Costello and Vu [11][2], and refinements and generalisations were proved by Basak and Rudelson [3] and Addario-Berry and Eslava [1]. In particular, the latter authors proved a sharp *hitting time* type result: if we consider the random graph *process* where we start with the empty graph on n vertices (or the empty bipartite graph with $n + n$ vertices) and add random edges one-by-one (respecting our bipartition, in the bipartite case), then whp at the very same moment where the last isolated vertex disappears our graph becomes nonsingular.

[1]There is a slight difference between a random symmetric Bernoulli matrix and the adjacency matrix of a random graph: namely, the adjacency matrix of any graph has zeroes on the diagonal. However, the same techniques usually apply to both settings, and we will not further concern ourselves with this detail.

[2]The Costello–Vu proof was only written for $\mathbb{G}(n, p)$, but it can be easily adapted to $\mathbb{G}(n, n, p)$; alternatively, see [14] for a very simple proof in the $\mathbb{G}(n, n, p)$ case.

Below the threshold $\log n/n$, it is natural to ask whether the only obstacles for singularity are "local dependencies" such as isolated vertices. In their aforementioned paper, Costello and Vu [11] actually proved that for $p \geqslant (1/2+\varepsilon)\log n/n$, whp the subgraph obtained from $\mathbb{G}(n,p)$ by deleting isolated vertices is non-singular. In follow-up work Costello and Vu [10] considered the regime where $p \geqslant c\log n/n$ for any constant $c > 0$; this necessitated the consideration of more sophisticated types of "local dependencies" than isolated vertices. The most obvious example is *cherries*: pairs of degree-1 vertices with the same neighbour. More recently, DeMichele, the first author, and Moreira [13] gave a combinatorial description of the rank of $G \sim \mathbb{G}(n,p)$ and $G \sim \mathbb{G}(n,n,p)$, in terms of a procedure that iteratively deletes local dependencies, which holds whp whenever $\lim_{n\to\infty} np = \infty$ (i.e., when p asymptotically dominates $1/n$).

The most challenging regime is where $p = c/n$ for constant c. An asymptotic for the typical rank of $\mathbb{G}(n,c/n)$ was conjectured by Bauer and Golinelli [5] (motivated by statistical physics considerations), and was later proved by Bordenave, Lelarge, and Salez [6]. In his lecture at the 2014 International Congress of Mathematicians [21] (also in [20]), Vu asked whether one can also give a precise combinatorial characterisation of the rank in this regime.

The main purpose of this work is to provide an answer to Vu's question, and the analogous question for $\mathbb{G}(n,n,c/n)$, exactly characterising the rank of sparse random graphs (and in the process, providing a linear-time algorithm to compute the rank).

To state our main theorem, we need to introduce the Karp–Sipser leaf removal algorithm, which was introduced in 1981 by Karp and Sipser [15] as a tool to study matchings in random graphs (in a paper which kickstarted the *differential equations method* for random graph processes; see [22]), but is now also of great importance in statistical physics, theoretical computer science, and random matrix theory (see for example [5, 4, 18, 6, 7]).

Definition 1 (Karp–Sipser leaf removal)**.** *Starting from a graph G, choose an arbitrary degree-1 vertex and delete it together with its neighbour. Repeat this "leaf-deletion" until no further degree-1 vertices remain. Let $i(G)$ be the number of isolated vertices in the resulting graph. If G is bipartite, let $i_1(G)$ and $i_2(G)$ be the number of isolated vertices on the two sides of the given bipartition V_1, V_2. Let* $\mathrm{core}_{\mathrm{KS}}(G)$ *be the graph of remaining non-isolated vertices (the* Karp–Sipser core*). One can check that $i(G)$ and $\mathrm{core}_{\mathrm{KS}}(G)$ (and $i_1(G), i_2(G)$, if G is bipartite) do not depend on the order that the leaf-deletions are performed (see for example the appendix of* [4]*).*

It is easy to check that a single step of leaf-removal decreases $\operatorname{rank} A(G)$ by exactly 2, and if G is bipartite, decreases $\operatorname{rank} B(G)$ by exactly 1. It is then easy to deduce that $\operatorname{rank} A(G) \leqslant n - i(G)$ if G has n vertices (i.e., $\operatorname{corank} A(G) \geqslant i(G)$), and if G is bipartite with n vertices on each side then $\operatorname{rank} B(G) \leqslant n - \max(i_1(G), i_2(G))$ (i.e., $\operatorname{corank} B(G) \geqslant \max(i_1(G), i_2(G))$). We will refer to these two bounds as the *Karp–Sipser bounds* for the rank of $A(G)$ and $B(G)$, respectively. We remark that there is a one-sided version of the Karp–Sipser bound for $B(G)$ (where leaves are

only removed from one of the two sides of our bipartite graph), sometimes called the *2-core bound* in the computer science and statistical physics literature [8, 2, 12].

The Karp–Sipser process takes care of "tree-like" local dependencies. In random graphs $\mathbb{G}(n,p)$ or $\mathbb{G}(n,n,p)$ with $np \to \infty$, these are whp the only types of dependencies that exist (see [13, 10]); that is, the Karp–Sipser core is nonsingular, so the Karp–Sipser bound is sharp.

However, in the case $p = O(1/n)$, there may be "cycle-like" local dependencies in the Karp-Sipser core, such as pairs of degree-2 vertices with the same neighbourhood. We capture dependencies of this type in the following definition.

Definition 2 (Special cycles)**.** *Say an induced cycle in a graph G is* special *if its length is divisible by 4, and if every second vertex has degree 2 in G. In particular, an* isolated cycle *is a cycle in which every vertex has degree exactly 2 (i.e., it is its own connected component), so isolated cycles with length divisible by 4 are special "in two different ways". Let $s(G)$ be the number of special cycles in G, where we count each isolated cycle twice.*

If G is bipartite, say an induced cycle in G is 1-special (respectively, 2-special) if its length is divisible by 4, and every vertex in V_1 (respectively, every vertex in V_2) has degree 2. Let $s_1(G)$ and $s_2(G)$ be the numbers of 1-special and 2-special cycles in G, respectively.

Our main theorem says that for $c \neq e$, the rank of a sparse random graph $\mathbb{G}(n,c/n)$ or $\mathbb{G}(n,n,c/n)$ can be described in terms of the Karp–Sipser bound and the special cycles within the Karp–Sipser core.

Theorem 3. *Fix a constant $c \neq e$.*

(A) *Let $G \sim \mathbb{G}(n,c/n)$. Then whp* $\operatorname{corank} A(G) = i(G) + s(\operatorname{core}_{\mathrm{KS}}(G))$.
(B) *Let $G \sim \mathbb{G}(n,n,c/n)$. Then whp*

$$\operatorname{corank} B(G) = \max\big(i_1(G) + s_1(\operatorname{core}_{\mathrm{KS}}(G)),\ i_2(G) + s_2(\operatorname{core}_{\mathrm{KS}}(G))\big).$$

Our proof of Theorem 3 involves a wide range of tools and ideas, both original and adapted from existing work. This includes analysis of degree-constrained random graphs and of the *Karp–Sipser leaf-removal algorithm*, robust analysis of random walks, spectral convergence machinery for locally convergent graphs, a "rank-boosting" technique, and some special-purpose notions of matrix pseudoinverses and "minimal kernel vectors". To try to give a rough impression of the most fundamental difficulty compared to previous work: note that the rank of a matrix can be interpreted as the size of its largest nonsingular submatrix. In the setting of most previous work, maximum nonsingular submatrices are in some sense "robustly" nonsingular (in particular, the corresponding subgraphs have good *expansion* properties, which turn out to play a crucial role) and one can use fairly lossy estimates. However, in our situation the largest nonsingular submatrices are in some sense "only barely nonsingular", with essentially the weakest possible expansion a nonsingular submatrix can have, and there is almost no room to make any kind of lossy approximation.

Once one has a characterisation of the rank in terms of explicit combinatorial structures, it becomes possible to prove further results about the rank via combinatorial tools. Indeed, as corollaries of our main theorem and its proof, we are able to show a number of additional theorems: we compute the asymptotic singularity probability of the 2-core, we obtain a very strong bound on the difference between the matching number and the rank, and we prove a central limit theorem for the rank of $\mathbb{G}(n, c/n)$.

References

[1] Louigi Addario-Berry and Laura Eslava, *Hitting time theorems for random matrices*, Combin. Probab. Comput. **23** (2014), 635–669.

[2] R. C. Alamino and D. Saad, *Typical kernel size and number of sparse random matrices over Galois fields: a statistical physics approach*, Phys. Rev. E (3) **77** (2008), 061123, 12.

[3] Anirban Basak and Mark Rudelson, *Sharp transition of the invertibility of the adjacency matrices of sparse random graphs*, arXiv:1809.08454.

[4] M. Bauer and O. Golinelli, *Core percolation in random graphs: a critical phenomena analysis*, The European Physical Journal B **24** (2001), 339–352.

[5] M. Bauer and O. Golinelli, *Exactly solvable model with two conductor-insulator transitions driven by impurities*, Physical Review Letters **86** (2001), 2621–2624.

[6] Charles Bordenave, Marc Lelarge, and Justin Salez, *The rank of diluted random graphs*, Ann. Probab. **39** (2011), 1097–1121.

[7] Amin Coja-Oghlan, Oliver Cooley, Mihyun Kang, Joon Lee, and Jean Bernoulli Ravelomanana, *The sparse parity matrix*, arXiv:2107.06123.

[8] Amin Coja-Oghlan, Alperen A. Ergür, Pu Gao, Samuel Hetterich, and Maurice Rolvien, *The rank of sparse random matrices*, Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, PA, 2020, pp. 579–591.

[9] Kevin P. Costello, Terence Tao, and Van Vu, *Random symmetric matrices are almost surely nonsingular*, Duke Math. J. **135** (2006), 395–413.

[10] Kevin P. Costello and Van Vu, *On the rank of random sparse matrices*, Combin. Probab. Comput. **19** (2010), 321–342.

[11] Kevin P. Costello and Van H. Vu, *The rank of random graphs*, Random Structures Algorithms **33** (2008), 269–285.

[12] Amir Dembo and Andrea Montanari, *Finite size scaling for the core of large random hypergraphs*, Ann. Appl. Probab. **18** (2008), 1993–2040.

[13] Patrick DeMichele, Margalit Glasgow, and Alexander Moreira, *On the rank, kernel, and core of sparse random graphs*, arXiv:2105.11718.

[14] Asaf Ferber, Matthew Kwan, and Lisa Sauermann, *Singularity of sparse random matrices: simple proofs*, arXiv:2011.01291.

[15] R. M. Karp and M. Sipser, *Maximum matching in sparse random graphs*, 22nd Annual Symposium on Foundations of Computer Science (sfcs 1981), IEEE, October 1981.

[16] J. Komlós, *On the determinant of* (0, 1) *matrices*, Studia Sci. Math. Hungar. **2** (1967), 7–21.

[17] J. Komlós, *On the determinant of random matrices*, Studia Sci. Math. Hungar. **3** (1968), 387–399.

[18] M. Mézard, F. Ricci-Tersenghi, and R. Zecchina, *Two solutions to diluted p-spin models and XORSAT problems*, J. Statist. Phys. **111** (2003), 505–533.

[19] Konstantin Tikhomirov, *Singularity of random Bernoulli matrices*, Ann. of Math. (2) **191** (2020), 593–634.

[20] Van Vu, *Some recent results on random matrices*, lecture at a workshop on Probabilistic Techniques and Applications, hosted by the Institute for Pure & Applied Mathematics (IPAM), `http://www.ipam.ucla.edu/abstract/?tid=8303&pcode=CMAWS1`, 2009.

[21] Van Vu, *Combinatorial problems in random matrix theory*, invited lecture at the International Congress of Mathmaticians (ICM) in Seoul, `https://www.youtube.com/watch?v=9REUYqf8EWA`, 2014.

[22] N.C. Wormald, *The differential equation method for random graph processes and greedy algorithms*, Lectures on Approximation and Randomized Algorithms, PWN, Warsaw, 1999, pp. 73–155.

On vertex Ramsey graphs with forbidden subgraphs

Michael Krivelevich

(joint work with Sahar Diskin, Ilay Hoshen, and Maksim Zhukovskii)

A graph G is said to be r-vertex Ramsey for a graph A is every r-coloring of $V(G)$ has a monochromatic copy of A. Which graphs can be vertex-Ramsey for a given graph A? A classical result due to Nešetřil and Rödl [1] states that given a finite family of graphs $\mathcal{F}$, a graph A and a positive integer r, if every graph $B \in \mathcal{F}$ has a 2-vertex-connected subgraph which is not a subgraph of A, then there exists an $\mathcal{F}$-free graph which is vertex r-Ramsey with respect to A.

Our main result shows that the above sufficient condition is also necessary. We say that B is an A-forest of size ℓ if $B = \cup_{i=1}^{\ell} B_i$, where for every $1 \leqslant i \leqslant \ell$, B_i is isomorphic to a subgraph of A, and for every $i \geqslant 2$, $\left|V(B_i) \cap V\left(\cup_{j=1}^{i-1} B_j\right)\right| \leqslant 1$. Clearly, B is an A-forest if and only if every 2-connected subgraph B_0 of B is a subgraph of A. We then have:

Theorem 1. *Let $\ell > 0$ be an integer. Let B be an A-forest of size ℓ. Let $r > 0$ be an integer such that $r \geqslant \ell\,(2(|V(A)|-1)(|V(B)|-2)+1)$, and let G be an r-vertex Ramsey graph with respect to A. Then G contains a copy of B.*

References

[1] J. Nešetřil, V. Rödl, *Partitions of vertices*, Commentationes Mathematicae Universitatis Carolinae, **17** (1976) 85–95.

Towards the Erdős–Hajnal Conjecture

Matija Bucić

(joint work with Pablo Blanco and Tung Nguyen, Alex Scott and Paul Seymour)

A graph G *contains* a graph H if H is isomorphic to an induced subgraph of G, and G is *H-free* otherwise. Erdős and Hajnal [6, 7] proposed in 1977 the following well-known conjecture, stating that unlike in arbitrary graphs where we can only guarantee a clique or a stable set of polylogarithmic size if we impose even a little bit of structure on the graph, by say forbidding some fixed graph H as an induced subgraph, we are suddenly guaranteed a polynomial-size clique or an independent set.

Conjecture 1. *For every graph H there exists $c > 0$ such that every H-free graph G contains a clique or an independent set of order at least $|G|^c$.*

The conjecture attracted a great deal of attention over the years with two large surveys [9, 2] focusing exclusively on it. The one due to Gyárfás from the late '90s [9] raised the question of trying to prove the conjecture for specific small graphs. There seem to be two major motivating factors behind this. The first one is that it is easier to think about the conjecture with a fixed small forbidden structure in mind and we may hope the ideas developed might apply in much more generality. The second is a classical result due to Alon, Pach and Solymosi [1] that the class of graphs satisfying the Erdős–Hajnal Conjecture is closed under the substitution operation, in particular, this implies that proving the conjecture for any new fixed graph immediately implies it for an infinite family of graphs. In particular, the last remaining case explicitly raised by Gyárfás is the five-vertex path, which is also the smallest open case of the Erdős–Hajnal Conjecture.

In terms of general results, the best result until now which applies to *all* graphs H dates back to the original paper of Erdős and Hajnal where they proved the following:

Theorem 2. *For every graph H there exists $c > 0$ such that every H-free graph G contains a clique or an independent set of order at least $2^{c\sqrt{\log |G|}}$.*

We prove the first numerical improvement of this result, which applies for *all* graphs H, since it was announced, over 45 years ago:

Theorem 3 (B., Nguyen, Scott, Seymour)**.** *For every graph H there exists $c > 0$ such that every H-free graph G contains a clique or an independent set of order at least $2^{c\sqrt{\log |G| \log\log |G|}}$.*

While it might seem like a modest improvement our new improved bound is a very natural "intermediate" point between the theorem and conjecture of Erdős and Hajnal. Indeed it was raised by Conlon, Fox and Sudakov as an intermediate goal in [5]. In addition, preceding the very recent resolution of Conjecture 1 in the case of $H = C_5$ by Chudnovsky, Scott, Seymour and Spirkl [4] the same group of authors together with Fox previously proved Theorem 3 in this special case in [3].

In a certain sense the main reason behind the bound in Theorem 3 appearing as a natural milestone when one tries to prove Conjecture 1 is rooted in the original approach of Erdős and Hajnal yielding Theorem 2. In fact, it arises as a natural "barrier" for this approach. Our second result manages to break this barrier in the smallest open case of the conjecture, namely when $H = P_5$, the five vertex path.

Theorem 4 (Blanco, B.)**.** *There exists $c > 0$ such that every P_5-free graph G contains a clique or an independent set of order at least $2^{c(\log |G|)^{2/3}}$.*

We also manage to extend the substitution result of Alon, Pach and Solymosi to apply for the class of graphs which satisfy the weaker bound above, in particular implying that we obtain the same improvement for an infinite family of graphs.

Over the years there have been a number of results closely related to the Erdős–Hajnal Conjecture (namely Conjecture 1). Our proof methods allow us to improve the bounds in a number of these results.

The first one of these is an often very useful theorem of Rödl [11] which states that:

Theorem 5. *For every graph H and all $\varepsilon > 0$, there exists $\delta > 0$ with the following property. For every H-free graph G, there exists $S \subseteq V(G)$ with $|S| \geqslant \delta|G|$ such that one of $G[S], \overline{G}[S]$ has at most $\varepsilon\binom{|S|}{2}$ edges.*

The second one is a "supersaturation" type strengthening of Theorem 5 due to Nikiforov [10]:

Theorem 6. *For every graph H and all $\varepsilon > 0$, there exists $\delta > 0$ such that if G is a graph containing less than $(\delta|G|)^{|H|}$ induced copies of H, then there exists $S \subseteq V(G)$ with $|S| \geqslant \delta|G|$ such that one of $G[S], \overline{G}[S]$ has at most $\varepsilon\binom{|S|}{2}$ edges.*

A very natural question which arises given this result and actually directly connects it to the Erdős–Hajnal Conjecture, is how large can we take δ as a function of ε? Rödl's original proof gave a tower-type bound because it used the regularity lemma, but Fox and Sudakov [8] made a significant improvement, showing a common generalisation of both Theorem 2 and Theorem 6:

Theorem 7. *There exists $c > 0$ such that for every graph H and all $\varepsilon \in (0, 1/2)$, setting $\delta = 2^{-c|H|\left(\log \frac{1}{\varepsilon}\right)^2}$ satisfies Theorem 6.*

We obtain a strengthening of Theorem 7, which is in turn, an improvement over Theorem 6 and Theorem 5. In addition, this result has our new general bound on the Erdős–Hajnal Conjecture, namely Theorem 3, as an immediate corollary.

Theorem 8 (B., Nguyen, Scott, Seymour)**.** *For every graph H there exists c such that, if $\varepsilon \in (0, 1/2)$ and*

$$\delta = 2^{-c\left(\log^2 \frac{1}{\varepsilon}\right)^2 / \log\log \frac{1}{\varepsilon}},$$

and G is a graph containing less than $(\delta|G|)^{|H|}$ induced copies of H, then there exists $S \subseteq V(G)$ with $|S| \geqslant \delta|G|$ such that one of $G[S], \overline{G}[S]$ has at most $\varepsilon\binom{|S|}{2}$ edges.

References

[1] N. Alon, J. Pach and J. Solymosi, "Ramsey-type theorems with forbidden subgraphs", *Combinatorica* **21** (2001), 155–170.

[2] M. Chudnovsky, "The Erdős–Hajnal conjecture—a survey", *Journal of Graph Theory* **75**, no. 2, (2014), 178–190.

[3] M. Chudnovsky, J. Fox, A. Scott, P. Seymour and S. Spirkl, "Towards Erdős–Hajnal for graphs with no 5-hole", *Combinatorica* **39** (2019), 983–991.

[4] M. Chudnovsky, A. Scott, P. Seymour, and S. Spirkl, *Erdős-Hajnal for graphs with no 5-hole*, Proceedings of the London Mathematical Society, to appear.

[5] D. Conlon, J. Fox and B. Sudakov, "Recent developments in graph Ramsey theory", in: *Surveys in Combinatorics 2015*, Cambridge University Press, 2015, 49–118.

[6] P. Erdős and A. Hajnal, "On spanned subgraphs of graphs", *Graphentheorie und Ihre Anwendungen* (Oberhof, 1977).

[7] P. Erdős and A. Hajnal, "Ramsey-type theorems", *Discrete Applied Mathematics* **25** (1989), 37–52.

[8] J. Fox and B. Sudakov, "Induced Ramsey-type theorems", *Advances in Mathematics* **219** (2008), 1771–1800.
[9] A. Gyárfás, "Reflections on a problem of Erdős and Hajnal", in: The Mathematics of Paul Erdős, (R. L. Graham and J. Nešetřil, eds.), *Algorithms and Combinatorics* **14**, Vol. II, Springer-Verlag, Heidelberg, (1997), 93–98.
[10] V. Nikiforov, "Edge distribution of graphs with few copies of a given graph", *Combin. Probab. Comput.* **15** (2006), 895–902.
[11] V. Rödl, "On universality of graphs with uniformly distributed edges", *Discrete Math.* **59** (1986), 125–134.

Hyperplane covers of finite spaces

István Tomon
(joint work with János Nagy, Péter Pál Pach)

Let p be a prime. What is the minimum number of hyperplanes in a covering $\mathcal{H}$ of the finite space $\mathbb{F}_p^n$? Clearly, without further assumptions on $\mathcal{H}$, the answer is p, as one can take the p translates of any hyperplane H, and this is minimal. Therefore, one would like to impose certain restrictions to avoid this triviality. Problems of this sort are extensively studied, one of the classical results in the area is due to Alon and Füredi [1]. They show that removing a single point of $\mathbb{F}_p^n$ changes the answer drastically, the minimal number of hyperplanes in a covering becomes exactly $(p-1)n$ (to be more precise, in [1] they considered coverings of $[k]^n\backslash\{(1,\dots,1)\}$ over $\mathbb{R}$, however, their arguments extend easily to the aforementioned finite field variant). Here, we consider a different variation of this problem.

A covering of a set with a collection of its subsets is *irredundant*, if no proper subcollection forms a covering. Given a hyperplane H in $\mathbb{F}_p^n$, we denote by $H^\perp$ the normal vector of H. That is, if H is given by the equation $\langle v,x\rangle + t = 0$ with some $v \in \mathbb{F}_p^n\backslash\{0\}$ and $t \in \mathbb{F}_p$, we have $H^\perp = v$ (note that v is unique only up to scaling, but this will cause no issues later). Equipped with these definitions, we are interested in the minimal number of hyperplanes in an irredundant covering $\mathcal{H}$ of $\mathbb{F}_p^n$ such that the vectors $\{H^\perp : H \in \mathcal{H}\}$ span the whole space. We denote this minimum by $f_p(n)$. The motivation for studying this function comes from surprising connections with several long-standing conjectures in linear algebra and group theory, which we discuss later.

Let us first make some simple observations about $f_p(n)$. To begin with, $f_p(1) = p$ and $f_p(2) = p+1$. While the former is trivial, the latter requires some explanation. Consider the $p+1$ lines in $\mathbb{F}_p^2$ going through the origin. These form an irredundant covering, and the normal vectors trivially span the whole space, showing that $f_p(2) \leqslant p+1$. On the other hand, as every line in $\mathbb{F}_p^2$ covers exactly p points, the only covering with p lines is the one which consists of the p translates of a fixed line, in which case the normal vectors only span a 1-dimensional subspace. Therefore, $f_p(2) \geqslant p+1$. In general, a blow-up type construction gives the upper bound $f_p(n) \leqslant \lceil \frac{pn}{2} \rceil + 1$, and this is the best upper bound we are aware of for every prime p and $n \geqslant 2$. From below, it is easy to show that $f_p(n) \geqslant n+1$. Indeed, if

the vectors $\{H^\perp : H \in \mathcal{H}\}$ span $\mathbb{F}_p^n$, then there are at least n of them. But if they are linearly independent, it is easy to show that $\mathcal{H}$ cannot be a covering.

One might notice that in case $p = 2$, the lower and upper bounds coincide, so we have $f_2(n) = n + 1$. The problem becomes interesting when $p \geqslant 3$. It turns out that proving a lower bound of the form $f_p(n) \geqslant (1 + \varepsilon_p)n$ with some $\varepsilon_p > 0$ is already difficult, with several surprising consequences.

- Given $A \subset \mathbb{F}$ and a multiset $B \subset \mathbb{F}_p^n$, say that B is an *A-basis* if every $v \in \mathbb{F}_p^n$ can be expressed as $v = \sum_{b \in B} c_b b$, where $c_b \in A$ for every $b \in B$. Also, B is an *additive basis*, if B is a $\{0, 1\}$-basis. The Additive Basis conjecture [2] of Alon, Linial and Meshulem states that for every p there exists some $k = k(p)$ such that union of k linear basis in $\mathbb{F}_p^n$ is an additive basis. This conjecture is wide open for every $p \geqslant 3$. It turns out that if $f_3(n) \geqslant (1 + \varepsilon)n$ for some $\varepsilon > 0$, then the Additive basis conjecture holds for $p = 3$.

  Szegedy [8] proposed a weakening of the Additive basis conjecture: there exists some $k' = k'(p)$ such that the union of k' linear basis is an $(\mathbb{F}_p \backslash \{0\})$-basis. The statement $f_p(n) \geqslant (1 + \varepsilon_p)n$ for some $\varepsilon_p > 0$ implies this conjecture.
- The Alon-Jaeger-Tarsi conjecture [3, 4] states the following. Let $p \geqslant 5$ and let $M \in \mathbb{F}_p^{n \times n}$ be an invertible matrix. Then there exists $x \in \mathbb{F}_p^n$ such that neither x, nor Mx has a zero coordinate. In [3], it was proved that the conjecture holds if p is a proper prime power, however, it remained open for every prime p until recently. Nagy and Pach [5] proved that the conjecture holds if $p \geqslant 67, p \neq 79$. Surprisingly, if p satisfies $f_p(n) > 2n$, then the Alon-Jaeger-Tarsi conjecture holds for p.
- Let G be an abelian group, and let $H_1 x_1, \ldots, H_k x_k$ be an irredundant covering of G with cosets, where $H_1, \ldots, H_k < G$. Pyber [7] proposed the conjecture that $|G : \bigcap_{i \in [k]} H_i| = 2^{O(k)}$. This conjecture is implied by the statement $f_p(n) = \Omega(n \log p)$.

Building on the ideas of [5], we establish the following lower bound in [6].

Theorem 1. *There exists $c > 0$ such that $f_p(n) \geqslant c \frac{\log p}{\log \log p} \cdot n$. Also, there exists $\varepsilon > 0$ such that if $p \geqslant 5$, then $f_p(n) \geqslant (1 + \varepsilon)n$.*

In my talk, I give a rough outline of the proof of this theorem. As one can see, Theorem 1 resolves the conjecture of Szegedy [8] about weak additive bases for $p \geqslant 5$, and it implies the Alon-Jaeger-Tarsi conjecture for sufficiently large primes. Actually, by studying properties of irredundant coverings (therefore, not by directly by looking at the function $f_p(n)$), we [6] can also establish the following strengthening of the weak additive basis conjecture: for every p, there exists $A \subset \mathbb{F}_p$ of size $O(\log p)$ such that the union of p basis is an A-basis in $\mathbb{F}_p^n$. In the case of coset covers of groups, we get the following bound. This improves an old result of Tomkinson [9] for abelian groups, and is only a bit short of the conjectured bound of Pyber.

Corollary 2. *Let G be an abelian group, and $H_1x_1, \dots, H_kx_k$ be an irredundant coset cover of G. Then*

$$|G : \bigcap_{i\in[k]} H_i| \leqslant 2^{O(k \log\log k)}.$$

References

[1] N. Alon, Z. Füredi, *Covering the cube with by affine hyperplanes*, European J. Combin. **14** (2) (1993), 79–83.

[2] N. Alon, N. Linial, and R. Meshulam, *Additive bases of vector spaces over prime fields*, J. Combin. Theory Ser. A **57** (1991), 203–210.

[3] N. Alon, and M. Tarsi, *A nowhere-zero point in linear mappings*, Combinatorica **9** (4) (1989), 393–395.

[4] F. Jaeger, *Nowhere-zero flow problems*, in "Selected Topics in Graph Theory" (L. Beineke and R. Wilson, Eds.), Vol. 3, pp. 91–95, Academic Press, London/New York, 1988.

[5] J. Nagy, and P. P. Pach, *The Alon-Jaeger-Tarsi conjecture via group ring identities*, preprint (2021), arXiv:2107.03956.

[6] J. Nagy, P. P. Pach, and I. Tomon, *Additive bases, coset covers, and non-vanishing linear maps*, preprint (2021), arXiv:2111.13658.

[7] L. Pyber, *How abelian is a finite group?* in: The Mathematics of Paul Erdős, vol. I, Springer-Verlag, Heidelberg, (1996), 372–384.

[8] B. Szegedy, *Coverings of abelian groups and vector spaces*, J. Combin. Theory Ser. A **114** (1) (2007), 20–34.

[9] M.J. Tomkinson, *Groups covered by finitely many cosets or subgroups*, Comm. Algebra **15** (4) (1987), 845–859.

Roots of random functions: Recent progress and open questions

Van Vu

One of the key question in mathematics is: *Where are roots of a function ?*

In this work, we focus on the case when the function in question is generated randomly. Random functions have been studied for almost 100 years in many fields of mathematics, from probability to analysis to combinatorics.

A function is typically written down in a base $f(x) = \sum_{i=0}^{n} a_i h_i(x)$, where $h_0, h_1, h_2, \cdots$ form a base.

Example. $h_i(x) = x^i$ (Taylor expansion), $h_i(x) \cos ix$ (Fourier expansion) etc.

Random functions. To obtain a random function, one sets a_i be independent random variables. For normalization, we write $a_i = c_i\xi_i$ where c_i are scalar coefficients and ξ are independent random variables with mean 0 and variance 1. These ξ_i are not necessarily idd.

Example. Consider the Taylor base $h_i(x) = x^i$; in this case we talk about random polynomials. Within this class, there are already many different ensembles which play important roles in different fields. For instance,

$c_i = 1$; Kac polynomials

$c_i = \frac{1}{\sqrt{i!}}$; Weyl polynomials

$c_i = \sqrt{\binom{n}{i}}$; Binomial polynomials.

While our study also treats complex roots, in the rest of the abstract we concentrate on real roots. Apparently, the number of real roots is a random variable.

Problem 1. *What can we say about its mean, variance, and limiting distribution of this real random variable ?*

The series of papers by Kac and Littlewood-Offord in the 1940s on this question is the starting point of the theory of random functions. However, after more than 80 years, many very basic questions are still open. Let us start with

Kac polynomials. The simplest ensemble when all $c_i = 1$

$$P_n(x) = \sum_{i=0}^{n} \xi_i x^i,$$

where ξ_i are i.i.d copies of a random variable ξ with mean 0 and variance 1. Let $N_{n,\xi}$ denote the number of real roots of P_n.

Theorem 2 (Littlewood-Offord; 1943)**.** *For ξ being Rademacher, Gaussian, or uniform on $[-1, 1]$, we have with probability $1 - o(1)$*

$$\frac{\log n}{\log \log n} \leqslant N_{n,\xi} \leqslant \log^2 n.$$

In the proof of this theorem, Littlewood and Offord introduced their famous anti-concentration inequality. At about the same time, Kac discovered a general formula for the expectation

$$\mathbb{E} N_{n,\xi} = \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} |y| p(t, 0, y) dy, \tag{1}$$

where $p(t, x, y)$ is the joint density for $(P_{n,\xi}(t), P'_n(t))$ at the point (x, y).

In the Gaussian case, the RHS can be computed explicitly and Kac showed

$$\mathbb{E} N_{n,Gauss} = \frac{1}{\pi} \int_{-\infty}^{\infty} \sqrt{\frac{1}{(t^2-1)^2} + \frac{(n+1)^2 t^{2n}}{(t^{2n+2}-1)^2}} dt = \frac{2}{\pi} \log n + o(\log n). \tag{2}$$

The RHS of Kac's formula is not easy (in many cases impossible) to compute if ξ is not Gaussian. In particular, the Rademacher (± 1) case was a real challenge. 10 years after Kac's paper, Erdos-Offord (1956) showed (using an entirely new argument) that

$$\mathbb{E} N_{n,Rademacher} = \frac{2}{\pi} \log n + o(\log n). \tag{3}$$

In the late 1960s and early 1970s, Ibragimov and Maslova successfully generalized Erdös-Offord's method to handle any general random ξ with mean 0 and variance 1, showing that

$$\mathbb{E}N_{n,\xi} = \frac{2}{\pi}\log n + o(\log n). \tag{4}$$

Thus, it takes 3 decades and the join effort of many leading mathematicians to settle this very first case. Now, we state the general problem

Problem 3. *Determine the expectation $\mathbb{E}N_{n,\xi}$ for general random polynomials or even general random functions.*

A notable fact here is that Kac's formula (or more advanced methods on Gaussian processes) allows us to obtain a precise answer in the case when ξ is Gaussian. Much less has been known for other variables. Until recently, there was no analogue of Erdös-Offord's result for other classes of random polynomals (listed above), as their proof used the fact that all $c_i = 1$ in a critical way. Now we introduce our new method

The Universality method. This is a method to prove asymptotics results concerning one ensemble by comparison to another ensemble whose behavior is known.

A simple example is the basic central limit theorem for the sum $S_n = \frac{1}{\sqrt{n}}\sum_i \xi_i$, of i.i.d random variables ξ_i with mean zero and unit variance. We can prove this theorem by first noticing that the statement holds for the case when ξ_i are Gaussian (in this case S_n is itself Gaussian). Next, we show that when we switch each Gaussian variable to a general ξ, the distribution of S_n does not deviate too much. This seems to be an overkill, as one can prove the CLT in more direct ways. However, for many subtle quantities, this is essentially the only efficient method at the moment. An early combinatorial application in this spirit is the Sandwiching argument (initiated by Kim and Vu, around 2003) which compares random regular graphs with Erdos-Renyi graphs by switching edges. In 2009, Tao and Vu successfully used the Universality approach for many problems in random matrix theory (leading to, along others, the Four moment theorem).

In 2015, Tao and Vu managed to use this method to study the number of real roots of Weil polynomial and Binomial polynomial with ξ being Rademacher. By a direct comparison with the Gaussian case, we obtained, for a general variable ξ

$$\mathbb{E}N_{n,\xi} = (1 + o(1))\mathbb{E}N_{n,Gauss}.$$

In both cases, the expectation is of order $\Theta(\sqrt{n})$. The method also works for Kac's polynomials (reproving Erdos-Offord result), but the details are quite technical, with a heavy use of Inverse Littlewood-Offord theory.

Moving from Taylor expansion to Fourier expansion, researchers have been considering functions of the form

$$F_n(x) = \sum_{i=0}^{n} c_i \xi_i \cos ix,$$

or, more generally,

$$F_n(x) = \sum_{i=1}^{n} c_i \xi_i \cos ix + d_i \eta_i \sin ix.$$

These are called random trigonometric polynomials. The behavior of the Kac polynomial, Weyl polynomial, and trigonometric polynomials are very different. The first has (typlically) $\Theta(\log n)$ real roots, the second $\Theta(\sqrt{n})$, and the third $\Theta(n)$. Moreover, the methods researchers used to treat these three ensembles have been usually very technical and completely different. Furthermore, the universality arguments from Tao-Vu's paper do not apply for random trigonometric polynomials, and relatively little was known beyond the Gaussian setting.

A new, unified, universality frame work. In 2017, O. Nguyen and Vu found a new universality argument which gives a unified treatment of all discussed classes of random functions and many others. Our result provides a general condition which guarantees that two ensembles can be compared. In all known cases, checking the validity of the condition is straightforward. This way, we deduced new proof of many complicated results in a few pages, via routine computation. We can also prove several new results. The proof of the master (universality) theorem itself is also new, much shorter and simpler than the proofs of the universality theorem in any individual case; see our paper (Roots of random functions: a framework for local universality, American Journal of Mathematics 144(01), 2022, 1-74) for more details.

For variance and limiting distributions, we only have results for few special ensembles. Most of the problem is still open, and is at the very heart of the theory.

Nearly all k-SAT functions are unate

Yufei Zhao

(joint work with József Balogh, Dingding Dong, Bernard Lidický, and Nitya Mani)

We establish the following result, originally conjectured by Bollobás, Brightwell, and Leader [4].

Theorem 1. *Fix $k \geqslant 2$. The number of k-SAT functions on n Boolean variables is $(1 + o(1))2^{n+\binom{n}{k}}$.*

Equivalently: a $1 - o(1)$ fraction of all k-SAT functions on n Boolean variables are unate.

Here a k-SAT function is a function $f\colon \{0,1\}^n \to \{0,1\}$ of the form $f(x_1, \dots, x_n) = C_1 \wedge C_2 \wedge \cdots \wedge C_m$, where each clause C_i has the form $z_1 \vee \cdots \vee z_k$ with $z_1, \dots, z_k \in \{x_1, \overline{x_1}, \dots, x_n, \overline{x_n}\}$. (This is the CNF version of the problem; it is also equivalent

to consider the DNF version.) We call such a function *unate* if it is monotone after first negating some subset of variables, or equivalently, each variable x_i appears only positively (as x_i) or only negatively (as $\overline{x_i}$). For fixed $k \geqslant 2$, an easy argument shows that the number of unate k-SAT functions is $(1 + o(1))2^{n+\binom{n}{k}}$.

Bollobás, Brightwell, and Leader [4] proved a weaker version of this conjecture for $k = 2$: the number of 2-SAT functions on n Boolean variables is $2^{(1+o(1))\binom{n}{2}}$. The conjecture for $k = 2$ was proved by Allen [1] and $k = 3$ by Ilinca and Kahn [6]. Our work [2, 5] settles the conjecture completely.

We obtain a slightly stronger conclusion. Here a *minimal* k-SAT formula is a formula where deleting any clause changes the function.

Theorem 2. *For $k \geqslant 2$. A $1 - o(1)$ fraction of all minimal k-SAT formulae on n Boolean variables are unate.*

Our work has two parts. The first part [5] (joint with Dong and Mani) reduces the Bollobás–Brightwell–Leader conjecture to a Turán problem on partially directed hypergraphs. I presented this part in an April 2022 Oberwolfach meeting (workshop 2217: Combinatorics, Probability and Computing). Balogh, who was also present at the meeting, and also Lidický, later joined the effort, and together with all the coauthors, we solved the Turán problem, thereby fully resolving the Bollobás–Brightwell–Leader conjecture.

Next we state our Turán result on partially directed hypergraphs. We refer to [5] for the definitions of a k-PDG and $\vec{T}_k$.

Theorem 3. *Fix $k \geqslant 2$. There exists $\theta > \log_2 3$ such that every $\vec{T}_k$-free n-vertex k-PDG with $\alpha\binom{n}{k}$ undirected edges and $\beta\binom{n}{k}$ directed edges satisfies*

$$\alpha + \theta\beta \leqslant 1 + o_{n\to\infty}(1).$$

We prove the inequality by providing a short sum-of-squares certificate. We were assisted by the use of the flag algebra method, which helped us solve the problem for small values of k, and from which we could extrapolate to general k by hand. This is an interesting application of the flag algebra method for which the computer played a crucial assistive role in finding the form of the answer, but ultimately the proof is quite clean and does not require substantial computation.

There several interesting open directions. What happens when k grows with n? Bollobás and Brightwell [3] conjectured that as long as $k = k(n) \leqslant (1/2 - c)n$ for any constant $c > 0$, the number of k-SAT functions is $2^{(1+o(1))\binom{n}{k}}$. Our proof method allows k to grow slowly with n, but the quantitative dependencies are far worse than linear in n. It would be interesting to investigate the threshold $k(n)$ below which a typical k-SAT function is unate.

In a different direction, what about sparser formulae? Consider a uniformly chosen minimal k-SAT formula with n variables and m clauses. For a fixed k, what is the threshold $m(n)$ above which a typical minimal k-SAT formula is unate? And when the typical formula is not unate, what is its typical structure? These questions are analogous to well-studied questions in random graphs concerning the typical structure of a triangle-free graph at various edge densities.

References

[1] Peter Allen, *Almost every 2-SAT function is unate*, Israel J. Math. **161** (2007), 311–346.
[2] József Balogh, Dingding Dong, Bernard Lidický, Nitya Mani, Yufei Zhao *Nearly all k-SAT functions are unate*, arXiv:2209.04894.
[3] Béla Bollobás and Graham R. Brightwell, *The number of k-SAT functions*, Random Structures Algorithms **22** (2003), 227–247.
[4] Béla Bollobás, Graham R. Brightwell, and Imre Leader, *The number of 2-SAT functions*, Israel J. Math. **133** (2003), 45–60.
[5] Dingding Dong, Nitya Mani, and Yufei Zhao, *Enumerating k-SAT functions*, ACM-SIAM Symposium on Discrete Algorithms (SODA'22), arXiv:2107.09233.
[6] L. Ilinca and J. Kahn, *The number of 3-SAT functions*, Israel J. Math. **192** (2012), 869–919.

Large Cliques in Graphs with High Chromatic Number

Penny Haxell
(joint work with Colter MacDonald)

The classical theorem of Brooks [2] tells us that the chromatic number $\chi(G)$ of every graph G with maximum degree $\Delta(G) \geqslant 3$ attains the simple greedy upper bound $\chi(G) = \Delta(G) + 1$ only when G contains a clique of size $\Delta(G) + 1$ (and hence $G = K_{\Delta(G)+1}$ if it is connected). The very natural question therefore arises: if G has chromatic number $\Delta(G)$, or more generally very close to $\Delta(G)$, what can we say about G? In particular, is it necessarily true that it must contain a large clique?

In 1977, Borodin and Kostochka [1] took a first step in this direction, and conjectured that if $\Delta(G) \geqslant 9$ and $\chi(G) = \Delta(G)$ then G should contain a clique with $\Delta(G)$ vertices. The condition $\Delta(G) \geqslant 9$ is necessary due for example to the graph $C_5(3)$ formed by replacing each vertex of a 5-cycle by a triangle, and each edge by a $K_{3,3}$. The maximum degree of $C_5(3)$ is 8, it does not contain a K_8 (or even a K_7), and it is not 7-colourable since it has 15 vertices and no independent set of size 3.

Despite substantial work on the problem by many authors, the Borodin-Kostochka Conjecture remains open in general. In their original paper, Borodin and Kostochka [1] showed the existence of a clique on $\lfloor(\Delta(G)+1)/2\rfloor$ vertices when $\Delta(G) \geqslant 7$. This was improved by Mozhan [18], who showed the existence of a clique on $\lfloor(2\Delta(G)+1)/3\rfloor$ vertices when $\Delta(G) \geqslant 10$. A breakthrough was achieved by Kostochka who improved the clique size to $\Delta(G) - 28$ in [16]. This line of work was continued in Mozhan's Ph.D thesis (which is not easily available, see [20]), where he proved the existence of a $(\Delta(G) - 3)$-clique for $\Delta(G) \geqslant 31$. The best known result is currently $\Delta(G) - 3$ for $\Delta(G) \geqslant 13$ due to Cranston and Rabern [7]. There has also been significant progress in proving the Borodin-Kostochka Conjecture for special classes of graphs, for example claw-free graphs [6] and graphs without induced copies of P_5 or C_4 [9]. Reed [20] proved the very strong result that the conjecture holds for all graphs of sufficiently large maximum degree (where the bound implied by the proof is about 10^{14}, but it is remarked in the paper that this could be brought down substantially, perhaps as low as 10^6 or even 10^3).

Graphs G for which $\chi(G) = \Delta(G) - t$ for t in the range $1 \leqslant t \leqslant 5$ were studied by Farzad, Molloy and Reed [8], who proved that if $\Delta(G)$ is at least some (large unspecified) constant, then such graphs must contain one of a very specific set of subgraphs, all of which contain a large clique. For example, in the simplest case $t = 1$, if $\chi(G) \geqslant \Delta(G) - 1$ then G contains a $(\Delta(G) - 1)$-clique or a $(\Delta(G) - 4)$-clique joined to a C_5 (which itself contains a $(\Delta(G) - 2)$-clique). When $t = 4$ the list consists of 420 subgraphs, and when $t = 5$ it grows to at least 17000.

Our main result is the following.

Theorem 1. *Let t be a nonnegative integer. Then every graph G with $\chi(G) = \Delta(G) - t$ and $\Delta(G) \geqslant 6t^2 + 20t + 16$ contains a clique of size $\Delta(G) - 2t^2 - 6t - 3$.*

(We remark that we do not attempt to obtain the best lower bound for $\Delta(G)$, and the expression $6t^2 + 20t + 16$ can be somewhat improved with extra work.) Theorem 1 generalizes the result of Cranston and Rabern [7] and Mozhan [18], and also in spirit the result of Farzad, Molloy and Reed [8], with a weaker lower bound for the clique size for $1 \leqslant t \leqslant 5$ but with a small and concrete lower bound condition on $\Delta(G)$. It also makes a step towards the resolution of a long-standing question of Reed [21], who addressed the more general problem of proving the existence of large cliques in graphs G in which $\chi(G)$ is large in terms of $\Delta(G)$. He conjectured that the maximum size $\omega(G)$ of a clique in any graph G should satisfy

$$\chi(G) \leqslant \lceil \frac{1}{2}(\Delta(G) + 1 + \omega(G)) \rceil.$$

(The graph $C_5(s)$ obtained by replacing each vertex of C_5 with a K_s is a tight example for every value of s.) Interpreting Reed's Conjecture for $\chi(G) = \Delta(G) - t$ gives that a clique of size at least $\Delta(G) - 2t - 2$ should exist in G. Our result provides a weaker form of this statement, with a quadratic expression in t instead of the linear quantity $2t + 2$.

This related conjecture has also generated much interest, for example, see [5, 11, 13, 14, 15, 19, 21, 22].

The main tool we use is the method of *Mozhan partitions*. This important technique was introduced by Mozhan in [18], and further developed and used extensively by Cranston and Rabern in e.g. [6, 7] to make significant progress on the Borodin-Kostochka Conjecture. In particular this machinery was key in their proof of the best current bound for the general version of the problem. Typically a graph will have many Mozhan partitions, and the method involves a scheme called a *move sequence* for exploring the set of all such partitions to find one that is particularly favourable. While we use this same basic approach in our proof, our choice of move sequence is quite different from that used by Cranston and Rabern in [7], and in particular provides a simpler proof of their result as well as the generalization given by Theorem 1. We remark that the specific case $t = 1$ was addressed in [17] using an argument parallel to that of [7], and gave a lower bound of $\Delta(G) - 17$ on the clique size.

References

[1] O. Borodin and A. Kostochka, On an upper bound of a graph's chromatic number depending on the graph's degree and density, *J. Combin. Theory Ser. B*, 23, (1977), 247-250.

[2] R. Brooks, On Colouring the Nodes of a Network, *Math. Proc. Cambridge Phil. Soc.*, 37 (1941), 194–197.

[3] P. Catlin, Embedding subgraphs and coloring graphs under extremal degree conditions, *Ph. D thesis*, (1976), The Ohio State University.

[4] P. Catlin, Another bound on the chromatic number of a graph, *Discrete Mathematics*, 24, (1978), 1–6.

[5] M. Chudnovsky, A. King, M. Plumettaz, and P. Seymour, A local strengthening of Reed's ω, Δ, χ conjecture for quasi-line graphs, *SIAM Journal on Discrete Mathematics*, 27(1), (2013), 95–108.

[6] D. Cranston and L. Rabern, Coloring claw-free graphs with $\Delta - 1$ colors, *SIAM Journal on Discrete Mathematics*, 27(1), (2013), 534–549.

[7] D. Cranston and L. Rabern. Graphs with $\chi = \Delta$ have big cliques, *SIAM Journal on Discrete Mathematics*, 29(4), (2015), 1792–1814.

[8] B. Farzad, M. Molloy and B. Reed, $(\Delta-k)$-critical graphs, *Journal of Combinatorial Theory, Series B*, 93, (2005), 173–185.

[9] U. Gupta and D. Pradhan, Borodin–Kostochka's Conjecture on (P_5, C_4)-free graphs, *Journal of Applied Mathematics and Computing*, 65, (2020), 1-8.

[10] P. Haxell, A note on vertex list-colouring, *Combinatorics, Probability and Computing*, 10, (2001), 345–347.

[11] T. Kelly and L. Postle, A local epsilon version of Reed's conjecture. *Journal of Combinatorial Theory, Series B*, 141, (2020), 181–222.

[12] A. King, Hitting all maximum cliques with a stable set using lopsided independent transversals, *Journal of Graph Theory*, 67(4), (2011), 300–305.

[13] A. King and B. Reed, Bounding χ in terms of ω and Δ for quasi-line graphs, *Journal of Graph Theory*, 59(3), (2008), 215–228.

[14] A. King and B. Reed, A short proof that χ can be bounded ϵ away from $\Delta + 1$ toward ω, *Journal of Graph Theory*, 81(1), (2016), 30–34.

[15] A. Kohl and I. Schiermeyer. Some results on Reed's conjecture about ω, Δ, and χ with respect to α, *Discrete Mathematics*, 310(9), (2010), 1429–1438.

[16] A. Kostochka. Degree, density and chromatic number of graphs. *Metody Diskret. Analiz.*, 35, (1980), 45–70.

[17] C. MacDonald, On finding large cliques when χ is near Δ, *MMath Thesis*, University of Waterloo, (2021).

[18] N. Mozhan. Chromatic number of graphs with a density that does not exceed two-thirds of the maximal degree, *Metody Diskretn. Anal*, 39, (1983), 52–65.

[19] L. Rabern, A note on Reed's conjecture, *SIAM Journal on Discrete Mathematics*, 22(2), (2008), 820–827.

[20] B. Reed, A Strengthening of Brooks Theorem, *J. Combin. Theory Ser. B*, 76, (1999), 136-149.

[21] B. Reed, ω, Δ and χ, *Journal of Graph Theory*, 27(4), (1998), 177–212.

[22] I. Schiermeyer, Chromatic number of P_5-free graphs: Reed's Conjecture, *Discrete Mathematics*, 339(7), (2016), 1940–1943.

On the Ryser–Brualdi–Stein conjecture

Richard Montgomery

The study of transversals in Latin squares dates back at least to the 18th century when Euler considered Latin squares which can be decomposed into full transversals [5]. A *Latin square of order* n is an n by n grid filled with n symbols, so that every symbol appears exactly once in each row and each column. A *transversal* of a Latin square of order n is a collection of cells in the grid which share no row, column or symbol, while a *full transversal* is a transversal with n cells.

Key examples of Latin squares include the addition tables of finite groups, which easily provide examples that, if n is even, then there are Latin squares of order n with no full transversal (e.g., the addition table for $\mathbb{Z}_2$). In 1967, Ryser [11] conjectured that there are no such Latin squares of order n when n is odd (see also [1]), while Brualdi [3] and Stein [13] later independently conjectured that every Latin square of order n has a transversal with at least $n-1$ cells. The following combined form of these conjectures, known as the Ryser–Brualdi–Stein conjecture, has become the most widely known open problem on transversals in Latin squares.

Conjecture 1 (The Ryser–Brualdi–Stein conjecture)**.** *Every Latin square of order n has a transversal with $n-1$ cells, and a full transversal if n is odd.*

Towards Conjecture 1, increasingly large transversals were shown to exist in any Latin square by Koksma [8], and Drake [4], before Brouwer, De Vries and Wieringa [2] and Woolbright [14] independently showed that every Latin square of order n has a transversal with at least $n-\sqrt{n}$ cells. In 1982, Shor [12] showed that a transversal with $n-O(\log^2 n)$ cells exists in any Latin square of order n, though the proof had an error that was only noticed and corrected by Hatami and Shor in 2008 [6]. This bound (essentially) stood for several decades until the breakthrough work of Keevash, Pokrovskiy, Sudakov and Yepremyan [7] in 2020, which showed that every Latin square of order n has a transversal with $n-O(\log n/\log\log n)$ elements.

In this talk, I will discuss the following result.

Theorem 2. *There is some $n_0 \in \mathbb{N}$ such that every Latin square of order $n \geqslant n_0$ contains a transversal with $n-1$ cells.*

The bound $O(\log n/\log\log n)$ in the result by Keevash, Pokrovskiy, Sudakov and Yepremyan [7] is a natural barrier, and it seems likely this is the best bound that can be achieved with methods that approach each Latin square in the same manner. Thus, for Theorem 2, we introduce the first techniques to identify and exploit the possible algebraic properties behind the entries in a Latin square.

To study transversals in Latin squares, it is common to work on an equivalent formulation in properly coloured bipartite graphs (see, for example, [7]). Let G be a complete bipartite graph with n vertices in each class which is properly coloured using n colours. To prove Theorem 2, it is equivalent to show that, when n is large, G contains a *rainbow matching* with $n-1$ edges. That is, a matching in which each edge has a different colour.

Under this equivalence, the possible algebraic properties behind the entries in a Latin square become the possible algebraic properties of the colouring of G. To identify some rough algebraic properties in the colouring, we consider pairs of colours c, d which, for practical purposes, we can consider to be equivalent when constructing a rainbow matching. Two colours c, d are considered to be equivalent if we can robustly find small rainbow matchings M_1 and M_2 in G with the same vertex class and with exactly the same colours, except M_1 uses the colour c and M_2 uses instead the colour d. If we are building a rainbow matching constructively, we can add M_1 to the growing matching. Then, instead of finding a rainbow matching using exactly the unused vertices and unused colours, we need only find a rainbow matching using exactly the unused vertices and all of the unused colours, except now we want such a matching using exactly one colour in $\{c, d\}$. If the colour c is used in this matching, then we can switch M_2 with M_1 to get a rainbow matching overall.

This is a small example of how colours c and d can be considered 'equivalent'. In the full proof, classes of colours are developed, where each class consists of colours for which any pair can be considered equivalent. These classes are developed so that they have a rough algebraic structure, inspired in part by extremal examples for Conjecture 1.

To prove Theorem 2, we then use a combination of the semi-random method and the absorption method. We use an implementation of the semi-random method in this setting from work of the author with Pokrovskiy and Sudakov [9], and therefore all the main novelty occurs in our use of absorption. Since its codification in 2008 as a general approach by Rödl, Ruciński and Szemerédi [10], absorption has been a critical tool in turning approximate results into exact results. We aim to set aside some special 'absorber' which can be extended into a rainbow matching with $(1-o(1))n$ edges in our properly coloured bipartite graph G using the semi-random method. The aim is that the absorber should have some special properties to turn this into a rainbow matching with $n-1$ edges.

However, the extremal examples showing a full transversal may not always exist (when n is even) demonstrate the challenge of using the absorption method in this setting. In these examples, the algebraic properties behind the equivalent colourings prevent the existence of the typical absorbers used for an application of the absorbing method. To prove Theorem 2, we instead use the colour classes that we develop to create an 'absorption structure' with a very restricted absorption property. We then introduce an 'addition structure' with much less restrictive requirements, which is able to identify the pair of vertices we will leave out of the rainbow matching, leaving a set of uncovered vertices and colours that satisfies the restrictive property required for absorption.

In summary, then, we identify some rough algebraic structure in the colouring of G and use this to construct an absorption structure and an addition structure, both of which we set aside. Then, we use the semi-random method method to find a large rainbow matching using the unused vertices and colours, before applying

the addition structure and then the absorption structure to turn this into a rainbow matching with $n-1$ edges.

References

[1] D. Best and I. M. Wanless. What did Ryser conjecture? *arXiv preprint arXiv:1801.02893*, 2018.

[2] A. E. Brouwer, A. de Vries, and R. Wieringa. A lower bound for the length of partial transversals in a Latin square. *Nieuw Archief Voor Wiskunde*, 26(2):330–332, 1978.

[3] R. A. Brualdi and H. J. Ryser. *Combinatorial matrix theory*. Cambridge University Press, 1991.

[4] D. A. Drake. Maximal sets of latin squares and partial transversals. *Journal of Statistical Planning and Inference*, 1(2):143–149, 1977.

[5] L. Euler. Recherches sur un nouvelle espéce de quarrés magiques. *Verhandelingen uitgegeven door het zeeuwsch Genootschap der Wetenschappen te Vlissingen*, pages 85–239, 1782.

[6] P. Hatami and P. W. Shor. A lower bound for the length of a partial transversal in a latin square. *Journal of Combinatorial Theory, Series A*, 115(7):1103–1113, 2008.

[7] P. Keevash, A. Pokrovskiy, B. Sudakov, and L. Yepremyan. New bounds for Ryser's conjecture and related problems. *arXiv preprint arXiv:2005.00526*, 2020.

[8] K. K. Koksma. A lower bound for the order of a partial transversal in a Latin square. *Journal of Combinatorial Theory*, 7(1):94–95, 1969.

[9] R. Montgomery, A. Pokrovskiy, and B. Sudakov. Decompositions into spanning rainbow structures. *Proceedings of the London Mathematical Society*, 119(4):899–959, 2019.

[10] V. Rödl, A. Ruciński, and E. Szemerédi. A Dirac-type theorem for 3-uniform hypergraphs. *Combinatorics, Probability and Computing*, 15(1-2):229–251, 2006.

[11] H. Ryser. Neuere probleme der kombinatorik. *Vorträge über Kombinatorik, Oberwolfach*, pages 69–91, 1967.

[12] P. W. Shor. A lower bound for the length of a partial transversal in a Latin square. *Journal of Combinatorial Theory, Series A*, 33(1):1–8, 1982.

[13] S. K. Stein. Transversals of Latin squares and their generalizations. *Pacific J. Math.*, 59:567–575, 1975.

[14] D. E. Woolbright. An $n \times n$ Latin square has a transversal with at least $n - \sqrt{n}$ distinct symbols. *Journal of Combinatorial Theory, Series A*, 24(2):235–237, 1978.

Rainbow matchings in hypergraphs

Lisa Sauermann

(joint work with Cosmin Pohoata, Dmitrii Zakharov)

Motivated by classical questions about transversals in Latin Squares such as the famous Ryser–Brualdi–Stein Conjecture, there has been a lot of work on finding rainbow matchings in properly edge-colored graphs (see e.g. [6, 11] and the references therein). In a properly edge-colored graph every color class is a matching, so these questions amount to finding a rainbow matching among a collection of matchings of different colors. Similar questions have also been studied in the setting of hypergraphs, which was the focus of this talk.

A matching in an r-uniform hypergraph is a collection of pairwise disjoint edges (and the size of the matching is the number of edges it consists of). Given matchings $M_1, \dots, M_N$ in some r-uniform hypergraph, where we think of each matching as colored in a different color, a *rainbow matching* is a matching consisting of edges

$e_1 \in M_{i_1}, \ldots, e_\ell \in M_{i_\ell}$ with distinct indices $i_1, \ldots, i_\ell \in \{1, \ldots, N\}$ (in other words, a matching consisting of edges with distinct colors).

The following problem goes back to Aharoni and Berger [1].

Problem 1. *Let $M_1, \ldots, M_N$ be matchings of size t in some r-uniform hypergraph. How large does N need to be (in terms of t and r), such that it is always possible to find a rainbow matching of size t? In other words, how large does N need to be such that it is always possible to find distinct indices $i_1, \ldots, i_t \in \{1, \ldots, N\}$ and pairwise disjoint edges $e_1 \in M_{i_1}, \ldots, e_t \in M_{i_t}$?*

We remark that there may be edges belonging to more than one of the matchings $M_1, \ldots, M_N$, in which case we can think of such edges having more than one color. To form a rainbow matching, one may choose which color to use these edges with.

Equivalently to Problem 1, one may ask about the maximum possible number of matchings of size t in some r-uniform hypergraph without a rainbow matching of size t. Denoting this maximum possible number by $F(r,t)$, the answer to Problem 1 is precisely $N = F(r,t) + 1$.

It is also natural to ask about Problem 1 with the additional restriction that the underlying r-uniform hypergraph is r-partite. This was in fact the original version of the problem proposed by Aharoni and Berger in [1]. Let $f(r,t)$ be the maximum possible number of matchings of size t in some r-partite r-uniform hypergraph without a rainbow matching of size t. We clearly have the inequality $f(r,t) \leqslant F(r,t)$.

In the case of uniformity $r = 2$, i.e. in the case of graphs, Aharoni and Berger [1] proved that $f(2,t) = 2t - 2$ (relying on previous ideas of Drisko [8]). For $F(2,t)$ the best known bounds are $2t - 2 = f(2,t) \leqslant F(r,t) \leqslant 3t - 3$ due to Aharoni–Berger–Chudnovsky–Howard–Seymour [2]. In general, Glebov–Sudakov–Szabó [10] conjectured that $f(r,t)$ is upper-bounded by a linear function of t for any fixed $r \geqslant 2$ (or stated more formally, that for any fixed $r \geqslant 2$ there is a constant $c(r)$ such that $f(r,t) \leqslant c(r) \cdot t$ holds for all t). Alon [4] had also already asked in 2011 whether this is true, based on an intriguing connection with the Erdős–Ginzburg–Ziv problem from additive combinatorics. For the fractional version of Problem 1 (where $M_1, \ldots, M_N$ are fractional matchings and one is looking for a rainbow fractional matching), such a bound was recently proved by Aharoni, Holzman and Jiang [3], using tools from topology.

Nevertheless, it turns out that the conjecture of Glebov–Sudakov–Szabó [10] is actually false: We showed that for any fixed uniformity $r \geqslant 3$, the functions $f(r,t)$ and $F(r,t)$ are in fact on the order of t^r (up to constant factors depending on r).

Theorem 2. *For any fixed uniformity $r \geqslant 3$, there exist positive constants c_r and C_r such that*

$$c_r t^r \leqslant f(r,t) \leqslant F(r,t) \leqslant C_r t^r$$

holds for all $t \geqslant 2$.

Note that for fixed uniformity $r \geqslant 3$, this theorem determines $f(r,t)$ and $F(r,t)$ up to constant factors. In other words, Theorem 2 essentially (up to constant

factors) solves Problem 1 (as well as its r-partite analogue) in the setting of fixed uniformity $r \geqslant 3$.

The best previous lower bounds for both $f(r,t)$ and $F(r,t)$ were linear in t. The best previous upper bound for both $f(r,t)$ and $F(r,t)$ was $(t-1)\binom{tr}{r}$ due to Munhá Correia, Sudakov and Tomon [7], stated in the following theorem. This upper bound is on the order of t^{r+1} for fixed r, so Theorem 2 improves both the known lower and upper bounds in this regime.

Theorem 3 ([7])**.** *For any $t \geqslant 2$ and $r \geqslant 2$, we have*

$$f(r,t) \leqslant F(r,t) \leqslant (t-1)\binom{tr}{r}.$$

This upper bound due to Munhá Correia, Sudakov and Tomon [7] improved upon previous upper bounds of Alon [4] and Glebov–Sudakov–Szabó [10]. Their proof of $F(r,t) \leqslant (t-1)\binom{tr}{r}$ is of a linear algebraic nature, using exterior power algebras.

Our proof of the upper bound in Theorem 2 is purely combinatorial, and is motivated by arguments that first appeared in the context of the famous sunflower problem [12, 5] and were later used in the resolution of the fractional Kahn–Kalai conjecture by Frankston–Kahn–Narayanan–Park [9]. For the lower bound in Theorem 2 we gave an explicit construction.

It is also interesting to study the functions $f(r,t)$ and $F(r,t)$ in the opposite regime, where $t \geqslant 2$ is fixed and r is large. For example, Alon [4] explicitly asked about studying the function $f(r,t)$ in the case of $t = 3$ and large r, i.e. about understanding the growth behavior of $f(r,3)$ as a function of r.

For fixed $t \geqslant 2$ and large r, the upper bound $(t-1)\binom{tr}{r}$ for $F(r,t)$ due to Munhá Correia, Sudakov and Tomon is (up to constant factors depending on t) on the order of $(t^t/(t-1)^{t-1})^r/\sqrt{r}$, i.e. it is exponential in r with base $t^t/(t-1)^{t-1}$ (which, for relatively large t, is roughly et). We proved that this upper bound is tight up to sub-exponential factors in r.

Theorem 4. *For any fixed $t \geqslant 2$, and any large r, we have*

$$F(r,t) \geqslant \left(\frac{t^t}{(t-1)^{t-1}}\right)^{r-O(\sqrt{r})}.$$

Here, the implicit constant in the O-notation may depend on t.

For $f(r,t)$, i.e. in the setting of r-partite graphs, an easy adaptation of the proof of Munhá Correia–Sudakov–Tomon yields $f(r,t) \leqslant (t-1)\cdot t^r$. Again, we proved that for fixed $t \geqslant 2$ and large r, this is tight up to sub-exponential factors in r.

Theorem 5. *For any fixed $t \geqslant 2$, and any $r \geqslant 2$, we have*

$$t^{r-O(\sqrt{r})} \leqslant f(r,t) \leqslant (t-1)\cdot t^r.$$

Again, the implicit constant in the O-notation may depend on t.

Theorem 5 in particular answers a question of Glebov–Sudakov–Szabó [10], asking whether $f(r,t)$ can be upper-bounded by a function of the form $\alpha_t\beta^r$ (where

α_t may depend on t, but β is an absolute constant). The lower bounds in Theorems 4 and 5 can be proved via a probabilistic argument with some ingredients from additive combinatorics. Note that Theorems 4 and 5 (together with the upper bound $F(r,t) \leqslant (t-1)\binom{tr}{r}$ due to Munhá Correia–Sudakov–Tomon) determine the functions $f(r,t)$ and $F(r,t)$ for fixed $t \geqslant 2$ up to lower-order terms (i.e. up to sub-exponential terms in r). Recall that in the opposite regime, for fixed $r \geqslant 3$, Theorem 2 determines these functions up to constant factors (depending on r).

References

[1] R. Aharoni and E. Berger, Rainbow matchings in r-partite r-graphs, *Electron. J. Combin.* **16**(1) (2009), R119.

[2] R. Aharoni, E. Berger, M. Chudnovsky, D. Howard, and P. Seymour: Large rainbow matchings in general graphs, *European J. Combin.* **79** (2019), 222–227.

[3] R. Aharoni, R. Holzman, and Z. Jiang, Rainbow fractional matchings, *Combinatorica* **39** (2019), pp. 1191–1202.

[4] N. Alon, Multicolored matchings in hypergraphs, *Mosc. J. Comb. Number Theory* **1** (2011), 3–10.

[5] R. Alweiss, L. Shachar, K. Wu, and J. Zhang, Improved bounds for the sunflower lemma, *Ann. of Math.* (2) **194** (2021), no. 3, 795–815.

[6] D. Chakraborti, and P.-S. Loh, Rainbow matchings in edge-colored simple graphs, preprint: `arXiv:2011.04650`.

[7] D. Munhá Correia, B. Sudakov, and I. Tomon, Flattening rank and its combinatorial applications, *Linear Algebra Appl.* **625** (2021), 113–125.

[8] A. Drisko, Transversals in row-latin rectangles, *J. Combin. Theory Ser. A* **84** (1998), 181–195.

[9] K. Frankston, J. Kahn, B. Narayanan, and J. Park, Thresholds versus fractional expectation-thresholds, *Ann. of Math.* (2) **194** (2021), no. 2, 475–495.

[10] R. Glebov, B. Sudakov, and T. Szabó, How many colors guarantee a rainbow matching?, *Electron. J. Combin.* **21**(1) (2014), P1.27.

[11] P. Keevash, A. Pokrovskiy, B. Sudakov, and L. Yepremyan, New bounds for Ryser's conjecture and related problems, *Trans. Amer. Math. Soc. Ser. B* **9** (2022), 288–321.

[12] S. Lovett, N. Solomon, and J. Zhang. From DNF compression to sunflower theorems via regularity. *34th Computational Complexity Conference (CCC 2019) , Leibniz International Proceedings in Informatics* **137** (2019), 5:1–5:14.

Ascending subgraph decompositions

Alexey Pokrovskiy

(joint work with Kyriakos Katsamaktsis, Shoham Letzter, Benny Sudakov)

Since $\binom{m}{2} = 1 + 2 + \cdots + m - 1$, every graph G with $\binom{m}{2}$ edges can be (edge)-decomposed into subgraphs $H_1, \dots, H_{m-1}$ such that $e(H_i) = i$. This talk was about studying whether any extra structure can be imposed in such a decomposition. When G is complete, then a well-known conjecture of Gyárfás predicts that the graphs H_i can be picked to be any prescribed trees. On the other hand Alavi et al. predicted that some additional structure can be imposed on the graphs $H_1, \dots, H_t$ even when G is a general graph with $\binom{m}{2}$ edges [1].

Conjecture 1 (Alavi, Boals, Chartrand, Erdos, Oellermann, [1]). *Every graph G with $\binom{m}{2}$ edges has a decomposition into subgraphs $H_1, \dots, H_{m-1}$ such that $e(H_i) = i$ and for each i, H_i is isomorphic to a subgraph of H_{i+1}.*

A decomposition of G as in the above conjecture is called an *ascending subgraph decomposition*. This conjecture was known to hold for a variety of host graphs G. Alavi et al. proved it for path/cycle forests and also bounded degree forests[1]. Gárfás, Faudree, and Schelp proved it for star forests. Fu and Hu proved it for regular graphs. Fu proved it for graphs with maximum degree $\leqslant m/2$. Additionally there's been a lot of interest in proving stronger conclusions for various host graphs G. For example, the result of Fu produces an isomorphic *matching* decomposition — this could be seen as a variant of edge-colouring. Ma, Zhou, Zhou proved that star forests with stars of size $\geqslant m$ have an ascending *star* decomposition [4]. This is equivalent to the following purely number-theoretic fact:

Theorem 2 (Ma, Zhou, Zhou, [4]). *Let $x_1, \dots, x_t \geqslant m$ be numbers summing to $\binom{m}{2}$. Then the interval $[m-1] = \{1, \dots, m-1\}$ can be partitioned into disjoint sets $S_1, \dots, S_t$ with $\sum S_i = x_i$ for all i.*

We point to Chapter 8 of the book [2] for a detailed survey of ascending decompositions. In this talk a proof of Conjecture 1 for sufficiently large m was presented.

Theorem 3 ([3]). *For sufficiently large m, every graph with $\binom{m}{2}$ has an ascending subgraph decomposition.*

The graphs in the ascending decomposition produced by this theorem have a very special structure — each H_i consists of a vertex disjoint union of one (potentially large) star S together with a lot of components of bounded size. An important intermediate step in the proof of constructing ascending decompositions in this theorem is to first construct isomorphic decompositions of graphs. In this direction the following lemma is proved.

Lemma 4 ([3]). *Let $k \in [m/100, 100m]$, and let G be a graph with $\leqslant m^2$ edges. Then all, except $o(m^2)$, edges of G can be decomposed into k graphs $H_1, \dots, H_k$ which are all isomorphic to each other.*

References

[1] Alavi, Y., Boals, A. J., Chartrand, G., Erdos, P., & Oellermann, O. R. (1987). *The ascending subgraph decomposition problem.* Congressus Numerantium, **58**(7), 14.

[2] Ali, A., Chartrand, G., & Zhang, P. (2021). *Irregularity in Graphs.* New York: Springer.

[3] Katsamaktsis, K, Letzter, S, Pokrovskiy, A, & Sudakov, B. *In preparation.*

[4] Ma, K., Zhou, H., & Zhou, J. (1994). *On the ascending star subgraph decomposition of star forests.* Combinatorica, **14**(3), 307-320.

From Sparse To Dense Random Graphs: Using Sparsity to Prove Results About Mean Field Models

David Gamarnik
(joint work with Yatin Dandy, Lenka Zdeborová)

Friendly bisections. Background and the main results

A partition of nodes $[n]$ of a simple graph into two subsets $A, B \subset [n], A \cup B = [n], A \cap B = \varnothing$ is called "friendly" if for each node $u \in [n]$ the number of neighbors of u on the same side as u is at least as large as the number of neighbors on the opposite side of u. Namely, $d_A(u) \geqslant d_B(u), \forall u \in A$ and $d_B(u) \geqslant d_A(u), \forall u \in B$, where $d_S(u)$ denotes the number of neighbors of u in the set $S \subset [n]$. We call a partition asymptotically friendly if the relation above holds for all by $o(n)$ many nodes (as such this definition is only applicable to *sequences* of graphs as opposed to a fixed graph). We study the existence of friendly or asymptotically friendly partitions in dense random graphs sequences $\mathbb{G}(n, 1/2)$ which are also bisections, in the sense that $|A|$ and $|B|$ are equal when n is even, or differ by at most 1, when n is odd. A conjecture by Füredy [Für88] also included as Problem 20 in Green's [Gre] list of 100 open problems, postulates the existence of such bisections with high probability (w.h.p.) as n increases. This conjecture was confirmed recently by Ferber et al [FKN$^+$21]:

Theorem 1. *Asymptotically friendly bisections exist in* $\mathbb{G}(n, 1/2)$ *w.h.p. as* $n \to \infty$.

The proof of the theorem above is constructive and is based on analyzing a carefully crafted stochastic process which ends in an asymptotically friendly bisection.

In this work we offer an alternative proof of Theorem 1 which is also a generalization of the theorem. To state the theorem we fix $h \geqslant 0$. A bisection (A, B) of $[n]$ is called h-friendly if

$$\frac{d_A(u) - d_B(u)}{\sqrt{n}} \geqslant h, \qquad \forall\ u \in A,$$
$$\frac{d_B(u) - d_A(u)}{\sqrt{n}} \geqslant h, \qquad \forall\ u \in B.$$

The definition of (A, B) being asymptotically h-friendly is similar. The scaling $\cdot/\sqrt{n}$ will be justified next. We establish the following result, which generalizes Theorem 1.

Theorem 2. *An asymptotically* h*-friendly partition of* $\mathbb{G}(n, 1/2)$ *exists w.h.p. if* $h < h^* \approx 0.175$ *and does not exist w.h.p. if* $h > h^*$.

The proof approach for Theorem 2 is unfortunately non-constructive. In fact we provide an evidence that a constructive approach for the case when h is strictly positive might not exist in the sense that polynomial time algorithms for finding h-friendly asymptotic partitions might not exist. We comment on in the next

section. For now we provide a brief overview of the proof approach. The proof is based on (a) establishing a similar result in a sparse random graph setting, namely for the graph $\mathbb{G}(n, c/n)$ where c is a sufficiently large constant, and (b) utilizing a powerful probabilistic invariance methods also known as Lindeberg's method which allows one to substitute one randomness (in our case symmetric Bernoulli with parameter 1/2) with another randomness (in our case biased Bernoulli with parameter c/n).

The proof of the existence of an asymptotically friendly bisections in sparse graphs $\mathbb{G}(n, c/n)$ (with an appropriate modification of the notion of the underlying asymptotics) is done following the methods in Gamarnik and Li [GL18]. There 2-dimensional large deviations method are used coupled with a configuration model of a random graph to argue the existence of an asymptotically friendly partition (though not called that way in the paper). The Lindeberg's method follows the lines used earlier for connecting optimization problems in sparse and dense random graphs, such as [DMS17],[Sen18],[Pan18],[CGPR19]. An interesting feature of our work is that all of the known to us prior papers on the invariance principle used random graphs coupled with the invariance method as a proof method in order to establish a result in sparse graphs, which is usually considerably harder. Namely, it followed the "dense-to-sparse" path as a proof technique. Somewhat surprisingly, in our setting the direction is the opposite "sparse-to-dense" one: while we did not succeed in proving the result for dense graphs $\mathbb{G}(n, 1/2)$ directly, we managed to prove it in sparse graphs (following the work already done in [GL18]) and then translate it to the case of dense graphs by the invariance. The advantage of working with sparse graphs, which is not readily available in the dense graph setting, is the existence of a configuration model which provides an important decoupling of the underlying probabilistic dependencies.

Algorithmic implications

We now briefly discuss potential algorithmic implications of our findings. As mentioned earlier the case $h = 0$ was addressed in Theorem 1 by Ferber et al [FKN$^+$21] constructively, specifically by building an algorithm which results in an asymptotically friendly bisection. This raises the question as to wether a similar algorithm exists for $h > 0$. While we do not provide a definitive answer to this question one way or the other, we provide below an indirect suggestion that fast (polynomial time) algorithms might not exist. This is done by studying the solution space geometry of friendly partitions and verifying in particular that the model exhibits an Overlap-Gap-Property (OGP) for some $0 < h < h^*$. The presence of OGP is a barrier to large classes of algorithms [Gam21] and in most studied models coincides with the onset of an apparent algorithmic hardness evidenced by failure to find a constructive solution. The theorem below states that the OGP indeed takes place in our setting for some h. For the purposes of giving a formal statement of the theorem it is convenient to think about bisections as functions $\sigma : [n] \to \{-1, 1\}$ with $A = h^{-1}(1)$ and $B = h^{-1}(-1)$. As a result σ is a bisection if $\sum_j \sigma(j) \in \{-1, 0, 1\}$,

and it is h-friendly if $(1/\sqrt{n})\sum_j J_{ij}\sigma_i\sigma_j \geqslant h$, where $J = (J_{ij}, i, j \in [n])$ is the adjacency matrix of the underlying graph.

Theorem 3 (Overlap Gap Property)**.** *There exists (numerically)* $0 < h_{\mathrm{OGP}} < h^*$ *such that* $\forall\ h \in (h_{\mathrm{OGP}}, h^*)\ \exists\ 0 < \nu_1 < \nu_2 < 1$ *with the following property: for every two* h*-friendly bisections* σ, τ*, it holds* $(1/n)d(\sigma,\tau) \in [0,\nu_1] \cup [\nu_2, 1]$*, w.h.p. as* $n \to \infty$*.*

Here $d(\cdot,\cdot)$ denotes the Hamming distance. Namely, every two friendly bisections have to be either "close" (at most distance $n\nu_1$) from each other or "far" (at least $n\nu_2$) distance from each other. The presence of this gap can be used to rule out classes of algorithms exhibiting some form of stability/noise insensitivity, as outlined in [Gam21]. Whether the algorithm constructed in [FKN$^+$21] exhibits such stability remains an interesting open question.

Admittedly, Theorem 3 provides some evidence of algorithmic non-existence only for some positive h, specifically for h which is at least the OGP threshold h_{OGP}. We do believe however that the problem might remain algorithmically hard for *all* strictly positive h. An evidence of this is found in the following theorem established by Behrens et al [BAKZ22], which we state informally.

Theorem 4 (Informally)**.** *OGP holds in random* r*-regular graphs* $\mathbb{G}_r(n)$ *for every* $h > 0$*.*

The proof uses the expander property of random regular graphs in a very important way and so it is not easily transferable to the Erdős-Rényi case $\mathbb{G}(n, c/n)$ or $\mathbb{G}(n, 1/2)$. The existence of such a translation and in general the presence of the OGP for all $h > 0$ is another interesting open question, which we leave for future research.

References

[BAKZ22] Freya Behrens, Gabriel Arpino, Yaroslav Kivva, and Lenka Zdeborová, *(dis) assortative partitions on random regular graphs*, arXiv preprint arXiv:2202.10379 (2022).

[CGPR19] Wei-Kuo Chen, David Gamarnik, Dmitry Panchenko, and Mustazee Rahman, *Suboptimality of local algorithms for a class of max-cut problems*, The Annals of Probability **47** (2019), no. 3, 1587–1618.

[DMS17] Amir Dembo, Andrea Montanari, and Subhabrata Sen, *Extremal cuts of sparse random graphs*, The Annals of Probability **45** (2017), no. 2, 1190–1217.

[FKN$^+$21] Asaf Ferber, Matthew Kwan, Bhargav Narayanan, Ashwin Sah, and Mehtaab Sawhney, *Friendly bisections of random graphs*, arXiv preprint arXiv:2105.13337 (2021).

[Für88] Z. Füredy, *Personal communication.*

[Gam21] David Gamarnik, *The overlap gap property: A topological barrier to optimizing over random structures*, Proceedings of the National Academy of Sciences **118** (2021), no. 41.

[GL18] David Gamarnik and Quan Li, *On the max-cut over sparse random graph*, Random Structures and Algorithms **52** (2018), no. 2, 219–262.

[Gre] Ben Green, *100 open problems*, Manuscript **1**.

[Pan18] Dmitry Panchenko, *On the k-sat model with large number of clauses*, Random Structures & Algorithms **52** (2018), no. 3, 536–542.

[Sen18] Subhabrata Sen, *Optimization on sparse random hypergraphs and spin glasses*, Random Structures & Algorithms **53** (2018), no. 3, 504–536.

A note on Pseudorandom Ramsey graphs

Jacques Verstraëte
(joint work with Dhruv Mubayi)

The *Ramsey number* $r(F,t)$ is the minimum N such that every F-free graph on N vertices has an independent set of size t. When $F = K_s$ we simply write $r(s,t)$ instead of $r(F,t)$. The current best available bounds for $r(s,t)$ as as follows: for $s \geqslant 3$, there exist constants $c_1(s), c_2(s) > 0$ such that

$$c_1(s)\frac{t^{\frac{s+1}{2}}}{(\log t)^{\frac{s+1}{2}-\frac{1}{s-2}}} \leqslant r(s,t) \leqslant c_2(s)\frac{t^{s-1}}{(\log t)^{s-2}}. \tag{1}$$

The upper bound is due to Ajtai, Komlós and Szemerédi [1], and the lower bound is due to Bohman and Keevash [7], using the *random K_s-free process*. For $s = 3$, the lower bound was proved in a celebrated paper of Kim [13] and the upper bound was proved by Shearer [19] with $c_2(3) = 1 + o(1)$. In particular, recent results of Bohman and Keevash [6] and Fiz Pontiveros, Griffiths and Morris [12] together with the bound of Shearer show as $t \to \infty$:

$$(\tfrac{1}{4} - o(1)) \cdot \frac{t^2}{\log t} \leqslant r(3,t) \leqslant (1 + o(1)) \cdot \frac{t^2}{\log t}. \tag{2}$$

In this note, we show that if certain density-optimal K_s-free pseudorandom graphs exist, then $r(s,t) = t^{s-1+o(1)}$. An (n,d,λ) *graph* is an n-vertex d-regular graph such that the absolute value of every eigenvalue of its adjacency matrix, besides the largest one, is at most λ. We refer the reader to Krivelevich and Sudakov [15] for a survey of (n,d,λ)-graphs. Sudakov, Szabo and Vu [22] show that a K_s-free (n,d,λ)-graph satisfies

$$\lambda = \Omega(d^{s-1}/n^{s-2}) \tag{3}$$

as $n \to \infty$. For $s = 3$, if G is any triangle-free (n,d,λ)-graph with adjacency matrix A, then

$$0 = \mathrm{tr}(A^3) \geqslant d^3 - \lambda^3(n-1). \tag{4}$$

If $\lambda = O(\sqrt{d})$, then this gives $d = O(n^{2/3})$ matching (3). Alon [2] constructed a triangle-free pseudorandom graph attaining this bound, and Conlon [9] more recently analyzed a randomized construction with the same average degree. The Alon-Boppana Bound [17, 18] shows that $\lambda = \Omega(\sqrt{d})$ for every (n,d,λ)-graph provided d/n is bounded away from 1. Sudakov, Szabo and Vu [22] raised the question of the existence of optimal pseudorandom K_s-free graphs for $s \geqslant 4$, namely (n,d,λ)-graphs achieving the bound in (3) with $\lambda = O(\sqrt{d})$ and $d = \Omega(n^{1-\frac{1}{2s-3}})$. We show that a positive answer to this question gives the exponent of the Ramsey numbers $r(s,t)$ via a short proof of the following general theorem, based on ideas of Alon and Rödl [4]:

Theorem 1. *Let d, n, t be positive integers and $t = \lceil 2n \log^2 n/d \rceil$. If there exists an F-free (n, d, λ)-graph and n is large enough, then*

$$r(F, t) = \Omega\left(\frac{n}{\lambda} \log^2 n\right). \tag{5}$$

In particular, if K_s-free (n, d, λ)-graphs exist with $d = \Omega(n^{1-\frac{1}{2s-3}})$ and $\lambda = O(\sqrt{d})$, then

$$r(s, t) = \Omega\Big(\frac{t^{s-1}}{(\log t)^{2s-4}}\Big). \tag{6}$$

Alon and Krivelevich [3] gave a construction of K_s-free (n, d, λ)-graphs with $d = \Omega(n^{1-1/(s-2)})$ and $\lambda = O(\sqrt{d})$ for all $s \geqslant 3$, and this was slightly improved by Bishnoi, Ihringer and Pepe [5] to obtain $d = \Omega(n^{1-1/(s-1)})$. This is the current record for the density of a K_s-free (n, d, λ)-graph with $\lambda = O(\sqrt{d})$. The problem of obtaining optimal K_s-free pseudorandom constructions in the sense (3) with $\lambda = O(\sqrt{d})$ for $s \geqslant 4$ seems difficult and is considered to be a central open problem in pseudorandom graph theory. The problem of determining the growth rate of $r(s, t)$ is classical and much older, and it wasn't completely clear whether the upper or lower bound in (1) was closer to the truth. Based on Theorem 1, it seems reasonable to conjecture that if $s \geqslant 4$ is fixed, then $r(s, t) = t^{s-1+o(1)}$ as $t \to \infty$.

Applying Theorem 1 when F is bipartite can give lower bounds on $r(F, t)$ that are better than those obtained from the F-free process. It is a wide open conjecture of Erdős that $r(C_4, t) \leqslant t^{2-\epsilon}$ for some $\epsilon > 0$, and the cycle complete Ramsey numbers $r(C_k, t)$ appear to be very difficult to determine – the best upper bounds are provided by Sudakov [21] for odd cycles and Caro, Li, Rousseau and Zhang [8] for even cycles. The best lower bound for fixed $\ell \geqslant 4$ is

$$r(C_\ell, t) = \Omega\left(\frac{t^{(\ell-1)/(\ell-2)}}{\log t}\right) \tag{7}$$

due to Bohman and Keevash [7] by analyzing the C_ℓ-free process. Theorem 1 gives $r(C_4, t) = \Omega(t^{3/2}/\log t)$ which matches (7), as well as results that exceed the previous best known bounds of (7) from the random C_ℓ-free process for certain values of ℓ:

Theorem 2. *There exists a constant $c > 0$ such that as $t \to \infty$,*

$$\begin{aligned} r(C_6, t) &\geqslant ct^{5/4}/\log^{1/2} t \\ r(C_{10}, t) &\geqslant ct^{9/8}/\log^{1/4} t \\ r(C_5, t) &\geqslant (1 - o(1))t^{11/8} \\ r(C_7, t) &\geqslant (1 - o(1))t^{11/9}. \end{aligned}$$

The last two statements appear to be the first instances of a graph F containing a cycle for which random graphs do not supply the right exponent for $r(F, t)$ as $t \to \infty$.

References

[1] Ajtai, Miklós; Komlós, János; Szemerédi, Endre A note on Ramsey numbers. J. Combin. Theory Ser. A 29 (1980), no. 3, 354–360.

[2] Alon, Noga Explicit Ramsey graphs and orthonormal labelings. Electron. J. Combin. 1 (1994), Research Paper 12, approx. 8 pp.

[3] Alon, Noga; Krivelevich, Michael(IL-TLAV) Constructive bounds for a Ramsey-type problem. Graphs Combin. 13 (1997), no. 3, 217–225.

[4] Alon, Noga; Rödl, Vojtěch Sharp bounds for some multicolor Ramsey numbers. (English summary) Combinatorica 25 (2005), no. 2, 125–141.

[5] Bishnoi, Anurag; Ihringer, Ferdinand; Pepe, Valentina A construction for clique-free pseudorandom graphs https://arxiv.org/abs/1905.04677

[6] Bohman, Tom; Keevash, Peter Dynamic concentration of the triangle-free process. The Seventh European Conference on Combinatorics, Graph Theory and Applications, 489–495, CRM Series, 16, Ed. Norm., Pisa, 2013.

[7] Bohman, Tom; Keevash, Peter The early evolution of the H-free process. Invent. Math. 181 (2010), no. 2, 291–336.

[8] Caro, Yair; Li, Yusheng; Rousseau, Cecil C; Zhang, Yuming Asymptotic bounds for some bipartite graph: complete graph Ramsey numbers. Discrete Math. 220 (2000), no. 1-3, 51–56.

[9] Conlon, David A sequence of triangle-free pseudorandom graphs. Combin. Probab. Comput. 26 (2017), no. 2, 195–200.

[10] Dudek, Andrzej; Retter, Troy; Rödl, Vojtěch On generalized Ramsey numbers of Erdős and Rogers. J. Combin. Theory Ser. B 109 (2014), 213–227.

[11] Erdős, P.; Szekeres, G. A combinatorial problem in geometry. Compositio Math. 2 (1935), 463–470.

[12] Fiz Pontiveros, Gonzalo; Griffiths, Simon; Morris, Robert The triangle-free process and the Ramsey number $R(3, k)$ Mem. Amer. Math. Soc., to appear.

[13] Kim, Jeong Han The Ramsey number R(3,t) has order of magnitude $t^2/\log t$. Random Structures Algorithms 7 (1995), no. 3, 173–207.

[14] Kostochka, Alexandr; Mubayi, Dhruv; Verstraëte, Jacques Hypergraph Ramsey numbers: triangles versus cliques. J. Combin. Theory Ser. A 120 (2013), no. 7, 1491–1507.

[15] Krivelevich, Michael; Sudakov, Benny Pseudo-random graphs. More sets, graphs and numbers, 199–262, Bolyai Soc. Math. Stud., 15, Springer, Berlin, 2006.

[16] Mubayi, Dhruv; Williford, Jason On the independence number of the Erdős-Rényi and projective norm graphs and a related hypergraph. J. Graph Theory 56 (2007), no. 2, 113–127.

[17] Nilli, A. On the second eigenvalue of a graph. Discrete Math. 91 (1991), no. 2, 207–210.

[18] Nilli, A. Tight estimates for eigenvalues of regular graphs. Electron. J. Combin. 11 (2004), no. 1, Note 9, 4 pp.

[19] Shearer, James Lower bounds for small diagonal Ramsey numbers. J. Combin. Theory Ser. A 42 (1986), no. 2, 302–304.

[20] Spencer, Joel Asymptotic lower bounds for Ramsey functions. Discrete Math. 20 (1977/78), no. 1, 69–76.

[21] Sudakov, Benny A note on odd cycle-complete graph Ramsey numbers. Electron. J. Combin. 9 (2002), no. 1, Note 1, 4 pp.

[22] Sudakov, Benny; Szaó, Tibor; Vu, Van H. A generalization of Turán's theorem. J. Graph Theory 49 (2005), no. 3, 187–195.

Improved Elekes–Szabó type estimates using proximity

József Solymosi
(joint work with Josh Zahl)

The Schwartz–Zippel lemma controls the size of the intersection of a Cartesian product and the zero-locus of a polynomial:

Theorem 1 (Schwartz–Zippel)**.** *Let F be a field, let $A_1, \dots, A_k$ be subsets of F of size N, and let f be a non-zero k-variate polynomial with coefficients in F. Then*

$$|(A_1 \times \dots \times A_k) \cap Z(f)| \leqslant (\deg f) N^{k-1}. \tag{1}$$

The bound (1) can be tight, for example, if $Z(f)$ is a union of parallel, axis-parallel hyperplanes. Motivated by questions in combinatorial geometry, Elekes and Szabó [11] investigated situations where Inequality (1) can be strengthened. They were interested in the situation where k and $\deg f$ are fixed, and N is large.

Definition 2. *Let F be an infinite field. We say a k-variate polynomial f with coefficients in F has* Schwartz–Zippel power saving *if there are constants $C, c > 0$ so that for all $N \geqslant 1$ and all subsets $A_1, \dots, A_k$ of F of size N, we have*

$$|(A_1 \times \dots \times A_k) \cap Z(f)| \leqslant C N^{d-c}, \quad d = \dim Z(f). \tag{2}$$

If $Z(f)$ is reducible, then f has Schwartz–Zippel power saving if and only if all of the maximal-dimension irreducible components of $Z(f)$ have Schwartz–Zippel power saving. Thus it makes sense to consider the case where f is irreducible. When $k = 2$, no polynomials have Schwartz–Zippel power saving. When $k = 3$, however, the situation is quite different. The following result of Elekes and Szabó [2, Theorem 3] shows that an irreducible polynomial $f \in \mathbb{C}[x, y, z]$ either must have Schwartz-Zippel power saving, or it must have a special structure.

Theorem 3 (Elekes–Szabó)**.** *Let $f \in \mathbb{C}[x, y, z]$ be irreducible. Then at least one of the following is true.*

(A) *f has Schwartz–Zippel power saving.*
(B) *After possibly permuting the coordinates x, y, z, we have $f(x, y, z) = g(x, y)$, for some bivariate polynomial g.*
(C) *f encodes additive group structure.*

If either of Items (B) or (C) hold, then Item (A) does not; in fact, for every N, there exist sets $A, B, C \subset \mathbb{C}$ of size N with

$$|(A \times B \times C) \cap Z(f)| \geqslant (N-2)^2/8.$$

Item (A) has already been defined, and item (B) is self-explanatory; geometrically, it says that $Z(f)$ is an axis-parallel cylinder above a curve.

We say f encodes additive group structure if for a generic point $p \in \mathbb{C}^3$, there is a (Euclidean) neighborhood U of p, a set $V \subset \mathbb{C}$, and analytic functions $\phi\colon U \to \mathbb{C}^3$ and $\psi\colon V \to \mathbb{C}$, so that $\psi \circ f \circ \phi(x, y, z) = x + y - z$. When $f(x, y, z)$ is of the special form $h(x, y) - z$, then the situation is particularly simple: f encodes additive structure of and only if h has the form $h(x, y) = a(b(x) + c(y))$ or $h(x, y) =$

$a(b(x)c(y))$ for univariate polynomials a, b, c. This special case was analyzed in an earlier work by Elekes and Rónyai [1].

We will restrict attention to three dimensions, and characteristic 0. We will be interested in quantitative versions of Theorem 3, and specifically, we wish to obtain explicit lower bounds on the size of the Schwartz–Zippel power saving. In this direction, Raz, Sharir, and de Zeeuw [5] strengthened Theorem 3 by establishing the explicit power saving $c = 1/6$ for Item (A). The proof in [5] generalized several related arguments that had been previously used to obtain the same power savings in certain special cases. In the other direction, Makhul, Roche-Newton, Warren, and de Zeeuw [3] obtained an upper bound on the size of the Schwartz–Zippel power saving by showing that the polynomial $f(x, y, z) = (x-y)^2 + x - z$ only has Schwartz–Zippel power saving $c = 1/2$. Our main result is a version of Theorem 3 in three dimensions for Cartesian products of real numbers, with power saving $c = 2/7$. In what follows, we identify points $x \in \mathbb{R}$ with the corresponding point $x + 0i \in \mathbb{C}$.

Theorem 4. *Let $f \in \mathbb{C}[x, y, z]$ be irreducible. Then at least one of the following is true.*

(A) *For all finite sets $A, B, C \subset \mathbb{R}$ with $|A| \leqslant |B| \leqslant |C|$, we have*

$$|(A \times B \times C) \cap Z(f)| \lesssim (|A||B||C|)^{4/7} + |B||C|^{1/2}, \tag{3}$$

where the implicit constant depends on the degree of f.

(B) *After possibly permuting the coordinates x, y, z, we have $f(x, y, z) = g(x, y)$, for some bivariate polynomial g.*

(C) *f encodes additive group structure.*

Specializing to the case $f(x, y, z) = h(x, y) - z$, we record the following corollary.

Corollary 5. *Let $h \in \mathbb{C}[x, y]$. Then exactly one of the following holds.*

(A) *For all finite sets $A, B \subset \mathbb{R}$ with $|A| \leqslant |B|$, we have*

$$|h(A \times B)| \gtrsim \min\left(|A|^{3/4}|B|^{3/4},\ |A|^2\right). \tag{4}$$

(B) *h has the special form $h(x, y) = a(b(x) + c(y))$ or $h(x, y) = a(b(x)c(y))$ for univatiate polynomials a, b, c.*

References

[1] G. Elekes and L. Rónyai. A combinatorial problem on polynomials and rational functions. *J. Combinat. Theory Ser. A*. 89:1–20, 2000.

[2] G. Elekes and E. Szabó. How to find groups? (and how to use them in Erdős geometry?) *Combinatorica*, 32(5):537–571, 2012.

[3] M. Makhul, O. Roche-Newton, A. Warren and F. de Zeeuw. Constructions for the Elekes-Szabó and Elekes-Rónyai problems. *Electron. J. Combin.* 27(1), Paper 1.57, 2020.

[4] O. E. Raz, M. Sharir, and J. Solymosi. Polynomials vanishing on grids: The Elekes-Rónyai problem revisited. *Amer. J. Math.*, 138: 1029–1065, 2016.

[5] O. E. Raz, M. Sharir, and F. de Zeeuw. Polynomials vanishing on Cartesian products: The Elekes-Szabó theorem revisited. *Duke Math. J.* 165(18):3517–3566, 2016.

Zeros and computational complexity of combinatorial polynomials

Alexander Barvinok

This a survey talk on combinatorial applications of the following simple result from complex analysis. In [1], I call it "the interpolation method", although a more suitable name would perhaps be "the extrapolation method".

Lemma 1. *Let $U \subset \mathbb{C}$ be a connected open set containing 0 and 1. Then there is a constant $\gamma = \gamma(U) > 0$ such that the following holds. Let*

$$g(z) = \sum_{k=0}^{n} c_k z^k, \quad n \geqslant 2,$$

be a polynomial such that $g(z) \neq 0$ for all $z \in U$. Then up to relative error $0 < \epsilon < 1$, the value of $g(1)$ is determined by the coefficients c_k with $k \leqslant \gamma(\ln n - \ln \epsilon)$ and can be computed from those coefficients in polynomial time.

Here we say that complex numbers $w_1, w_2 \neq 0$ approximate each other within relative error $\epsilon > 0$ if we can write $w_1 = e^{z_1}$ and $w_2 = e^{z_2}$ with $|z_1 - z_2| \leqslant \epsilon$.

Sketch of proof. First, we consider a special case when

$$U = \mathbb{D}_\beta = \{z : \ |z| < \beta\} \quad \text{for some} \quad \beta > 1$$

is a disc. Let $f(z) = \ln g(z)$ and let

$$T_m(z) = f(0) + \sum_{k=1}^{m} \frac{f^{(k)}(0)}{k!} z^k$$

be the Taylor polynomial of $f(z)$. It is not hard to show that

$$|f(z) - T_m(z)| \ \leqslant \ \frac{n}{\beta^m(\beta-1)(m+1)} \quad \text{provided} \quad |z| \leqslant 1$$

and that the values of $f^{(k)}(0)$ can be computed from c_k for $k = 1, \ldots, m$ in polynomial time. To ensure that $T_m(1)$ approximates $f(1)$ within additive error ϵ and hence $e^{T_m(1)}$ approximates $g(1)$ within relative error ϵ, it suffices to take $m = O_\beta(\ln n - \ln \epsilon)$. For a general set U, we construct an auxiliary disc $\mathbb{D}_\beta$ for some $\beta > 1$ and a polynomial map $\phi : \mathbb{D}_\beta \longrightarrow U$ such that $\phi(0) = 0$ and $\phi(1) = 1$ and apply the analysis of the special case to the composition $g(\phi(z))$, see Section 2.2 of [1] for details. □

As follows from the proof, the constant $\gamma = \gamma(U)$ in Lemma 1 is

$$\gamma(U) = O\left(\frac{1}{\beta - 1}\right)$$

where $1 < \beta < 2$ is the radius of the disc $\mathbb{D}_\beta$ for which there is a holomorphic map $\phi : \mathbb{D}_\beta \longrightarrow U$ such that $\phi(0) = 0$ and $\phi(1) = 1$ (the map ϕ exists because of the Riemann Conformal Mapping Theorem). One interesting case is when

$$U = \mathbb{C} \setminus \{z \in \mathbb{R} : \ z < -\delta\}$$

for some $0 < \delta < 1$. In this case, one can choose

$$\beta = 1 + \Omega\left(\frac{1}{\sqrt{\delta}}\right) \quad \text{and} \quad \phi(z) = \frac{\rho}{(1-\xi z)^2} - \rho \quad \text{where} \quad \xi = 1 - \sqrt{\frac{\rho}{1+\rho}}$$

for some $\rho = \rho(\delta)$. Then, from the Heilmann–Lieb Theorem [7], we obtain the following corollary.

Theorem 2. *For a graph $G = (V, E)$, let $m_k = m_k(G)$ be the number of matchings $M \subset E$ with k edges. Then, up to relative error $0 < \epsilon < 1$, the total number*

$$\sum_{k=0}^{|V|/2} m_k$$

of matchings in G, is determined by the numbers m_k with

$$k \leqslant \gamma\sqrt{\Delta}\ln\frac{|V|}{\epsilon}$$

and can be computed from those numbers in polynomial time. Here $\Delta = \Delta(G)$ is the largest degree of a vertex in G and $\gamma > 0$ is an absolute constant.

As similar result is obtained for the number of independent sets in claw-free graphs, if one uses the Chudnovsky–Seymour Theorem [6], see [2] for details. The complexity of the resulting deterministic algorithm for counting matchings in graphs roughly matches that of the algorithm of Bayati et al. [4], which is based on correlation decay. See also [8] for counting matchings and other subgraphs of logarithmic size in bounded degree graphs in polynomial (as opposed to quasi-polynomial) time.

The bottleneck of the method is in establishing zero-free regions. Here is a result pertaining to permanents of complex matrices.

Theorem 3. *Let $A = (a_{ij})$ be an $n \times n$ complex matrix such that*

$$|1 - a_{ij}| \ \leqslant\ 0.5 \quad \textit{for all} \quad i, j.$$

Then $\operatorname{per} A \neq 0$.

Sketch of proof. Using the Laplace row expansion for the permanent, we prove by induction on n the following stronger statement: If A and B are $n \times n$ complex matrices satisfying the condition of the theorem that differ in at most one row, then $\operatorname{per} A \neq 0$, $\operatorname{per} B \neq 0$ and the angle between $\operatorname{per} A$ and $\operatorname{per} B$, considered as vectors in $\mathbb{R}^2 = \mathbb{C}$, does not exceed $\pi/2$, see Section 3.6 of [1] for details. □

It follows that the permanent of $n \times n$ complex matrix $A = (a_{ij})$ can be approximated within relative error $0 < \epsilon < 1$ in quasi-polynomial $n^{O(\ln n - \ln \epsilon)}$ time provided $|1 - a_{ij}| \leqslant 0.49$ for all i, j. Similar results can be obtained for other classes of permanents.

For example, we can allow greater variance of the real part of a_{ij}, provided we restrict the imaginary part:

If

$$\delta \ \leqslant\ \Re\, a_{ij} \ \leqslant\ 1 \quad \text{and} \quad |\Im\, a_{ij}| \ \leqslant \frac{1}{2}\delta^2 \quad \text{for some} \quad 0 < \delta < 1,$$

then $\operatorname{per} A \neq 0$. Consequently $\operatorname{per} A$ can be approximated within relative error $0 < \epsilon < 1$ in quasi-polynomial $n^{O(\ln n - \ln \epsilon)}$ time provided a_{ij} are real and satisfy $\delta \leqslant a_{ij} \leqslant 1$ for some $\delta > 0$, fixed in advance, see Section 3.7 of [1].

Other examples of matrices amenable to this method include diagonally dominant matrices [3] and also matrices where the ℓ^1-distance of each row to the row of 1's does not exceed γn for some absolute constant $\gamma > 0$, see Section 5.5 of [1].

It is not clear whether the constant 0.5 in Theorem 3 can be improved. We have

$$\operatorname{per} B = 0 \quad \text{where} \quad B = \begin{pmatrix} \frac{1+i}{2} & \frac{1-i}{2} \\ \frac{1-i}{2} & \frac{1+i}{2} \end{pmatrix} = 0,$$

and hence 0.5 cannot be replaced by $\sqrt{2}/2 \approx 0.71$. Furthermore, Boris Bukh noticed [5] that

$$\operatorname{per}(B \otimes J_{2m+1}) = 0,$$

where J_{2m+1} is the $(2m+1) \times (2m+1)$ matrix filled by 1's and hence there are complex matrices $A = (a_{ij})$ of an arbitrarily large size such that $\operatorname{per} A = 0$ and $|1 - a_{ij}| = \sqrt{2}/2$ for all i, j.

References

[1] A. Barvinok, *Combinatorics and Complexity of Partition Functions*, Algorithms and Combinatorics, **30**, Springer, Cham, 2016.

[2] A. Barvinok, Approximating real-rooted and stable polynomials, with combinatorial applications, *Online J. Anal. Comb.* No. **14** (2019), Paper No. 8, 13 pp.

[3] A. Barvinok, Computing permanents of complex diagonally dominant matrices and tensors, *Israel J. Math.* **232** (2019), no. 2, 931–945.

[4] M. Bayati, D. Gamarnik, D. Katz, C. Nair, and P. Tetali, Simple deterministic approximation algorithms for counting matchings, in *STOC'07–Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, 122–127, ACM, New York, 2007.

[5] B. Bukh, *personal communication*, 2015.

[6] M. Chudnovsky and P. Seymour, The roots of the independence polynomial of a clawfree graph, *J. Combin. Theory Ser. B* **97** (2007), no. 3, 350–357.

[7] O.J. Heilmann and E.H. Lieb, Theory of monomer-dimer systems, *Comm. Math. Phys.* **25** (1972), 190–232.

[8] V. Patel and G. Regts, Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials, *SIAM J. Comput.* **46** (2017), no. 6, 1893–1919.

The existence of subspace designs

Ashwin Sah

(joint work with Peter Keevash and Mehtaab Sawhney)

A widely circulated problem in the 1970s asked for vector space analogues of combinatorial designs, whereby combinatorial designs could be considered as designs in vector spaces over the 'field with one element'. This problem arose during an exciting time in the history of combinatorial designs, when Wilson [14] proved the graph case of the Existence Conjecture (a problem posed by Steiner in the 19th century, eventually resolved by Keevash [8]). In an early article on the general

algebraic problem, Cameron [3] gave his 'commentary' on the combined efforts of many researchers, including Petrenjuk, Wilson, Ray-Chaudhuri, Noda, Bannai, Delsarte, Goethals, and Seidel. Cameron remarked that subspace 1-designs (spreads) are 'common', but there were no known non-trivial subspace t-designs with $t > 1$.

This problem has recently seen considerable progress, following a renewed interest due to its connections with Network Coding and advances in techniques, including computational methods for finding explicit examples and probabilistic methods for obtaining general results. To discuss progress on the problem to date we require the following definitions. Let $\mathbb{F}_q$ be a finite field of order q. Let $\mathrm{Gr}_q(n, r)$ denote the set of r-dimensional subspaces ('r-spaces') of the n-dimensional vector space $\mathbb{F}_q^n$. An $(n, s, r, \lambda)_q$-design consists of a subset of $\mathrm{Gr}_q(n, s)$, called blocks, such that each r-space is contained in exactly λ blocks. This definition captures the established meaning of 'subspace design' in Combinatorics and in Network Coding, although we remark that there is also a large literature in Theoretical Computer Science on a similar but weaker notion of 'subspace design' (replace 'exactly' by 'at most') introduced by Guruswami and Xing [6].

There are some parallels between the histories of subspace designs and combinatorial designs. Indeed, for combinatorial designs it was a longstanding open problem, resolved by Teirlinck [12], to show the existence of non-trivial (n, s, r, λ)-designs for all r and some λ (where 'non-trivial' means that $s > r$ and not all s-sets are blocks). Similarly, the existence of non-trivial $(n, s, r, \lambda)_q$-designs for all r and some λ was a longstanding open problem, resolved much more recently by Fazeli, Lovett, and Vardy [4]. This general result was preceded by various explicit constructions; for details of these we refer to the survey by Braun, Kiermaier, and Wassermann [2]. While Teirlinck used an explicit construction, the construction in [4] is probabilistic (adapting a method of Kuperberg, Lovett, and Peled [9]), and requires $\lambda \geqslant q^{Crn}$.

The parallels continue for Steiner systems, where for many years after Teirlinck's result the existence of (n, s, r)-designs with $s > r \geqslant 3$ was only known in sporadic cases, and the existence of any examples for $r \geqslant 6$ was unknown until the general result of [8]. The situation for $(n, s, r)_q$-designs was even more dire, and was highlighted by Kalai [7] as one of the most important open problems remaining in Design Theory. It was conjectured by Metsch [10] that no such designs with $s > r > 1$ exist. This was recently disproved by Braun, Etzion, Östergård, Vardy, and Wassermann [1], who developed improved computational methods to find $(13, 3, 2)_2$-designs. However, there were no known examples for any other parameters, let alone any general results.

We remedy this situation by completely answering the question: we show the existence of $(n, s, r)_q$-designs for any prime power q and $s > r \geqslant 1$. Moreover, our result is analogous to Keevash's, in that we show the existence of $(n, s, r)_q$-designs for all sufficiently large n satisfying the necessary 'divisibility conditions'. Here recall the Gaussian q-binomial $\begin{bmatrix} n \\ k \end{bmatrix}_q$, the number of $\mathbb{F}_q$-subspaces of $\mathbb{F}_q^n$ of dimension

k, also given by the formula

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{(q^n-1)\cdots(q^1-1)}{(q^k-1)\cdots(q^1-1)(q^{n-k}-1)\cdots(q^1-1)}.$$

Definition 1. *Let q be a prime power and let $\mathbb{F}_q^n$ be the n-dimensional vector space over $\mathbb{F}_q$. For $s > r$ and $\lambda \geqslant 1$, an $(n,s,r,\lambda)_q$-design is a multicollection $\mathcal{S}$ of s-dimensional subspaces such that every r-dimensional subspace is contained in exactly 1 space in $\mathcal{S}$. We say it is* simple *if there are no repeated s-spaces.*

Theorem 2. *Fix q,s,r. For $n \geqslant n_2(q,s)$ such that $\begin{bmatrix} s-i \\ r-i \end{bmatrix}_q \mid \begin{bmatrix} n-i \\ r-i \end{bmatrix}_q$ for all $0 \leqslant i \leqslant r-1$ there is an $(n,s,r)_q$-design.*

Additionally, one can prove an analogue for "sufficiently pseudorandom" collections of r-dimensional subspaces, similar in spirit to [8, Theorem 1.10] (with certain q-analogues of pseudorandomness conditions).

We also prove a counting version as a simple corollary of the proof.

Corollary 3. *Under the assumptions of Theorem 2, for $n \geqslant n_3(q,s)$ the number of $(n,s,r)_q$-designs is*

$$\left((1 \pm q^{-c_3(r,s)n})\frac{\begin{bmatrix} n-r \\ s-r \end{bmatrix}_q}{\exp(\begin{bmatrix} s \\ r \end{bmatrix}_q - 1)}\right)^{\begin{bmatrix} n \\ r \end{bmatrix}_q / \begin{bmatrix} s \\ r \end{bmatrix}_q}.$$

The situation when $\lambda > 1$ is very similar, with a few added considerations regarding simplicity and the approximate covering step.

Theorem 4. *Fix q,s,r,λ. For $n \geqslant n_4(q,s,\lambda)$ such that $\begin{bmatrix} s-i \\ r-i \end{bmatrix}_q \mid \lambda\begin{bmatrix} n-i \\ r-i \end{bmatrix}_q$ for all $0 \leqslant i \leqslant r-1$ there is a simple $(n,s,r,\lambda)_q$-design.*

1. New techniques: absorption in rigid algebraic scenarios

Classic methods such as the Rödl nibble for hypergraph matchings or more recent results can easily be seen to give an "approximate" version, i.e., a collection of s-dimensional spaces which cover $1-o(1)$ fraction of the r-spaces exactly once, and the remainder is uncovered. Therefore, as is typical, the key issue is dealing with the remainder. The most general form of this is the idea of *absorption*. One sets aside some structure before attempting to solve a decomposition problem. Then, after approximately decomposing everything else, the remainder is small enough so that it can be handled in conjunction with the absorbing structure (akin to a sponge absorbing water).

For the problem of constructing Steiner systems and other designs, traditional "random" absorbers are not sufficient due to the sparsity of usable "local switches" to work with in such structures. The work of Keevash [8] introduced a powerful idea of *randomized algebraic constructions* to use as templates to construct Steiner systems in general.

In our situation, algebraic structure is already inherently present and hence we are more restricted in various ways. One can still perform approximate decompositions via the Rödl nibble (or the more modern technique of random removal

processes), since this can be seen purely from a hypergraph matching perspective, and we develop a framework for working with notions of embedding, pseudorandomness, and "typicality" in q-analogues of hypergraphs, which we call *q-systems*.

The template, a special set of r-dimensional subspaces coming from a collection of s-spaces, is formed via a randomized algebraic process as in [8], but we are focused on making entire vector spaces play nice with respect to each other. Furthermore, one must ensure the template is sufficiently generic to work with and the necessary algebraic constructions may not exist over $\mathbb{F}_q$ directly. Thus, we must pass to a field extension $L/\mathbb{F}_q$ and put an L-space structure on $\mathbb{F}_q^n$. In cases where n is not divisible by any small number, this is not directly possible and we embed multiple incompatible L-structures on vector spaces of finite codimension.

For the absorption process, after creating an approximate decomposition the remainder (or *leave*) is covered using a "bounded signed integral decomposition" by understanding certain associated lattices, and then using a "subspace exchange process" to massage this integral decomposition into a form amenable to absorption using the template structure. The latter bears similarity to the "clique exchange" of [8], although the q-analogue and multiple $L/\mathbb{F}_q$-structures pose various new technical difficulties.

However, the integral decomposition is significantly hampered by the rigidity of the subspace setting. For Steiner systems, the key associated lattice is defined by relatively simple divisibility conditions (due to work of Graver and Jurkat [5] and of Wilson [13]) and it in fact has a particularly natural "bounded" generating set to work with, formed by certain "octahedral" structures (see e.g. [8, Section 5]). However, work of Ray-Chaudhuri and Singhi [11] shows that lattices associated to $(n, s, r)_q$-designs are not nearly so nice. As a result, we work with a greedily designed bounded approximate generating set, and introduce a way to boost this approximate behavior by using multiple copies to "cover gaps".

References

[1] Michael Braun, Tuvi Etzion, Patric R. J. Östergård, Alexander Vardy, and Alfred Wassermann, *Existence of q-analogs of Steiner systems*, Forum Math. Pi **4** (2016), e7, 14.

[2] Michael Braun, Michael Kiermaier, and Alfred Wassermann, *q-analogs of designs: subspace designs*, Network coding and subspace designs, Signals Commun. Technol., Springer, Cham, 2018, pp. 171–211.

[3] Peter J. Cameron, *Generalisation of Fisher's inequality to fields with more than one element*, Combinatorics (Proc. British Combinatorial Conf., Univ. Coll. Wales, Aberystwyth, 1973), London Math. Soc. Lecture Note Ser., No. 13, Cambridge Univ. Press, London, 1974, pp. 9–13.

[4] Arman Fazeli, Shachar Lovett, and Alexander Vardy, *Nontrivial t-designs over finite fields exist for all t*, J. Combin. Theory Ser. A **127** (2014), 149–160.

[5] J. E. Graver and W. B. Jurkat, *The module structure of integral designs*, J. Combinatorial Theory Ser. A **15** (1973), 75–90.

[6] Venkatesan Guruswami and Chaoping Xing, *List decoding Reed-Solomon, algebraic-geometric, and Gabidulin subcodes up to the Singleton bound*, STOC'13—Proceedings of the 2013 ACM Symposium on Theory of Computing, ACM, New York, 2013, pp. 843–852.

[7] Gil Kalai, *Designs exist! [after Peter Keevash]*, Astérisque (2016), Exp. No. 1100, 399–422.

[8] Peter Keevash, *The existence of designs*, arXiv:1401.3665.

[9] Greg Kuperberg, Shachar Lovett, and Ron Peled, *Probabilistic existence of regular combinatorial structures*, Geom. Funct. Anal. **27** (2017), 919–972.
[10] Klaus Metsch, *Bose-Burton type theorems for finite projective, affine and polar spaces*, Surveys in combinatorics, 1999 (Canterbury), London Math. Soc. Lecture Note Ser., vol. 267, Cambridge Univ. Press, Cambridge, 1999, pp. 137–166.
[11] D. K. Ray-Chaudhuri and N. M. Singhi, *q-analogues of t-designs and their existence*, Linear Algebra Appl. **114/115** (1989), 57–68.
[12] Luc Teirlinck, *Nontrivial t-designs without repeated blocks exist for all t*, Discrete Math. **65** (1987), 301–311.
[13] Richard M. Wilson, *The necessary conditions for t-designs are sufficient for something*, Utilitas Math. **4** (1973), 207–215.
[14] Richard M. Wilson, *An existence theory for pairwise balanced designs. III. Proof of the existence conjectures*, J. Combinatorial Theory Ser. A **18** (1975), 71–79.

Generating random d-regular graphs quickly: reciprocal rejection sampling

Oliver Riordan
(joint work with Nick Wormald)

Probably the third most natural random graph model, after the binomial and uniform random graphs $G(n,p)$ and $G(n,m)$, is the *random d-regular graph* $G_{n,d}$, which is simply a d-regular graph on a given set $[n]$ of n vertices chosen uniformly at random from the set of all such graphs. In contrast to the other two models mentioned, the task of generating a sample from this model is highly non-trivial, especially when d grows as a function of n. Our aim here is to describe a simple new idea, *reciprocal rejection sampling*, that is one ingredient in efficient generation of random d-regular graphs, allowing (together with many previous ideas and one further twist) rapid generation when $d = o(\sqrt{n})$. Concretely, we obtain a relatively simple algorithm that is linear time in the size $O(dn)$ of the output, improving on the previous bound of $O(dn + d^4)$ obtained by Arman, Gao and Wormald [1].

As in previous work on this problem, the basic strategy is to first generate a random d-regular *multi*-graph $G^*_{n,d}$ according to the configuration model of Bollobás [2]. This has the property that, conditional on the numbers of loops, double-edges, triple-edges and so on, it is uniformly distributed on the set of all d-regular multigraphs with these parameters. So, as in previous work, it suffices to 'push' the uniform distribution step-by-step from one such set to another, decreasing the relevant parameter. The key (most difficult) step is to eliminate double-edges. Suppressing the dependence on n and d in the notation, let $\mathcal{G}_k$ be the set of d-regular graphs on $[n]$ with exactly k double-edges, no loops, and no triple-edges (or higher multiplicity edges). Then the key tool is an algorithm that, starting with a uniformly random $G \in \mathcal{G}_k$, either 'fails' with small probability η_k, or outputs a uniformly random $G \in \mathcal{G}_{k-1}$.

As in much previous work (see in particular the key paper [4]), we do this using switchings, replacing a double-edge xy and two single-edges ab and cd in G with edges ax, xc, by and yd. More precisely, a *potential switching* consists of a double-edge xy (with a given order on its vertices) and two single-edges ab and

cd, and a potential switching is a *switching* if the edges involved are distinct and vertex-disjoint, and in G none of the edges ax, xc, by or yd is present. In that case the resulting graph G' is indeed an element of $\mathcal{G}_{k-1}$. Note that from any $G \in \mathcal{G}_k$ there are exactly $8ks^2$ potential switchings, where $s = dn/2 - 2k$ is the number of single-edges.

For $G' \in \mathcal{G}_k$ let $d(G')$ be the number of switchings from $\mathcal{G}_k$ that result in G', i.e., the in-degree of G' in the bipartite graph on $\mathcal{G}_k$ and $\mathcal{G}_{k-1}$ with an edge (G, G') for each graph $G \in \mathcal{G}_k$ and each valid switching from G resulting in G'. Algorithmically, given G we select a uniformly random potential switching (which is easy to do). If this is not an actual switching we (in the first formulation) return 'fail'. If it is, with probability $d_0/d(G')$ we accept the result, where d_0 is a lower bound on $\min\{d(G') : G' \in \mathcal{G}_{k-1}\}$; otherwise, we return 'fail'. This acceptance probability exactly corrects for the variation in the number of possible ways to arrive at a given G', resulting in an exactly uniform distribution on $\mathcal{G}_{k-1}$.

So far we have described ideas from previous work. The key new idea is that we can carry out the acceptance/rejection step *without calculating* $d(G')$. This may seem surprising at first, but here is the key idea. Given a function $\theta : \mathcal{G}_{k-1} \to [0, 1]$, by a *$\theta$-tester* we mean an algorithm that, given $G' \in \mathcal{G}_{k-1}$ as input, efficiently (here in time $O(1)$) returns 'yes' with probability $\theta(G')$, and 'no' otherwise. Note that the tester has access to G' but (typically) does not have time to actually calculate the desired probability $\theta(G')$. A trivial, but key, observation (the reciprocal part) is that, given a θ-tester for some $\theta \leqslant 1/2$, we may easily construct a $1/(1 + \theta)$ tester. Simply repeatedly call the θ-tester until the first time 'no' is returned, and then return 'yes' or 'no' according to whether this is after an odd or even number of calls: the overall probability of 'yes' is

$$(1 - \theta) + \theta^2(1 - \theta) + \theta^4(1 - \theta) + \cdots = 1/(1 + \theta).$$

The next key ingredient is a θ-tester for some θ proportional to $d(G') - d_0$; again it is far from clear that one exists. Fortunately one does, due to the structure of the set of *reverse switchings*, which are simply switchings viewed backwards from $G' \in \mathcal{G}_{k-1}$. To describe these it is easiest to work in the configuration model, or equivalently with 'half-edges', of which there are exactly d incident with each vertex. (In other words, we give the two edges in a double-edge distinct identities.) In these terms an edge aa' consists of two paired half-edges (treated as ordered). By a *potential reverse switching* we simply mean a list of 4 edges aa', bb', cc', dd' such that a and b are distinct half-edges in the same vertex, and c and d are distinct half-edges in the same vertex. We write $\mathcal{P}$ for the set of potential reverse switchings, noting that $|\mathcal{P}| = (nd(d-1))^2$.

Now $P \in \mathcal{P}$ corresponds to a valid reverse switching if and only if $a'a$, bb' and $c'c$, dd' are vertex-disjoint two-edge paths, with certain edges missing in G' (corresponding to those to be added in after deleting $a'a$, bb', $c'c$ and dd'). We now define certain simple 'bad' sets $\mathcal{B}_i \subset \mathcal{P}$, such that the set of valid switchings is exactly $\mathcal{P} \setminus \bigcup_i \mathcal{B}_i$. The key is that we can do this in such a way that (a) each set $\mathcal{B}_i$ has a certain size b_i that is the same for all $G' \in \mathcal{G}_{k-1}$, and (b) given G' we can quickly sample a uniformly random element of $\mathcal{B}_i$. For example, one such

set is the set of all P in which a' and c' are in the same vertex, which has size exactly $nd^2(d-1)^2$, and which can be sampled by first choosing this common vertex v, then choosing a' and c' independently from the half-edges incident with v, and working out from there. From this point it is not too hard to construct the required tester: we set $d_0 = |\mathcal{P}| - \sum_i b_i$, so

$$d(G') - d_0 = \sum b_i - |\bigcup \mathcal{B}_i| = \sum_i |\mathcal{B}_i^*|$$

where $\mathcal{B}_i^* = \mathcal{B}_i \cap \bigcup_{j<i} \mathcal{B}_j$. Since we can sample from $\mathcal{B}_i$, we can construct a $(|\mathcal{B}_i^*|/b_i)$-tester, and these can be combined without difficulty.

Of course, there are further difficulties to be overcome to make the algorithm work. A key one is that the rejection probability η_k needs to be small enough, namely $O(1/d^2)$ to allow for the $O(d^2)$ steps from a typical $\mathcal{G}_k$ to $\mathcal{G}_0$. For the 'reciprocal rejection' outlined above, this is indeed the case. However there is also the 'forward rejection' step where a potential switching that is not an actual switching is discarded. Here the rejection probability is $\Theta(d/n)$, which is too large. A clever but rather complicated solution to this was found by Gao and Wormald [3], involving different types of switchings and forward and backward steps. An alternative, new, solution is to carry out certain forward potential switchings that are not valid switchings, namely those that create exactly one new double-edge, but then immediately carry out another switching to eliminate this new double-edge. This leads to two types of forward switching (1-step and 2-step) from $\mathcal{G}_k$ to $\mathcal{G}_{k-1}$. It turns out that the reciprocal rejection method can be applied to both, leading to exact uniform sampling.

Finally, we note that while the method in principle sounds fairly general, it does not seem so easy to find situations in which it can be applied, due to the requirement (a) in particular on the 'bad' sets. For example, our method does not obviously work for random graphs with a given non-regular degree sequence. It would be interesting to find other applications, and even more so to see whether the method can be modified to work in such more general settings.

References

[1] A. Arman, P. Gao and N. Wormald, Fast uniform generation of random graphs with given degree sequences, *Random Structures Algorithms* **59** (2021), 291–314.

[2] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European J. Combin.* **1** (1980), 311–316.

[3] P. Gao and N. Wormald, Uniform generation of random regular graphs, *SIAM J. Comput.* **46** (2017), 1395–1427.

[4] B.D. McKay and N.C. Wormald, Uniform generation of random regular graphs of moderate degree, *J. Algorithms* **11** (1990), 52–67.

Unit distances and distinct distances (in general norms)

Noga Alon
(joint work with Matija Bucić, Lisa Sauermann)

A *d-norm* X is a real normed space of dimension d. Let $U(n, X)$ denote the maximum possible number of unit distances determined by a set of n distinct points in X. Let $D(n, X)$ denote the minimum possible number of distinct distances determined by n distinct points in X. We prove the following two main results.

Theorem 1. *(i) For every $d \geqslant 2$ there is a d-norm X so that for every n,*

$$U(n, X) \leqslant \frac{1}{2} dn \log_2 n.$$

In fact, this holds for all d-norms but a meagre set in the Bair space of d-norms.
(ii) For every $d \geqslant 2$ and all $n > n_0(d)$, and for every d-norm X,

$$U(n, X) \geqslant \frac{1}{2}(d - 1 - o(1))n \log_2 n$$

where the $o(1)$-term tends to 0 as n tends to infinity.

It is clear that for every d and every d-norm X, $U(n, X) \leqslant n - 1$ as shown by a set of n points along an arithmetic progression on a line.

Theorem 2. *For every d and every $n \geqslant n_0(d)$, and for every d-norm X but a meagre set, $D(n, X) = (1 - o(1))n$ where the $o(1)$-term tends to 0 as n tends to infinity.*

The results are motivated by old questions of Erdős [2] concerning the above problems in the Euclidean plane, and by questions of Ulam and Erdős [3] about general norms. The unit distance problem in the Euclidean plane is wide open, see [6], [8] for the best known bounds. For Euclidean spaces it is easy in all dimensions $d \geqslant 4$, where the function is quadratic in n.

The distinct distances problems in the Euclidean plane is better understood, and the correct value is known up to a $\sqrt{\log n}$ factor, as proved by Guth and Katz [4]. For higher dimensional Euclidean spaces the correct exponent is not known, see [7] for the known bounds.

Our results above essentially settle the problems of estimating the minimum possible value of $U(n, X)$ and the maximum possible value of $D(n, X)$ for a d-norm X. The results settle, in a strong form, problems and conjectures of Brass, of Matoušek, and of Brass, Moser and Pach. See [1], Chapter 5, and [5] for the precise formulation of the problems solved.

The proofs combine combinatorial, geometric and probabilistic ideas with tools from Linear Algebra, Algebraic Topology and Algebraic Geometry.

References

[1] P. Brass, W. Moser and J. Pach, *Research Problems in Discrete Geometry*, Springer, New York, 2005. xii+499 pp

[2] P. Erdős, On sets of distances of n points, Amer. Math. Monthly 53 (1946), 248–250.

[3] P. Erdős, Problems and results in combinatorial geometry, Discrete geometry and convexity (New York, 1982), Ann. New York Acad. Sci., vol. 440, New York Acad. Sci., New York, 1985, pp. 1–11.

[4] L. Guth and N. H. Katz, On the Erdős distinct distances problem in the plane, Ann. of Math. (2) 181 (2015), no. 1, 155–190.

[5] J. Matoušek, The number of unit distances is almost linear for most norms, Adv. Math. 226 (2011), no. 3, 2618–2628.

[6] J. Spencer, E. Szemerédi and W. Trotter, Jr., Unit distances in the Euclidean plane, Graph Theory and Combinatorics (Cambridge, 1983), Academic Press, London, 1984, pp. 293–303.

[7] J. Solymosi and V. H. Vu, Near optimal bounds for the Erdős distinct distances problem in high dimensions, Combinatorica 28 (2008), no. 1, 113–125.

[8] L. A. Székely, Crossing numbers and hard Erdős problems in discrete geometry, Combin. Probab. Comput. 6 (1997), no. 3, 353–358.

Problem Session

Nati Linial (chair)

Reinhard Diestel

Erdős [2] asked whether there is a function $f\colon \mathbb{N}^2 \to \mathbb{N}$ such that every finite graph G of chromatic number at least $f(r,k)$ has a subgraph of chromatic number at least k that contains no cycle of length at most r. The following problem weakens this by providing a local alternative:

Question 1. *Is there a function $f\colon \mathbb{N}^2 \to \mathbb{N}$ such that every finite graph G of chromatic number at least $f(r,k)$ satisfies one of the following two conditions:*

(1) *G has a subgraph of chromatic number at least k and girth greater than r;*

(2) *The r-local covering G_r of G has chromatic number at least k.*

Here, the *r-local* covering G_r of G is its covering space with characteristic subgroup generated by the walks in G from any fixed base point to a cycle of length at most r, round it, and back along the access path [1] (Fig. 1).

Both statements, (1) and (2), imply $\chi(G) \geqslant k$. They may thus be considered as a purely global, or purely local, reason for $\chi(G) \geqslant k$ if 'global' and 'local' are taken to mean involving cycles of length greater than or at most r, respectively. My motivation for (2) is also that while it is strong enough to imply $\chi(G) \geqslant k$ – notice that any ℓ-colouring of G with $\ell < k$ lifts to an ℓ-colouring of G_r – it is weaker than other natural local reasons for $\chi(G) \geqslant k$, such as the following ones.

One natural local reason for $\chi(G) \geqslant k$ is that $K_k \subset G$. It was proved by Rödl [3] that the assertion of our problem holds for $r = 3$ with (2) replaced by $K_k \subset G$, but the proof does not seem to generalise to larger r. One might try to weaken $K_k \subset G$ to the statement that G has a subgraph G'' with $\chi(G'') \geqslant k$ whose cycle space is generated by cycles of length at most r. However this would still be stronger

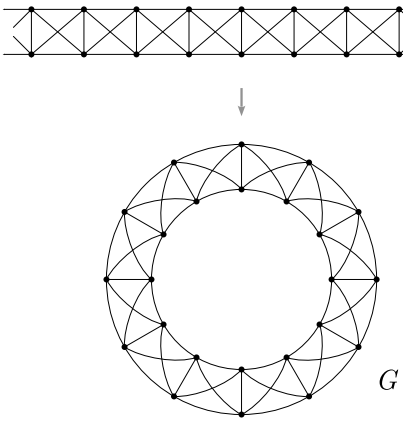

FIGURE 1. The r-local covering of G for $3 \leqslant r \leqslant 11$

than (2) since any such subgraph G'' of G lifts to G_r [1], so that $\chi(G_r) \geqslant k$ too. Conversely, (2) does not readily imply the existence of such a subgraph G'' of G.

References

[1] R. Diestel, R. Jacobs, P. Knappe, and J. Kurkofka. Graph decompositions. arXiv:2207.04855.
[2] P. Erdős, Problems and results in chromatic graph theory, in (F. Harary ed.) *Proof Techniques in Graph Theory* (Academic Press, 1969), 27–35.
[3] V. Rödl. On the chromatic number of subgraphs of a given graph. *Proc. AMS* **64** (1977), 370–371.

Mehtaab Sawhney

The following is a discrepancy theory question asked in an online algorithms setting.

Question 2. *Consider $v_1, \ldots, v_n \in \mathbb{R}^n$ with $\|v_i\|_2 \leqslant 1$. Does there exist a randomized online algorithm assigning $\varepsilon_i \in \{\pm 1\}$ such that $\|\sum \varepsilon_i v_i\|_\infty \leqslant C\sqrt{\log n}$ whp?*

Online means that we decide ε_i based on $v_1, \ldots, v_i$ having decided the previous signs. The Komlós conjecture is this problem without the online algorithm portion, for which the best known bound is $C\sqrt{\log n}$ due to Banaszczyk [1] (and the conjecture is a constant, which is not attainable for online settings). This was made algorithmic by Bansal, Dadush, Garg, and Lovett [2]. This $C\sqrt{\log n}$ can be done with $\varepsilon_i \in \{\pm 1, 2\}$ due to Liu, Sah, and Sawhney [3].

References

[1] Wojciech Banaszczyk, *Balancing vectors and Gaussian measures of n-dimensional convex bodies*, Random Structures Algorithms **12** (1998), 351–360.

[2] Nikhil Bansal, Daniel Dadush, Shashwat Garg, and Shachar Lovett, *The Gram-Schmidt walk: a cure for the Banaszczyk blues*, STOC'18—Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, ACM, New York, 2018, pp. 587–597.
[3] Yang P. Liu, Ashwin Sah, and Mehtaab Sawhney, *A Gaussian fixed point random walk*, 13th Innovations in Theoretical Computer Science Conference, LIPIcs. Leibniz Int. Proc. Inform., vol. 215, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2022, pp. Art. No. 101, 10.

Matija Bucić

Question 3. *Let M_n be the minimum number such that given M_n transpositions in S_n, one can always find a sequence of distinct elements of this set $\pi_1, \dots, \pi_k$ such that $\pi_1 \cdots \pi_k = 1$. What is M_n? This is called the additive dimension of the set of transpositions.*

What is known is the following upcoming result.

Theorem 4 (Alon, Bucić, Sauermann, Zakharov, Zamir 2023+)**.** *Any n vertex properly edge-colored graph with $10^8 n \log n \log\log n$ edges has a rainbow cycle.*

This is tight up to the $\log\log n$ factor. Applying this to a suitable Cayley graph shows that $M_n = O(n \log^2 n)$, and there is a construction for $M_n = \Omega(n \log n / \log\log n)$. For the theorem above, a known lower bound is $\Omega(n \log n)$, and this result addresses a question of Keevash, Mubayi, Sudakov, and Verstraete [1].

References

[1] Peter Keevash, Dhruv Mubayi, Benny Sudakov, and Jacques Verstraëte, *Rainbow Turán problems*, Combin. Probab. Comput. **16** (2007), 109–126.

Peter Allen

The following guarantees embedding of graphs with maximum degree in a host given certain pseudorandomness conditions, irrespective of the density regime.

Theorem 5 (Allen, Böttcher, Davies, Hng, Skokan 2021)**.** *Given k, Δ there exists $f_k(\Delta)$ and $\varepsilon > 0$ such that the following holds. Consider n large, $p > 0$ (potentially dependent on n), and G an n-vertex k-uniform hypergraph with the following property: for all F with at most $f_k(\Delta)$ vertices, then*

$$\mathbb{P}(\varphi\colon V(F) \to V(G) \text{ is a homomorphism}) = (1 \pm \varepsilon) p^{e(F)},$$

where φ is a uniform vertex map, and the same when we condition φ to map any specific vertex of F to any specific vertex of G.

Then G contains any n-vertex graph H with maximum degree $\Delta(H) \leqslant \Delta$.

One would like to understand how much is needed for this statement to hold.

Question 6. *What is $f_k(\Delta)$?*

We remark that $f_k(\Delta) = \Omega(\Delta)$ is known by random construction.

Ehud Friedgut

The parity set has the property that any subcube is half in the set and half out of the set. However, modulo 3 conditions will only work on sufficiently large subcubes. Can we find examples on a slice?

Question 7. *A sub-balanced slice of dimension* $2k$ *is the following set: we fix some* $n-k$ *coordinates* 1*, some* $n-k$ *coordinates* 0*, and we consider all possibilities for the remaining* $2k$ *coordinates which are half and half between* $0, 1$*.*

For all indicator functions f *on the slice with* $\mathbb{E}f = 1/2$*, does there exist a sub-balanced slice of dimension* $\log^* n$ *(or growing) on which* $\mathbb{E}f > 0.51$*?*

Jacques Verstraete

Let G be a graph of edge-density $p \in [0,1]$. The *positive discrepancy* of a graph G is defined by

$$\mathrm{disc}^+(G) = \max\left\{e(X) - p\binom{|X|}{2} \colon X \subseteq V(G)\right\}$$

and the *negative discrepancy* is

$$\mathrm{disc}^-(G) = \max\left\{p\binom{|X|}{2} - e(X) \colon X \subseteq V(G)\right\}.$$

The *discrepancy* of G is $\max\{\mathrm{disc}^+(G), \mathrm{disc}^-(G)\}$. Erdős, Goldreich, Pach and Spencer [2] proved that there exists an absolute constant $c > 0$ such that if G is an n-vertex graph of density p satisfying $p(1-p) \geqslant 1/n$, then

$$\mathrm{disc}(G) \geqslant c\sqrt{p(1-p)}n^{3/2}.$$

This is tight as shown by considering the random graph $G_{n,p}$. Bollobás and Scott [1] proved further that for some absolute constant $b > 0$,

$$\mathrm{disc}^+(G) \cdot \mathrm{disc}^-(G) \geqslant bp(1-p)n^3.$$

We propose the following conjecture:

Conjecture 8. *For all* $\varepsilon > 0$*, there exists a constant* $a = a(\varepsilon) > 0$ *such that for any* n*-vertex graph* G *of density* p *satisfying* $|p - (i-1)/i| > \varepsilon$ *for every* $i \in \mathbb{Z}^+$*,*

$$\mathrm{disc}^+(G) \geqslant a\sqrt{p(1-p)}n^{3/2}.$$

Complete n-vertex balanced multipartite graphs with i parts have density about $(i-1)/i$, but their positive discrepancy is $O(n)$, which explains why we require separation between p and these densities in the conjecture. In particular, the conjecture could be extended to say that if $p_0 < 1/2$ is a constant and $1/n \leqslant p \leqslant p_0$, then for any n-vertex graph G of density p, the positive discrepancy of G is $\Omega(p^{1/2}n^{3/2})$ as $n \to \infty$. A similar conjecture could be made for uniform hypergraphs.

References

[1] B. Bollobás, and A. Scott, Discrepancy in graphs and hypergraphs, pp 33–56 in: More Sets, Graphs and Numbers. Bolyai Society Mathematical Studies **15**, Springer, 2006.
[2] P. Erdős, M. Goldberg, J. Pach and J. Spencer, Cutting a graph into two dissimilar halves, Journal of Graph Theory **12**, 121–131 (1988).

Charlotte Knierim

Given a digraph D, consider a weighting $\omega\colon E(D) \to \mathbb{R}_{\geqslant 0}$ so that $\omega^+(v) \geqslant 1$ for all $v \in V(D)$.

Theorem 9 (Knierim, Larcher, Martinsson, Noever 2020 [1])**.** *In the above setup, we can find a cycle with $\omega(C) \geqslant \log\log n/\log n$, and this is tight.*

For the tightness, take an ℓ-ary tree for k generations, and adding edges from leaves to all vertices along the path to it from the root. Then put a weight of $1/\ell$ and $1/k$ appropriately to satisfy the condition and optimize in k and ℓ.

Question 10. *What happens if D is Eulerian? What happens if D is strongly 2-connected? Concretely, if D is Eulerian, is there a cycle of weight $\Omega(1)$?*

References

[1] Charlotte Knierim, Maxime Larcher, Anders Martinsson, and Andreas Noever, *Long cycles, heavy cycles and cycle decompositions in digraphs*, J. Combin. Theory Ser. B **148** (2021), 125–148.

Jacob Fox

This question arises out of current work of Conlon, Fox, Pham, and Yepremyan.

Question 11. *Let $f(n)$ be the maximum size (number of elements) of a monochromatic (ignoring 0) linear subspace guaranteed in any 2-coloring of the nonzero elements of $\mathbb{F}_2^n$. Is $f(n) = O(n)$?*

By Ramsey methods such as the Finite Union Theorem, $f(n) \to \infty$ as $n \to \infty$ ($\log^* n$ lower bound, perhaps more recent work gives finitely many logarithms). Additionally, $f(n) = O(n \log n)$ by considering a random coloring (and it does not do better). Noga Alon has a conjecture about Ramsey Cayley graphs that implies this.

Karim Adiprasito

Question 12. *Given a graph G and $\mathcal{C}$ a collection of cycles, let $\|\mathcal{C}\|_1 = \sum_{\gamma\in\mathcal{C}} |\gamma|$ be the total length. Given a field F and $\mathcal{C}$, glue in the corresponding 2-disks into the cell complex underlying the graph, and consider the homology over F. Let $\alpha_F(G)$ be the minimum $\|\mathcal{C}\|_1$ such that the F-homology vanishes after this gluing. Let $\alpha_{\mathbb{Z}}(G)$ be the same for $\mathbb{Z}$-homology, and $\alpha_{\pi_1}(G)$ for vanishing the first homotopy group. Let $\alpha_\infty(G)$ be the same but for vanishing all homotopy groups. We have $\alpha_F(G) \leqslant \alpha_{\mathbb{Z}}(G) \leqslant \alpha_{\pi_1}(G) \leqslant \alpha_\infty(G)$. Is it true that $\alpha_{\pi_1}(G)/\alpha_{\mathbb{F}_2}(G)$ is unbounded? For the strongest form, one would ask for G of bounded degree.*

For very partial progress, one can force $\alpha_{\pi_1}(G) - \alpha_{\mathbb{F}_2}(G) \to \infty$ by taking a (bounded degree) triangulation of $\mathbb{RP}^2$ and taking disjoint unions.

Julia Böttcher

Question 13. *Do there exist n-vertex graphs G_n which is $\{C_3, C_5, \ldots, C_{2k-1}\}$-free (as subgraphs) with chromatic number $\chi(G_n) \geqslant 4$ and minimum degree $\delta(G_n) \geqslant f(k)n$ where $f(k)$ is at least linear in $1/k$?*

It is known for $f(k) = 3/(2k^2+k+1)$ and it is known to be false for $f(k) = 1/(2k)$ when $k > 7000$ [1].

References

[1] Julia Böttcher, Nóra Frankl, Domenico Mergoni Cecchelli, Olaf Parzcyk, and Jozef Skokan, *Graphs with large minimum degree and no small odd cycles are* 3*-colorable*, arXiv:2302.01875.

Noga Alon

The following is conjectured in forthcoming work of Alon, Bucić, Sauermann, Zakharov, and Zamir.

Question 14. *For every $k \geqslant 2$ does there exist $c(k)$ so that for all groups G with $|G| = n$ and $S \subseteq G$ with $|S| \geqslant c(k) \log n$, there are $T_1, \ldots, T_k \subseteq S$ pairwise disjoint and an ordering of each T_i so that each product is the same?*

For abelian G it is still not known, but $k = 2$ is trivial and $k = 3$ follows from the Erdős-Szemerédi sunflower problem for $k = 3$ (similar to cap-set). For all k and abelian G it is true with an extra $\log \log n$ factor due to what is currently known about the Sunflower Conjecture [1]. If every element in S is forced to have order 2 then it is known up to $\log \log n$ for nonabelian.

References

[1] Ryan Alweiss, Shachar Lovett, Kewen Wu, and Jiapeng Zhang, *Improved bounds for the sunflower lemma*, Ann. of Math. (2) **194** (2021), 795–815.

Reporter: Ashwin Sah, Mehtaab Sawhney

Participants

Prof. Dr. Karim Adiprasito
Einstein Institute for Mathematics
The Hebrew University of Jerusalem
Edmond J. Safra Campus
Givat Ram
Jerusalem 9190401
ISRAEL

Prof. Dr. Peter Allen
Department of Mathematics
London School of Economics
Houghton Street
London WC2A 2AE
UNITED KINGDOM

Prof. Dr. Noga Alon
Department of Mathematics
Princeton University
Office: 706 Fine Hall
Washington Street
Princeton NJ 08544
UNITED STATES

Prof. Dr. József Balogh
Department of Mathematical Sciences
University of Illinois
Office: 233 B Illini Hall
1409 West Green Street
Urbana IL 61801
UNITED STATES

Prof. Dr. Alexander Barvinok
Department of Mathematics
University of Michigan
4066 East Hall
Ann Arbor, Michigan 48109-1043
UNITED STATES

Dr. Julia Böttcher
Department of Mathematics
London School of Economics
Houghton Street
London WC2A 2AE
UNITED KINGDOM

Dr. Matija Bucić
Fine Hall 1203
Princeton University and
Institute for Advanced Study
Washington Road
Princeton NJ 08540
UNITED STATES

Prof. Dr. Maria Chudnovsky
Department of Mathematics
Princeton University
211 Fine Hall
68 Washington Road
Princeton NJ 08544
UNITED STATES

Prof. Dr. David Conlon
Department of Mathematics
California Institute of Technology
1200 E California Boulevard
Pasadena CA 91125
UNITED STATES

Prof. Dr. Reinhard Diestel
Mathematisches Seminar
Universität Hamburg
Bundesstraße 55
20146 Hamburg
GERMANY

Dr. Asaf Ferber
Department of mathematics
University of California
Irvine CA 92697
UNITED STATES

Prof. Dr. Jacob Fox
Department of Mathematics
Stanford University
Building 380
Stanford CA 94305
UNITED STATES

Prof. Dr. Ehud Friedgut
Department of Mathematics
The Weizmann Institute of Science
234 Herzl Street
P.O. Box 26
Rehovot 7610001
ISRAEL

Prof. Dr. David Gamarnik
Massachusetts Institute of Technology
Sloan School of Management
E 62 - 563
100 Main Street
Cambridge MA 02139
UNITED STATES

Prof. Dr. Penny E. Haxell
Department of Combinatorics and Optimization
University of Waterloo
Waterloo ON N2L 3G1
CANADA

Dr. Annika Heckel
Matematiska Institutionen
Uppsala Universitet
Box 480
751 06 Uppsala
SWEDEN

Dr. Matthew Jenssen
School of Mathematics
The University of Birmingham
Edgbaston
Birmingham B15 2TT
UNITED KINGDOM

Prof. Dr. Mihyun Kang
Institut für Diskrete Mathematik
Technische Universität Graz
Steyrergasse 30
8010 Graz
AUSTRIA

Prof. Dr. Peter Keevash
Mathematical Institute
University of Oxford
Andrew Wiles Building
Radcliffe Observatory Quarter
Woodstock Road
Oxford OX2 6GG
UNITED KINGDOM

Charlotte Knierim
CAB G17
Department of Computer Science
ETH Zürich
Universitätstrasse 6
8092 Zürich
SWITZERLAND

Prof. Dr. Daniel Král
Faculty of Informatics
Masaryk University
Botanická 68 A
602 00 Brno
CZECH REPUBLIC

Prof. Dr. Michael Krivelevich
School of Mathematical Sciences
Sackler Faculty of Exact Sciences
Tel Aviv University
Ramat Aviv, Tel Aviv 6997801
ISRAEL

Dr. Matthew Kwan
IST Austria
Am Campus 1
3400 Klosterneuburg
AUSTRIA

Dr. Shoham Letzter
Department of Mathematics
University College London
Gower Street
London WC1E 6BT
UNITED KINGDOM

Dr. Noam Lifshitz
Einstein institute of mathematics,
The Hebrew University of Jerusalem
Givat Ram
Jerusalem 9190401
ISRAEL

Prof. Dr. Nathan Linial
School of Computer Science and Engineering
The Hebrew University of Jerusalem
Givat Ram
Jerusalem 9190401
ISRAEL

Prof. Dr. Tomasz Łuczak
Faculty of Mathematics and Computer Science
Adam Mickiewicz University
ul. Uniwersytetu Poznańskiego 4
61-614 Poznań
POLAND

Dr. Richard Montgomery
Mathematics Institute
University of Warwick
Gibbet Hill Road
Coventry CV4 7AL
UNITED KINGDOM

David Munhá Correia
Department of Mathematics
ETH Zürich
Rämistrasse 101
8092 Zürich
SWITZERLAND

Dr. Bhargav Narayanan
Department of Mathematics
Rutgers University
Busch Campus, Hill Center
New Brunswick NJ 08854-8019
UNITED STATES

Prof. Dr. Sergey Norin
Department of Mathematics and Statistics
McGill University
805 Sherbrooke Street West
Montréal QC H3A 2K6
CANADA

Prof. Dr. János Pach
Renyi Institute
Realtanoda u. 13-15
Budapest H-1364 POB 127
HUNGARY

Dr. Will Perkins
Department of Mathematics, Statistics and Computer Science
University of Illinois at Chicago
851 South Morgan Street
Chicago IL 60607-7045
UNITED STATES

Prof. Dr. Oleg Pikhurko
Mathematics Institute
University of Warwick
Coventry CV4 7AL
UNITED KINGDOM

Dr. Alexey Pokrovskiy
Department of Mathematics,
University College London
Gordon Street
London WC1H 0AY
UNITED KINGDOM

Prof. Dr. Oliver Riordan
Mathematical Institute
Oxford University
Woodstock Road
Oxford OX2 6GG
UNITED KINGDOM

Prof. Dr. Vojtěch Rödl
Department of Mathematics and
Computer Science
Emory University
400 Dowman Drive
30322 Atlanta GA 30322
UNITED STATES

Ashwin Sah
Massachusetts Institute of Technology
Department of Mathematics
77 Massachusetts Avenue
Cambridge MA 02139
UNITED STATES

Dr. Julian Sahasrabudhe
Dept. of Math. and Math. Stats.
(DPMMS), University of Cambridge
Wilberforce road
Cambridge CB3 0WA
UNITED KINGDOM

Prof. Dr. Wojciech Samotij
School of Mathematical Sciences
Tel Aviv University
Ramat Aviv, Tel Aviv 6997801
ISRAEL

Prof. Dr. Lisa Sauermann
Massachusetts Institute of Technology
Department of Mathematics
Building 2
77 Massachusetts Avenue
Cambridge MA 02139
UNITED STATES

Mehtaab Sawhney
Department of Mathematics
Massachusetts Institute of
Technology
77 Massachusetts Avenue
Cambridge, MA 02139-4307
UNITED STATES

Prof. Dr. Mathias Schacht
Fachbereich Mathematik
Universität Hamburg
Bundesstraße 55
20146 Hamburg
GERMANY

Dr. Bjarne Schülke
Department of Mathematics
California Institute of Technology
1200 E California Boulevard
Pasadena CA 91125
UNITED STATES

Prof. Dr. Alex Scott
Mathematical Institute
University of Oxford
Radcliffe Observatory Quarter
Andrew Wiles Building
Woodstock Road
Oxford OX2 6GG
UNITED KINGDOM

Prof. Dr. Asaf Shapira
Department of Mathematics
School of Mathematical Sciences
Tel Aviv University
P.O. Box 39040
Ramat Aviv, Tel Aviv 6997801
ISRAEL

Prof. Dr. József Solymosi
Department of Mathematics
University of British Columbia
1984 Mathematics Road
Vancouver V6T 1Z2
CANADA

Dr. Sophie Spirkl
Dept. of Combinatorics and Optimization
University of Waterloo
200 University Ave W
Waterloo ON N2L 3G1
CANADA

Prof. Dr. Benjamin Sudakov
Department of Mathematics
ETH Zürich, HG G 65.1
Rämistrasse 101
8092 Zürich
SWITZERLAND

Prof. Dr. Tibor Szabó
Institut für Mathematik
Freie Universität Berlin
Arnimallee 6
14195 Berlin
GERMANY

Dr. Istvan Tomon
Departement Mathematik
ETH - Zentrum
Rämistrasse 101
8092 Zürich
SWITZERLAND

Prof. Dr. Jacques Verstraëte
Department of Mathematics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0112
UNITED STATES

Prof. Dr. Van H. Vu
Department of Mathematics
Yale University
Box 20 82 83
New Haven, CT 06520
UNITED STATES

Prof. Dr. Uli Wagner
Institute of Science and Technology Austria (ISTA)
Am Campus 1
3400 Klosterneuburg
AUSTRIA

Prof. Dr. Yufei Zhao
Massachusetts Institute of Technology
77 Massachusetts Avenue
P.O. Box Room 2-271
Cambridge MA 02139
UNITED STATES