Mathematisches Forschungsinstitut Oberwolfach

Report No. 22/2022

DOI: 10.4171/OWR/2022/22

Combinatorics, Probability and Computing

Organized by
Michael Krivelevich, Tel Aviv
Robert Morris, Rio de Janeiro
Oliver Riordan, Oxford
Angelika Steger, Zürich

24 April – 30 April 2022

Abstract. The main theme of this workshop was the use of probabilistic methods in combinatorics and theoretical computer science. This area is evolving extremely quickly, with the introduction of powerful new methods and the development of increasingly sophisticated techniques, and there have been a number of very significant breakthroughs in the area in recent years. The workshop emphasized several of these recent breakthroughs, which include applications of probabilistic techniques to extremal questions, and of combinatorial techniques to areas of discrete probability theory, such as random matrices and planar percolation.

Mathematics Subject Classification (2020): 05-xx, 60-xx.

Introduction by the Organizers

The meeting was very well attended with 53 in-person participants from around the globe, including the US, Israel, Canada, Brazil, and many European countries, and 7 online participants from the US, Australia and Germany. In addition, many excellent mathematicians who would have loved to participate could not be invited, for lack of space. The program consisted of 8 main lectures, 19 shorter talks, and a problem session, with plenty of time for discussion.

There have been a number of very significant breakthroughs in probabilistic combinatorics in recent years, and the workshop highlighted several of these. Perhaps the most spectacular of these breakthroughs was the proof of the Kahn–Kalai conjecture by Jinyoung Park and Huy Tuan Pham, both of whom spoke (remotely) about their work during the workshop. Their theorem implies that the threshold of any increasing property is at most $O(\log n)$ times the "expectation threshold"

which is a natural lower bound. Pham presented the proof of a related conjecture of Talagrand, which also follows from their method.

There were a number of other major highlights of the workshop, including the proof of the classical Ryser–Brualdi–Stein conjecture for large even n, which was announced by Richard Montgomery in his talk. More precisely, Montgomery used a random process together with a novel kind of algebraic absorbing structure to show that if n is sufficiently large, then every Latin square of order n has a transversal with $n-1$ cells.

Patrick Morris presented his recent proof of a beautiful conjecture of Krivelevich, Sudakov and Szabó, on the threshold for the existence of a triangle-factor in a sparse pseudorandom graph. Roughly speaking, his theorem shows (together with a construction of Alon) that as soon as one can guarantee the existence of a triangle, one can also find a triangle-factor. He also obtained bounds for general clique-factors, but for larger cliques it is not known (and it is a major open problem to determine) whether or not matching constructions exist.

Julian Sahasrabudhe and Matthew Jenssen both gave talks about different aspects of their stunning recent work (with Marcelo Campos and Marcus Michelen) on symmetric random matrices. Sahasrabudhe spoke about their proof of an exponential upper bound on the singularity probability of a uniformly-chosen $n \times n$ symmetric random matrix with entries in $\{-1, 1\}$, while Jenssen discussed a very recent extension of that work, bounding the least singular value of a larger class of random matrices. In the proofs of these results they introduced some exciting new techniques for controlling approximate negative correlation between random variables, and for "reusing" randomness.

Katherine Staden presented her proof (with Peter Keevash) of Ringel's tree packing conjecture, which was first posed in 1963. The conjecture states that, for any tree T with n edges, the complete graph with $2n+1$ vertices can be decomposed into $2n+1$ copies of T. Staden and Keevash actually proved a much more general result, about decompositions of pseudorandom graphs into trees of appropriate sizes. Their proof is very technically impressive, and builds on the proof (by the same authors) of the "generalised Oberwolfach problem".

There were many other excellent talks during the week. To mention just a few: Vincent Tassion described a beautiful proof (found with Laurin Köhler-Schindler) of an extremely general Russo–Seymour–Welsh theorem, only using symmetry and positive association; Shoham Letzter presented an almost-sharp bound on the maximum possible number of edges in an r-uniform hypergraph on n vertices that contains no tight cycles; Matija Bucic presented a breakthrough (found with Richard Montgomery) on an old conjecture of Erdős and Gallai, showing that every graph with n vertices can be decomposed into $O(n \log^\star n)$ cycles and edges; Lisa Sauermann described a method (found with Asaf Ferber and Matthew Kwan) of showing that certain list-decodable Reed–Solomon codes exist; Anita Liebenau presented some of her recent results with Nina Kamčev and Natasha Morrison on "uncommon" systems of equations; Ashwin Sah and Mehtaab Sawhney each described different aspects of their recent work (joint with Matthew Kwan and

Michael Simkin) on Steiner Triple Systems; Marcelo Campos presented some new bounds on the number of sumsets; and Benny Sudakov presented a proof (found with Nemanja Draganić and David Munhá Correia) of a conjecture of Erdős on the pancyclicity of Hamiltonian graphs.

As always, and on behalf of all participants, the organizers would like to thank the staff and the director of the Mathematisches Forschungsinstitut Oberwolfach for providing such a stimulating and inspiring atmosphere.

Acknowledgement: The MFO and the workshop organizers would like to thank the National Science Foundation for supporting the participation of junior researchers in the workshop by the grant DMS-1641185, "US Junior Oberwolfach Fellows". Moreover, the MFO and the workshop organizers would like to thank the Simons Foundation for supporting Liana Yepremyan in the "Simons Visiting Professors" program at the MFO.

Workshop: Combinatorics, Probability and Computing

Table of Contents

Abstracts

New bounds on the Erdős-Gallai Cycle Decomposition Conjecture

Matija Bucić
(joint work with Richard Montgomery)

When is it possible to decompose a graph into edge-disjoint subgraphs with certain properties? Many classical problems in extremal combinatorics fall within this framework, and together with the natural generalisation to hypergraphs they have found some remarkable applications both within and beyond extremal combinatorics. The particular case where we seek to decompose a graph into cycles has a long history, dating back to the 18th century and Euler's result on the existence of Euler tours. As Veblen later observed, Euler's result immediately implies that any graph with even vertex degrees (i.e., any *Eulerian* graph) has a decomposition into cycles, while it is immediate that if a graph has a vertex of odd degree then it can not be decomposed into cycles. This result is known as Veblen's theorem and it arose naturally in connection to his algebraic approach for attacking the Four-Colour Theorem in 1912.

This result gives, however, no non-trivial bound on the number of cycles needed for a decomposition. That $O(n)$ cycles might be sufficient is easily seen to be equivalent to the following conjecture of Erdős and Gallai [4] dating back to 1966, which is one of the major open problems on graph decompositions.

Conjecture 1. *Any n-vertex graph can be decomposed into $O(n)$ cycles and edges.*

As we have noted, Conjecture 1 is equivalent to conjecturing that every n-vertex Eulerian graph can be decomposed into $O(n)$ cycles, however, the optimal implicit constants in these conjectures, if they hold, seem likely to be different. For the Eulerian problem, Hajós conjectured that $\frac{n}{2}$ cycles should be sufficient, while the best known lower bound for the number of cycles and edges required in Conjecture 1 is $(\frac{3}{2} - o(1))n$, as observed by Erdős in 1983 [3], improving on a previous construction of Gallai [4].

Despite a lot of work on related problems over the years direct progress towards the Erdős-Gallai Conjecture has only been made within the last decade. The previous related results, which we discuss first, mostly focused on the analogous path decomposition problem and the covering version of the Erdős-Gallai conjecture.

Path decompositions. In the 1960's, Gallai [10] posed the analogous path decomposition version of Conjecture 1. In particular, he conjectured that any connected n-vertex graph can be decomposed into at most $\frac{n+1}{2}$ paths. Lovász [10] in 1968 proved that any graph can be decomposed into at most $n-1$ paths. This follows easily from his complete solution to the problem of how many paths *or* cycles one needs to decompose an n-vertex graph, to which the answer is $\lfloor \frac{n}{2} \rfloor$. Currently the best general bound in the path decomposition problem is due independently to Dean and Kouider and Yan, who showed that any graph can be decomposed into at most $\lfloor \frac{2n}{3} \rfloor$ paths. Gallai's path decomposition result is known to hold for

quite a few special classes of graphs, with connected planar graphs being the most recent addition to the list, this result is due to Blanché, Bonamy and Bonichon.

Covering problems. Another interesting direction which has attracted a lot of attention is a covering version of Conjecture 1, in which we do not insist that the cycles we find should be disjoint, only that together they contain all the edges of the host graph. In 1985, Pyber proved the covering version of the Erdős-Gallai conjecture, showing that the edges of any n-vertex graph can be covered with $n-1$ cycles and edges. The analogous covering version of Gallai's conjecture, raised by Chung in 1980, has been settled first approximately by Pyber in 1996 and then completely by Fan in 2002, who showed that the edges of any connected graph can be covered by $\lceil \frac{n}{2} \rceil$ paths. The covering version of Hajós' conjecture was solved by Fan, who showed that any n-vertex Eulerian graph can be covered by at most $\lfloor \frac{n-1}{2} \rfloor$ cycles, settling another conjecture of Chung.

Results on the Erdős-Gallai conjecture. In more recent years, the Erdős-Gallai conjecture (along with more accurate results on the implicit bounds) has been shown to hold for two large specific classes of graphs – random graphs and graphs with linear minimum degree. The conjecture was first established for a typical binomial random graph $G(n,p)$ (for any $p = p(n)$) by Conlon, Fox and Sudakov [1]. Korándi, Krivelevich, and Sudakov [9] found the correct leading constant here, showing that $(\frac{1}{4} + \frac{p}{2} + o(1))n$ cycles and edges are typically sufficient to decompose $G(n,p)$. For constant edge probability p, Glock, Kühn, and Osthus [6] were even able to determine with high probability the exact minimum number of cycles and edges required to decompose a (quasi)random graph. On the other hand, Conjecture 1 was first shown to hold for graphs with linear minimum degree again by Conlon, Fox and Sudakov [1]. Very recently, the asymptotically correct bound of $(\frac{3}{2} + o(1))n$ cycles and edges has been proved by Girão, Granet, Kühn, and Osthus [5] for large graphs with linear minimum degree.

For almost 50 years, the best known bound in the general case of the Erdős-Gallai conjecture (as observed by Erdős and Gallai) came from a simple argument involving the iterative removal of a longest cycle, which shows that an n-vertex graph can always be decomposed into $O(n \log n)$ cycles and edges. In 2014, Fox, Conlon and Sudakov [1] made the first major breakthrough on this problem, showing that such a decomposition with only $O(n \log\log n)$ cycles and edges always exists. Here will we give the following improvement on this bound, where $\log^\star n$ is the iterated logarithm function.

Theorem 2. *Any n-vertex graph can be decomposed into $O(n \log^\star n)$ cycles and edges.*

A key new idea introduced in [1] was the use of graph expansion. One of the new insights we bring to the table is the use of a much weaker, sublinear, version of expansion. The main benefit of this is that it is far less costly to reduce the general problem to decomposing expanders, albeit at a cost of being much more difficult to work with. The first to use a similar notion were Komlós and Szemerédi [7, 8] although since then it has been modified in a number of ways and has recently

found some remarkable applications. In addition to introducing an alternative notion of robustness inherent to this notion of expansion, we develop a range of new tools, which we hope will prove useful elsewhere. In particular, we show that inducing our expander onto a random linear-sized vertex set is likely to produce an expander. This is particularly difficult as the expansion we use is sublinear, and proving it involves a carefully designed random process. Other highlights include a similar result but while randomly removing edges, the (almost) decomposition of any graph into expander, finding a 'connective skeleton' (a sparse subgraph of an expander retaining some expansion properties), and our use of the Aharoni-Haxell hypergraph matching theorem as a 'probability booster'. The actual application of these tools to the Erdős-Gallai conjecture begins by almost decomposing the initial arbitrary graph into robust sublinear expanders before decomposing each of these expanders carefully into paths (using an aforementioned result of Lovász) while avoiding a 'connective skeleton' which has been carefully found to have few edges while being very well-connected so that in particular it can be used to connect these paths into edge-disjoint cycles. All remaining edges lie within the connective skeletons, which are sparse, so there will be very few such edges and this will allow us, via a careful iterative argument, to prove Theorem 2.

References

[1] D. Conlon, J. Fox, and B. Sudakov. Cycle packing. *Random Structures & Algorithms*, 45(4):608–626, 2014.

[2] P. Erdős. Some unsolved problems in graph theory and combinatorial analysis. In *Combinatorial Mathematics and its Applications (Proc. Conf., Oxford, 1969)*, pages 97–109. Academic Press, London, 1971.

[3] P. Erdős. On some of my conjectures in number theory and combinatorics. In *Proceedings of the fourteenth Southeastern conference on combinatorics, graph theory and computing*, volume 39, pages 3–19, 1983.

[4] P. Erdős, A. W. Goodman, and L. Pósa. The representation of a graph by set intersections. *Canadian Journal of Mathematics*, 18:106–112, 1966.

[5] A. Girão, B. Granet, D. Kühn, and D. Osthus. Path and cycle decompositions of dense graphs. *Journal of the London Mathematical Society*, 2021.

[6] S. Glock, D. Kühn, and D. Osthus. Optimal path and cycle decompositions of dense quasirandom graphs. *Journal of Combinatorial Theory, Series B*, 118:88–108, 2016.

[7] J. Komlós and E. Szemerédi. Topological cliques in graphs. *Combinatorics, Probability and Computing*, 3(2):247–256, 1994.

[8] J. Komlós and E. Szemerédi. Topological cliques in graphs II. *Combinatorics, Probability and Computing*, 5(1):79–90, 1996.

[9] D. Korándi, M. Krivelevich, and B. Sudakov. Decomposing random graphs into few cycles and edges. *Combin. Probab. Comput.*, 24(6):857–872, 2015.

[10] L. Lovász. On covering of graphs. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 231–236. Academic Press New York, 1968.

The number of sumsets in $\mathbb{Z}_n$

Marcelo Campos

Given a prime n and sets $A, B \subset \mathbb{Z}_n$ we may define their sumset as

$$A + B := \{a + b \ : a \in A, b \in B\}.$$

This object has been intensely studied over the years and is one of the main objects of study of additive combinatorics. In this context, a natural question that arises is how many subsets of $\mathbb{Z}_n$ are of the form $A + A$ for some $A \subset \mathbb{Z}_n$? Green and Ruzsa [1] gave a first answer to this question.

Theorem 1 (Green, Ruzsa, 2004)**.** *Let n be a prime. There are*

$$2^{n/3+o(n)}$$

subsets of $\mathbb{Z}_n$ of the form $A + A$ for some $A \subset \mathbb{Z}_n$.

In a different direction Alon, Graville and Ubis [2] proved various results about counting sets of the form $A + B$, for instance they proved the following.

Theorem 2 (Alon, Granville, Ubis, 2010)**.** *Let n be a prime. There are*

$$2^{n/2+o(n)}$$

subsets of $\mathbb{Z}_n$ of the form $A + B$ for some $A, B \subset \mathbb{Z}_n$, with $|A|, |B| \to \infty$.

Notice that for each $\{0\} \subset A \subset \{0\} \cup [n/3, 2n/3]$ the set $A + A$ is different, so there are at least $2^{n/3}$ choices for a set of the form $A + A$. In fact, for any arithmetic progression P, with $|P| = n/3$, and x such that $(x + P) \cap (P + P) = \varnothing$, we an take all $\{x\} \subset A \subset P \cup \{x\}$. In [1] the authors provide this construction, show it provides $\Omega(n^2 2^{n/3})$ sets of the form $A + A$, since one has n^2 choices for P. Green and Ruzsa suggest this lower bound might be tight, and in [2] the authors also ask for a tight bound on the number of sets of the form $A + A$. One of my main contributions in this work is providing an answer to these questions up to a constant factor.

Theorem 3. *Let n be a prime. There are*

$$\Theta\left(n^{8/3} 2^{n/3}\right)$$

subsets of $\mathbb{Z}_n$ of the form $A + A$ for some $A \subset \mathbb{Z}_n$.

The factor of $n^{8/3}$ instead of n^2 is somewhat surprising, the lower bound in this theorem comes from a construction similar to the one in [1]. Take an arithmetic progression P, now of length $\frac{n - \log_2 n}{3}$, and two points $x, y \in \mathbb{Z}_n$ such that for each $S \subset P$ the set $(\{x, y\} + S) \setminus P + P$ is different. In this case one can prove that there are $\Omega(n^{8/3} 2^{n/3})$ choices for $A + A$ of this form, where $\Omega(n)$ choices come from choosing x, y, $\Omega(n^2)$ choices from choices of P and $\Omega(n^{-1/3} 2^{n/3})$ from choices of S. On the way to prove Theorem 3 we also prove a typical structure result, showing that almost all sets of the form $A + A$ come from a construction of this type.

Another direction one can generalize these theorems is to count sumsets of a given size. Notice that for a progression P, of length $\frac{m-k}{2}$, and $x \in \mathbb{Z}_n$ such that $(x+P) \cap (P+P) = \varnothing$, there are

$$\Omega\left(\binom{\frac{m-k}{2}}{k}\right)$$

choices for sets $B = A + A$, with $|B| = m$ and $|A| = k$, and $\{x\} \subset A \subset P \cup \{x\}$. This construction also gives essentially the tight bound.

Theorem 4. *Let n be a prime, $k \geqslant (\log n)^8$, $(5-c)k \leqslant m \leqslant \frac{2}{3}n + k$. There are*

$$\Theta\left(n^3\binom{\frac{m-k}{2}}{k}\right)$$

size m subsets of $\mathbb{Z}_n$ of the form $A + A$ for some $A \subset \mathbb{Z}_n$, with $|A| = k$.

We are also able to provide a suitable sharp typical structure result in this case.

The proof of these results uses a container theorem proved in previous work by the author [3], using the asymmetric container lemma. This container theorem needs to be combined with a suitable removal lemma for sets with bounded doubling that was proved by Shao [4]. Additionally instead of applying the container lemma normally, one needs first to use a stability theorem and find a large arithmetic progression inside $A + A$ to then be able to apply a container theorem on a 'smaller scale'. After applying the container theorem combined with a suitable removal lemma, the problem reduces to a combinatorial problem about counting neighborhoods in graphs.

References

[1] B. Green, I. Ruzsa, *Counting sumsets and sum-free sets modulo a prime*, Studia Sci. Math. Hungarica **41** (2004), 285–293.

[2] N. Alon, A. Granville, A. Ubis, *The number of sumsets in a finite field*, Bull. London Math. Soc **42** (2010): 784–794.

[3] M. Campos, *On the number of sets with a given doubling constant*, Israel Journal of Mathematics **236** (2020): 711–726.

[4] X. Shao, *On an almost all version of the Balog-Szemeredi-Gowers theorem*, Discrete Analysis (2019).

Sumset estimates in higher dimensions

David Conlon

(joint work with Jeck Lim)

Given two subsets A, B of an abelian group, the sumset $A + B$ is defined by

$$A + B = \{a + b : a \in A, b \in B\}$$

and the difference set $A - B$ is defined similarly. We discuss two recent results [1, 2] about the estimation of sumsets in $\mathbb{R}^d$.

First, we show that

$$|A-A| \geqslant \left(2d-2+\frac{1}{d-1}\right)|A| - (2d^2-4d+3)$$

for any sufficiently large finite subset A of $\mathbb{R}^d$ that is not contained in a translate of a hyperplane. By a construction of Stanchescu [4], this is best possible and thus resolves an old question first raised by Uhrin [5].

Second, we show that if $\mathcal{L}_1$ and $\mathcal{L}_2$ are linear transformations from $\mathbb{Z}^d$ to $\mathbb{Z}^d$ satisfying certain mild conditions, then, for any finite subset A of $\mathbb{Z}^d$,

$$|\mathcal{L}_1 A + \mathcal{L}_2 A| \geqslant \left(|\det(\mathcal{L}_1)|^{1/d} + |\det(\mathcal{L}_2)|^{1/d}\right)^d |A| - o(|A|).$$

This result corrects and confirms the two-summand case of a conjecture of Bukh and is best possible up to the lower-order term for many choices of $\mathcal{L}_1$ and $\mathcal{L}_2$. As an application, we prove a lower bound for $|A+\lambda\cdot A|$ when A is a finite set of real numbers and λ is an algebraic number. In particular, when λ is of the form $(p/q)^{1/d}$ for some $p, q, d \in \mathbb{N}$, each taken as small as possible for such a representation, we show that

$$|A+\lambda\cdot A| \geqslant (p^{1/d}+q^{1/d})^d |A| - o(|A|).$$

This is again best possible up to the lower-order term and extends a recent result of Krachun and Petrov [3] which treated the case $\lambda=\sqrt{2}$.

References

[1] D. Conlon and J. Lim, Difference sets in $\mathbb{R}^d$, preprint available at arXiv:2110.09053 [math.CO].

[2] D. Conlon and J. Lim, Sums of linear transformations, preprint available at arXiv:2203.09827 [math.CO].

[3] D. Krachun and F. Petrov, On the size of $A+\lambda A$ for algebraic λ, preprint available at arXiv:2010.00119 [math.CO].

[4] Y. V. Stanchescu, An upper bound for d-dimensional difference sets, *Combinatorica* **21** (2001), 591–595.

[5] B. Uhrin, Some useful estimations in the geometry of numbers, *Period. Math. Hungar.* **11** (1980), 95–103.

Improved lower bounds on the depth of polynomial size Boolean circuits for the independent set problem

David Gamarnik

(joint work with Aukosh Jagannath, Alex Wein)

We establish order $\log n/\log\log n$ lower bound on the depth of polynomial size Boolean circuits producing independent sets with certain proximity to optimality. There was a series of prior such bounds for this problem, and the state of the art is order o(log n/loglog n), obtained by Rossman [3].

Boolean circuit is a class of algorithms described as a directed graph where all nodes except bottom layer are associated with standard logical operations $\wedge, \vee, \neg$ [1],[2]. Two fundamental complexity measures associated with Boolean

circuits are the size, namely the number of nodes in the directed graph, and the depth, namely the length of the longest path. One of the fundamental goals in the algorithmic complexity theory is establishing lower bounds on the size and depth of circuits for solving various algorithmic decision and search problems. One of the highlights of the theory is that the size of any poly-size circuit which solves the parity of n bits problem is $\Theta(\log n/\log\log n)$. This is based on the celebrated Hastad's Switching Lemma. The progress on similar bounds for combinatorial optimization problems was significantly slower. A series of progressively stronger bounds resulted in the bound by Rossman [3] (see also [4] and [5]), who showed that a poly-size Boolean circuit detecting the presence of a clique in an n-node graph should have depth at least order $\log n/(k_n^2 \log\log n)$, where k_n is the clique size.

In this work we show that a poly-size Boolean circuit producing an independent set larger than half optimum should have depth at least order $\log n/\log\log n$, thus improving over the bound in [3]. Strictly speaking though our bounds are not comparable, since the bound by Rossman is for the decision problem and our bounds is for the search problem (namely the problem of actually exhibiting the independent set itself).

Our formal result is stated as follows. We denote by $\mathcal{C}(n, p(n), \rho)$ the family of all $m = \binom{n}{2}$-input, n-output Boolean circuits C which satisfy the following properties:

(a) The depth of C is at most $p(n)$.
(b) For every graph $y \in \{0,1\}^m$, the output $C(G)$ is an independent set in y which satisfies

$$|C(y)| \geqslant \rho \max_{I \in \mathcal{I}(y)} |I|.$$

Here y encodes the adjacency matrix of a graph with m edges and $\mathcal{I}$ is the set of all independent sets in the graph y. In other words, this is a family of circuits which, given a graph as an input, is required to produce an independent set in this graph with value that is at least a multiplicative constant ρ away from optimality. We note that this family is non-empty, as it is possible to implement an exhaustive search type algorithm that finds a largest independent set as a Boolean circuit with depth bound that does not depend on n. Our main result below states that if the depth $p(n)$ is at most roughly $\log n/\log\log n$, any circuit in the family must have super-polynomial size.

Theorem 1. *Fix any $\rho > 1/2$ and $\alpha > 0$. Let*

$$p(n) = \frac{\log n}{(1+\alpha)\log\log n}.$$

Then for all sufficiently large n, and all $C \in \mathcal{C}(n, p(n), \rho)$, the size $s(C)$ of the circuit C satisfies

$$s(C) \geqslant n^{(\log n)^{\alpha/3}}.$$

Our proof is obtained by analyzing the performance of Boolean circuits on sparse Erdös-Rényi graphs $\mathbb{G}(n, d/n)$ and using the Overlap Gap Property (OGP). Roughly speaking, the OGP says that for every pair of sufficiently large independent sets in $\mathbb{G}(n, d/n)$ their intersection normalized by the number of nodes is either at most ν_1 or at least ν_2 for some $\nu_1 < \nu_2$. We show that poly-size circuits with depth smaller than one stated in the theorem can be run twice so that to produce pairs of independent sets with "prohibited" intersection size in (ν_1, ν_2), thus obtaining a contradiction.

References

[1] Sanjeev Arora and Boaz Barak. *Computational complexity: a modern approach*. Cambridge University Press, 2009.

[2] Noga Alon and Joel H Spencer. *The probabilistic method*. John Wiley & Sons, 2004.

[3] Benjamin Rossman, *On the constant-depth complexity of k-clique*, Proceedings of the fortieth annual ACM symposium on Theory of computing, 2008, pp. 721–730.

[4] Benjamin Rossman. *Average-case complexity of detecting cliques*. PhD thesis, Massachusetts Institute of Technology, 2010.

[5] Yuan Li, Alexander Razborov, and Benjamin Rossman. On the AC0 complexity of subgraph isomorphism. *SIAM Journal on Computing*, 46(3):936–971, 2017.

Cliques in the random graph process and the hitting time of clique factors

Annika Heckel

(joint work with Marc Kaufmann, Noela Müller, Matija Pasch)

Given n, r, π, let $H_r(n, \pi)$ denote the *random r-uniform hypergraph* where each of the $\binom{n}{r}$ possible hyperedges of size r is present independently with probability π. The usual random graph $G(n, p)$ is then given by the special case $H_2(n, p)$.

In 1979 Shamir asked the following natural question: How large does $\pi = \pi(n)$ need to be for $H_r(n, \pi)$ to contain a perfect matching, that is, a collection of n/r vertex-disjoint hyperedges? (Here and in the following, we implicitly assume $r|n$ whenever necessary.)

A seemingly closely related question, posed by Rućinski [7] and Alon and Yuster [1], is: For which $p = p(n)$ does the random graph $G(n, p)$ contain an r-clique factor, that is, a collection of n/r vertex-disjoint cliques of size r?

These questions remained some of the most prominent open problems in random (hyper-)graph theory until they were solved up to constant factors by Johansson, Kahn and Vu in their seminal paper [3]. It had long been assumed that the main (or only) obstacle in finding a perfect matching in $H_r(n, \pi)$ were *isolated vertices*, that is, vertices not contained in any hyperedge. In the clique factor setting, the corresponding obstacle are vertices not contained in any r-clique. Let

$$\pi_0 = \pi_0(r) = \frac{\log n}{\binom{n-1}{r-1}} \quad \text{and} \quad p_0 = p_0(r) = \pi_0^{1/\binom{r}{2}}$$

then π_0 and p_0 are known to be sharp thresholds for the properties 'minimum degree at least 1' and 'every vertex is covered by an r-clique', respectively. Johansson, Kahn and Vu [3] showed that π_0 and p_0 are (weak) thresholds for the existence of a perfect matching in $H_r(n,\pi)$ and the existence of an r-clique factor in $G(n,p)$, respectively.

Recently, Jeff Kahn [4] proved that π_0 is in fact a *sharp* threshold for the existence of a perfect matching in $H_r(n,\pi)$. Indeed, he was able to confirm the conjecture that isolated vertices are essentially the only obstacle, and therein answer Shamir's question, in the strongest possible sense: Let $h_1,\dots,h_{\binom{n}{r}}$ be a uniformly random order of the $\binom{n}{r}$ possible hyperedges on the vertex set $V=[n]$, and set

$$H_t=(V,\{h_1,\dots,h_t\}),$$

which defines the *random hypergraph process* $\mathcal{H}=(H_t)_{t=0}^{\binom{n}{r}}$. Let

$$T=\min\{t: H_t \text{ has no isolated vertices}\}$$

be the hitting time of the last isolated vertex disappearing. In [5], Kahn proved that this hitting time whp[1] coincides with the hitting time for a perfect matching.

Theorem 1 ([5]). *Let $n\in r\mathbb{Z}$, then with high probability,*

$$H_T \text{ has a perfect matching.}$$

Can we get similarly strong answers to the clique factor question? Here the question seems much harder because, unlike hyperedges in the random hypergraph, cliques do not appear independently from each other. However, for sharp thresholds, it has indeed been possible to *reduce* the clique factor problem to the perfect matching problem, using the following coupling result of Riordan (for $r\geqslant 4$) and the first author (for $r=3$):

Theorem 2 ([6, 2]). *Let $r\geqslant 3$. There are constants $\varepsilon(r),\delta(r)>0$ such that, for any $p=p(n)\leqslant n^{-2/r+\varepsilon}$, letting $\pi=p^{\binom{n}{2}}(1-n^{-\delta})$, we may couple the random graph $G=G(n,p)$ with the random r-uniform hypergraph $H=H_r(n,\pi)$ so that, whp, for every hyperedge in H there is a copy of K_r in G with the same vertex set.*

Corollary 3. *p_0 is a sharp threshold for the existence of an r-clique factor.*

In the same spirit, we wish to directly transfer Theorem 1 to the random graph process setting, showing its clique factor analogue. Let $\mathcal{G}=(G_t)_{t=0}^{\binom{n}{2}}$ be the random graph process, which is given by the random hypergraph process in the special case $r=2$. Define

$$T_G=\min\{t:\ \text{every vertex in } G_t \text{ is contained in at least one } r\text{-clique}\}$$

as the hitting time of an r-clique cover. Then to transfer Kahn's hitting time result to the clique factor setting, we need to find a copy of H_T within the cliques of G_{T_G}.

[1]We say that a sequence of events $(E_n)_n$ holds with high probability (whp) if $\mathbf{P}(E_N)\to\infty$ as $n\to\infty$.

Theorem 4. *Let $r \geqslant 4$. We may couple the random graph process $\mathcal{G} = (G_t)_{t=0}^{\binom{n}{2}}$ with the random r-uniform hypergraph process $\mathcal{H} = (H_t)_{t=0}^{\binom{n}{r}}$ so that, whp, for every hyperedge in H_T there is a clique in G_{T_G} on the same vertex set.*

Corollary 5. *Let $r \geqslant 4$ and $n \in r\mathbb{Z}$, then with high probability,*

$$G_{T_G} \text{ has an } r\text{-clique factor.}$$

One might wonder whether the coupling in Theorem 4 is such that up to time T and T_G, the hyperedges appear in the same order in $\mathcal{H}$ as they do as r-cliques in $\mathcal{G}$. The answer to this is 'almost' — we can match up the order of the hyperedges and cliques as long as they are edge-disjoint, but hyperedges which share two vertices with another hyperedge may appear at a different time in $\mathcal{H}$ as they do as cliques in $\mathcal{G}$ (but with high probability only about $\log^2 n$ such hyperedges exist in H_T).

The proof starts with the coupling of the random graph and hypergraph from Theorem 2 with p and π slightly larger than p_0 and π_0, respectively. We then carefully couple uniform orders of the edges and hyperedges in such a way that we are able to control most pairs of hyperedges sharing two vertices, and so that the hitting times T and T_G match. We then require a series of careful further couplings in order to treat the remaining pairs of hyperedges sharing two vertices.

Theorem 4 is stated for the case $r \geqslant 4$ because for $r = 3$ similar technical issues arise as in Theorem 2, which in that case were later resolved with additional arguments in [2]. It seems likely that a similar approach can be used to adapt the proof of Theorem 4 to the case $r = 3$.

References

[1] Noga Alon and Raphael Yuster. Threshold functions for h-factors. *Combinatorics, Probability and Computing*, 2(2):137–144, 1993.

[2] Annika Heckel. Random triangles in random graphs. *Random Structures & Algorithms*, 2021.

[3] Anders Johansson, Jeff Kahn, and Van Vu. Factors in random graphs. *Random Structures & Algorithms*, 33(1):1–28, 2008.

[4] Jeff Kahn. Asymptotics for Shamir's problem. *arXiv preprint arXiv:1909.06834*, 2019.

[5] Jeff Kahn. Hitting times for Shamir's problem. *arXiv preprint arXiv:2008.01605*, 2020.

[6] Oliver Riordan. Random cliques in random graphs. *arXiv preprint arXiv:1802.01948*, 2018.

[7] Andrzej Ruciński. Matching and covering the vertices of a random graph by copies of a given graph. *Discrete mathematics*, 105(1-3):185–197, 1992.

The least singular value of a random symmetric matrix

Matthew Jenssen

(joint work with Marcelo Campos, Marcus Michelen, Julian Sahasrabudhe)

Let A be a $n \times n$ random symmetric matrix whose entries on and above the diagonal $(A_{i,j})_{i \leqslant j}$ are i.i.d. with mean 0 and variance 1. This matrix model, sometimes called the Wigner matrix ensemble, was introduced in the 1950s in the seminal work of Wigner [13], who established the famous "semi-circular law" for the eigenvalues of such matrices.

Here we study the lower tail of the *least singular value* of A, which we denote by $\sigma_{\min}(A)$. We prove a bound on the lower tail of $\sigma_{\min}(A)$ for all random symmetric matrices A with i.i.d. *subgaussian* entries which is optimal up to constants. This confirms a folklore conjecture, explicitly stated by Vershynin in [12].

Theorem 1. *Let ζ be a subgaussian random variable with mean 0 and variance 1 and let A be a $n \times n$ random symmetric matrix whose entries above the diagonal $(A_{i,j})_{i \leqslant j}$ are independent and distributed according to ζ. Then for every $\epsilon \geqslant 0$,*

$$\mathbb{P}_A(\sigma_{\min}(A) \leqslant \epsilon n^{-1/2}) \leqslant C\epsilon + e^{-cn},$$

where $C, c > 0$ depend only on ζ.

This theorem is sharp up to the value of the constants $C, c > 0$. We note that the special case $\epsilon = 0$ tells us that the singularity probability of any random symmetric A with subgaussian entry distribution is exponentially small, generalising our previous work [3] on the case where ζ is uniform in $\{-1, 1\}$. A key ingredient in the proof of Theorem 1 is a new result on *repeated* eigenvalues of A, which in particular resolves a conjecture of Nguyen, Tao and Vu [9].

Theorem 2. *Let ζ be a subgaussian random variable with mean 0 and variance 1 and let A be a $n \times n$ random symmetric matrix where $(A_{i,j})_{i \leqslant j}$ are independent and distributed according to ζ. Then*

$$\mathbb{P}(A \text{ has a repeated eigenvalue }) \leqslant e^{-cn},$$

where $c > 0$ is a constant depending only on ζ.

We in fact prove a more refined version Theorem 2 which gives an upper bound on the probability that two eigenvalues of A fall into an interval of length $\epsilon n^{-1/2}$.

Previous results on the lower tail of $\sigma_{\min}(A)$ were obtained by Nguyen [8], Vershynin [12] and Jain, Sah and Sawhney [7] with the state of the art being the following bound from [7]: $\mathbb{P}_A(\sigma_{\min}(A) \leqslant \epsilon n^{-1/2}) \leqslant C\epsilon^{1/8} + e^{-c\sqrt{n}}$.

For large ϵ, e.g. $\epsilon \geqslant n^{-c}$, another very different and powerful set of techniques have been developed for this problem. The works of Tao and Vu [10, 11], Erdős, Schlein and Yau [4, 5, 6], and specifically Bourgade, Erdős, Yau and Yin [1] tell us that $\mathbb{P}(\sigma_{\min}(A) \leqslant \epsilon n^{-1/2}) \leqslant \epsilon + o(1)$, thus obtaining the correct dependence on ϵ asymptotically. We note however that it appears these techniques are limited to the large ϵ regime and different ideas are required for $\epsilon < n^{-C}$, and certainly for ϵ as small as $e^{-\Theta(n)}$. We refer the reader to our full paper [2] for more discussion.

References

[1] P. Bourgade, L. Erdős, H.-T. Yau, and J. Yin. Fixed energy universality for generalized Wigner matrices. *Comm. Pure Appl. Math.*, 69(10):1815–1881, 2016.

[2] M. Campos, M. Jenssen, M. Michelen, and J. Sahasrabudhe. The least singular value of a random symmetric matrix. *https://arxiv.org/abs/2203.06141*, 2022.

[3] M. Campos, M. Jenssen, M. Michelen, and J. Sahasrabudhe. The singularity probability of a random symmetric matrix is exponetially small. *https://arxiv.org/abs/2105.11384*, 2021.

[4] L. Erdős, B. Schlein, and H.-T. Yau. Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. *Ann. Probab.*, 37(3):815–852, 2009.
[5] L. Erdős, B. Schlein, and H.-T. Yau. Local semicircle law and complete delocalization for Wigner random matrices. *Comm. Math. Phys.*, 287(2):641–655, 2009.
[6] L. Erdős, B. Schlein, and H.-T. Yau. Wegner estimate and level repulsion for Wigner random matrices. *International Mathematics Research Notices*, 2010(3):436–479, 2010.
[7] V. Jain, A. Sah, and M. Sawhney. On the smallest singular value of symmetric random matrices. *arXiv preprint arXiv:2011.02344*, 2020.
[8] H. H. Nguyen. Inverse Littlewood-Offord problems and the singularity of random symmetric matrices. *Duke Math. J.*, 161(4):545–586, 2012.
[9] H. Nguyen, T. Tao, and V. Vu. Random matrices: tail bounds for gaps between eigenvalues. *Probability Theory and Related Fields*, 167(3):777–816, 2017.
[10] T. Tao and V. Vu. Random matrices: universality of local eigenvalue statistics. *Acta Math.*, 206(1):127–204, 2011.
[11] T. Tao and V. Vu. Random matrices: a general approach for the least singular value problem. *preprint*, 2008.
[12] R. Vershynin. Invertibility of symmetric random matrices. *Random Structures Algorithms*, 44(2):135–182, 2014.
[13] E. P. Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, 67(2):325–327, 1958.

The largest product-free subsets of the alternating groups

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

A subset A of a group G is called product-free if there is no solution to $a = bc$ with a, b, c all in A. It is easy to see that the largest product-free subset of the symmetric group S_n is obtained by taking the set of all odd permutations, i.e. $S_n \backslash A_n$, where A_n is the alternating group. By contrast, it is a long-standing open problem to find the largest product-free subset of A_n. This problem has been recently highlighted by Ben Green [4], who credits Edward Crane for the conjectured extremal examples, which are families of the form

$$F_I^x := \{\pi : \pi(x) \in I, \pi(I) \cap I = \varnothing\}$$

and their inverses. Writing μ for uniform measure on A_n and $|I| = t\sqrt{n}$, one can calculate $\mu(F_I^x) \approx te^{-t^2} n^{-1/2}$, which suggests the conjecture that the maximum measure should be $\Theta(n^{-1/2})$, and more precisely that it should be $\sim 1/\sqrt{2en}$. Improving earlier bounds of Kedlaya [6] and Gowers [3], the conjecture was proved up to logarithmic factors by Eberhard [1], who showed that any product-free $A \subseteq A_n$ has $\mu(A) = O(n^{-1/2} \log^{7/2} n)$. Our main result here answers the question completely, as follows.

Theorem 1. *Suppose n is sufficiently large and $A \subseteq A_n$ is a product-free subset of maximum size. Then A or A^{-1} is some F_I^x.*

We also obtain the following '99% stability result', showing that any large product-free subset of A_n is essentially contained in an F_I^x or its inverse. Our

stability result holds for sets whose measure is much smaller than the extremal family.

Theorem 2. *Suppose n is sufficiently large and $A \subseteq A_n$ is a product-free set with $\mu(A) \geqslant n^{-0.66}$. Then there is some F_I^x such that $\mu(A \backslash F_I^x) < n^{-0.66}$ or $\mu(A^{-1} \backslash F_I^x) < n^{-0.66}$.*

We also prove that if $A \subseteq A_n$ is a product-free set with size very close to the maximum then A or A^{-1} is contained in some F_I^x.

Theorem 3. *There exists an absolute constant c such that if n is sufficiently large and $A \subseteq A_n$ is a product-free set with $\mu(A) > \max_{I,x} \mu(F_I^x) - \frac{c}{n}$ then there is some F_I^x such that $A \subseteq F_I^x$ or $A^{-1} \subseteq F_I^x$.*

We also study the following '1% stability' problem. Suppose that $A \subseteq A_n$ is a product-free set of density $> 1/n^C$ for an absolute constant C. What can be said about the structure of A?

The structures appearing in the answer to this problem are as follows. A set of the form $\mathcal{D}_{i \to j} = \{\sigma \in A_n : \sigma(i) = j\}$ is called a *dictator*. Let $D_1, \ldots D_t$ be distinct dictators that have a nonempty intersection; we call their intersection a *t-umvirate*. Equivalently, a t-umvirate corresponds to ordered sets $I = (i_1, \ldots, i_t), J = (j_1, \ldots, j_t)$ and is given by letting $\mathcal{U}_{I \to J}$ be the set of permutations that send each i_k to j_k.

Our next theorem shows that any product-free set that is somewhat dense has some local structure. This is analogous to the strong local structure exhibited by the extremal families F_I^x, which have $\Theta(1)$ density inside each dictator $1_{x \to i}$ with $i \in I$, as when $|I| = \Theta(\sqrt{n})$ a random permutation sends I to its complement with probability $\Theta(1)$. We show that a similar, albeit weaker, phenomenon holds for product-free sets with any polynomial density that can be much smaller than that in the extremal examples: such sets have a density bump inside a t-umvirate.

Theorem 4. *Fix $r \in \mathbb{N}$, suppose n is sufficiently large, and $A \subseteq A_n$ is product-free with$\mu(A) > n^{-r}$. Then there exists some t-umvirate with $t \leqslant 4r$ in which A has density at least $n^{t/4}\mu(A)$.*

Our approach can be viewed as an non-abelian analogue of Roth's bound for sets of integers with no three-term arithmetic progression, whereby we improve the earlier approaches of Gowers and Eberhard by establishing a form of the 'Structure versus Randomness' dichotomy. We achieve this by exploiting some recent theory developed by Filmus, Kindler, Lifshitz and Minzer [2] for global hypercontractivity on the symmetric group (building on the corresponding theory for product spaces developed by Keevash, Lifshitz, Long and Minzer [5]), and by also developing some further theory of the 'Cayley operators' associated to global subsets.

The main steps in the proof of our main theorem are as follows.

(1) Achieving dictatorial structure: We show that A has large density inside many dictators. In fact, we show that in some sense, the product freeness of A is completely explained by its densities inside dictators.

(2) Achieving star structure: We then upgrade our dictatorial structure into a tighter star structure. We find some S that is either a star or an inverse star such that $|A\backslash S|$ is small and A has significant density in each restriction defined by S.
(3) Bootstrapping: Using the approximate star structure, we deduce our exact results from further stability analysis showing that any small deviation from the structure leads to a suboptimal configuration.

References

[1] Sean Eberhard, Product mixing in the alternating group, Discrete Analysis (2016).
[2] Yuval Filmus, Guy Kindler, Noam Lifshitz and Dor Minzer, Hypercontractivity on the symmetric group, arXiv:2009.05503.
[3] W. Tim Gowers, Quasirandom Groups, Combinatorics, Probability and Computing (2008).
[4] Ben Green, 100 open problems, manuscript.
[5] Peter Keevash, Noam Lifshitz, Eoin Long and Dor Minzer, Global hypercontractivity and its applications, arXiv:2103.04604.
[6] Kiran S. Kedlaya, Large product-free subsets of finite groups, J. Combinatorial Theory Series A (1997).

Enumerating matroids, linear spaces and clique-decompositions

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

Matroids (also sometimes known as *combinatorial geometries*) are fundamental objects that abstract the combinatorial properties of linear independence in vector spaces. Specifically, a matroid consists of a ground set E and a collection $\mathcal{I}$ of subsets of E called independent sets[1]. The defining properties of a matroid are that:

- the empty set is independent (that is, $\varnothing \in \mathcal{I}$);
- subsets of independent sets are independent (if $A' \subseteq A \subseteq E$ and $A \in \mathcal{I}$, then $A' \in \mathcal{I}$);
- if A and B are independent sets, and $|A| > |B|$, then an independent set can be constructed by adding an element of $A\backslash B$ to B (there is $a \in A\backslash B$ such that $B \cup \{a\} \in \mathcal{I}$).

Observe that any finite set of elements in a vector space (over any field) naturally gives rise to a matroid, though most matroids do not arise this way. The *rank* of a matroid is the maximum size of an independent set.

Enumeration of matroids is a classical topic, though the state of our knowledge is rather incomplete. Some early upper and lower bounds on the total number of matroids on a ground set of size n were obtained in the 1970s by Piff and Welsh [15], Piff [14] and Knuth [8], and these bounds were improved only recently

[1]Instead of defining a matroid by its collection of independent sets, some authors prefer to define a matroid by some other (equivalent) data, such as its collection of *flats*, its collection of *hyperplanes*, or its *rank function*. See for example [19, 12] for a more thorough introduction to matroids and their various definitions.

by Bansal, Pendavingh, and van der Pol [2]. It is also of interest to enumerate matroids of fixed rank: let $m(n,r)$ be the number of rank-r matroids on a ground set of size n. It is trivial to see that $m(n,1) = 2^n - 1$, and it is also possible to prove the exact identity $m(n,2) = b(n+1) - 2^n$, where $b(m)$ is the mth Bell number (which counts the number of partitions of an m-element set). This identity seems to have been first observed by Acketa [1].

For $r \geqslant 3$, an exact expression for $m(n,r)$ in terms of well-known functions does not seem to be possible[2], but after some exciting recent developments, rather precise asymptotic expressions have become available. First, Pendavingh and van der Pol [13] observed that (for constant $r \geqslant 1$) the lower bound $m(n,r) \geqslant (e^{1-r}n + o(n))^{n^{r-1}/r!}$ follows from Keevash's breakthrough work [6, 7] on existence and enumeration of *combinatorial designs*. They also proved an upper bound of the form $m(n,r) \leqslant (en + o(n))^{n^{r-1}/r!}$. Even more recently, van der Hofstad, Pendavingh and van der Pol [17] closed the gap for all $r \neq 3$, proving that $m(n,r) = (e^{1-r}n + o(n))^{n^{r-1}/r!}$ for constant $r \geqslant 4$. In the remaining case $r = 3$ they were able to prove $m(n,3) \leqslant (ne^{1+\beta} + o(n))^{n^2/6} \approx (1.4n)^{n^2/6}$, where $-0.67 < \beta < -0.65$ is the solution to a certain variational problem. In this paper, we close the gap completely in this case $r = 3$.

Theorem 1.

$$m(n,3) = \Big(\frac{1+\sqrt{3}}{2}e^{\sqrt{3}/2-3}n + o(n)\Big)^{n^2/6} \approx (0.16169n)^{n^2/6}.$$

In particular, Theorem 1 shows that the lower bound $m(n,3) \geqslant (e^{-2}n + o(n))^{n^2/6}$ obtainable from Keevash's results is far from sharp. This confirms a conjecture in [17] (and disproves the earlier [18, Conjecture 8.2.9]).

In fact, Theorem 1 is really a corollary of the following theorem, estimating the number of *linear spaces* on a set of n points. In incidence geometry, a linear space on a point set P is a collection of subsets of at least two points of P (called *lines*) such that each pair of points lies in a unique line (see for example [3, 16] for more on linear spaces). For reasons that will become clear in a moment, we denote the number of linear spaces on a set of n points by $p(n,3)$.

Theorem 2.

$$p(n,3) = \Big(\frac{1+\sqrt{3}}{2}e^{\sqrt{3}/2-3}n + o(n)\Big)^{n^2/6} \approx (0.16169n)^{n^2/6}.$$

We remark that one may also be interested in linear spaces in which no line has exactly 2 points (these are called *proper* linear spaces). It should be possible to adapt our proof to show that the expression in Theorem 2 is also a valid estimate for the number of proper linear spaces on a set of n points (though this would require some rather deep machinery due to Keevash [6] and McKay and Wormald [11]).

[2]Though, a lot of computational work has been done for small n, r, and there are many conjectures about the relations between the different $m(n,r)$; see for example [5] and the index for matroids on the On-Line Encyclopedia of Integer Sequences [4].

To explain the connection between Theorems 2 and 1 we need to make a few more definitions. A *d-partition* (or *generalised partition of type d*) of a ground set E is a collection of subsets of E (called *parts*) each having size at least d, such that every subset of d elements of E is contained in exactly one of the parts. So, a 1-partition is an ordinary partition, and a 2-partition is a linear space. For any $r \geqslant 2$, there is a correspondence between the set of $(r-1)$-partitions of E and the set of so-called *paving matroids* of rank r on the ground set E. Namely, a paving matroid of rank r is a matroid for which its set of *hyperplanes* (maximal subsets with rank $r-1$) form an $(r-1)$-partition of its ground set. See for example [19, Section 3] for more details.

For $r \geqslant 2$ let $p(n,r)$ be the number of paving matroids of rank r, or equivalently the number of $(r-1)$-partitions, on a ground set of size n. Given the above correspondence, we trivially have $p(n,r) \leqslant m(n,r)$, and it was proved by Pendavingh and van der Pol [13, Theorem 3] that $p(n,r) \leqslant m(n,r) \leqslant p(n,r)^{1+O(1/n)}$ for constant r. So Theorem 1 is a direct consequence of Theorem 2. One may also wish to use the language of graph theory: note that a linear space on a set of n points is precisely equivalent to a *clique-decomposition* of the complete graph K_n (meaning, a decomposition of the edges of K_n into nonempty cliques of arbitrary sizes).

References

[1] Dragan M. Acketa, *On the enumeration of matroids of rank* 2, Univ. u Novom Sadu Zb. Rad. Prirod.-Mat. Fak. **8** (1978), 83–90.

[2] Nikhil Bansal, Rudi A. Pendavingh, and Jorn G. van der Pol, *On the number of matroids*, Combinatorica **35** (2015), 253–277.

[3] Lynn Margaret Batten and Albrecht Beutelspacher, *The theory of finite linear spaces*, Cambridge University Press, Cambridge, 1993.

[4] OEIS Wiki Contributors, *Index to OEIS: Matroids, sequences related to*, `http://oeis.org/wiki/Index_to_OEIS:_Section_Mat#matroid`.

[5] W. M. B. Dukes, *On the number of matroids on a finite set*, Sém. Lothar. Combin. **51** (2004/05), Art. B51g, 12.

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

[7] Peter Keevash, *Counting designs*, J. Eur. Math. Soc. (JEMS) **20** (2018), 903–927.

[8] Donald E. Knuth, *The asymptotic number of geometries*, J. Combinatorial Theory Ser. A **16** (1974), 398–400.

[9] Matthew Kwan, *Almost all Steiner triple systems have perfect matchings*, Proc. Lond. Math. Soc. (3) **121** (2020), 1468–1495.

[10] Nathan Linial and Zur Luria, *An upper bound on the number of Steiner triple systems*, Random Structures Algorithms **43** (2013), 399–406.

[11] Brendan D. McKay and Nicholas C. Wormald, *Asymptotic enumeration by degree sequence of graphs of high degree*, European J. Combin. **11** (1990), 565–580.

[12] James Oxley, *Matroid theory*, second ed., Oxford Graduate Texts in Mathematics, vol. 21, Oxford University Press, Oxford, 2011.

[13] Rudi Pendavingh and Jorn van der Pol, *Enumerating matroids of fixed rank*, Electron. J. Combin. **24** (2017), Paper No. 1.8, 28.

[14] M. J. Piff, *An upper bound for the number of matroids*, J. Combinatorial Theory Ser. B **14** (1973), 241–245.

[15] M. J. Piff and D. J. A. Welsh, *The number of combinatorial geometries*, Bull. London Math. Soc. **3** (1971), 55–56.

[16] Ernest E. Shult, *Points and lines*, Universitext, Springer, Heidelberg, 2011.
[17] Remco van der Hofstad, Rudi Pendavingh, and Jorn van der Pol, *The number of partial Steiner systems and d-partitions.*
[18] J.G. van der Pol, *Large matroids: enumeration and typical properties*, Ph.D. thesis, Mathematics and Computer Science, September 2017, Proefschrift.
[19] D. J. A. Welsh, *Matroid theory*, L. M. S. Monographs, No. 8, Academic Press [Harcourt Brace Jovanovich, Publishers], London-New York, 1976.

Hypergraphs with no tight cycles

Shoham Letzter

1. Introduction

A basic fact in graph theory asserts that the maximum number of edges in a graph on n vertices that has no cycles is $n-1$ (and maximisers are trees). It is natural to consider an analogous question for hypergraphs: what is the maximum numebr of edges in an r-uniform hypergraph (henceforth *r-graph*) that contains no cycles? Unlike the graph case, for $r \geqslant 3$ there are various natural notions of cycles in r-graphs. Here we focus on the notion of tight cycles. An r-uniform *tight cycle* is a sequence $v_1, \dots, v_k$ of at least $r+1$ distinct vertices satisfying that $v_i \dots v_{i+r-1}$ is an edge for every $i \in [k]$ (indices are added modulo r). Denote by $f_r(n)$ the maximum possible number of edges in an r-graph on n vertices that contains no tight cycles. So we will be interested in estimating $f_r(n)$ for $r \geqslant 3$.

Consider the r-graph on vertex set $[n]$ whose edges are r-subsets of $[n]$ that contain 1. This hypergraph has $\binom{n-1}{r-1}$ edges and contains no tight cycles (indeed, it is easy to see that the edges of any tight cycle cannot be covered by a single vertex). In particular, $f_r(n) \geqslant \binom{n-1}{r-1}$. Sós and, independently, Verstraëte (see [4, 6]) conjectured that this bound is tight for large enough n. This was refuted by Huang and Ma [2], who showed that, for $r \geqslant 3$, there exists $c = c(r) > 1$ such that $f_r(n) \geqslant c \cdot \binom{n-1}{r-1}$ for large n. A clever construction due to B. Janzer [3] implies the best known lower bound to date, namely that $f_r(n) \geqslant \Omega(n^{r-1} \log n / \log\log n)$ for $r \geqslant 3$.

Notice that $f_r(n) \leqslant \binom{n}{r}$, trivially. An old paper of Erdős [1] implies the following improvement of this trivial bound: $f_r(n) = O(n^{r-2^{-(r-1)}})$. For $r = 3$, this was improved by Verstraëte (unpublished) who showed $f_3(n) = O(n^{5/2})$. Recently, these bounds were significantly improved by Sudakov and Tomon [5], who proved that $f_r(n) \leqslant n^{r-1} e^{O(\sqrt{\log n})}$ for $r \geqslant 3$, establishing that $f_r(n) = n^{r-1+o(1)}$. Our contribution is to sharpen the error term $e^{O(\sqrt{\log n})}$ in Sudakov and Tomon's bound, as follows.

Theorem 1. *Let $r \geqslant 3$. Then $f_r(n) = O(n^{r-1}(\log n)^5)$.*

2. Proof sketch

Our proof builds on ideas Sudakov and Tomon's work [5]. They introduce the notions of r-line-graphs, which are graphs that correspond naturally to r-partite r-graph, and expansion in such graphs. They show that, given a dense enough r-partite r-graph $\mathcal{H}$, the r-line-graph that corresponds to $\mathcal{H}$ contains a dense expander G. Next, they define σ-paths and σ-cycles, which correspond to tight paths and cycles in the original hypergraph $\mathcal{H}$. It thus suffices to show that every r-line-graph which is a dense expander contains a σ-cycle. Sudakov and Tomon are not able to prove this. Instead, they show that every expander contains either a σ-cycle or a very dense subgraph, and proceed via a density increment argument.

Our main contribution is to show that every r-line-graph which is a dense expander indeed contains a σ-cycle. A key step in our proof is to show that in such an expander G, for every vertex $x \in V(G)$, almost every other vertex $y \in V(G)$ can be reached from x via a short σ-path $P(x, y)$ in a 'robust' way, meaning that no vertex in the underlying r-graph $\mathcal{H}$ (except for those in the edge corresponding to x) meets too many of the paths $P(x, y)$. If the robustness requirement is dropped, we obtain a lemma from [5]. To prove the robust version, we use the non-robust version from [5] as a black box, along with another lemma from the same paper, which asserts that the removal of a small number of vertices from the underlying r-graph $\mathcal{H}$ does not ruin expansion.

To find a σ-cycle, let $P(x, y)$ be paths as above, defined for almost every $x, y \in V(G)$. Note that while we are guaranteed that, for every $x \in V(G)$, no vertex v of $\mathcal{H}$ meets too many paths $P(x, y)$, we do not have any control over the number of times v meets a path $P(x, y)$, for a given y. Nevertheless, since the paths $P(x, y)$ are short, for every $y \in V(G)$ there are few vertices in $\mathcal{H}$ that meet many path $P(x, y)$; denote the set of such vertices in $\mathcal{H}$ by $F(y)$. Using tools mentioned above, for every y and almost every x there is a short σ-path $Q(y, x)$ from y to x that avoids $F(y)$. To complete the proof, we note that the robustness implies that for almost every $x, y \in V(G)$ the path $Q(y, x)$ is defined, and there are linearly many $z \in V(G)$ for which $P(x, z)P(z, y)$ is a σ-path from x to y. Using robustness and the choice of $Q(y, x)$, the concatenation $P(x, z)P(z, y)Q(y, x)$ is a σ-cycle for almost every $x, y \in V(G)$ and linearly many $z \in V(G)$.

References

[1] P. Erdős, *On extremal problems of graphs and generalized graphs*, Israel J. Math. **2** (1964), 183–190.

[2] H. Huang and J. Ma, *On tight cycles in hypergraphs*, SIAM J. Discr. Math. **33** (2019), 230–237.

[3] B. Janzer, *Large hypergraphs without tight cycles*, arXiv:2012.07726 (2020).

[4] D. Mubayi, O. Pikhurko, and B. Sudakov, *Hypergraph Turán problem: Some open questions*, AIM workshop problem lists, manuscript, 2011.

[5] B. Sudakov and I. Tomon, *The extremal number of tight cycles*, Int. Math. Res. Not. (2021), to appear.

[6] J. Verstraëte, *Extremal problems for cycles in graphs*, Recent trends in combinatorics, Springer, 2016, pp. 83–116.

Uncommon and Sidorenko systems of equations

Anita Liebenau
(joint work with Nina Kamčev, Natasha Morrison)

In 1962, Erdős [4] conjectured that the number of monochromatic k-cliques K_k in any two-colouring of the edges of K_n is asymptotically minimised by a random two-colouring. A graph H with this property is called *common*, i.e. if the number of monochromatic copies of H in any two-colouring of the edges of K_n, is asymptotically minimised by a random two-colouring. In 1980, Burr and Rosta [1] extended Erdős' conjecture to arbitrary H. This conjecture was independently disproved by Thomason [17], who also disproved Erdős' conjecture by showing that K_4 is uncommon, and by Sidorenko [16]. Later, Jagger, Šťovíček and Thomason [9] proved that common graphs are in fact quite rare by showing that any graph containing K_4 as a subgraph, is uncommon. Although the conjectures of Erdős and of Burr and Rosta are false in this very strong sense, the desire to characterise common and uncommon graphs continues to this day, see for example see [8, 10] and the references therein.

The phenomenon of a random colouring minimising the number of monochromatic structures was also investigated in the arithmetic setting. In 1996, Graham, Rödl and Ruciński [7] asked about the minimal proportion of monochromatic Schur triples, i.e. solutions to $x + y = z$, in two-colourings of $[n]$. This number was independently shown [3, 13, 15] to be far below the number expected in a random colouring. In contrast to this, in the finite Abelian group setting, Cameron, Cilleruelo and Serra [2] showed that the random colouring minimises the number of monochromatic solutions to any equation in an odd number of variables.

Inspired by the above-mentioned results for graphs and arithmetic structures, Saad and Wolf [14] initiated a systematic study of the number of monochromatic solutions to linear patterns in more generality. From now on, we work over the finite field $\mathbb{F}_q$, where q is a prime power, following e.g. [5, 14]. Consider a linear map $L : (\mathbb{F}_q^n)^k \to (\mathbb{F}_q^n)^m$ with coefficients in $\mathbb{F}_q$. Say that L is *common* if the density of monochromatic solutions to the system of linear equations $L(\mathbf{x}) = 0$ in any two-colouring of $\mathbb{F}_q^n$ is asymptotically minimised by the expected density of solutions in a random two-colouring.

Earlier work focused on systems consisting of a single equation $a_1x_1 + \cdots + a_kx_k = 0$ with coefficients $a_i \in \mathbb{F}_q^\times = \mathbb{F}_q \backslash \{0\}$. As mentioned above, whenever k is odd, such an equation is common [2]. For even k, Saad and Wolf [14] proved that the equation is common whenever the coefficients can be partitioned into pairs, each pair summing to zero. They conjectured that this sufficient condition is also necessary, which was confirmed by Fox, Pham and Zhao [5]. Hence homogeneous equations of this form are completely characterised.

Fox, Pham and Zhao [5] asked about a similar characterisation for common *systems* of equations, hoping that it might lead to a better understanding of the analogous properties for graphs and hypergraphs, but they note that they do not have a guess for such a characterisation. We make significant progress towards

a classification for common systems of two or more linear equations. As a first step, we show that in many cases, the presence of a particular small subsystem is sufficient for a system to be uncommon. This can be seen as an analogue of the result of Jagger, Šťovíček and Thomason [9] mentioned above, showing that any graph containing a copy of K_4 is uncommon.

In order to state our main results, we now introduce some definitions. Let L be a collection of m linear forms $L_1, \dots, L_m$ in k variables with coefficients in $\mathbb{F}_q$. For an ℓ-variable system L', we say that L *induces* L' *as a subsystem* if there exists a subset $\{i_1, \dots, i_\ell\} \subseteq [k]$ such that $L(x_1, \dots, x_k) = 0$ implies that $L'(x_{i_1}, \dots, x_{i_\ell}) = 0$. A system L is called an $(m \times k)$-*system* if the rows of the coefficient matrix of L are linearly independent over $\mathbb{F}_q$. Finally, L is *degenerate* if it induces the equation $x_i - x_j = 0$, for some $i \neq j$, and *non-degenerate* otherwise.

Theorem 1. *Let q be an odd prime power, let $2 \leqslant m < k$ be integers and let L be a non-degenerate $(m \times k)$-system over $\mathbb{F}_q$. If L induces a (2×4)-system, then L is uncommon.*

We remark that an $(m \times k)$-system always satisfies $m \leqslant k$, and that $m = k$ implies that L is trivially common. Furthermore, the non-degeneracy condition on the subsystem (implied by the non-degeneracy of L) is required, as, for example, the system $x_1 = x_2 = \ldots = x_k$ is common. The commonness of a degenerate system is determined by the commonness of a non-degenerate subsystem.

For any prime $p > 3$ and any power q of p, an arithmetic progression of length four (4-AP) is a non-degenerate (2×4)-system over $\mathbb{F}_q$, and so we obtain the following corollary which resolves a question of Saad and Wolf [14].

Corollary 2. *Let q be an odd prime power, let $2 \leqslant m < k$ be integers and let L be a non-degenerate $(m \times k)$-system over $\mathbb{F}_q$. If L induces an arithmetic progression of length four, then L is uncommon.*

Since an arithmetic progression of length $k \geqslant 4$ induces a 4-AP, our result implies that any k-AP, and any system inducing a k-AP is uncommon. We remark that Corollary 2 has been independently proved by Versteegen [18].

Earlier results on the uncommonness of 4APs [14, 19] used geometric intuition relying on strong structural properties of arithmetic progressions, but perhaps surprisingly we do not utilise these properties here. Excitingly, Theorem 1 applies to *any* non-degenerate (2×4)-system. Not only does this determine the uncommonness of a large and general family of systems, but as we only use weak conditions on the structure of the solution space, there is reason to believe that our techniques could be used to characterise other general families of systems.

Theorem 1 is obtained as a consequence of a much more general (and more technical) result. We refer the reader to Theorem 3.1 in [11] for the exact statement of the theorem. Roughly speaking, this result provides a sufficient condition for a system to be uncommon based solely on particular 'critical' subsystems which all have rank at most two.

Finally, we introduce a parameter that provides some guidance towards a characterisation. The *length* of an equation E is the number of variables in E with

non-zero coefficients. Given a system L, let $s(L)$ denote the minimal length of an equation induced by L.

Theorem 3. *Let q be a prime power, let $2 \leqslant m < k$ be integers and let L be an $(m \times k)$-system over $\mathbb{F}_q$ such that $s(L)$ is even. If every equation of length $s(L)$ induced by L is uncommon, then L is uncommon.*

We have also made progress towards the related notion of a *Sidorenko* system of equations. In particular, we show in [12] that $s(L)$ being even is a necessary condition for a system to be Sidorenko. We also provide a large class of non-trivial Sidorenko systems of equations.

References

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

[2] P. Cameron, J. Cilleruelo, and O. Serra. On monochromatic solutions of equations in groups. *Rev. Mat. Iberoam.*, 23(1):385–395, 2007.

[3] B.A. Datskovsky. On the number of monochromatic Schur triples. *Adv. in Appl. Math.*, 31(1):193–198, 2003.

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

[5] J. Fox, H.T. Pham, and Y. Zhao. Common and Sidorenko Linear Equations. *Q. J. Math.*, 72(4) 1223–1234, 2021.

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

[7] R. Graham, V. Rödl, and A. Ruciński. On Schur properties of random subsets of integers. *J. Number Theory*, 61(2):388–408, 1996.

[8] R. Hancock, D. Kral, M. Krnc, and J. Volec. Towards characterizing locally common graphs. *arXiv:2011.02562*, 2020.

[9] C. Jagger, P. Šťovíček, and A. Thomason. Multiplicities of subgraphs. *Combinatorica*, 16(1):123–141, 1996.

[10] D. Král', J.A. Noel, S. Norin, J. Volec, and F. Wei. Non-bipartite k-common graphs. Combinatorica (2022), 1–28.

[11] N. Kamčev, A. Liebenau, and N. Morrison. On uncommon systems of equations *arxiv:2106.08986*, 2021.

[12] N. Kamčev, A. Liebenau, and N. Morrison. Towards a characterisation of Sidorenko systems. *arXiv:2107.14413*, 2021.

[13] A. Robertson and D. Zeilberger. A 2-coloring of $[1, n]$ can have $(1/22)n^2 + O(n)$ monochromatic Schur triples, but not less! *Electron. J. Combin.*, 5:Research Paper 19, 4, 1998.

[14] A. Saad and J. Wolf. Ramsey multiplicity of linear patterns in certain finite abelian groups. *Q. J. Math.*, 68(1):125–140, 2017.

[15] T. Schoen. The number of monochromatic Schur triples. *European J. Combin.*, 20(8):855–866, 1999.

[16] A. F. Sidorenko. Cycles in graphs and functional inequalities. *Mat. Zametki*, 46(5):72–79, 104, 1989.

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

[18] L. Versteegen. Linear configurations containing 4-term arithmetic progressions are uncommon. *arXiv:2106.06846*, 2021.

[19] J. Wolf. The minimum number of monochromatic 4-term progressions in $\mathbb{Z}_p$. *J. Comb.*, 1(1):53–68, 2010.

A proof of the Ryser-Brualdi-Stein conjecture for large even n

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.

Conjecture 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 prove Theorem 2, we use a combination of the semi-random method and the absorption method. We use an implementation of the semi-random method in this setting in [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. That is, we aim to set aside some special 'absorber' which can be extended into a transversal with $(1-o(1))n$ cells in a Latin square of order n using the semi-random method. The aim is that the absorber should have some special properties to turn this into a transversal with $n-1$ cells.

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 entries in the Latin square prevent the existence of the typical absorbers used for an application of the absorbing method. Latin squares arising as the addition tables of groups are good examples of such algebraic properties, but other extremal examples show these properties can be more complicated still.

To prove Theorem 2, we introduce methods to group the symbols of the Latin square into colour classes with some limited algebraic properties. We then use this properties to create both an 'absorption structure' as well as introducing an 'addition structure' which will identify which row and column should be left out of the $n-1$ transversal that we create.

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.

Triangle factors in pseudorandom graphs

Patrick Morris

We say a graph G contains a *triangle factor* if there is a collection of vertex disjoint triangles which completely cover the vertex set of G. As a natural generalisation of a perfect matching in a graph, triangle factors are a fundamental object in probabilistic, extremal and algorithmic graph theory. However, unlike perfect matchings, it is not easy to verify whether a graph G contains a triangle factor or not. Indeed Schaeffer [8] proved that determining if a graph on $n \in 3\mathbb{N}$ vertices contains a triangle factor is an NP-complete problem. Given that we can not hope for a nice characterisation of graphs which contain triangle factors, there has been a focus on providing sufficient conditions. One classical such theorem is due to Corrádi and Hajnal [4] who showed that a triangle factor is guaranteed if the host graph is sufficiently dense.

Theorem 1. *If G is a graph on $n \in 3\mathbb{N}$ vertices with minimum degree $\delta(G) \geqslant 2n/3$, then G contains a triangle factor.*

This theorem is tight, as can be seen, for example, by taking G to be a complete graph with a clique of size $n/3+1$ removed to leave an independent set of vertices, say I. All examples verifying the tightness of Theorem 1 share some features with the graph given here. For example they contain much larger independent sets than almost all graphs of this density. Therefore, one might hope to capture more graphs having a triangle factor by adding a condition that precludes the atypical behaviour of the extremal examples.

This naturally leads us to the notion of *pseudorandom graphs*, which are, roughly speaking, graphs which imitate random graphs of the same density. The study of pseudorandom graphs, initiated in the 1980s by Thomason [14], has become a central and vibrant field at the intersection of Combinatorics and Theoretical Computer Science. One way of imposing pseudorandomness is through the spectral notion of the eigenvalue gap. Here we focus on so-called (n, d, λ)-graphs G which are d-regular n-vertex graphs with *second eigenvalue* λ. By second eigenvalue, what is actually meant is the second largest eigenvalue of the adjacency matrix in absolute value. It turns out that this parameter λ controls the pseudorandomness of the graph G, with smaller values of λ giving graphs that have stronger pseudorandom properties.

It follows from simple linear algebra that for an (n, d, λ)-graph, one has that $\lambda \leqslant d$ always and moreover, as long as d is not too close to n, say $d \leqslant 2n/3$, one has that $\lambda = \Omega(\sqrt{d})$. Thus, we think of (n, d, λ)-graphs with $\lambda = \Theta(\sqrt{d})$ as being *optimally pseudorandom*.

A prominent theme in the study of pseudorandom graphs has been to give conditions on the parameters, n, d and λ which guarantee certain properties of an (n, d, λ)-graph. For example, it follows easily from the so-called expander mixing lemma that any (n, d, λ)-graph G with $\lambda < d^2/n$ contains a triangle. In particular, any optimally pseudorandom graph with $d = \omega(n^{2/3})$ must contain a triangle.

Moreover, this condition is tight due to a triangle-free construction of an (n, d, λ)-graph due to Alon [2] with $d = \Theta(n^{2/3})$ and $\lambda = \Theta(n^{1/3})$. In general, finding optimal conditions for subgraph appearance in (n, d, λ)-graphs seems very hard. Indeed the only tight conditions that are known are those for fixed size odd cycles [3, 9]. With respect to spanning structures, it is only perfect matchings that have been well understood [9].

In this talk we present an answer to what has become one of the central problems in this area, by giving a tight condition for an (n, d, λ)-graph to contain a triangle factor.

Theorem 2. *There exists $\varepsilon > 0$ such that any (n, d, λ)-graph with $n \in 3\mathbb{N}$ and $\lambda \leqslant \varepsilon d^2/n$, contains a triangle factor.*

Theorem 2 was conjectured by Krivelevich, Sudakov and Szabó [10] in 2004. Focusing solely on optimally pseudorandom graphs, that is, setting $\lambda = \Theta(\sqrt{d})$, Theorem 2 gives that any optimally pseudorandom graph with $d = \omega(n^{2/3})$ contains a triangle factor. Comparing this to Theorem 1, we see that imposing pseudorandomness, which is easy to compute via the second eigenvalue, allows us to capture much sparser graphs which are guaranteed to contain a triangle factor.

Let us note that Theorem 2 is tight due to the construction of Alon discussed above. Indeed, one of the reasons that the conjecture has attracted so much attention is that it marks a distinct difference between the behaviour of random graphs and that of (optimally) pseudorandom graphs. In random graphs, we know that triangles appear at density roughly $p = n^{-1}$, whilst for triangle factors the threshold is considerably denser, namely $p = n^{-2/3}(\log n)^{1/3}$ [7]. On the other hand, there exists triangle-free, optimally pseudorandom graphs with density roughly $n^{-1/3}$, but any pseudorandom graph whose density is a constant factor larger than this is guaranteed to have not only a triangle but a triangle factor. Furthermore, it follows from Theorem 2 and (the proof of) a result of Han, Kohayakawa, Person and the author [5] that even more is true.

Corollary 3. *There exists $\varepsilon > 0$ such that any (n, d, λ)-graph G with $\lambda \leqslant \varepsilon d^2/n$ is* 2-universal. *That is, given any graph F on at most n vertices, with maximum degree* 2*, G contains a copy of F.*

Theorem 2 concludes a body of work towards the conjecture of Krivelevich, Sudakov and Szabó and the proof of the theorem builds upon the many beautiful ideas of various authors, which have arisen in this study. The first step towards the conjecture was given by Krivelevich, Sudakov and Szabó [10] themselves, who showed that $\lambda \leqslant \varepsilon d^3/(n^2 \log n)$ for some sufficiently small ε is enough to guarantee a triangle factor. This was improved by Allen, Böttcher, Hàn, Kohayakawa and Person [1] to $\lambda \leqslant \varepsilon d^{5/2}/n^{3/2}$ and they could also prove that with this condition one can guarantee the appearance of the square of a Hamilton cycle. Nenadov [12] then got very close to the conjecture, showing that $\lambda \leqslant \varepsilon d^2/(n \log n)$ guarantees a triangle factor. Concentrating solely on optimally pseudorandom graphs, these results read that having degree $d = \omega(n^{4/5}(\log n)^{2/5})$, $\omega(n^{3/4})$ and $\omega((n \log n)^{2/3})$ respectively, guarantees the existence of a triangle factor.

In a different direction, one can fix the condition that $\lambda \leqslant \varepsilon d^2/n$ for some small $\varepsilon > 0$ and prove the existence of other structures giving evidence for a triangle factor. Again, this was initiated by Krivelevich, Sudakov and Szabó [10] who proved that with this condition, one can guarantee the existence of a fractional triangle factor. Another interesting result of Sudakov, Szabó and Vu [13] showed that when we have $\lambda \leqslant \varepsilon d^2/n$, we have many triangles and these are well distributed in the (n, d, λ)-graph. Indeed they proved a Turán result showing that one needs to delete at least half the edges of such an (n, d, λ)-graph in order to eliminate all the triangles. Another more recent result due to Han, Kohayakawa and Person [6] shows that $\lambda \leqslant \varepsilon d^2/n$ guarantees the existence of a *near triangle factor*; that there are vertex disjoint triangles covering all but $n^{647/648}$ vertices of such an (n, d, λ)-graph.

Our proof of Theorem 2 incorporates discrete algorithmic techniques, probabilistic methods, fractional relaxations and linear programming duality and the method of absorption. For details of the proof, we refer to the full paper of this result [11].

References

[1] P. Allen, J. Böttcher, H. Hàn, Y. Kohayakawa, and Y. Person. Powers of Hamilton cycles in pseudorandom graphs. *Combinatorica*, 37(4):573–616, 2017.

[2] N. Alon. Explicit Ramsey graphs and orthonormal labelings. *Electronic Journal of Combinatorics*, 1:R12, 8pp, 1994.

[3] N. Alon and N. Kahale. Approximating the independence number via the ϑ-function. *Mathematical Programming*, 80(3):253–264, 1998.

[4] K. Corrádi and A. Hajnal. On the maximal number of independent circuits in a graph. *Acta Mathematica Academiae Scientiarum Hungarica*, 14(3-4):423–439, 1963.

[5] J. Han, Y. Kohayakawa, P. Morris, and Y. Person. Finding any given 2-factor in sparse pseudorandom graphs efficiently. *Journal of Graph Theory*, 96(1), 87–108, 2021.

[6] J. Han, Y. Kohayakawa, and Y. Person. Near-optimal clique-factors in sparse pseudorandom graphs. *Combinatorics, Probability & Computing*, 30(4), 570–590, 2021.

[7] A. Johansson, J. Kahn, and V. Vu. Factors in random graphs. *Random Structures & Algorithms*, 33(1):1–28, 2008.

[8] R. M. Karp. On the computational complexity of combinatorial problems. *Networks*, 5(1):45–68, 1975.

[9] M. Krivelevich and B. Sudakov. Pseudo-random graphs. In *More sets, graphs and numbers*, volume 15 of *Bolyai Soc. Math. Stud.*, pages 199–262. Springer, Berlin, 2006.

[10] M. Krivelevich, B. Sudakov, and T. Szabó. Triangle factors in sparse pseudo-random graphs. *Combinatorica*, 24(3):403–426, 2004.

[11] P. Morris. Clique factors in pseudorandom graphs. arXiv preprint, arXiv:2101.05092, 2021.

[12] R. Nenadov. Triangle-factors in pseudorandom graphs. *Bulletin of the London Mathematical Society*, 51(3):421–430, 2019.

[13] B. Sudakov, T. Szabó, and H. Van Vu. A generalization of Turán's theorem. *Journal of Graph Theory*, 49(3):187–195, 2005.

[14] A. Thomason. Random graphs, strongly regular graphs and pseudorandom graphs. In *Surveys in combinatorics*, volume 123 of *London Math. Soc. Lecture Note Ser.*, pages 173–195. Cambridge Univ. Press, Cambridge, 1987.

Down-set thresholds

Bhargav Narayanan
(joint work with Benjamin Gunby, Xiaoyu He)

We elucidate the relationship between the threshold and the expectation-threshold of a down-set. Qualitatively, our main result demonstrates that there exist down-sets with polynomial gaps between their thresholds and expectation-thresholds; in particular, the logarithmic gap predictions of Kahn–Kalai [1] and Talagrand [2] about up-sets do not apply to down-sets. Quantitatively, we show that any collection $\mathcal{G}$ of graphs on $[n]$ that covers the family of all triangle-free graphs on $[n]$ satisfies the inequality $\sum_{G\in\mathcal{G}} \exp(-\delta e(G^c)/\sqrt{n}) < 1/2$ for some universal $\delta > 0$, and this is essentially best-possible; in particular, the this shows that the expectation-threshold of the down-set of triangle-free graphs on $[n]$ is $\Omega(1/\sqrt{n})$ while its threshold is $\Theta(1/n)$.

References

[1] J. Kahn and G. Kalai, *Thresholds and expectation thresholds*, Combin. Probab. Comput. **16** (2007), 495–502.
[2] M. Talagrand, *Are many small sets explicitly small?*, Proceedings of the 2010 ACM International Symposium on Theory of Computing, 2010, 13–35.

A proof of the Kahn-Kalai Conjecture

Jinyoung Park
(joint work with Huy Tuan Pham)

Given a finite set X, write 2^X for the power set of X. For $p \in [0,1]$, let μ_p be the product measure on 2^X given by $\mu_p(A) = p^{|A|}(1-p)^{|X\setminus A|}$. In this paper $\mathcal{F} \subseteq 2^X$ always denotes an *increasing property*, meaning that if $B \supseteq A \in \mathcal{F}$, then $B \in \mathcal{F}$. It is a well-known fact that $\mu_p(\mathcal{F})(:= \sum_{A\in\mathcal{F}} \mu_p(A))$ is strictly increasing in p for any $\mathcal{F} \neq \varnothing, 2^X$. The *threshold*, $p_c(\mathcal{F})$, is then the unique p for which $\mu_p(\mathcal{F}) = 1/2$. In this paper, we resolve a conjecture of Kahn and Kalai [12], reiterated by Talagrand [16], relating the threshold and the "expectation-threshold" (definition is below).

Following [16], we say $\mathcal{F}$ is *p-small* if there is $\mathcal{G} \subseteq 2^X$ such that

$$\mathcal{F} \subseteq \langle \mathcal{G} \rangle := \bigcup_{S\in\mathcal{G}} \{T : T \supseteq S\} \tag{1}$$

and

$$\sum_{S\in\mathcal{G}} p^{|S|} \leqslant 1/2.$$

We say that $\mathcal{G}$ is a *cover* of $\mathcal{F}$ if (1) holds. The *expectation-threshold* of $\mathcal{F}$, $q(\mathcal{F})$, is defined to be the maximum p such that $\mathcal{F}$ is p-small. Observe that $q(\mathcal{F})$ is a trivial lower bound on $p_c(\mathcal{F})$, since

$$\mu_p(\mathcal{F}) \leqslant \mu_p(\langle \mathcal{G} \rangle) \leqslant \sum_{S\in\mathcal{G}} p^{|S|}. \tag{2}$$

Note that, with X_p the random variable whose distribution is μ_p, the right-hand side of (2) is $\mathbb{E}[|\{S \in \mathcal{G} : S \subseteq X_p\}|]$.

Given an increasing property $\mathcal{F}$, write $\ell(\mathcal{F})$ for the size of a largest minimal element of $\mathcal{F}$. Our main theorem resolves the following conjecture of Kahn and Kalai [12].

Theorem 1 (The Kahn-Kalai Conjecture)**.** *There is a universal constant K such that for every finite set X and increasing property $\mathcal{F} \subseteq 2^X$,*

$$p_c(\mathcal{F}) \leqslant Kq(\mathcal{F}) \log \ell(\mathcal{F}).$$

Roughly speaking, Theorem 1 says that for any increasing property, the threshold is never far from its trivial lower bound given by the expectation threshold.

Thresholds have been a central subject in the study of random discrete structures since its initiation by Erdős and Rényi [4, 5], the study of which has flourished in random graph theory, computer science [13, 16], and statistical physics [9]. The definition of the threshold above is finer than the original Erdős-Rényi notion, according to which $p^* = p^*(n)$ is **a** threshold for $\mathcal{F} = \mathcal{F}_n$ if $\mu_p(\mathcal{F}) \to 0$ when $p \ll p^*$ and $\mu_p(\mathcal{F}) \to 1$ when $p \gg p^*$. That $p_c(\mathcal{F})$ is always an Erdős-Rényi threshold follows from [3], in which it was observed that *every* increasing $\mathcal{F}$ admits a threshold in the Erdős-Rényi sense. While much work has been done identifying thresholds for specific properties (see [2, 10]), another intensively studied direction in the study of thresholds is "sharpness" of thresholds: we refer interested readers to [7, 8].

To emphasize the strength of Theorem 1, we point out that, in [12], Kahn and Kalai wrote that "It would probably be more sensible to conjecture that it is not true." The expectation threshold is the most naive (and often the easiest) approach to estimating the threshold, and Theorem 1 says that for *every* increasing property, its threshold is only within a logarithmic factor of this naive estimate. In particular, many important works in this area have been on thresholds for *specific properties,* and Theorem 1 easily implies some of those very hard results on the location of thresholds, for example, the appearance of perfect matchings in random r-uniform hypergraphs [11], and the appearance of a given bounded degree spanning tree in a random graph [15]. For more discussion on the significance and applications of this theorem, we refer the readers to [6], in which a weaker ("fractional") version of Theorem 1 was proved.

Part of our proof is inspired by the algorithm in [1] and the analysis of the algorithm in [6, 14], though our implementation is significantly different from the ideas in those papers. In particular, our analysis completely avoids the use of "spread," which was a key ingredient in the aforementioned papers. We remark that our proof technique can also be adapted to simplify the proof of the main lemma of [6].

References

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

[2] B. Bollobás, *Random Graphs,* second ed., Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge, 2001.
[3] B. Bollobás and A. Thomason, *Threshold functions,* Combinatorica **7** (1987), 35–38.
[4] P. Erdős and A. Rényi, *On random graphs I,* Publ Math Debrecen **6** (1959), 290–297.
[5] P. Erdős and A. Rényi, *On the evolution of random graphs,* Publ Math Inst Hungar Acad Sci **5** (1960), 17–61.
[6] K. Frankston, J. Kahn, B. Narayanan, and J. Park, *Thresholds versus fractional expectation-thresholds,* Ann. of Math. (2) **194** (2021), no. 2, 475–495.
[7] E. Friedgut, *Sharp thresholds of graph properties, and the k-sat problem,* J. Amer. Math. Soc. **12** (1999), 1017–1054, With an appendix by J. Bourgain.
[8] E. Friedgut, *Park-Hunting for sharp thresholds,* Random Structures Algorithms **26** (2005), 37–51.
[9] G. R. Grimmett, *The random-cluster model,* Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Math. Sciences], vol. **333**, Springer-Verlag, Berlin, 2006.
[10] S. Janson, T. Łuczak, and A. Rucinski, *Random graphs,* Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.
[11] A. Johansson, J. Kahn, and V. Vu, *Factors in random graphs,* Random Structures Algorithms **33** (2008), 1–28.
[12] J. Kahn and G. Kalai, *Thresholds and expectation thresholds,* Combin. Probab. Comput. **16** (2007), 495–502.
[13] G. Kalai, *Boolean functions: influence, threshold and noise,* European Congress of Mathematics, 85–110, Eur. Math. Soc., Zürich, 2018.
[14] J. Kahn, B. Narayanan, and J. Park, *The threshold for the square of a Hamilton cycle,* Proc. Amer. Math. Soc. **149** (2021), no. 8, 3201–3208.
[15] R. Montgomery, *Spanning trees in random graphs,* Adv. Math. **356** (2019), 106793, 92.
[16] M. Talagrand, *Are many small sets explicitly small?,* STOC'10 – Proceedings of the 2010 ACM International Symposium on Theory of Computing, 13–35, ACM, New York, 2010.

On a conjecture of Talagrand on selector processes and a consequence on positive empirical processes

Huy Tuan Pham
(joint work with Jinyoung Park)

The study of suprema of stochastic processes is a central interest in probability theory with influential applications in related areas. We refer the readers to [4, 8] for extensive discussions of various aspects of this subject. Through many fundamental developments, one now has fairly good understanding of the suprema of centered Gaussian processes[1]. In particular, one can associate each Gaussian process $(Z_t)_{t\in T}$ indexed by a set T with a metric on T given by $d(t,s) := \mathbb{E}[(Z_t - Z_s)^2]$, and Talagrand's celebrated generic chaining bound and majorizing measure theorem [3, 5, 8] determine the expectation of the supremum $\sup_{t\in T} Z_t$ up to a constant by a quantity depending only on the metric space (T,d). Via this fundamental result, one can obtain deep insights and characterizations of the suprema of Gaussian processes. One important example is Theorem 1 below, which gives a nice geometric characterization of large suprema of Gaussian processes: such event must be contained in a union of half-spaces whose sum of measures is small.

[1]Following [8], we always assume Gaussian processes are centered, i.e., $\mathbb{E}Z_t = 0$ for all $t \in T$.

Theorem 1 (Talagrand, Theorem 2.12.2 in [8]). *There exists $L > 0$ such that the following holds. Let g be an M-dimensional standard Gaussian vector. For $\mathcal{T} \subseteq \mathbb{R}^M$, consider the process $Z_t = \langle t, g\rangle$ for $t \in \mathcal{T}$. Then one can find a sequence of half-spaces H_k of $\mathbb{R}^M$ with*

$$\left\{\sup_{t\in\mathcal{T}} Z_t \geqslant L\mathbb{E}\sup_{t\in\mathcal{T}} Z_t\right\} \subset \bigcup_{k\geqslant 1} H_k,$$

and

$$\sum_{k\geqslant 1} \mathbb{P}(H_k) \leqslant \frac{1}{2}.$$

Our main contribution in this paper is the proofs of a conjecture of Talagrand on selector processes (Theorem 2; originally [6, Problem 4.1], [7, Conjecture 5.7] and [8, Research Problem 13.2.3]) and a result on positive empirical processes (Theorem 3; a question of Talagrand [9] and a problem posed in [6]), which are analogous to Theorem 1. We first quickly state our main results, and then provide more context, definitions, and motivations for Talagrand's questions.

Given a finite set X, write 2^X for the power set of X. For $p \in [0,1]$, let μ_p be the product measure on 2^X given by $\mu_p(A) = p^{|A|}(1-p)^{|X\setminus A|}$. We use X_p for the random variable whose distribution is μ_p. For $S \subseteq X$, define the *upset generated by* S to be $\langle S\rangle := \{T : T \supseteq S\}$. Following [7], we say $\mathcal{F} \subseteq 2^X$ is *p-small* if there is $\mathcal{G} \subseteq 2^X$ such that

$$\mathcal{F} \subseteq \langle \mathcal{G}\rangle := \bigcup_{S\in\mathcal{G}} \langle S\rangle \tag{1}$$

and

$$\sum_{S\in\mathcal{G}} p^{|S|} \leqslant 1/2. \tag{2}$$

We say $\mathcal{G}$ is a *cover* of $\mathcal{F}$ if (1) holds.

Our first main result is the Bernoulli-p analog of Theorem 1.

Theorem 2. *There exists $L > 0$ such that the following holds. Consider any $0 < p < 1$, any finite set X and any collection Λ of sequences $\boldsymbol{\lambda} = (\lambda_i)_{i\in X}$ with $\lambda_i \geqslant 0$. Then the family*

$$\left\{S \subseteq X : \sup_{\boldsymbol{\lambda}\in\Lambda} \sum_{i\in S} \lambda_i \geqslant L\mathbb{E}\sup_{\boldsymbol{\lambda}\in\Lambda} \sum_{i\in X_p} \lambda_i\right\}$$

is p-small.

In [7], Talagrand explains the meaning of the above theorem this way: Conjecture 5.7 (now Theorem 2) shows that "if you are given a selector process, and would like to prove that, within a multiplicative factor, the quantity $\mathbb{E}\sup_{\boldsymbol{\lambda}\in\Lambda}\sum_{i\in X_p}\lambda_i \leqslant M$ for a constant M, there is in the end **no other way** than to find the witnesses that the set $\{S \subseteq X : \sup_{\boldsymbol{\lambda}\in\Lambda}\sum_{i\in S}\lambda_i \geqslant LM\}$ is small." In the same place, Talagrand suggests that this result "provides fundamental information."

Our second main result is the analog of Theorem 1 for positive empirical processes.

Theorem 3. *There exists $L > 0$ such that the following holds. For any $N > 0$ and i.i.d. random variables $Y_1, \dots, Y_N$ distributed according to a Borel probability measure on a Polish space $\mathbb{T}$, and any finite collection $\mathcal{F}$ of functions $f : \mathbb{T} \to \mathbb{R}_{\geqslant 0}$ with $\mathcal{F} \subseteq L^\infty(\mathbb{T})$, consider the positive empirical process $Z_f = \frac{1}{N}\sum_{i=1}^N f(Y_i)$. Assume that $0 < \mathbb{E}[\sup_{f\in\mathcal{F}} Z_f] < \infty$. Then one can find a collection $\mathcal{C}$ of pairs (g,t) where g is a nonnegative function on $\mathbb{T}$ and $t > 0$, so that with $E_{g,t} := \{Z_g \geqslant t\}$, we have*

$$\left\{\sup_{f\in\mathcal{F}} Z_f \geqslant L\mathbb{E}\sup_{f\in\mathcal{F}} Z_f\right\} \subset \bigcup_{(g,t)\in\mathcal{C}} E_{g,t},$$

and

$$\sum_{(g,t)\in\mathcal{C}} \mathbb{P}(E_{g,t}) \leqslant \frac{1}{2}.$$

We remark that the conclusion of Theorem 3 readily extends to cases where $\mathcal{F}$ is not necessarily finite, for example when $\mathcal{F}$ is a totally bounded infinite subset of $L^\infty(\mathbb{T})$.

The full version of our paper can be found at [2]. Our proofs of Theorem 2 and Theorem 3 are also closely connected to and motivate our recent resolution of the Kahn-Kalai conjecture [1].

References

[1] J. Park and H. T. Pham, *A proof of the Kahn-Kalai conjecture,* preprint.

[2] J. Park and H. T. Pham, *On a conjecture of Talagrand on selector processes and a consequence on positive empirical processes,* preprint.

[3] M. Talagrand, *Regularity of gaussian processes,* Acta Math. **159** (1987), 99–149.

[4] M. Talagrand, *Donsker classes and random geometry,* Ann. Probab. **15**(4) (1987), 1327–1338.

[5] M. Talagrand, *A simple proof of the majorizing measure theorem,* Geom. Funct. Anal. **2** (1992), 118–125.

[6] M. Talagrand, *Selector processes on classes of sets,* Probab. Theory Relat. Fields **135** (2006), 471–486.

[7] M. Talagrand, *Are many small sets explicitly small?,* Proceedings of the 2010 ACM International Symposium on Theory of Computing (2010), 13–35.

[8] M. Talagrand, *Upper and Lower Bounds for Stochastic Processes,* Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge (A Series of Modern Surveys in Mathematics book series) vol. **60**, Springer, 2021.

[9] M. Talagrand, personal communication, 2022.

A random Hall-Paige Conjecture

Alexey Pokrovskiy
(joint work with Alp Müyesser)

A complete mapping of a group G is a bijection $\phi\colon G \to G$ such that $x \mapsto x\phi(x)$ is also bijective (this is equivalent to asking that the multiplication table of the group has a transversal i.e. a set of $|G|$ entries that don't repeat rows, columns, or symbols). Hall and Paige [4] conjectured in 1955 that a finite group G has a complete mapping whenever $\prod_{x\in G} x \in G'$ (where G' is the commutator subgroup of the group). They proved that this is a necessary condition, so the main part of the conjecture is to prove that "$\prod_{x\in G} x \in G' \implies G$ has a complete mapping". This was confirmed in 2009 by Wilcox [6], Evans [3], and Bray [1] with a proof using the classification of finite simple groups. Recently, Eberhard, Manners, and Mrazović [2] found an alternative proof of the conjecture for sufficiently large groups using ideas from analytic number theory. Their proof gives a very precise estimate on the number of complete mappings that each group has.

In this talk, we presented a third proof of the conjecture using a different set of techniques, this time coming from probabilistic combinatorics. This proof only works for sufficiently large groups, but generalizes the conjecture in a new direction. This direction is to find complete mappings not just in full groups but also in their subsets. Specifically:

Theorem 1 (Müyesser-Pokrovskiy, [5])**.** *Let G be a group of order n. Let $p \geqslant n^{-1/10^{100}}$. Let $R^1, R^2, R^3 \subseteq G$ be p-random subsets, sampled independently. Then, with high probability, the following holds.*

Let $X, Y, Z \subseteq G$ be subsets satisfying the following properties.

- $|X\Delta R^1| + |Y\Delta R^2| + |Z\Delta R^3| \leqslant p^{10^{10}} n/\log(n)^{10^{15}}$
- $|X| = |Y| = |Z|$
- $\prod X \prod Y (\prod Z)^{-1} \in G'$

Then, there exists a bijection $\phi\colon X \to Y$ such that $x \mapsto x\phi(x)$ is a bijection from X to Z.

Using Theorem 1, it is possible to settle the following open problems for sufficiently large groups.

(1) It is possible to confirm a conjecture of Snevily by characterising large subsquares of multiplication tables of finite groups that admit transversals. Previously, this characterisation was known only for abelian groups of odd order (by a combination of papers by Alon; Dasgupta, Károlyi, Serra, and Szegedy; Arsovski).

(2) It is possible to characterize the abelian groups that can be partitioned into zero-sum sets of specific sizes. This solves a problem of Tannenbaum, and confirms a conjecture of Cichacz.

(3) It is possible to characterize harmonious groups, that is, groups with an ordering in which the product of each consecutive pair of elements is distinct. This solves a problem of Evans.

(4) It is possible to characterize the groups with which any path can be assigned a cordial labelling. In the case of abelian groups, this confirms a conjecture of Patrias and Pechenik.

References

[1] John N. Bray, Qi Cai, Peter J. Cameron, Pablo Spiga, and Hua Zhang. The Hall–Paige conjecture, and synchronization for affine and diagonal groups. *Journal of Algebra*, 545:27–42, 2020.
[2] Sean Eberhard, Freddie Manners, and Rudi Mrazović. An asymptotic for the Hall-Paige conjecture. *arXiv preprint arXiv:2003.01798*, 2020.
[3] Anthony Evans. The admissibility of sporadic simple groups. *J. Algebra*, 321(1):105–116, 2009.
[4] M. Hall and L.J. Paige. Complete mappings of finite groups. *Pacific J. of Math.*, 5:541–549, 1955.
[5] A. Muyesser, A. Pokrovskiy*A random Hall-Paige Conjecture*, arXiv:2204.09666 (2022).
[6] S. Wilcox. Reduction of the Hall-Paige conjecture to sporadic simple groups. *J. Algebra*, 321(5):1407–1428, 2009.

High-girth Steiner triple systems

Ashwin Sah

(joint work with Matthew Kwan, Mehtaab Sawhney, and Michael Simkin)

We prove a 1973 conjecture of Erdős on the existence of Steiner triple systems with arbitrarily high girth. Specifically, given $g \in \mathbb{N}$, any sufficiently large $n \equiv 1, 3 \pmod 6$ admits a 3-graph on n vertices so that every pair is in a unique hyperedge and so that for all $4 \leqslant j \leqslant g$, any j vertices contains at most $j - 3$ hyperedges. We also prove an analogue for Latin squares conjectured by Linial.

The techniques build on the iterative absorption framework for designs introduced by Glock, Kühn, Lo, and Osthus [3] as well as prior approximate results of Bohman and Warnke [1] as well as Glock, Kühn, Lo, and Osthus [2]. At a high level, the idea is to use explicit combinatorial constructions to form a finite absorbing structure that can only "fix" very restricted types of defects, and then to iteratively transform the decomposition problem into smaller and smaller decomposition problems until it is of bounded size, able to be solved using the absorber. Specifically, we start with the entire vertex set of size n and attempt to cover all edges outside of smaller and smaller localized subsets. To simplify, this iterative transformation process relies on a triangle-removal process (or high-girth triangle removal process due to [1, 3] in our setting) and a "cover-down" step which runs a disjoint matching process to use to use all remaining edges outside of the new vertex set.

Unfortunately, in the setting of high-girth Steiner systems, the issue of "hybrid configurations" spanning multiple levels of this iteration poses severe difficulties. As an initial issue, a positive fraction of the remaining triangles are expected to be forbidden after each step. This effect multiplies as the iteration continues, leading to a "constraint focusing" issue. This is mitigated by (a) reducing the number of steps of iteration via an efficient (high-girth) absorber construction which allows

one to have a polynomial-size (instead of bounded-size) final vertex set, and (b) an initial sparsification which mitigates the focusing effect.

To analyze the most general hybrid configurations that can arise, we adapt ideas of Bohman and Warnke [1], creating a "weight-system" framework for analyzing moments of random processes and allowing for a retrospective analysis approach to assuring various quantities remain bounded throughout our multi-stage random procedure. Finally, to ensure the high-girth process runs far enough to use in iterative absorption, we introduce tools to regularize guiding auxiliary data.

References

[1] Tom Bohman and Lutz Warnke, *Large girth approximate Steiner triple systems*, J. Lond. Math. Soc. (2) **100** (2019), 895–913.

[2] Stefan Glock, Daniela Kühn, Allan Lo, and Deryk Osthus, *On a conjecture of Erdős on locally sparse Steiner triple systems*, Combinatorica **40** (2020), 363–403.

[3] Stefan Glock, Daniela Kühn, Allan Lo, and Deryk Osthus, *The existence of designs via iterative absorption: hypergraph F-designs for arbitrary F*, Mem. Amer. Math. Soc. (to appear).

The singularity probability of a random symmetric matrix

Julian Sahasrabudhe

(joint work with Marcelo Campos, Matthew Jenssen, Marcus Michelen)

In the late 1950s Wigner [20], initiated the study of the eigenvalues of random symmetric matrices (sometimes called *Wigner random matrices* after him) and proved the, now classical, "semi-circular law" for the distribution of the eigenvalues of such matrices. Here we study the probability that such a matrix is singular.

Let A_n be drawn uniformly from all $n \times n$ symmetric matrices with entries in $\{-1, 1\}$. We consider the quantity

$$\mathbb{P}(\det(A_n) = 0).$$

A natural lower bound comes from the event that two rows are equal up to sign and indeed, a moment's thought reveals that

$$\mathbb{P}(\det(A_n) = 0) \geqslant (1 + o(1))n^2 2^{-n}.$$

One attractive feature of this problem is that it is widely believed that this lower bound is indeed represents *the truth*. That is, it is believed that this lower bound, or at least $2^{-n(1+o(1))}$, is also a corresponding *upper-bound*. However, establishing this "obvious" conjecture has proven to be a fascinating challenge and has lead to much interesting mathematics.

We first mention the great success that has been attained on the analogous problem for the related *iid* model. For this, we let B_n be drawn uniformly from all $n \times n$ matrices with all entries in $\{-1, 1\}$; that is, we have dropped the symmetry condition. Here the problem got its start with the work of Komlós in the 60s [10, 11] who showed that B_n is non-singular with probability $1 - o(1)$. This was then greatly imporved 20 years later in a beautiful paper of Kahn, Komlós and

Szemerédi [9]. Further progress was made by Tao and Vu [15, 16, 17] Bourgain, Vu and Wood [2], and Rudelson and Vershynin [13], until the problem was essentially resolved by Tikhomirov [18] who showed that

$$\mathbb{P}(\det(B_n) = 0) = 2^{-n(1+o(1))},$$

in a breakthrough paper.

Progress on the singularity probability for the symmetric model A_n has come somewhat more recently and more slowly. The problem of showing that A_n is almost surely non-singular goes back, at least, to Weiss in the early 1990s but was not resolved until 2005 by Costello, Tao and Vu [5], who obtained the bound

$$\mathbb{P}(\det(A_n) = 0) \leqslant n^{-1/8+o(1)}.$$

The first super-polynomial bounds were obtained by Nguyen [12] and, simultaneously, Vershynin [19], the latter obtaining a bound of the form $\exp(-n^c)$. In 2019, a more combinatorial perspective was introduced by Ferber, Jain, Luh and Samotij [7] and applied by Ferber and Jain [6] to show

$$\mathbb{P}(\det A_n = 0) \leqslant \exp(-cn^{1/4}(\log n)^{1/2}).$$

In a similar spirit, Campos, Mattos, Morris and Morrison [4] then improved this bound to

$$\mathbb{P}(\det A_n = 0) \leqslant \exp(-cn^{1/2}),$$

by proving a "rough" inverse Littlewood-Offord theorem, inspired by the theory of hypergraph containers (see [1, 14]). This bound was then improved by Jain, Sah and Sawhney [8], who improved the exponent to $-cn^{1/2}\log^{1/4} n$, and, simultaneously, by the authors of this paper [3] who improved the exponent to $-c(n\log n)^{1/2}$.

It turns out that the convergence of these results to the exponent of $-c(n\log n)^{1/2}$ is no coincidence and, in fact, represents a natural barrier for the problem. This "barrier" was first pointed out in [4], where the authors noticed that all previous papers, for the core of their proofs, work only with the top-half of the matrix (i.e. the half above the diagonal), and assume the worst case in the rest. The authors then argue that any argument of this form cannot do better than a bound of $\exp(-c(n\log n)^{1/2})$. Therefore, to get beyond this obstruction, somehow the randomness of the matrix must "reused".

In this talk, I discuss our recent proof of an exponential upper-bound on the singularity probability of a symmetric random matrix, thereby breaking-though this barrier and giving the optimal bound, up to the constant in the exponent.

Theorem 1. *Let A_n be uniformly drawn from all $n \times n$ symmetric matrices with entries in $\{-1, 1\}$. Then*

$$\mathbb{P}(\det(A_n) = 0) \leqslant e^{-cn}, \tag{1}$$

where $c > 0$ is an absolute constant.

In this talk I sketch the main new combinatorial technique of this paper that allows us to "reuse" the randomness of our matrix by pushing some of the randomness onto the random selection of a potential kernel vector from our discretized

sphere. I also touch on another main challenge of this paper which is a result that shows the "approximate negative correlation" of certain kinds of linear events.

References

[1] J. Balogh, R. Morris, and W. Samotij. Independent sets in hypergraphs. *Journal of the American Mathematical Society*, 28(3):669–709, 2015.

[2] J. Bourgain, V. H. Vu, and P. M. Wood. On the singularity probability of discrete random matrices. *J. Funct. Anal.*, 258(2):559–603, 2010.

[3] M. Campos, M. Jensen, M. Michelen, and J. Sahasrabudhe. Singularity of random symmetric matrices revisited. *Proc. of Amer. Math. Soc.*, 2021.

[4] M. Campos, L. Mattos, R. Morris, and N. Morrison. On the singularity of random symmetric matrices. *Duke Mathematical Journal*, 170(5):881–907, 2021.

[5] K. P. Costello, T. Tao, and V. Vu. Random symmetric matrices are almost surely nonsingular. *Duke Math. J.*, 135(2):395–413, 2006.

[6] A. Ferber and V. Jain. Singularity of random symmetric matrices—a combinatorial approach to improved bounds. *Forum Math. Sigma*, 7:Paper No. e22, 29, 2019.

[7] A. Ferber, V. Jain, K. Luh, and W. Samotij. On the counting problem in inverse Littlewood–Offord theory. *Journal of the London Mathematical Society*.

[8] V. Jain, A. Sah, and M. Sawhney. On the smallest singular value of symmetric random matrices. *Combinatorics, Probability and Computing*, page 1–22, 2021.

[9] J. Kahn, J. Komlós, and E. Szemerédi. On the probability that a random ± 1-matrix is singular. *J. Amer. Math. Soc.*, 8(1):223–240, 1995.

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

[11] J. Komlós. On the determinant of random matrices. *Studia Sci. Math. Hungar.*, 1968.

[12] H. H. Nguyen. Inverse Littlewood-Offord problems and the singularity of random symmetric matrices. *Duke Math. J.*, 161(4):545–586, 2012.

[13] M. Rudelson and R. Vershynin. The Littlewood-Offord problem and invertibility of random matrices. *Adv. Math.*, 218(2):600–633, 2008.

[14] D. Saxton and A. Thomason. Hypergraph containers. *Inventiones mathematicae*, 201(3):925–992, 2015.

[15] T. Tao and V. Vu. On random ± 1 matrices: singularity and determinant. *Random Structures Algorithms*, 28(1):1–23, 2006.

[16] T. Tao and V. Vu. On the singularity probability of random Bernoulli matrices. *J. Amer. Math. Soc.*, 20(3):603–628, 2007.

[17] T. Tao and V. H. Vu. Inverse Littlewood-Offord theorems and the condition number of random discrete matrices. *Ann. of Math. (2)*, 169(2):595–632, 2009.

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

[19] R. Vershynin. Invertibility of symmetric random matrices. *Random Structures Algorithms*, 44(2):135–182, 2014.

[20] E. P. Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, 67(2):325–327, 1958.

List-decoding for Reed–Solomon codes

Lisa Sauermann

(joint work with Asaf Ferber, Matthew Kwan)

Reed–Solomon codes are a family of error-correcting codes that have been studied intensively in many different contexts since they were introduced in [3]. As the parameters of the code, consider a prime power q and integers $1 \leqslant k < n \leqslant q$.

Then, for n distinct *evaluation points* $a_1, \ldots, a_n \in \mathbb{F}_q$, the $[n,k]$-Reed–Solomon code with these evaluation points is defined to be the set of *codewords*

$$\mathcal{C}^{(k)}_{a_1,\ldots,a_n} := \{(f(a_1), \ldots, f(a_n)) \mid f \in \mathbb{F}_q[x], \deg f < k\} \subseteq \mathbb{F}_q^n.$$

One reason for the great interest in Reed–Solomon codes is that they behave optimally with respect to the classical *unique decoding* problem, having an optimal trade-off between rate and distance. The *rate* of a code $\mathcal{C} \subseteq \mathbb{F}_q^n$ is defined to be $\log_q |\mathcal{C}|/n$, and the *distance* of $\mathcal{C}$ is defined to be the minimum Hamming distance between a pair of distinct codewords $\gamma, \gamma' \in \mathcal{C}$ (recall that the Hamming distance between γ and γ' is the number of positions in which γ and γ' disagree). Every $[n,k]$-Reed–Solomon code has rate k/n and distance $n-k+1$. By the Singleton bound [6], this is the highest possible rate for any code of this distance.

An important generalization of unique decoding is the notion of *list-decoding*, and properties of Reed–Solomon codes with respect to list-decodability are much less understood. Roughly speaking, while the unique encoding problem demands that the original codeword can be uniquely reconstructed from a noisy signal, for the list-decoding problem we are satisfied with a short list of candidate codewords for a noisy signal. List-decodability was first introduced by Elias [1] and Wozencraft [7] in the 1950s, and has since been used in several different areas of theoretical computer science. Regarding list-decodability of Reed–Solomon codes specifically, there are applications in complexity theory and the theory of pseudorandomness. The problem of understanding the (combinatorial) list-decodability of Reed–Solomon codes has been raised by many researchers over the last two decades, and there has been a lot of recent work investigating this problem [2, 4, 5]. Still, a lot of questions remain open.

In order to formally define what it means for a code to be list-decodable, we need to introduce some more definitions. Given $r \in (0,1)$, and $\beta \in \mathbb{F}_q^n$, the *Hamming ball* of (relative) radius r centered at β is defined as

$$B_r(\beta) := \{\gamma \in \mathbb{F}_q^n \mid \gamma[i] = \beta[i] \text{ for at least } (1-r)n \text{ positions } 1 \leqslant i \leqslant n\}$$

(here, by $\gamma[i]$ we denote the symbol in the i-th position of $\gamma \in \mathbb{F}_q^n$). In other words, this Hamming ball consists of all points $\gamma \in \mathbb{F}_q^n$ that differ in at most rn coordinates from β.

A code $\mathcal{C} \subseteq \mathbb{F}_q^n$ is called (r,L)-*list-decodable* (for some radius $r \in (0,1)$ and some *list size* $L \in \mathbb{N}$) if we have $|\mathcal{C} \cap B_r(\beta)| \leqslant L$ for all $\beta \in \mathbb{F}_q^n$. In other words, $\mathcal{C}$ is (r,L)-list-decodable if each Hamming ball of (relative) radius r in $\mathbb{F}_q^n$ contains at most L codewords from $\mathcal{C}$. Note that for list size $L = 1$, the setting of (r,L)-list-decodability precisely corresponds to the classical unique decoding setting.

One may now ask about the best possible trade-off between the rate of an (r,L)-list-decodable code and the radius r and list size L. It is not hard to show that for fixed $r \in (0,1)$ and $L \in \mathbb{N}$ with $r \geqslant L/(L+1)$, every (r,L)-list-decodable code $\mathcal{C} \subseteq \mathbb{F}_q^n$ must have rate $o(1)$ as $n \to \infty$. So let us from now on assume that $r < L/(L+1)$. Shangguan and Tamo [5] proved the following upper bound for the rate of any (r,L)-list-decodable code $\mathcal{C} \subseteq \mathbb{F}_q^n$ (where n is large).

Theorem 1 (Shangguan and Tamo [5])**.** *For fixed list size $L \in \mathbb{N}$ and radius $0 < r < L/(L+1)$, and large n, any (r, L)-list-decodable linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$ has rate at most*

$$1 - \frac{L+1}{L} \cdot r + o(1).$$

This result can be viewed as a generalization of the classical Singleton bound [6] for unique decoding, and can be proved with a similar counting argument. Shangguan and Tamo [5] conjectured that the rate in their general upper bound can be attained for Reed–Solomon codes (similarly to the fact that Reed–Solomon codes behave optimally with respect to unique decoding).

Conjecture 2 (Shangguan and Tamo [5])**.** *For fixed list size $L \in \mathbb{N}$ and radius $0 < r < L/(L+1)$, and large n, there exist (r, L)-list-decodable Reed–Solomon codes with rate*

$$1 - \frac{L+1}{L} \cdot r - o(1).$$

Shangguan and Tamo [5] proved their conjecture for $L = 2$ and $L = 3$, but for larger L the conjecture is still wide open. We prove the conjecture for general $L \in \mathbb{N}$ up to some constant-factor-loss in the rate.

Theorem 3. *For fixed list size $L \in \mathbb{N}$ and radius $0 < r < L/(L+1)$, and large n, there exist (r, L)-list-decodable Reed–Solomon codes with rate*

$$\frac{1}{10}\left(1 - \frac{L+1}{L} \cdot r\right).$$

The constant 10 is not optimized in our proof, and can be improved to any constant $c > 2$. Our proof shows, similarly to previous works on this topic [2, 5], that for q sufficiently large with respect to n, for random evaluation points $a_1, \ldots, a_n \in \mathbb{F}_q$, the Reed–Solomon code $\mathcal{C}^{(k)}_{a_1,\ldots,a_n}$ (with $k = \lceil (1 - \frac{L+1}{L} \cdot r)/10 \rceil$) will likely be (r, L)-list-decodable (unfortunately, the proof does not give an explicit choice of such evaluation points). Our result requires a bound of $q \geqslant n^{1.25}$ for the field size (which can be improved to $q \geqslant n^c$ for any $c > 1$ at the expense of increasing the constant 10). In contrast, the previous works [2, 5] required q to be exponential in n.

In the case of radius $r = 1 - \varepsilon$ for small $\varepsilon > 0$, Theorem 3 gives the following corollary.

Theorem 4. *For fixed $\varepsilon \in (0, 1)$, and large n, there exist $(1-\varepsilon, \lceil 3/\varepsilon \rceil)$-list-decodable Reed–Solomon codes with rate at least $\varepsilon/15$.*

By the list-decoding capacity theorem, any $(1 - \varepsilon, L)$-list-decodable Reed–Solomon code (with list size L less than exponential in n) can have rate at most $\varepsilon + o(1)$. Theorem 4 matches this rate upper bound up to the constant the factor 15. In previous work, Guo, Li, Shangguan, Tamo, and Wootters [2] showed a similar result with rate $\Omega(\varepsilon/\log(1/\varepsilon))$, matching the list-decoding capacity upper bound up to a logarithmic factor. They stated that their "motivating question is

whether or not RS codes can be list-decoded up to radius $1-\varepsilon$ with rates $\Omega(\varepsilon)$", and Theorem 4 answers this question affirmatively.

Our proof of Theorem 3 also show a more general result about list-decodability of random puncturings of a given code with very large distance, but this talk was only focused on Reed–Solomon codes.

References

[1] P. Elias, *List decoding for noisy channels*, Wescon Convention Record, Part 2, Institute of Radio Engineers, pp. 99–104, 1957.

[2] Z. Guo, R. Li, C. Shangguan, I. Tamo, and M. Wootters, *Improved List-Decodability and List-Recoverability of Reed-Solomon Codes via Tree Packings*, 2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS), 2022, pp. 708–719.

[3] I. S. Reed and G. Solomon, *Polynomial codes over certain finite fields*, J. Soc. Indust. Appl. Math. **8** (1960), 300–304.

[4] A. Rudra and M. Wootters, *Every list-decodable code for high noise has abundant near-optimal rate puncturings*, 46th Annual ACM Symposium on Theory of Computing (STOC 2014), pages 764–773, 2014.

[5] C. Shangguan and I. Tamo, *Combinatorial list-decoding of Reed-Solomon codes beyond the Johnson radius*, 52nd Annual ACM Symposium on Theory of Computing (STOC 2020), pp. 538–551, 2020.

[6] R. Singleton, *Maximum distance q-nary codes*, IEEE Trans. Inform. Theory **10** (1964), 116–118.

[7] J. M. Wozencraft, *List decoding*, In Quarterly Progress Report, Research Laboratory of Electronics, MIT, vol. 48, pages 90–95, 1958.

On the Threshold for Steiner Triple Systems

Mehtaab Sawhney

(joint work with Ashwin Sah, Michael Simkin)

We consider the following problem: for what p does the random r-uniform hypergraph, $\mathbb{G}^{(r)}(n,p)$ on n vertices contain an (n,q,r) design? Recall that an (n,q,r) design is a system of sets of size q such that every set of size r is contained in exactly 1 set of size q and $\mathbb{G}^{(r)}(n,p)$ is the random r-uniform hypergraph where every set r-edge is contained with probability p.[1]

This question contains a number of questions which themselves been sources of intense study. First, the deterministic question for $p=1$ of the mere existence of (n,q,r)-designs was a long standing open question which was famously resolved by Keevash [1] via a new method of Randomized Algebraic Constructions. Subsequently a second proof, was given by Glock, Kühn, Lo, and Osthus [2] based on the method of iterative absorption. Second, for $q=2$ and $r=1$, note that an $(n,2,1)$ design is simply a perfect matching and therefore the question specializes to the threshold for $\mathbb{G}(n,p)$ containing a perfect matching. This was resolved in foundational work of Erdős and Rényi [5] with a hitting time result being proved in later work of Bollobás and Thomason [4]. The more general question of when

[1]Note that for this question to make sense, it is necessary to restrict to n which satisfy the necessary divisibility conditions for an (n,q,r) design to exist.

$\mathbb{G}^{(r)}(n,p)$ contains an $(n,r,1)$ design was an infamous open problem, going under the name of Shamir's problem. The threshold for containing an $(r,1)$-design was first resolved up to constant factors in seminal work of Johansson, Kahn, and Vu [3] and then later with a sharp constant and a hitting time result being proved in difficult work of Kahn [7, 6].

Our main result is resolving the threshold for containing a $(3,2)$-design up to a sub-polynomial factor.

Theorem 1. *There exists $C>0$ such that*

$$\lim_{\substack{n\to\infty \\ n\equiv 1,3 \mod 6}} \mathbb{P}[\mathbb{G}^{(3)}(n,\exp(C(\log n)^{3/4})n^{-1}) \text{ contains a } (3,2)\text{-design}] = 1.$$

We note that by coupon-collector that the threshold for containing a $(3,2)$-design (or a Steiner triple system) is $\gtrsim (\log n)/n$ and therefore the above theorem, with results of Friedgut [10], imply a sharp threshold sequence for the property of containing a Steiner triple system. We also prove the analogous result for Latin squares.

Our result relies on recent breakthrough work due to Frankston, Kahn, Narayan, and Park [8] resolving the Fractional Expectation Threshold vs Threshold conjecture. (We note that more recently, the full Expectation Threshold vs Threshold conjecture, was resolved in breakthrough work of Park and Pham [9].) The crucial ingredient we require is the following result regarding *spread*-measures.

Theorem 2 (From [8, Theorem 1.6])**.** *Consider a finite ground set Z and fix a nonempty collection of subsets $\mathcal{H}\subseteq 2^Z$. Let μ be a probability measure on $\mathcal{H}$. For $q>0$ we say that μ is q-spread if for every set $S\subseteq Z$:*

$$\mu(\{A\in\mathcal{H}: S\subseteq A\}) \leqslant q^{|S|}.$$

There exists a constant $C>0$ such that the following holds. Suppose that there exists a q-spread probability measure on $\mathcal{H}$. Then a random binomial subset of Z where each element is sampled with probability $\min(Cq\log|Z|,1)$ contains an element of $\mathcal{H}$ as a subset with probability at least $3/4$.

The key task therefore is to produce a sufficiently spread measure on the set of Steiner triple systems. In previous applications of [8, Theorem 1.6], the measure μ was taken to be uniform over all possible copies of the embedded structure. However results regarding the enumeration of Steiner triple systems are too poor to directly pursue this strategy (see [8, Section 8.D] for further discussion).

Instead our construction of a spread measure of Steiner triple systems follows from a delicate boostrapping procedure. Roughly the proof proceeds in four steps. First, we set aside a sparse set of triangles of suitable spread on top of a set of vertices X. Second, using tools from iterative absorption, we cover all edges outside of X with spread $|X|^{-1+o(1)}$. Third by induction, we cover the edges within X with a certain spread. Finally, using the triangles set aside initially we "delocalize" the spread with certain gadgets and a careful analysis reveals that this procedure allows one to iteratively construct measures of improved spread.

References

[1] P. Keevash, The existence of designs, preprint available at arXiv:1401.3665 [math.CO].
[2] S. Glock, D. Kühn, A. Lo, and D. Osthus, *The existence of designs via iterative absorption: hypergraph F-designs for arbitrary F*, Mem. Amer. Math. Soc. (to appear).
[3] A. Johansson, J. Kahn, and V. Vu. *Factors in random graphs*. Random Structures Algorithms 33 (2008), no. 1, 1–28.
[4] B. Bollobás and A. Thomason. *Random graphs of small order*. Random graphs '83 (Poznań, 1983), 47–97, North-Holland Math. Stud., 118, Ann. Discrete Math., 28, North-Holland, Amsterdam, 1985.
[5] P. Erdős., A. Rényi. *On the existence of a factor of degree one of a connected random graph*. Acta Math. Acad. Sci. Hungar. 17 (1966), 359–368.
[6] J. Kahn. *Hitting times for Shamir's problem*. Trans. Amer. Math. Soc. 375 (2022), no. 1, 627–668.
[7] J. Kahn, *Asymptotics for Shamir's Problem* preprint available at arXiv:1909.06834 [math.CO].
[8] K. Frankston, J. Kahn, B. Narayanan, and J. Park, *Thresholds versus fractional expectation-thresholds,* Ann. of Math. (2) **194** (2021), no. 2, 475–495.
[9] J. Park and H. T. Pham, *A proof of the Kahn-Kalai conjecture,* preprint available at arXiv:2203.17207 [math.CO].
[10] E. Friedgut. *Hunting for sharp thresholds*. Random Structures Algorithms 26 (2005), no. 1-2, 37–51.

Induced subgraphs of induced subgraphs of large chromatic number

Alex Scott

(joint work with António Girão, Freddie Illingworth, Emil Powierski, Michael Savery, Youri Tamitegama and Jane Tan)

It is a fundamental problem of graph theory to understand what structures must appear in graphs of large chromatic number. A straightforward reason for a graph to have large chromatic number is the presence of a large clique. However, this is not necessary as there are examples of triangle-free graphs with large chromatic number (see, for instance, Tutte [3] and Zykov [10] among many classical constructions).

If a graph has large chromatic number and small clique number, it is reasonable to ask whether it contains induced subgraphs with large chromatic number which are 'simple' in some way. A natural conjecture, attributed to Esperet (see [9]), asked whether all graphs with large chromatic number contain either a large clique or a triangle-free induced subgraph with large chromatic number (the non-induced version of this was proven by a beautiful argument of Rödl [7]). This conjecture was recently disproved by Carbonero, Hompe, Moore and Spirkl [2], who found a surprising new twist on a construction of Kierstead and Trotter [5], and proved the following.

Theorem 1 (Carbonero, Hompe, Moore and Spirkl [2])**.** *There are graphs of arbitrarily large chromatic number that contain neither a K_4 nor an induced triangle-free subgraph of chromatic number greater than four.*

Let us write $\omega(H)$ for the *clique number* (the maximum number of vertices in a complete subgraph) of a graph H and $\chi(H)$ for the chromatic number of H. We say that a graph G is *F-free* if it does not contain an induced copy of F.

Briański, Davies and Walczak [1] extended the result of Carbonero, Hompe, Moore and Spirkl to cliques of prime order, in an ingenious paper proving the following.

Theorem 2 (Briański, Davies and Walczak [1])**.** *For every prime p, there are K_{p+1}-free graphs G of arbitrarily large chromatic number such that every K_p-free induced subgraph H of G satisfies $\chi(H) \leqslant \omega(H)^{\omega(H)^2}$.*

As a consequence, they showed that there are classes of graphs that are χ-bounded but not polynomially χ-bounded, disproving another conjecture of Esperet [4] (see [8] and [9] for definitions and related discussion).

Theorems 1 and 2 show that there are K_{p+1}-free graphs of large chromatic number in which every induced subgraph of large chromatic number contains a copy of K_p. It is natural to ask whether anything is true for other graphs F. In other words, is there a graph G with small clique number and large chromatic number such that every induced subgraph of G with large chromatic number contains an induced copy of F? Perhaps surprisingly, our main theorem answers this in the affirmative for every nontrivial graph F. Indeed, G can be taken to have the same clique number as F.

Theorem 3. *For every graph F with at least one edge, there is a constant c_F and graphs G of arbitrarily large chromatic number and the same clique number as F such that every F-free induced subgraph of G is c_F-colourable.*

Theorem 3 gives a graph G with the same clique number as F. In the case when F is triangle-free, it is natural to ask whether we can take our graphs G to have the same girth as F. We conjecture that this is the case.

Conjecture 4. *For every graph F with at least one cycle, there exists a constant b_F and graphs G of arbitrarily large chromatic number and the same girth as F such that every F-free induced subgraph of G is b_F-colourable.*

It would already be interesting to prove this in the special case where F is the 5-cycle. We prove a weaker version of the conjecture, where girth is replaced by *odd girth*, the length of the shortest odd cycle.

Theorem 5. *For every nonbipartite graph F, there is a constant c'_F and graphs G of arbitrarily large chromatic number and the same odd girth as F such that every F-free induced subgraph of G is c'_F-colourable.*

The proofs of Theorems 3 and 5 follow a common framework, building on the arguments of Carbonero, Hompe, Moore and Spirkl [2], and Briański, Davies and Walczak [1]; but the constructions are much more delicate and require additional ideas. In the construction of Theorem 3, it is convenient to begin with the oriented Zykov graph (which was also used in [1] and [2]). However, for Theorem 5, we introduce an oriented version of graphs of Nešetril and Rödl [6]. A key ingredient

of the proof, which may be of independent interest, is to show that every cycle in these graphs has a large number of direction changes. Another new ingredient in both proofs is the use of B_h-sets to control the appearance of the graph F.

References

[1] M. Briański, J. Davies and B. Walczak. Separating polynomial χ-boundedness from χ-boundedness. arXiv:2201.08814.

[2] A. Carbonero, P. Hompe, B. Moore and S. Spirkl. A counterexample to a conjec- ture about triangle-free induced subgraphs of graphs with large chromatic number. arXiv:2201.08204.

[3] B. Descartes. k-chromatic graphs without triangles. American Mathematical Monthly, 61(5):352–353, 1954.

[4] L. Esperet. Graph colorings, flows and perfect matchings. Habilitation thesis, Uni- versité Grenoble Alpes, 2017.

[5] H. A. Kierstead and W. T. Trotter. Colorful induced subgraphs. Discrete Mathe- matics, 101:165–169, 1992.

[6] J. Nešetřil and V. Rődl. Partitions of vertices. Commentationes Mathematicae Universitatis Carolinae, 17(1):85–95, 1976. 22.

[7] V. Rödl. On the chromatic number of subgraphs of a given graph. Proceedings of the American Mathematical Society, 64:370–371, 1977

[8] A. Scott. Graphs of large chromatic number. In: Proceedings of the ICM 2022, to appear.

[9] A. Scott and P. Seymour. A survey of χ-boundedness. Journal of Graph Theory, 95(3):473–504, 2020.

[10] A. Zykov. On some properties of linear complexes (in Russian). Matematicheskiĭ Sbornik, 24:163–188, 1949.

Weakly Saturated Hypergraphs and a Conjecture of Tuza

Asaf Shapira

(joint work with M. Tyomkyn)

Typical problems in extremal combinatorics, e.g. the Turán and Ramsey problems, ask how large or small a discrete structure can be, assuming it possesses certain properties. While in many cases it seems hopeless to obtain full solutions to these problems, one would at least like to know that these extremal functions are "well behaved". For example, it is natural to ask if the quantities $\mathrm{ex}(n,H)/n^r$ and $R(n)^{1/n}$ tend to a limit. While it is easy to see that the first quantity indeed tends to a limit [18], it is a famous open problem of Erdős [9, 10, 12] to prove that the second one does so as well. Our aim in this paper is to prove that another well studied extremal function is well behaved.

For a fixed r-graph H, an r-graph $G=(V,E)$ is called *H-saturated* if it does not contain a copy of H but for any edge $e \in \binom{V}{r}\backslash E(G)$ adding e to G creates a copy of H. We let $\mathrm{sat}(n,H)$ denote the smallest number of edges in an H-saturated r-graph on n vertices. Let K_t^r denote the complete r-graph on t vertices; when $r=2$ (i.e. when dealing with graphs) we use K_t instead of K_t^2. The problem of determining $\mathrm{sat}(n,K_t)$ was raised by Zykov [28] in the 1940's and studied in the 1960's by Erdős, Hajnal and Moon [11] who showed that $\mathrm{sat}(n,K_t)=\binom{n}{2}-\binom{n-t+2}{2}$. Their result was later generalized by Bollobás [5] who showed that $\mathrm{sat}(n,K_t^r)=$

$\binom{n}{r} - \binom{n-t+r}{r}$. It is worth noting that the proof in [5] introduced the (equivalent and) highly influential *Two Families Theorem*.

We say that G is *weakly H-saturated* if the edges of $\binom{V}{r}\backslash E(G)$ admit an ordering $e_1, \dots, e_k$ such that for each $i = 1, \dots, k$ the r-graph $G_i := G \cup \{e_1, \dots, e_i\}$ contains a copy of H containing the edge e_i. We refer to the sequence $e_1, \dots, e_k$ as a *saturation process*. Define $\mathrm{wsat}(n, H)$ to be the smallest number of edges in a weakly H-saturated r-graph on n vertices. Note that we may automatically assume that any G realizing $\mathrm{wsat}(n, H)$ is H-free, as otherwise we could remove an edge from a copy of H in G to obtain a smaller weakly H-saturated r-graph. Hence weak saturation can be viewed as an extension of the notion of (ordinary) saturation. The problem of determining $\mathrm{wsat}(n, H)$ was first introduced in 1968 by Bollobás [6] who conjectured that $\mathrm{wsat}(n, K_t) = \mathrm{sat}(n, K_t)$. This was proved independently by Frankl [15] and Kalai [16, 17] using the skewed variant of Bollobás's Two Families Theorem (a related statement for matroids was proven earlier by Lovász [19]) and further extended by Alon [1] and Blokhuis [4]. This result, which has several other equivalent formulations, is amongst the most classical and important results of extremal combinatorics. See e.g. the discussions in [2, 20, 23, 25].

At this point it is natural to ask if for every H there is a constant C_H so that

$$\mathrm{wsat}(n, H) = (C_H + o(1))n. \tag{1}$$

Such a result was obtained in 1985 by Alon [1], who proved that for graphs the function $\mathrm{wsat}(n, H)$ is (essentially) subadditive, implying that $\mathrm{wsat}(n, H)/n$ tends to a limit, by Fekete's subadditivity lemma [14].

Much less was known when H is an r-graph with $r \geqslant 3$. Bollobás's construction from [5] gives a simple bound of

$$\mathrm{wsat}(n, H) \leqslant \mathrm{sat}(n, H) = O_H(n^{r-1}).$$

A more refined result was obtained by Tuza [27] who introduced the following key definition. The *sparseness* of an r-graph H, denoted $s(H)$, is the smallest size of a vertex set $W \subseteq V$ contained in precisely one edge of H; note that $1 \leqslant s(H) \leqslant r$ for every non-empty r-graph H. It was proved in [27] that for every r-graph H there are two positive reals c_H and C_H such that

$$c_H \cdot n^{s-1} \leqslant \mathrm{wsat}(n, H) \leqslant C_H \cdot n^{s-1}. \tag{2}$$

It was further conjectured in [27] that the more refined bound $\mathrm{wsat}(n, H) = C_H \cdot n^{s-1} + O(n^{s-2})$ holds for every r-graph of sparseness s. See also the recent survey [8] on saturation problems where this conjecture is further discussed. Since such a result is not known even for graphs (i.e. when $r = s = 2$), Tuza [27] asked if one can improve upon (2) by showing that for every r-graph we have $\mathrm{wsat}(n, H) = C_H \cdot n^{s-1} + o(n^{s-1})$ where $s = s(H)$. Prior to this work, such a result was only known for $r = 2$ by Alon's result (1). In this paper we fully resolve Tuza's problem for all r-graphs.

Theorem 1. *For every r-graph H there is $C_H > 0$ such that*

$$\lim_{n\to\infty} \mathit{wsat}(n, H)/n^{s-1} = C_H,$$

where $s = s(H)$ is the sparseness of H. In particular, for every r-graph H there is $C'_H \geqslant 0$ such that

$$\lim_{n\to\infty} wsat(n, H)/n^{r-1} = C'_H.$$

Proof and paper overview: It is natural to ask why Alon's [1] one-paragraph proof of Theorem 1 for $s = 2$ is hard to extend to $s > 2$. Perhaps the simplest reason is that one cannot hope to show that in these cases the function $\mathrm{wsat}(n, H)$ is subadditive since a subadditive function is necessarily of order $O(n)$, while we know from (2) that when $s \geqslant 3$ the function $\mathrm{wsat}(n, H)$ is of order at least n^2. The main novelty in this paper is in finding a direct and efficient way to use an m-vertex r-graph witnessing the fact that $\mathrm{wsat}(m, H)$ is small, in order to build arbitrarily large n-vertex r-graphs witnessing the fact that $\mathrm{wsat}(n, H)$ is small. One of the main tools we use to construct such an example is Rödl's approximate designs theorem [24] which enables us to efficiently combine many examples of size m into one of size n. Rödl's result would only allow us to construct a saturation process generating part of the edges of K_n^r. To complete this saturation process we would also need another set of gadgets.

References

[1] N. Alon. An extremal problem for sets with applications to graph theory. *J. Combin. Theory Ser. A*, 40(1):82–89, 1985.

[2] L. Babai and P. Frankl. *Linear algebra methods in combinatorics*. University of Chicago, 1988.

[3] N.C. Behague. Hypergraph saturation irregularities. *The Electronic Journal of Combinatorics*, 25:#P2.11, 2018.

[4] A. Blokhuis. Solution of an extremal problem for sets using resultants of polynomials. *Combinatorica*, 10(4):393–396, 1990.

[5] B. Bollobás. On generalized graphs. *Acta Mathematica Academiae Scientiarum Hungarica*, 16:447–452, 1965.

[6] B. Bollobás. Weakly k-saturated graphs. In *Beiträge zur Graphentheorie (Kolloquium, Manebach, 1967)*, pages 25–31. Teubner, Leipzig, 1968.

[7] D. Chakraborti and P-S. Loh. Minimizing the numbers of cliques and cycles of fixed size in an F-saturated graph. *European Journal of Combinatorics*, 90:103185, 2020.

[8] B.L. Currie, J.R. Faudree, R.J. Faudree, and J.R. Schmitt. A survey of minimum saturated graphs. *The Electronic Journal of Combinatorics*, 1000:DS19–Oct, 2021.

[9] P. Erdős. Problems and results in graph theory. *The theory and applications of graphs (Kalamazoo, MI, 1980)*, pages 331–341, 1981.

[10] P. Erdős. Some new problems and results in graph theory and other branches of combinatorial mathematics. In *Combinatorics and Graph Theory*, pages 9–17. Springer, 1981.

[11] P. Erdős, A. Hajnal, and J. W. Moon. A problem in graph theory. *The American Mathematical Monthly*, 71(10):1107–1110, 1964.

[12] P. Erdős. On the combinatorial problems which I would most like to see solved. *Combinatorica*, 1(1):25–42, 1981.

[13] R. J. Faudree, R. J. Gould, and M. S. Jacobson. Weak saturation numbers for sparse graphs. *Discuss. Math. Graph Theory*, 33(4):677–693, 2013.

[14] M. Fekete. Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten. *Mathematische Zeitschrift*, 17(1):228–249, 1923.

[15] P. Frankl. An extremal problem for two families of sets. *European J. Combin.*, 3(2):125–127, 1982.

[16] G. Kalai. Weakly saturated graphs are rigid. In *Convexity and graph theory (Jerusalem, 1981)*, volume 87 of *North-Holland Math. Stud.*, pages 189–190. North-Holland, Amsterdam, 1984.
[17] G. Kalai. Hyperconnectivity of graphs. *Graphs Combin.*, 1(1):65–79, 1985.
[18] G. Katona, T. Nemetz, and M. Simonovits. On a problem of turán in the theory of graphs. *Mat. Lapok*, 15:228–238, 1964.
[19] L. Lovász. Flats in matroids and geometric graphs. In *Combinatorial surveys (Proc. Sixth British Combinatorial Conf., Royal Holloway Coll., Egham, 1977)*, pages 45–86, 1977.
[20] J. O'Neill and J. Verstraete. A generalization of the Bollobás set pairs inequality. *arXiv preprint arXiv:1812.00537*, 2018.
[21] O. Pikhurko. The minimum size of saturated hypergraphs. *Combinatorics, Probability and Computing*, 8(5):483–492, 1999.
[22] O. Pikhurko. Results and open problems on minimum saturated hypergraphs. *Ars Comb.*, 72:111–127, 07 2004.
[23] J.-E. Pin. On two combinatorial problems arising from automata theory. In *North-Holland Mathematics Studies*, volume 75, pages 535–548. Elsevier, 1983.
[24] V. Rödl. On a packing and covering problem. *European Journal of Combinatorics*, 6(1):69–78, 1985.
[25] A. Scott and E. Wilmer. Combinatorics in the exterior algebra and the Bollobás Two Families Theorem. *arXiv preprint arXiv:1907.06019*, 2019.
[26] Z. Tuza. A generalization of saturated graphs for finite languages. *Tanulmányok–MTA Számitástech. Automat. Kutató Int. Budapest,(185)*, pages 287–293, 1986.
[27] Z. Tuza. Asymptotic growth of sparse saturated structures is locally determined. *Discrete Math.*, 108(1-3):397–402, 1992. Topological, algebraical and combinatorial structures. Frolík's memorial volume.
[28] A. Zykov. On some properties of linear complexes. *Matematicheskii sbornik*, 66(2):163–188, 1949. In Russian.

Ringel's tree packing conjecture in quasirandom graphs

Katherine Staden
(joint work with Peter Keevash)

The following conjecture was posed by Ringel [13] in 1963.

Ringel's Conjecture. For any tree T with n edges, the complete graph K_{2n+1} has a decomposition into $2n+1$ copies of T.

I discuss our 2020 proof of this conjecture for large n, via the following theorem which is a generalisation to decompositions of quasirandom graphs into trees of the appropriate size. For the statement we use the following quasirandomness definition: we say that a graph G on n vertices is (ξ, s)-*typical* if every set S of at most s vertices has $((1\pm\xi)d(G))^{|S|}n$ common neighbours, where $d(G) = e(G)\binom{n}{2}^{-1}$ is the density of G.

Theorem 1. *There is $s \in \mathbb{N}$ such that for all $0 < p \leqslant 1$ there exist ξ, n_0 such that for any $n \geqslant n_0$ such that $p(n-1)/2 \in \mathbb{Z}$ and any tree T of size $p(n-1)/2$, any (ξ, s)-typical graph G on n vertices of density p can be decomposed into n copies of T.*

The case $p = 1$ of Theorem 1 establishes Ringel's conjecture for large n, a result also obtained in 2020 independently by Montgomery, Pokrovskiy and Sudakov [12] by different methods, along the lines of their proof of an asymptotic version in [11]. They show that certain edge-colourings of K_{2n+1} contain a rainbow copy of T, such that the required T-decomposition can be obtained by cyclically shifting this rainbow copy. This approach is specific to the complete graph, and does not apply to the more general setting of quasirandom graphs as in Theorem 1.

Ringel's conjecture was well-known as one of the major open problems in the area of *graph packing*. In a graph packing problem, one is given a host graph G and a guest graph F and the task is to fit as many edge-disjoint copies of F into G as possible. If the size (number of edges) of F divides that of G, it may be possible to find a perfect packing, or *F-decomposition* of G. More generally, given a family $\mathcal{F}$ of graphs of total size equal to the size of G, we seek a partition of (the edge set of) G into copies of the graphs in $\mathcal{F}$.

These problems have a long history, going back to Euler in the eighteenth century. The flavour of the problem depends very much on the size of F. The earliest results concern F of fixed size, in which case F-decompositions can be naturally interpreted as combinatorial designs.

There is also a large literature on F-decompositions where the number of vertices of F is comparable with, or even equal to, that of G. Classical results of this type are Walecki's 1882 decompositions of K_{2n} into Hamilton paths, and of K_{2n+1} into Hamilton cycles. There are many further results on Hamilton decompositions of more general host graphs.

Much of the literature on F-decompositions for large F concerns decompositions into trees. Besides Ringel's conjecture, the other major open problem of this type is a conjecture of Gyárfás [6], saying that K_n should have a decomposition into any family of trees $T_1, \dots, T_n$ where each T_i has i vertices. Both conjectures have a large literature of partial results; the most significant of these include the following. Joos, Kim, Kühn and Osthus [7] proved both conjectures for bounded degree trees. Ferber and Samotij [5] and Adamaszek, Allen, Grosu and Hladký [2] obtained almost-perfect packings of almost-spanning trees with maximum degree $O(n/\log n)$. These results were generalised by Allen, Böttcher, Hladký and Piguet [4] to almost-perfect packing of spanning graphs with bounded degeneracy and maximum degree $O(n/\log n)$. This was extended in [3] to perfect packings provided linearly many of the graphs are slightly smaller than spanning and have linearly many leaves. Very recently, Gyárfás's conjecture for maximum degree $O(n/\log n)$ trees was proved in [1]. The above results mainly use randomised embeddings, for which a maximum degree bound $O(n/\log n)$ is necessary for concentration of probability. While the results of Montgomery, Pokrovskiy and Sudakov [10, 11] also use probabilistic methods, they are able to circumvent the maximum degree barrier by methods such as the cyclic shifts mentioned above.

Our proof proceeds via a rather involved embedding algorithm, in which the various subroutines are analysed by a wide range of methods, some of which are adaptations of existing methods (particularly from [10] and [3], and also our own

recent methods in [8] for the 'generalised Oberwolfach problem', which are in turn based on [9]), but most of which are new, including a method for allocating high degree vertices and their edges via partitioning and edge-colouring arguments and a method for approximate decompositions based on a series of matchings in auxiliary hypergraphs.

References

[1] P. Allen, J. Böttcher, D. Clemens, J. Hladký, D. Piguet and A. Taraz, *The tree packing conjecture for trees of almost linear minimum degree*, arXiv:2106.11720 (2021).

[2] A. Adamaszek, P. Allen, C. Grosu and J. Hladký, *Almost all trees are almost graceful*, Random Struct. Algo. **56** (4) (2020), 948–987.

[3] P. Allen, J. Böttcher, D. Clemens and A. Taraz, *Perfectly packing graphs with bounded degeneracy and many leaves*, arXiv:1906.11558 (2019).

[4] P. Allen, J. Böttcher, J. Hladký and D. Piguet, *Packing degenerate graphs*, Adv. Math. **354** (2019).

[5] A. Ferber and W. Samotij, *Packing trees of unbounded degrees in random graphs*, J. Lond. Math. Soc. **99** (2016), 653–677.

[6] A. Gyárfás and J. Lehel, *Packing trees of different order into K_n*, in Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Volume 1, (1978), 463–469.

[7] F. Joos, J. Kim, D. Kühn and D. Osthus, *Optimal packings of bounded degree trees*, J. European Math. Soc. **21** (2019), 3573–3647.

[8] P. Keevash and K. Staden, *The generalised Oberwolfach problem*, J. Combin. Theory B **152** (2022), 281–318.

[9] P. Keevash, *The existence of designs II*, arXiv:1802.05900 (2018).

[10] R. Montgomery, A. Pokrovskiy and B. Sudakov, *Decompositions into spanning rainbow structures*, Proc. London Math. Soc. **119** (2019), 899–959.

[11] R. Montgomery, A. Pokrovskiy and B. Sudakov, *Embedding rainbow trees with applications to graph labelling and decomposition*, J. European Math. Soc. **22** (2020), 3101–3132.

[12] R. Montgomery, A. Pokrovskiy and B. Sudakov, *A proof of Ringel's conjecture*, Geom. Func. Anal. **31** (2021), 663–720.

[13] G. Ringel, *Theory of graphs and its applications*, in Proc. Symp. Smolenice (1963).

Erdős's conjecture on the pancyclicity of Hamiltonian graphs

Benny Sudakov

(joint work with Nemanja Draganić, David Munhá Correia)

Hamiltonicity is one of the most central notions in graph theory, and it has been extensively studied by numerous researchers. The problem of deciding Hamiltonicity of a graph is NP-complete, but a central theme in Combinatorics is to derive sufficient conditions for this property. One example is the classical Dirac's theorem [8], which states that every n-vertex graph with minimum degree at least $n/2$ contains a Hamilton cycle. A closely related notion is that of pancyclicity - a graph is said to be *pancyclic* if it contains all cycles of length from 3 up to n.

In 1973, Bondy [3] stated his celebrated meta-conjecture that any non-trivial condition which implies that a graph is Hamiltonian should also imply that it is pancyclic (up to a certain collection of simple exceptional graphs). As an example, Bondy [2] himself first showed that every n-vertex graph with minimum degree of at least $n/2$ is either pancyclic or isomorphic to the complete bipartite graph

$K_{n/2,n/2}$, thus strengthening Dirac's theorem. This meta-conjecture sparked a lot of research which led to various appealing results and methods. For example, Bauer and Schmeichel [1], relying on previous results of Schmeichel and Hakimi [15] showed that the sufficient conditions for Hamiltonicity given by Bondy [4], Chvátal [6] and Fan [10] all imply pancyclicity, up to a certain small family of exceptions. Also, Jackson and Ordaz [11] conjectured that if a graph G has connectivity $\kappa(G)$ strictly larger than its independence number $\alpha(G)$, then G must be pancyclic. This was motivated by the classical theorem of Chvátal and Erdős [6] that a graph with $\kappa(G) \geqslant \alpha(G)$ must be Hamiltonian. An approximate form of the conjecture was proven by Keevash and Sudakov [12], who showed that $\kappa(G) \geqslant 600\alpha(G)$ is already sufficient for pancyclicity.

Bondy's meta-conjecture deals with conditions for Hamiltonicity which imply pancyclicity. In a similar vein, one can ask the following natural question: Let G be a Hamiltonian graph; under which assumptions can we guarantee that G is also pancyclic or more generally, that it has many cycle lengths? An example of a problem of this type was introduced by Jacobson and Lehel at the 1999 conference "Paul Erdős and His Mathematics", where they asked for the minimal number of cycle lengths in a k-regular n-vertex Hamiltonian graph. They conjectured (see Verstraëte [16] for a stronger conjecture) that already when $k \geqslant 3$, there are $\Omega(n)$ many lengths. Recently, it was shown by Bucić, Gishboliner and Sudakov [5] that any Hamiltonian graph with minimum degree at least 3 has $n^{1-o(1)}$ different cycle lengths; previously, the best known lower bound was of order $\sqrt{n}$ [14].

In this paper, we consider another problem in the area. Around the same time when Bondy stated his meta-conjecture, in 1972, Erdős [9] put forward the following question: Given an n-vertex Hamiltonian graph with independence number $\alpha(G) \leqslant k$, how large does n have to be in terms of k in order to guarantee that G is pancyclic? He proved that it is enough to have $n = \Omega(k^4)$ and conjectured that already $n = \Omega(k^2)$ should be enough. A simple construction shows that this is best possible. Let $C_1, \ldots, C_k$ be disjoint cliques of size $2k-2$, and let each C_i have two distinguished vertices a_i and b_i. Let G be the graph obtained by connecting a_i and b_{i+1} by an edge (with addition modulo k). One can easily see that this graph, which has $2k(k-1)$ vertices, is Hamiltonian and its independence number is k. However, it does not contain a cycle of length $2k-1$ (and thus, it is not pancyclic), since every cycle must be either a subgraph of one of the cliques C_i, or contain at least two vertices from each clique. The former type of cycles all have length at most $2k-2$, whereas the latter have length at least $2k$.

Since then, there has been several improvements upon Erdős's initial result – by Keevash and Sudakov [12] who proved that $n = \Omega(k^3)$ is enough, by Lee and Sudakov [13] who improved it to $n = \Omega(k^{7/3})$, and finally by Dankovics [7] who showed that $n = \Omega(k^{11/5})$ suffices. In this paper we resolve the conjecture of Erdős, showing that if a Hamiltonian graph G has $n = \Omega(k^2)$ vertices and $\alpha(G) \leqslant k$, then G is pancyclic.

Theorem 1. *Every Hamiltonian graph G with $\alpha(G) \leqslant k$ and at least ck^2 vertices is pancyclic, for a large enough absolute constant c.*

The main part of the proof is to show that we can find a cycle of length $n-1$, as one can then use this to obtain pancyclicity using the already existing techniques of Keevash and Sudakov. Our proof is quite short and relies on a new idea of finding a certain structure which we call a special matching, which has the property that any edge added between an appropriate pair of vertices in the special matching, creates a cycle of length $n-1$.

References

[1] Douglas Bauer and Edward Schmeichel. Hamiltonian degree conditions which imply a graph is pancyclic. *Journal of Combinatorial Theory, Series B*, 48(1):111–116, 1990.

[2] J Adrian Bondy. Pancyclic graphs i. *Journal of Combinatorial Theory, Series B*, 11(1):80–84, 1971.

[3] J Adrian Bondy. Pancyclic graphs: recent results, infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), vol. i, 181–187. In *Colloq. Math. Soc. János Bolyai*, volume 10.

[4] J Adrian Bondy. *Longest paths and cycles in graphs of high degree*. Department of Combinatorics and Optimization, University of Waterloo, 1980.

[5] Matija Bucić, Lior Gishboliner, and Benny Sudakov. Cycles of many lengths in Hamiltonian graphs. In *Extended Abstracts EuroComb 2021*, pages 366–372. Springer, 2021.

[6] Václav Chvátal. On Hamilton's ideals. *Journal of Combinatorial Theory, Series B*, 12(2):163–168, 1972.

[7] Attila Dankovics. Low independence number and Hamiltonicity implies pancyclicity. *Journal of Graph Theory*, 95(2):181–191, 2020.

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

[9] Paul Erdős. Some problems in graph theory. In *Hypergraph Seminar*, pages 187–190. Springer, 1972.

[10] Geng-Hua Fan. New sufficient conditions for cycles in graphs. *Journal of Combinatorial Theory, Series B*, 37(3):221–227, 1984.

[11] Bill Jackson and Oscar Ordaz. Chvátal-Erdős conditions for paths and cycles in graphs and digraphs. a survey. *Discrete mathematics*, 84(3):241–254, 1990.

[12] Peter Keevash and Benny Sudakov. Pancyclicity of Hamiltonian and highly connected graphs. *Journal of Combinatorial Theory, Series B*, 100(5):456–467, 2010.

[13] Choongbum Lee and Benny Sudakov. Hamiltonicity, independence number, and pancyclicity. *European Journal of Combinatorics*, 33(4):449–457, 2012.

[14] Kevin G Milans, Florian Pfender, Dieter Rautenbach, Friedrich Regen, and Douglas B West. Cycle spectra of Hamiltonian graphs. *Journal of Combinatorial Theory, Series B*, 102(4):869–874, 2012.

[15] Edward F Schmeichel and S Louis Hakimi. A cycle structure theorem for Hamiltonian graphs. *Journal of Combinatorial Theory, Series B*, 45(1):99–107, 1988.

[16] Jacques Verstraëte. Extremal problems for cycles in graphs. In *Recent trends in combinatorics*, pages 83–116. Springer, 2016.

Crossing probabilities in planar percolation

Vincent Tassion
(joint work with Laurin Köhler-Schindler)

We consider percolation on the square lattice. The most famous percolation model, called Bernoulli percolation, is defined as follows. Fix a parameter $p \in [0,1]$. For each edge of the square lattice, toss independently a biased coin: the edge is open with probability p and closed otherwise. We are interested in the connectivity properties of the resulting random graph, obtained by keeping only the open edges.

These connectivity properties are encoded by so-called **crossing probabilities** (probabilities that certain rectangles are crossed from left to right). In the eighties, Russo, Seymour and Welsh obtained general bounds on these crossing probabilities for Bernoulli percolation, which rapidly became a central tool to analyze the critical behaviour of the model.

Theorem 1 (RSW theory [Rus81, SW78])**.** *Consider Bernoulli percolation on the square lattice at parameter $p = 1/2$. Fix $\rho > 0$. There exists $c = c(\rho) > 0$ such that,*

$$\forall n \geqslant 1 \quad c \leqslant \mathbf{P}\Big[\overset{\rho n}{\boxed{}} n \Big] \leqslant 1 - c.$$

The proof has two steps: the first step relies on self-duality of the model at $p = 1/2$, which implies that squares are crossed with probability $1/2$. The second step is the derivation of bounds for rectangle crossings. The derivation of bounds on rectangle crossings from bounds on square crossings does not follow from a trivial argument, and is at the heart of RSW theory.

For many years, this powerful Russo-Seymour-Welsh (RSW) theory only applied to a very restricted class of models with strong independence properties. In the last 20 years, this RSW lemma has been successfully extended to other percolation models, such as Voronoi percolation [BR06a, Tas16] or FK percolation [BD12, DST17], leading to important new results and applications (computation of the critical point, derivation of fractal properties of the critical regime, description of the near-critical regime,...).

Almost all the models with a rigorously established RSW theory satisfy two properties:

[Symmetry] The percolation measure is invariant under the symmetries of $\mathbb{Z}^2$ (translations, $\pi/2$-rotations, horizontal and vertical reflections).

[Positive Association] Increasing events (in particular crossing events) are positively correlated.

In this talk, we discuss a new result establishing that this two properties are sufficient to obtain a RSW property. This implies most of the previously known RSW results.

Theorem 2. [KST22] *Let* $\mathbf{P}$ *be a planar percolation measure satisfying [Symmetry] and [Positive Association]. Then*

$$\left(\mathbf{P}\left[\boxed{\overset{n}{}}\, n\right] \geqslant c\right) \Rightarrow \left(\mathbf{P}\left[\boxed{\overset{2n}{}}\, n\right] \geqslant c'\right),$$

where $c' = f(c)$ *independent of* n.

References

[BD12] V. Beffara and H. Duminil-Copin. The self-dual point of the two-dimensional random-cluster model is critical for $q \geqslant 1$. *Probab. Theory Related Fields*, 153(3-4):511–542, 2012.

[BR06a] B. Bollobás and O. Riordan. The critical probability for random Voronoi percolation in the plane is 1/2. *Probab. Theory Related Fields*, 136(3):417–468, 2006.

[KST22] Laurin Köhler-Schindler and Vincent Tassion. Crossing probabilities for planar percolation, 2022+. To appear in *Duke Math. J.*

[Rus81] L. Russo. On the critical percolation probabilities. *Z. Wahrscheinlichkeit.*, 56(2):229–237, 1981.

[SW78] P. D. Seymour and D. J. A. Welsh. Percolation probabilities on the square lattice. In *Annals of Discrete Mathematics*, volume 3, pages 227–245. Elsevier, 1978.

[Tas16] V. Tassion. Crossing probabilities for Voronoi percolation. *Ann. Probab.*, 44(5):3385–3398, 2016.

[DST17] H. Duminil-Copin, V. Sidoravicius, and V. Tassion. Continuity of the phase transition for planar random-cluster and Potts models with $1 \leqslant q \leqslant 4$. *Comm. Math. Phys.*, 349(1):47–107, 2017.

Enumerating k-SAT functions

Yufei Zhao

(joint work with Dingding Dong and Nitya Mani)

How many k-SAT functions on n boolean variables are there? What does a typical such function look like?

Specifically, how many functions $f\colon \{0,1\}^n \to \{0,1\}$ can be written as

$$f(x_1, \ldots, x_n) = C_1 \vee C_2 \vee \cdots \vee C_m$$

where each clause C_i has the form $z_1 \wedge \cdots \wedge z_k$ with $z_1, \ldots, z_k \in \{x_1, \overline{x_1}, \ldots, x_n, \overline{x_n}\}$? (This is the DNF version; it is equivalent to the CNF version of the problem.)

A function is *unate* if it is monotone after first negating some subset of variables. Equivalently, a unate function is one with a formula where each variable x_i appears only positively (as x_i) or negatively (as $\overline{x_i}$) but not both. For fixed $k \geqslant 2$, an easy argument shows that the number of unate k-SAT functions is $(1 + o(1))2^{n+\binom{n}{k}}$.

Conjecture 1 (Bollobás, Brightwell, and Leader [3])**.** *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 *variables are unate.*

Bollobás, Brightwell, and Leader 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].

Bollobás and Brightwell [2] conjectured that if $k = k(n)$ is allowed to increase with n, as long as $k \leqslant (1/2 - c)n$ for some constant $c > 0$, the number of k-SAT functions on n-variables is $2^{(1+o(1))\binom{n}{k}}$. The situation is different for $k \geqslant n/2$.

In our work, we reduce Conjecture 1, for each fixed k, to a Turán problem on partially directed k-uniform hypergraphs. We also prove the $k = 4$ case of the conjecture via extremal graph theoretic arguments. A brute-force computation by Nitya Mani and Edward Yu also verifies the $k = 5$ case of the conjecture.

A *partially directed k-graph* (k-*PDG* for short) is obtained by starting with a k-uniform hypergraph and "orienting" some of the edges. To *orient* an edge means to pick a vertex in the edge. Below we write

$$\alpha = \frac{\#\text{undirected edges}}{\binom{n}{k}} \quad \text{and} \quad \beta = \frac{\#\text{directed edges}}{\binom{n}{k}}.$$

A *subgraph* of a k-PDG is obtained by allowing deletion of vertices, edges, and orientations.

As an example of our main theorem, we show that the following statement implies Conjecture 1 for $k = 2$.

Theorem 2. *For all sufficiently large n, every n-vertex 2-PDG without the following as a subgraph*

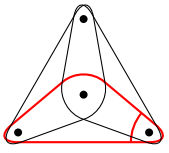
$\vec{T}_2 =$  , *edges* $= \{12, 1\check{3}, 23\}$

has

$$\alpha + (\log_2 3)\beta \leqslant 1.$$

Likewise, the following statement implies Conjecture 1 for $k = 3$.

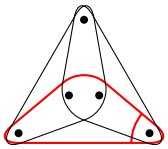
Theorem 3. *For all sufficiently large n, every n-vertex 3-PDG without the following subgraph*

$\vec{T}_3 =$  *edges* $= \{123, 12\check{4}, 134\}$

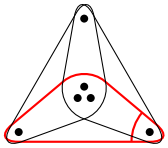
has

$$\alpha + (\log_2 3)\beta \leqslant 1.$$

More generally, we define $\vec{T}_k$ as the k-PDG obtained by starting with $\vec{T}_2$ and then adding $k-2$ common vertices to all three edges, e.g.,

$\vec{T}_4 =$  edges $= \{1234, 123\breve{5}, 1245\}$,

and

$\vec{T}_5 =$  edges $= \{12345, 1234\breve{6}, 12356\}$.

Conjecture 4. *Fix $k \geqslant 2$. For all sufficiently large n (depending on k), every n-vertex k-PDG without $\vec{T}_k$ as a subgraph satisfies*

$$\alpha + (\log_2 3)\beta \leqslant 1.$$

One of the our main results is the following.

Theorem 5. *For each fixed $k \geqslant 2$, Conjecture 4 implies Conjecture 1.*

Furthermore, we prove Conjecture 4 for $k = 4$; the proof applies a result of Füredi and Melaki [5] on the minimum triangular edge density in a graph of fixed edge density. Also, in an appendix by Nitya Mani and Edward Yu (see arXiv-v2 of [4]), a brute-force computation verifies Conjecture 4 for $k = 5$.

We also show that by enlarging the set of forbidden subgraphs from $\vec{T}_k$ to some special finite set $\mathcal{F}_k$, the extremal claim $\alpha + (\log_2 3)\beta \leqslant 1$ becomes essentially equivalent to Conjecture 1. For the precise statement as well as the construction of this forbidden set $\mathcal{F}_k$, see our full paper [4].

References

[1] Peter Allen, *Almost every 2-SAT function is unate*, Israel J. Math. **161** (2007), 311–346.

[2] Béla Bollobás and Graham R. Brightwell, *The number of k-SAT functions*, Random Structures Algorithms **22** (2003), 227–247.

[3] Béla Bollobás, Graham R. Brightwell, and Imre Leader, *The number of 2-SAT functions*, Israel J. Math. **133** (2003), 45–60.

[4] Dingding Dong, Nitya Mani, and Yufei Zhao, *Enumerating k-SAT functions*, ACM-SIAM Symposium on Discrete Algorithms (SODA) 2022, arXiv:2107.09233.

[5] Zoltán Füredi and Zeinab Maleki, *The minimum number of triangular edges and a symmetrization method for multiple graphs*, Combin. Probab. Comput. **26** (2017), 525–535.

[6] L. Ilinca and J. Kahn, *The number of 3-SAT functions*, Israel J. Math. **192** (2012), 869–919.

The phase transition in the random *d*-process

Lutz Warnke

One of the most interesting features of Erdős–Rényi random graphs is the 'percolation phase transition', where the global structure intuitively changes from only small components to a single giant component plus small ones.

In this talk we discuss the percolation phase transition in the random d-process, which is a time-evolving random graph model with bounded degrees: starting with an empty graph on n vertices, new random edges are added step-by-step so that the maximum degree remains at most d. For fixed $d \geqslant 3$, we (i) show that the d-process undergoes a giant component phase transition, and (ii) determine the asymptotic size of the giant component just after the phase transition. For $d = 2$, we also show that the giant component has a non-trivial distribution at the end of the 2-process. These results verify a conjecture of Balinska and Quintas from 1990, and solve a problem of Wormald from 1997.

The proofs deal with the subtle edge–dependencies in the d-process by a careful mix of different techniques. In particular, the core argument is based on tracking a large system of $O(d^4)$ many random variables via the differential equation method: these variables are used as input to suitable branching process approximation arguments, which in turn require an asymptotic analysis of the associated unusually large system of $O(d^4)$ many differential equation.

Based on two sets of joint works in progress: one with Nick Wormald, and one with Laura Eslava.

Problem Session

Nathan Linial (chair)

Alexey Pokrovskiy

Definition 1. *A closed walk W in a graph G has a* stacking *if one can draw W in $G \times \mathbb{R}$ without a crossing.*

Example 1. *If G has vertices label by $[5]$ and edges $\{(1,2),(2,3),(1,3),(3,4),(4,5),(5,3)\}$. Then $W = 1234531$ has a stacking while $W = 1231231$ has no stacking.*

The following is a result of Louder and Witten.

Theorem 2. *A closed walk W has a stacking in a graph G if and only if $W \neq kW'$ for any walk W' and $k \geqslant 2$.*

Problem 3. *Find an elementary proof of Theorem 2.*

Proposed by Ashwin Sah

Problem 4. *Let*

$$t_p(F,G) = p^{-e(F)}|G|^{-|F|}\hom(F,G),$$

the p-normalized homomorphism count of F within graph G. Is it possible to construct a sequence of graphs G_n with $|G_n| \to \infty$' with edge densities $p_n = |G_n|^{-o(1)}$ such that:

(1) $a_F = \lim_{n\to\infty} t_{p_n}(F, G_n)$ *exists for every graph F*
(2) $a_{C_4} = 1$
(3) $a_{K_3} = 1$
(4) $\sup_F a_F^{1/e_F} < \infty$
(5) $a_F \neq 1$ *for some F*

Noga Alon

Problem 5. *Given graphs G_1 and G_2 on the vertex set $[n]$ define $G_1 \oplus G_2 = ([n], E_1 \oplus E_2)$. Let $\mathcal{F}$ be a family of graphs on the vertex set $[n]$ such that for any distinct pair of graph F_1 and F_2 we have that $F_1 \oplus F_2$ is not a clique. Prove (or disprove) that*

$$|\mathcal{F}| = o\left(2^{\binom{n}{2}}\right).$$

Remark. This can be thought of as the first open case of polynomial density Hales–Jewett.

Mehtaab Sawhney

Problem 6. *Consider the random triangle removal process in K_n run for $n^2/6 - n^{1.99}$ steps, and let the output set be $\mathbf{T}_n$. Does there exist a constant C and a sequence of events $\mathcal{E}_n$ which occur with probability $1 - o(1)$ such that for any k distinct triangles $T_1, \dots, T_k$ of K_n we have*

$$\mathbb{P}[T_1, \dots, T_k \in \mathbf{T}_n | \mathcal{E}_n] \leqslant (C/n)^k?$$

That is, will the resulting random set of triangles be $O(1/n)$-spread?

Remark. We allow for situations where k grows with n, even as much as $k = n^2/12$, say. If we only require the estimate for $k \leqslant n^{1/10}$ the statement is relatively straightforward.

David Conlon

Problem 7. *Does there exist a constant $c = c_0$ such that:*

- *Fix $\varepsilon > 0$. If n is sufficiently large, any (n, d, λ)-graph G with $\lambda < (c_0 - \varepsilon)d^2/n$ contains a triangle.*
- *There exist arbitrarily large (n, d, λ)-graphs G with $\lambda < (c_0 + \varepsilon)d^2/n$ which have no triangles.*

Furthermore, given the existence of such a constant c_0, prove that for $\varepsilon > 0$ and n sufficiently large, any (n, d, λ)-graph G with $\lambda < (c_0 - \varepsilon)d^2/n$ contains many triangles or a large triangle factor (scaled appropriately).

Remark. The first and second bullet points are known with differing constants due to Alon's construction of quasirandom triangle-free graphs [1], and the interest is in a "sharp threshold".

References

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

Lior Gishboliner

Problem 8. *Given prime p and arbitrary $a \in \mathbb{F}_p \backslash \{0\}$, does there exist a set S such that S avoids nontrivial solutions to $x + ay = (a+1)z$ with $|S| \geqslant p^{1-o(1)}$?*

Remark. For fixed a with respect to p (e.g. $a = 2$) the result follows from a modification of the Behrend construction, but this gets worse with the height of a and a uniform bound is desired.

Matthew Kwan

This problem is no longer open; it was solved by Matthew Kwan and Lisa Sauermann during the workshop.

Jacob Fox, Matthew Kwan, and Hunter Spink [1] posed the following conjecture.

Problem 9. *Fix a constant d and let S be a sphere in $\mathbb{R}^d$. Fix nonzero vectors $a_1, \dots, a_n \in \mathbb{R}^d \backslash \{0\}$, let $(\xi_1, \dots, \xi_n) \in \{-1, 1\}^n$ be a uniformly random ± 1 vector, and let $X = \xi_1 a_1 + \cdots + \xi_n a_n$. Then $\Pr(X \in S) = O(1/\sqrt{n})$.*

More generally, for which interesting sets S do we have a bound like $\Pr(X \in S) = O(1/\sqrt{n})$? Note that this is the best we can hope for: even if S is a single point we can have $\Pr(X \in S) = \Theta(1/\sqrt{n})$.

References

[1] Jacob Fox, Matthew Kwan, and Hunter Spink, *Geometric and o-minimal Littlewood-Offord problems.*

Dhruv Mubayi

Problem 10. *Given a (large) graph G with edge density ξ and $n \to \infty$, determine the asymptotic maximum for the number of induced C_4's in G.*

We have the following conjecture of Liu, Mubayi, and Reiher [1].

Conjecture 11. *For $\xi > 1/2$, the maximum number of induced C_4's is achieved asymptotically by the construction for the triangle density problem.*

Theorem 12 ([1, Theorem 1.6])**.** *We have that Theorem 11 is true for $\xi = 1 - 1/k$ for integers k.*

References

[1] Xizhi Liu, Dhruv Mubayi, and Christian Reiher, *The feasible region of induced graphs.*

Yufei Zhao

The following problem stems from work of Conlon, Fox, Sudakov, and Zhao [1].

Problem 13. *Consider a graph G with $p = e(G)/\binom{n}{2}$, G being C_4-free, $p \gtrsim n^{-1/2}$, and $|e_G(A,B) - p|A||B|| = o(pn^2)$ for all $A, B \subseteq [n]$. Is it true that the number of copies of the Peterson graph in G is at least (up to a constant factor) the number of copies of the Peterson graph in $G(n,p)$?*

Remark. A similar question has also been posed by Jacques Verstraëte.

References

[1] David Conlon, Jacob Fox, Benny Sudakov, and Yufei Zhao, *The regularity method for graphs with few 4-cycles*, J. Lond. Math. Soc. (2) **104** (2021), 2376–2401.

Jozsef Balogh

Let G be a graph. A *standard clique edge-weighting* is an assignment w of weights to the edges as follows. Let e be an edge and r be the order of the largest clique containing e in G. Then

$$w(e) := \frac{r}{2(r-1)}.$$

We extend the definition of weight w to the entire graph as the sum over all edges

$$w(G) := \sum_{e \in E(G)} w(e).$$

Let $T(n,r)$ be the r-partite Turán graph on n vertices. Notice that the weights are set up such that for all $r \geqslant 2$,

$$\lim_{n\to\infty} \frac{w(T(n,r))}{n^2} = \frac{1}{4}.$$

Conjecture 14. *For every n-vertex graph G and its standard clique edge-weighting w,*

$$w(G) \leqslant \Big(\frac{1}{4} + o(1)\Big)n^2.$$

Moreover, for every $\varepsilon > 0$ there exists n_0 such that if $w(G) = (1/4 + o(1))n^2$ and $n \geqslant n_0$, then G is within edit distance at most εn^2 from some Turán graph.

The conjecture has been proved for K_5-free graphs by Balogh and Lidicky. Shortly after the workshop Domagoj Bradač proved the above conjecture in general.

Marcelo Campos

The following problem is due to Alon and Granville.

Problem 15. *Can one find an efficient algorithm such that given $B \subseteq \mathbb{Z}/n\mathbb{Z}$ determine whether or not $B = A + A$ for $A \subseteq \mathbb{Z}/n\mathbb{Z}$?*

Natasha Morrison

Fix a pair of equations

$$a_1x_1 + \ldots + a_kx_k = 0$$
$$b_1x_1 + \ldots + b_kx_k = 0.$$

We say this pair of equations is common with respect to 2-coloring if for all n and $f : \mathbb{F}_q^n \to [0,1]$ that

$$\mathbb{E}_{\substack{a_1x_1+\ldots+a_kx_k=0 \\ b_1x_1+\ldots+b_kx_k=0}}\left(\prod_i f(x_i) + \prod_i (1-f(x_i))\right) \geqslant 2^{1-k}.$$

Conjecture 16. *For all $k \geqslant 6$, k even, such that the dimension of the solution space of $a_1x_1 + \ldots + a_kx_k = 0, b_1x_1 + \ldots + b_kx_k = 0$ is $k-2$, and q sufficiently large, we have that the pair of equations is* uncommon.

Lutz Warnke

In a landmark paper from 1987, Shamir and Spencer [5] proved that the chromatic number $\chi(G_{n,p})$ of an n-vertex binomial random graph $G_{n,p}$ is typically contained in an interval of length at most $\omega\sqrt{n}$, where $\omega = \omega(n)$ is an arbitrary function with $\omega \to \infty$ as $n \to \infty$, as usual. For constant edge-probabilities $p \in (0,1)$, Alon noticed in the 1990s that this concentration interval length can be slightly improved to $\omega\sqrt{n}/\log n$, see [2, Excercise 7.9.3] and Scott's note [4]. In the sparse case where $p = p(n) \to 0$ vanishes as $n \to \infty$., i.e., for edge-probabilities of form $p = n^{-\alpha}$ with $\alpha \in (0, 1/2)$, Shamir and Spencer proved in their 1987 paper that $\chi(G_{n,p})$ is typically contained in an interval of length at most $\omega\sqrt{np}\log n$, and a modern inspection of their proof reveals that length $\omega\sqrt{np}$ suffices. Very recently Surya and Warnke [6] extended Alon's logarithmic improvement to the sparse case (by avoiding large deviation inequalities such as Janson's inequality, instead relying on more robust Chernoff bound based arguments), obtaining the following result.

Theorem 17 ([6])**.** *Let $\omega = \omega(n) \to \infty$ as $n \to \infty$ be an arbitrary function, and let $\delta \in (0,1)$ be a constant. If the edge-probability $p = p(n)$ satisfies $n^{-1/2+\delta} \ll p \leqslant 1-\delta$, then there exists a function $\Lambda_{n,p}$ such that the chromatic number $\chi(G_{n,p})$ of the random graph $G_{n,p}$ satisfies*

$$\lim_{n\to\infty} \mathbb{P}\left(|\chi(G_{n,p}) - \Lambda_{n,p}| \leqslant \frac{\omega\sqrt{np}}{\log n}\right) = 1.$$

Given the importance of the chromatic number $\chi(G_{n,p})$ for the development of random graph theory, it is of great interest to understand how sharp this concentration boundi is, i.e., if it is best possible up to $(\log n)^{O(1)}$ or $n^{o(1)}$ factors, say.

Problem 18. *Is the upper bound of Theorem 17 on the concentration of the chromatic number $\chi(G_{n,p})$ of the random graph $G_{n,p}$ close to best possible or not?*

There is some evidence in both directions. On the one hand, it might be close to best possible, since (a) it is the best bound one can hope for using the martingale based Shamir-Spencer proof framework [5], (b) the form of the $\omega\sqrt{np}/\log n$ bound

is consistent with the fact that, for $n^{-1} \ll p \ll n^{-1/2-\delta}$, the chromatic number is concentrated on two different values [1] due to work of Alon and Krivelevich [1], and (c) it is sharp (up to logarithmic factors) for uniform edge-probability $p = 1/2$ due to the work of Heckel and Riordan [3]. On the other hand, the bound of Theorem 17 might be far from best possible since (d) if one transfers the coloring heuristic of the Heckel and Riordan [3] to the sparse case, then with some hand-waving one arrives at the conclusion that the concentration should perhaps rather roughly be of form $\sqrt{np^3} \cdot (\log n)^{O(1)}$, say. In any case, we believe that a solution to Problem 18 will most likely advance the field by new ideas or proof techniques.

References

[1] N. Alon and M. Krivelevich. The concentration of the chromatic number of random graphs. *Combinatorica* **17** (1997), 303–313.

[2] N. Alon and J. Spencer. *The Probabilistic Method*, 2nd ed. Wiley-Interscience, New York (2000).

[3] A. Heckel and O. Riordan. How does the chromatic number of a random graph vary?. Preprint (2021). `arXiv:2103.14014`

[4] A. Scott. On the concentration of the chromatic number of random graphs. Explanatory note (2008). `arXiv:0806.0178`

[5] E. Shamir and J. Spencer. Sharp concentration of the chromatic number on random graphs $G_{n,p}$. *Combinatorica* **7** (1987), 121–129.

[6] E. Surya and L. Warnke On the concentration of the chromatic number of random graphs. Preprint (2022). `arXiv:2201.00906`

Andrzej Ruciński

Fix an alphabet A with $|A| = k \geqslant 2$ and consider words $w \in A^n$.

Definition 19. *Define a pair of subwords w_1, w_2 of w to be twins if they are identical and disjoint.*

Example 2. *The overline and underlined words in $\underline{1}\overline{10}12\underline{0}\overline{5}\underline{\overline{5}}$ are twins.*

Let $t(w)$ be the length of the longest pair of twins in w and

$$t_{n,k} = \min_{\substack{w \in A^n \\ |A|=k}} t(w).$$

Work of Axenovich, Person, and Puzynina [1] proved that $t(n,2) = n/2 - o(n)$ which immediately implies that $t(n,3) \geqslant n/3 - o(n)$ (by looking at the pair of most common letters). This was improved by Bukh and Zhou [2] to prove that $t(n,3) \geqslant .34n$ and they also proved that $t(n,4) \leqslant 0.493n$. Finally Dudek, Grytczuk, and Ruciński improved the bound on $t(n,3)$ by proving that asymptotically almost surely (a.a.s.) that $t(W_3(n)) \geqslant 0.411n$ where $W_k(n)$ is a uniformly random work of length n given an alphabet of size k.

Question 20. *Is it asymptotically almost surely true that $t(W_3(n)) = n/2 - o(n)$?*

Question 21. *Is $t(n,3) = n/2 - o(n)$?*

Question 22. *For every k, is it asymptotically almost surely true that $t(W_k(n)) = t(k,n) + o(n)$?*

References

[1] Maria Axenovich, Yury Person, and Svetlana Puzynina, *A regularity lemma and twins in words*, J. Combin. Theory Ser. A **120** (2013), 733–743.
[2] Boris Bukh and Lidong Zhou, *Twins in words and long common subsequences in permutations*, Israel J. Math. **213** (2016), 183–209.

Peter Allen

Conjecture 23. *Given an n-vertex graph G which is Δ-regular with $\Delta \geqslant 3$ and G is Ramanujan.*

Let $N = 10^{10^{10^{\Delta}}} n$ and let p be the threshold probability p for which with probability $1/2$ any two coloring of $G(N,p)$ must contain a copy of G.

Then the size Ramsey number of G is $\Theta(pn^2)$.

Remark. Informally, this question is asking whether for large sparse graphs which are expanders, the size Ramsey number is controlled by the construction coming from random graphs.

Reporter: Matija Bucić

Participants

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
08544 Princeton NJ 08544
UNITED STATES

Dr. Ryan Alweiss
Department of Mathematics
Princeton University
Fine Hall
Washington Road
Princeton, NJ 08544-1000
UNITED STATES

Prof. Dr. Paul Balister
Mathematical Institute
Oxford University
Woodstock Road
Oxford OX1 3LB
UNITED KINGDOM

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

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
08540 Princeton
UNITED STATES

Marcelo Campos
Instituto de Matematica Pura e
Aplicada - IMPA
Jardim Botanico
Estrada Dona Castorina, 110
22460 Rio de Janeiro, RJ 320
BRAZIL

Prof. Dr. Amin Coja-Oghlan
Fakultät für Informatik
Technische Universität Dortmund
Otto-Hahn-Straße 12
44227 Dortmund
GERMANY

Prof. Dr. David Conlon
Department of Mathematics
California Institute of Technology
1200 E California Boulevard
Pasadena CA 91125
UNITED STATES

Prof. Dr. David Gamarnik
Massachusetts Institute of Technology
Sloan School of Management
E 62 - 563
100 Main Street
Cambridge MA 02139
UNITED STATES

Lior Gishboliner
Department of Mathematics
ETH Zurich
Zurich, Switzerland
8092 Zürich
SWITZERLAND

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

Prof. Dr. Svante Janson
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. Jeff Kahn
Department of Mathematics
Rutgers University
Hill Center, Busch Campus
110 Frelinghuysen Road
Piscataway NJ 08854-8019
UNITED STATES

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

Prof. Dr. Michael Krivelevich
School of Mathematical Sciences
Sackler Faculty of Exact Sciences
Tel Aviv University
Ramat Aviv, Tel Aviv 6997801
ISRAEL

Dr. Gal Kronenberg
Mathematical Institute
Oxford University
Andrew Wiles Building
Woodstock Road
Oxford OX2 6GG
UNITED KINGDOM

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. Anita Liebenau
School of Mathematics and Statistics
UNSW Sydney
Rm. 6105
Sydney NSW 2052
AUSTRALIA

Prof. Dr. Nathan Linial
School of Computer Science and Engineering
The Hebrew University of Jerusalem
Givat Ram
Jerusalem 9190401
ISRAEL

Dr. Eoin Patrick Long
School of Mathematics and Statistics
The University of Birmingham
Edgbaston
Birmingham B15 2TT
UNITED KINGDOM

Prof. Dr. Eyal Lubetzky
Courant Institute of Mathematical Sciences
New York University
251 Mercer Street
New York NY 10012
UNITED STATES

Prof. Dr. Tomasz Łuczak
Faculty of Mathematics and Computer Science
Adam Mickiewicz University
ul. Uniwersytetu Poznańskiego 4
61-614 Poznań
POLAND

Leticia Mattos
Institut für Mathematik
Freie Universität Berlin
Arnimallee 3
14195 Berlin
GERMANY

Dr. Richard Montgomery
Mathematics Institute
University of Warwick
Gibbet Hill Road
Coventry CV4 7AL
UNITED KINGDOM

Dr. Patrick Morris
Universitat Politecnica de Catalunya (UPC)
Correr de Jordi Girona 29
08034 Barcelona
SPAIN

Prof. Dr. Robert Morris
IMPA
Instituto Matematica Pura e Aplicada
Estrada Dona Castorina, 110
Jardim Botanico
Rio de Janeiro 22460-320
BRAZIL

Dr. Natasha Morrison
Dept. of Mathematics & Statistics
University of Victoria
P.O. Box PO BOX 1700 STN
Victoria BC V8W 2Y2
CANADA

Prof. Dr. Dhruv Mubayi
Dept. of Mathematics, Statistics and Computer Science, M/C 249
University of Illinois at Chicago
851 S. Morgan Street
Chicago, IL 60607-7045
UNITED STATES

Dr. Bhargav Narayanan
Department of Mathematics
Rutgers University
Busch Campus, Hill Center
New Brunswick NJ 08854-8019
UNITED STATES

Prof. Dr. Konstantinos Panagiotou
Mathematisches Institut
Universität München
Theresienstraße 39
80333 München
GERMANY

Dr. Jinyoung Park
Department of Mathematics
Stanford University
450 Jane Stanford Way
Stanford, CA 94305
UNITED STATES

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

Huy Tuan Pham
Department of Mathematics
Stanford University
Stanford, CA 94305-2125
UNITED STATES

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. Andrzej Rucinski
Faculty of Mathematics & Computer
Science
Adama Mickiewicz University
ul. Uniwersytetu POznanskiego 4
61-614 Poznań
POLAND

Ashwin Sah
Department of Mathematics
Massachusetts Institute of
Technology
77 Massachusetts Avenue
Cambridge, MA 02139-4307
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

Dr. Deniz Sarikaya
Vrije Universiteit Brussel
Centre for Logic and Philosophy of
Science
Pleinlaan 2
1050 Brussels
BELGIUM

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

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

Dr. Katherine Staden
Faculty of Mathematics & Computing
The Open University
Walton Hall
Milton Keynes MK7 6AA
UNITED KINGDOM

Prof. Dr. Angelika Steger
Institut für Theoretische Informatik
ETH Zürich, CAB G 37.2
Universitätstrasse 6
8092 Zürich
SWITZERLAND

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

Prof. Dr. Vincent Tassion
Departement Mathematik
ETH-Zentrum
Rämistrasse 101
8092 Zürich
SWITZERLAND

Dr. Lutz P. Warnke
Department of Mathematics
University of California, San Diego
9500 Gilman Drive
La Jolla CA 92093-0112
UNITED STATES

Prof. Dr. Nicholas Wormald
School of Mathematics
Monash University
Room 455
9 Rainforest Walk
Clayton, VIC 3800
AUSTRALIA

Dr. Liana Yepremyan
Department of Mathematics
Emory University
201 Dewman Drive
Atlanta GA 30322
UNITED STATES

Prof. Dr. Yufei Zhao
Massachusetts Institute of Technology
77 Massachusetts Avenue
P.O. Box Room 2-271
Cambridge MA 02139
UNITED STATES