Workshop on Probabilistic and Extremal Combinatorics
(19w5009)

Penny Haxell (University of Waterloo)
Michael Krivelevich (Tel Aviv University)
Benny Sudakov (ETH Zurich)

September 1-6, 2019

1 Overview of the Field

Combinatorics, also known as Discrete Mathematics, is a fundamental mathematical discipline which studies
the nature of discrete mathematical objects, such as permutations, graphs and set-families. Encompassing
some of the most basic and natural mathematical questions, Combinatorics is probably as old as the human
ability to count. During the twentieth century and the first two decades of the present century, the field has
experienced remarkable growth, transitioning from a collection of mostly isolated results relying on ad hoc
arguments, to a major branch of mathematics with its own theory, problems and techniques. Modern Combi-
natorics has well-established and fruitful connections to numerous scientific fields, including Number Theory,
Discrete Geometry, Statistical Physics, Evolutionary Biology and, most notably, Theoretical Computer Sci-
ence, with which it is intimatively intertwined. Combinatorial methods and ideas have been exported to many
of the above disciplines, in some cases leading to outstanding results.

The workshop focused on two major branches of Combinatorics: Extremal Combinatorics and Proba-
bilistic Combinatorics. These two highly-related fields have flourished in recent decades, producing several
breakthroughs and many fascinating results, and leading to the creation of a vibrant community comprising
both senior and young researchers. We now briefly overview each of these two disciplines.

Extremal Combinatorics encompasses several important areas of research, such as Extremal Graph (and
Hypergraph) Theory, Extremal Set Theory, and Ramsey Theory. Extremal Graph Theory typically asks to
maximize/minimize some graph parameter over all graphs satisfying a given property. This field originated in
the classical theorems of Mantel (1907) and Turán (1941), which determined the maximum possible number
of edges of an $n$-vertex graph containing no clique of a given size. Since then, questions of this type have
been studied for other natural combinatorial objects, such as uniform hypergraphs, tournaments and matrices.
As a testament to the field’s importance, it suffices to say that the study of extremal-(hyper)graph-theoretic
questions has lead to major developments in Combinatorics, like the Regularity Method.

Ramsey Theory is the branch of Combinatorics dealing with the fundamental phenomenon, perhaps best
described by Motzkin’s famous quote, that “complete disorder is impossible”. In other words, every large
enough system, chaotic as it may be, must contain relatively large structured subsystems. Concrete mani-
festations of this idea have been studied in many settings, starting with F. Ramsey’s original theorem which
states that no matter how one colors the edges of a large enough complete uniform hypergraph, one is always
guaranteed to find a large monochromatic complete subhypergraph (which constitutes the structured subsys-
tem). Due to the fundamental nature of such questions, they have found uses in many combinatorial studies,
as well as in other branches of mathematics.
Probabilistic Combinatorics studies probability spaces of discrete structures. This field originated in the 1940s with the revolutionary realization that probabilistic reasoning can be applied to deterministic problems. A key example of this is a so-called existence proof, which goes as follows. In order to prove the existence of a combinatorial structure of a given type having some desirable properties, one can define a suitable probability distribution over objects of the given type, and show that the probability that a randomly drawn object possesses the prescribed properties is strictly larger than zero. Such a proof guarantees the existence of a certain object without explicitly describing it. This Probabilistic Method, as it came to be known, has been successfully used to find examples and counterexamples to multiple longstanding conjectures, and has since become a standard tool in the combinatorialist’s toolkit. This success has in turn led to the study of the typical properties of various combinatorial objects, such as random graphs and matrices. Probabilistic Combinatorics has also had a profound impact on Computer Science, serving as the mathematical foundation for the design of randomized algorithms and the study of randomness in computation, and inspiring such fundamental notions as quasi-randomness and graph expansion. Finally, the combination of probabilistic techniques with other sophisticated tools has recently led to striking results, like Keevash’s celebrated resolution of the 150-year old problem on the existence of designs.

2 Ramsey theory

2.1 Jacob Fox Hypergraph Ramsey numbers

In this talk, we solve a longstanding open problem of Erdős and Hajnal on off-diagonal hypergraph Ramsey numbers, and discuss a variety of related problems. A $k$-uniform hypergraph, or simply $k$-graph, $G = (V, E)$ consists of a vertex set $V$ and an edge set $E \subseteq \binom{V}{k}$. Ramsey’s theorem states that for any $k$-graphs $H_1$ and $H_2$, there is a positive integer $N$ such that any $k$-graph $G$ of order $N$ contains $H_1$ (as a subgraph) or its complement $\overline{G}$ contains $H_2$. The Ramsey number $r(H_1, H_2)$ is the smallest such $N$, and the main goal of graph Ramsey theory is to estimate $r(H_1, H_2)$, especially when one or both of $H_i$ is a complete graph $K_n$.

In 1972, Erdős and Hajnal [19] posed the problem of determining the minimum independence number of a 3-graph on $N$ vertices in which there are at most two edges among any four vertices, and showed that the answer is between $\Omega(\log N/\log \log N)$ and $O(\log N)$. In the language of Ramsey numbers, this is equivalent to

$$2^{\Theta(n)} \leq r(K_4^{(3)} \setminus \text{e}, K_n^{(3)}) \leq n^{O(n)},$$

where $K_4^{(3)} \setminus \text{e}$ is the 3-graph on four vertices and three edges. This problem has received considerable attention during the half-century since it was posed, see, e.g., [17]. We solve this Erdős-Hajnal problem and show that the upper bound in (1) is tight. That is, $r(K_4^{(3)} \setminus \text{e}, K_n^{(3)}) = n^{\Theta(n)}$. We prove this result with a carefully designed probabilistic construction, and entropy inequalities are used in the analysis. We further extend this result to prove similarly tight bounds for off-diagonal 3-uniform Ramsey numbers of link hypergraphs versus complete hypergraphs. For details, see [21].

[Based on joint work with Xiaoyu He.]

2.2 Dhruv Mubayi Polynomial to exponential transition in Ramsey theory

We give two new results in Ramsey theory, one for graphs and one for hypergraphs.

Graphs. For fixed $s \geq 3$, we prove that if optimal $K_s$-free pseudorandom graphs exist, then the Ramsey number $r(s, t) = t^{s-1+o(1)}$ as $t \to \infty$. Our method also improves the best lower bounds for $r(C_t, t)$ obtained by Bohman and Keevash [9] from the random $C_t$-free process by polylogarithmic factors for all odd $\ell \geq 5$ and $\ell \in \{6, 10\}$. For $\ell = 4$ it matches their lower bound from the $C_4$-free process. These results all follow from the following theorem.

Theorem 1 ([34]). Let $F$ be a graph, $n, d, \lambda$ be positive integers with $d \geq 1$ and $\lambda > 1/2$ and let $t = \lceil 2n \log^2 n/d \rceil$. If there exists an $F$-free $(n, d, \lambda)$-graph, then

$$r(F, t) > \frac{n}{20\lambda} \log^2 n.$$

(2)
We also prove, via a different approach, that \( r(C_5, t) > (1 + o(1))t^{11/8} \) and \( r(C_7, t) > (1 + o(1))t^{11/9} \). These improve the exponent of \( t \) in the previous best results and appear to be the first examples of graphs \( F \) with cycles for which such an improvement of the exponent for \( r(F, t) \) is shown over the bounds given by the random \( F \)-free process and random graphs. This is a joint work with Jacques Verstraëet.

**Hypergraphs.** Given \( s \geq k \geq 3 \), let \( h^{(k)}(s) \) be the minimum \( t \) such that there exist arbitrarily large \( k \)-uniform hypergraphs \( H \) whose independence number is at most polylogarithmic in the number of vertices and in which every \( s \) vertices span at most \( t \) edges. Erdős and Hajnal [19] conjectured (1972) that \( h^{(k)}(s) \) can be calculated precisely using a recursive formula and Erdős offered $500 for a proof of this. For \( k = 3 \) this has been settled for many values of \( s \) including powers of three but it was not known for any \( k \geq 4 \) and \( s \geq k + 2 \). We settle the conjecture for all \( s \geq k \geq 4 \) [33]. This is a joint work with Alexander Razborov.

### 3 Geometric combinatorics

#### 3.1 Hao Huang Covering cubes by hyperplanes

Note that the vertices of the \( n \)-dimensional cube \( \{0, 1\}^n \) can be covered by two affine hyperplanes \( x_1 = 1 \) and \( x_1 = 0 \). However if we leave one vertex uncovered, then suddenly at least \( n \) affine hyperplanes are needed. This was a classical result of Alon and Füredi [3], followed from the Combinatorial Nullstellensatz.

To see that \( n \) planes suffice, one can just take all \( x_i = 1 \) for \( i = 1, \cdots, n \).

In this talk, we consider the following natural generalization of the Alon–Füredi theorem: what is the minimum number of affine hyperplanes such that the vertices in \( \{0, 1\}^n \setminus \{\vec{0}\} \) are covered at least \( k \) times, and \( \vec{0} \) is uncovered? We conjecture that for fixed integer \( k \), and sufficiently large \( n \), the minimum number of affine hyperplanes needed is \( n + \binom{k}{2} \). We prove this conjecture for \( k \leq 3 \), using a punctured version of the Combinatorial Nullstellensatz.

We also develop an analogue of the Lubell–Yamamoto–Meshalkin inequality for subset sums, and solve the problem asymptotically for fixed \( n \) and \( k \to \infty \). We show that in this setting, the minimum number of affine hyperplanes needed is \( (1 + \frac{1}{2} + \cdots + \frac{1}{n} + o(1))k \), solving the fractional version of the cube covering problem.

Joint work with [16] with Alexander Clifton (Emory University)

#### 3.2 Peter Keevash Isoperimetric stability

A prominent open problem at the interface of Geometry, Analysis and Combinatorics is to understand the stability of isoperimetric problems. The meta-problem is to characterise sets whose boundary is close to the minimum possible given their volume; there are many specific problems obtained by giving this a precise meaning. For the discrete cube, the two natural notions are vertex boundary and edge boundary; the descriptions of the minimisers for both notions are classical results of Extremal Combinatorics. In joint work with Eoin Long, we have recently obtained corresponding stability results, namely that approximate minimisers for the edge boundary are close to disjoint unions of cubes, and approximate minimisers for the vertex boundary are close to generalised Hamming balls.

There is a large prior literature on the edge boundary, including earlier work of Ellis and of Ellis, Keller and Lifshitz that gives much more precise stability results for families that are very close to minimising the edge boundary. Also, Keller and Lifshitz independently used very different methods to prove a somewhat stronger version of my result with Long. In the setting of the vertex boundary, nothing was known prior to our work, no doubt due to the apparent intractability of extracting structural information from the only known proof method (compression, which alters the structure). Remarkably, we were able to obtain structural information via local stability arguments that keep tight control of structure over a long process of decompression.

Independently, Przykucki and Roberts recently gave a very different argument and obtained another stability result that complements ours (it is applicable to a different range of parameters).

A much wider context for the edge boundary is its interpretation (in the setting of biased measures on the cube) as the influence of Boolean functions and its connection (via the Margulis-Russo formula) to the sharp threshold phenomenon, which has led to fundamental insights and applications spanning a variety of areas of Mathematics (Gaussian Geometry and Isoperimetry), Computer Science (Computational Complexity) and Economics (Social Choice). One fundamental question addressed by a conjecture of Kahn and Kalai is to
characterise properties with coarse thresholds / functions of small influence. Results of Friedgut, Bourgain and Hatami say (roughly) that such properties / functions exhibit some kind of ‘junta-like’ behaviour, meaning that one can get a significant density increase by fixing the values of a small set of coordinates. However, all of these results only apply to the ‘dense setting’ where the initial density is bounded away from 0 and 1. In joint work with Lifshitz, Long and Minzer, we prove such results in the ‘sparse setting’ (i.e. any initial density) that establish a variant of the Kahn-Kalai Conjecture and a sharp form of Bourgain’s Theorem. Our main tool is a new hypercontractive inequality for quasirandom boolean functions. We also give applications of our sharp threshold result to Extremal Combinatorics (via the ‘junta method’), including proofs of two conjectures in a range of parameters that is within a constant factor of being optimal, namely the Huang-Loh-Sudakov Conjecture on cross matchings and the Furedi-Jiang-Seiver Conjecture on the Turán numbers of linear paths in hypergraphs.

Joint work with Noam Lifshitz, Eoin Long and Dor Minzer

3.3 Lisa Sauermann On the size of subsets of $F_p^n$ without $p$ distinct elements summing to zero

For given positive integers $m$ and $n$, what is the minimum integer $s$ such that among any $s$ points in the integer lattice $\mathbb{Z}^n$ there are $m$ points whose centroid is also a lattice point in $\mathbb{Z}^n$? This problem is called the Erdős-Ginzburg-Ziv problem, because its answer can be interpreted as the so-called Erdős-Ginzburg-Ziv constant $s(\mathbb{Z}_m^n)$ of $\mathbb{Z}^n_m$. In 1961, Erdős, Ginzburg and Ziv [18] proved a result essentially stating that $s(\mathbb{Z}_m^1) = 2m - 1$, and since then Erdős-Ginzburg-Ziv constants have been studied intensively.

Alon and Dubiner [2] showed that for a fixed dimension $n$, the quantity $s(\mathbb{Z}_m^n)$ grows linearly with $m$. On the other hand, they raised the question of finding good upper bounds for $s(\mathbb{Z}_m^n)$ when $m$ is fixed and $n$ is large. The most interesting case here is when $m = p \geq 3$ is a prime (since this case implies bounds for general $m$).

This problem of bounding $s(\mathbb{F}_p^n)$ for a fixed prime $p \geq 3$ and large $n$ is essentially equivalent to bounding the maximum size of a subset of $\mathbb{F}_p^n$ that does not contain $p$ distinct elements summing to zero. Asking about this maximum size is also a natural question in itself, and the special case of $p = 3$ is actually equivalent to the famous cap-set problem (which asks for the maximum size of a subset of $\mathbb{F}_3^n$ without a three-term arithmetic progression).

The main result discussed in this talk is the following new upper bound for the maximum size of a subset of $\mathbb{F}_p^n$ without $p$ distinct elements summing to zero for any fixed prime $p \geq 5$ and large $n$. For $p$ and $n$ going to infinity, this bound is of the form $p^{(1/2) + o(1)n}$, whereas all previously known upper bounds were of the form $p^{(1-o(1))n}$ (with $p^n$ being a trivial bound). Hence this new bound is a qualitative improvement over the previous bounds.

**Theorem 2.** Let $p \geq 5$ be a fixed prime. Then for any positive integer $n$ and any subset $A \subseteq \mathbb{F}_p^n$ which does not contain $p$ distinct elements $x_1, \ldots, x_p \in A$ with $x_1 + \cdots + x_p = 0$, we have

$$|A| < C_p \cdot (2\sqrt{p})^n.$$  

Here, $C_p$ is a constant only depending on $p$.

The proof of Theorem 2 uses the so-called multi-colored sum-free theorem which is a consequence of the Croot-Lev-Pach polynomial method. This method and its consequences were already applied by several authors to prove bounds for the maximum size of a subset of $\mathbb{F}_p^n$ without $p$ distinct elements summing to zero. However, using some new combinatorial ideas, we significantly improve their bounds.

3.4 Yufei Zhao Equiangular lines with a fixed angle

Solving a longstanding problem on equiangular lines, we determine, for each given fixed angle and in all sufficiently large dimensions, the maximum number of lines pairwise separated by the given angle.

**Theorem 3.** Fix $\alpha \in (0, 1)$. Let $k$ be the minimum number of vertices of a graph whose adjacency matrix has spectral radius exactly $(1 - \alpha)/(2\alpha)$, and set $k = \infty$ if no such graph exists.

The maximum number $N_\alpha(d)$ of equiangular lines in $\mathbb{R}^d$ with common angle $\arccos \alpha$ satisfies
(a) $N_{\alpha}(d) = \lfloor k(d-1)/(k-1) \rfloor$ for all sufficiently large $d > d_0(\alpha)$ if $k < \infty$.

(b) $N_{\alpha}(d) = d + o(d)$ as $d \to \infty$ if $k = \infty$.

Significant progress on made recently on the this problem in [7], where the following key step of the solution was introduced.

**Theorem 4.** For every $\alpha \in (0, 1)$, there exists some $\Delta = \Delta(\alpha)$ so that for every set of equiangular lines in $\mathbb{R}^d$ with common angle $\arccos \alpha$, one can choose a set $S$ of unit vectors, with one unit vector in the direction of each line in the equiangular set, so that each unit vector in $S$ has inner product $-\alpha$ with at most $\Delta$ other vectors in $S$.

Our key new insight is a new result in spectral graph theory, which shows that the multiplicity of the second largest eigenvalue of the adjacency matrix of a connected bounded degree graph is sublinear.

Joint work with Zilin Jiang, Jonathan Tidor, Yuan Yao, and Shengtong Zhang.

## 4 Extremal Set Theory

### 4.1 Shoham Letzter Minimum saturated families

We call a family $\mathcal{F}$ of subsets of $[n]$ $s$-saturated if it contains no $s$ pairwise disjoint sets, and moreover no subset of $[n]$ can be added to $\mathcal{F}$ while preserving this property.

More than 40 years ago, Erdős and Kleitman [20] conjectured that an $s$-saturated family of subsets of $[n]$ has size at least $(1 - 2^{-(s-1)} \cdot 2^n$. It is easy to show that every $s$-saturated family has size at least $\frac{1}{2} \cdot 2^n$, but, as was mentioned by Frankl and Tokushige [22], even obtaining a slightly better bound of $(1/2 + \varepsilon) \cdot 2^n$, for some fixed $\varepsilon > 0$, seems difficult. In [12], we proved such a result, showing that every $s$-saturated family of subsets of $[n]$ has size at least $(1 - 1/s) \cdot 2^n$.

This lower bound is a consequence of a multipartite version of the problem, in which we seek a lower bound on $|\mathcal{F}_1| + \ldots + |\mathcal{F}_s|$ where $\mathcal{F}_1, \ldots, \mathcal{F}_s$ are families of subsets of $[n]$, such that there are no $s$ pairwise disjoint sets, one from each family $\mathcal{F}_i$, and furthermore no subset of $[n]$ can be added to any of the families while preserving this property. We show that $|\mathcal{F}_1| + \ldots + |\mathcal{F}_s| \geq (s - 1) \cdot 2^n$, which is tight, as can be seen, for example, by taking $\mathcal{F}_1$ to be empty, and letting the remaining families be the families of all subsets of $[n]$.

We give two short proofs of this result, one using algebraic methods, and one based on correlation inequalities.

## 5 Extremal graph theory

### 5.1 Matija Bucić Nearly-linear monotone paths in edge-ordered graphs

How long a monotone path can one always find in any edge-ordering of the complete graph $K_n$? This appealing question was first asked by Chvátal and Komlós [15] in 1971, and has since attracted the attention of many researchers, inspiring a variety of related problems. The prevailing conjecture is that one can always find a monotone path of linear length, but until now the best known lower bound was $n^{2/3 - o(1)}$ due to Milans [32]. We almost close this gap, proving that any edge-ordering of the complete graph contains a monotone path of length $n^{1-o(1)}$. Our main results are the following.

**Theorem.** In any edge-ordering of the complete graph $K_n$, there is a monotone path of length

$$f(K_n) \geq \frac{n}{2^{O((\sqrt{\log n \log \log n})}}} = n^{1-o(1)}.$$

The altitude $f(G)$ of a graph $G$ is defined as the maximum $k$ such that every edge-ordering of $G$ has a monotone path of length $k$.

**Theorem.** Let $G$ be a graph with $n$ vertices and average degree $d \geq 2$. Then

$$f(G) \geq \frac{d}{2^{O((\sqrt{\log d \log \log n})}}.$$
This is a joint work with Matthew Kwan, Alexey Pokrovskiy, Benny Sudakov, Tuan Tran, and Adam Zsolt Wagner.

5.2 Shagnik Das How redundant is Mantel’s Theorem?

One of the all-time combinatorial classics, Mantel’s Theorem [30] asserts that if one forbids an \( n \)-vertex graph \( G \) from containing any triangle, then \( G \) can have at most \( \frac{n^2}{3} \) edges. This fundamental theorem has inspired a great deal of research, including the foundational work of Turán. Turán’s Theorem [43] describes the largest \( K_r \)-free graphs, in particular showing that they have \( (1 - \frac{1}{r^2}) \binom{n}{2} \) edges.

In this talk we explored an extension of these theorems suggested by Gil Kalai. In the above results, we forbade any of the potential \( r \)-cliques from appearing in our graph \( G \). Suppose instead that our resources are somewhat more limited, in the sense that, for some \( 0 \leq m \leq \binom{n}{r} \), we can only forbid \( m \) of the \( r \)-cliques from appearing in \( G \). What choice of forbidden cliques then minimises the maximum number of edges \( G \) can have? In particular, how many cliques need to be forbidden to recover the bounds from Mantel’s and Turán’s Theorems?

Some asymptotic bounds in this direction, showing that a cubic number of forbidden cliques are necessary and sufficient to recover the classic extremal bounds, were earlier obtained by Allen, Böttcher, Hladký and Piguet [1]. Focussing on the case \( r = 3 \), we presented some precise results beyond this initial range, determining how many additional edges can be achieved when fewer triangles are forbidden. At the other end of the spectrum, when there are only quadratically many forbidden triangles, we discussed the connection of this problem to combinatorial designs, explaining how Steiner Triple Systems can be used to construct sets of forbidden triangles that are hard to avoid.

This is a joint work with Ander Lamaison and Tuan Tran.

5.3 Daniel Kráľ Extremal problems concerning cycles in tournaments

The conjecture of Linial and Morgenstern [28] asserts that, among all tournaments with a given density \( d \) of cycles of length three, the density of cycles of length four is minimized by a random blow-up of a transitive tournament with all but one parts of equal sizes, i.e., a tournament with the structure similar to graphs appearing in the Erdős-Rademacher problem on triangles in graphs with a given edge density. We prove this conjecture for \( d \geq 1/36 \) using methods from spectral graph theory, and demonstrate that the structure of extremal examples is more complex than expected and give its full description for \( d \geq 1/16 \).

The talk was based on the paper [14].

5.4 Tibor Szabó On the quasirandomness of the projective norm graph

The projective norm graphs \( NG(q,t) \) provide tight constructions in the Turán problem for complete bipartite graphs \( K_{t,s} \) when \( s > (t - 1)! \) (15). The \( K_{t,s} \)-freeness of \( NG(q,t) \) is a very much atypical property: in a random graph with the same edge density a positive fraction of \( t \)-tuples are involved in a copy of \( K_{t,s} \). Yet, projective norm graphs are random-like in various other senses. Most notably their second eigenvalue is of the order of the square root of the degree, which, through the Expander Mixing Lemma, implies further quasirandom properties concerning the density of small enough subgraphs. In our papers [8, 31] we explore how far this quasirandomness goes. In particular we make progress on what subgraphs it must contain, beyond the eigenvalue methods of the Expander Mixing Lemma. Our results extend those of Alon and Shikelman [6] and Ma et al [29] on generalized Turán numbers, moreover imply that the \( K_{4,7} \)-free projective norm graph \( NG(q,4) \) does contain \( K_{4,6} \) for every prime power \( q \geq 5 \).

The main contribution of our proof is the estimation, and sometimes determination, of the number of solutions \( X \in \mathbb{F}_{q^r} \) of the norm equation system \( N_i(X + A_i) = a_i, \ (i = 1, \ldots, r) \) for various constants \( A_i \in \mathbb{F}_{q^r} \) and \( a_i \in \mathbb{F}_{q} \), where \( N_i : \mathbb{F}_{q^r} \to \mathbb{F}_{q} \) denotes the norm function, defined by \( N_i(A) = A^{q^{r-1} + q^{r-2} + \cdots + 1} \). In connection with norm graphs it has been proved before [25] that for \( r = t \) and arbitrary prime power \( q \), the number of solutions is at most \( t! \). Here we investigate this problem for arbitrary \( t \) when \( r \leq 3 \). We find that for \( r = 2 \) and arbitrary \( t \geq 3 \), the number of solutions is the expected \( (1 + o(1))t^{q^2-2} \), except in a very special case of \( t = 3 \), when it is two times that. We characterize the solutions in this special case via a surprising connection to classic planar Singer difference sets [41]. This connection can then be utilized to show that for
the critical case of \( r = 3 = t \), there do exist parameters with \( 6 = 3! \) solutions. Finally for \( r = 3 \) and arbitrary \( t \) we use elementary methods to bound the number of solutions by \( 6 d^{t-3} \), which was only known for \( t = 3 \), using algebraic geometry.

Joint work with Tomas Bayer, Tamás Mészáros, and Lajos Rónyi.

6 Extremal hypergraph theory

6.1 Noga Alon Traces of Hypergraphs

What is the largest number of distinct projections (traces) onto \( k \) coordinates guaranteed in every family of \( m \) binary vectors of length \( n \)?

This fundamental combinatorial question received a considerable amount of attention, has applications in theoretical Computer Science, Geometry, Machine Learning, Probability and more, and is wide open for most settings of the parameters.

In joint work with Moshkovitz and Solomon [4] we found an asymptotic solution of the question for linear \( k \) and sub-exponential \( m \), greatly improving earlier estimates.

In particular, for every constants \( r > 1 \) and \( 0 < a < 1 \), when \( n \) is large, every family of \( m = n^r \) binary vectors of length \( n \) has at least \( \Theta(n^C) \) projections on some set of \( k = an \) coordiantes, where \( C = \frac{r+1}{2}\log(1+a) - \log(1+a) \). This is tight up to the hidden polylogarithmic term in the \( \Theta \) notation.

6.2 David Conlon Improved bounds for the Brown–Erdős–Sós problem

Let \( f_r(n, v, e) \) be the maximum number of edges in an \( r \)-uniform hypergraph on \( n \) vertices which contains no induced subgraph with \( v \) vertices and at least \( e \) edges. The Brown–Erdős–Sós problem [10, 11] of determining \( f_r(n, v, e) \) is a central question in extremal combinatorics, with surprising connections to a number of seemingly unrelated areas. For example, the result of Ruzsa and Szemerédi [38] that \( f_3(n, 6, 3) = o(n^2) \) implies Roth’s theorem on the existence of 3-term arithmetic progressions in dense subsets of the integers.

As a generalisation of this result, it is conjectured that

\[
f_r(n, e(r-k)+k+1,e) = o(n^k)
\]

for any fixed \( r > k \geq 2 \) and \( e \geq 3 \). The best progress towards this conjecture, due to Sárközy and Selkow [39, 40], says that

\[
f_r(n, e(r-k)+k+\lfloor \log e \rfloor, e) = o(n^k),
\]

where the \( \log \) is taken base two. Building on an idea of Solymosi and Solymosi [42], we improve the Sárközy–Selkow bound, showing that \( \lfloor \log e \rfloor \) can be replaced with \( O(\log e / \log \log e) \). Even for the special case where \( r = 3 \) and \( k = 2 \), the proof relies on applying the hypergraph removal lemma at all uniformities.

Joint work with Lior Gishboliner, Yevgeny Levanzov and Asaf Shapira

6.3 Asaf Ferber Resilience for perfect matchings in random hypergraphs

Let \( m_d(k,n) \) be the smallest integer \( m \) for which every \( k \)-uniform hypergraph on \( n \) vertices and with minimum \( d \)-degree at least \( m \) contain a perfect matching (we always assume that \( n \) is divisible by \( k \)). The problem of determining \( m_d(k,n) \) has attracted a lot of attention in the last few decades. Quite surprisingly, this problem is still wide open, and have led to the development of many techniques and interesting connections with other areas (see, e.g., the surveys [37, 44] and the references therein).

Recently, there has been interest in extending extremal theorems to random environments. In our setting, we are interested in finding minimum degree conditions for the existence of a perfect matching in subgraphs of a typical random hypergraph \( H_{n,p}^k \). We prove that for \( p \geq C \max\{n^{-k/2+\epsilon}, n^{-k/d+1}\} \), a typical \( H_{n,p}^k \) is such that every subgraph \( G \subseteq H \) with minimum \( d \)-degree at least \( (1+o(1))m_d(k,n) \) contains a perfect matching. Note that this holds for all \( d, k \) even in cases when the exact value of \( m_d(k,n) \) is unknown. Our proof is based on a new “non-constructive” absorbers technique which we believe is of independent interest.

Joint work with Matthew Kwan.
6.4 Alexandr Kostochka *Super-pancyclic hypergraphs and bipartite graphs*

We find Dirac-type sufficient conditions for a hypergraph \( H \) with few edges to be hamiltonian. It is convenient to use the language of the bipartite graphs that are the incidence graphs of hypergraphs. For integers \( n, m, \) and \( \delta \) with \( \delta \leq m \), we denote by \( \mathcal{G}(n, m, \delta) \) the set of all bipartite graphs with partition \((X, Y)\) such that \(|X| = n \geq 2, |Y| = m \) and for every \( x \in X \), \( d(x) \geq \delta \). In 1981 Jackson proved that if a graph \( G \in \mathcal{G}(n, m, \delta) \) satisfies \( n \leq \delta \) and \( m \leq 2\delta - 2 \), then it contains a cycle that covers \( X \). He also conjectured that if \( G \in \mathcal{G}(n, m, \delta) \) is 2-connected, \( m \leq 3\delta - 5 \) and \( n \leq \delta \), then the same conclusion holds.

We prove this conjecture and describe the extremal graphs for Jackson’s Theorem. We also show that under the conditions of Jackson’s Theorem, the significantly stronger conclusion holds: \( G \) is X-super-pancyclic, i.e., for every \( A \subset X \) with \( |A| \geq 3 \), \( G \) has a cycle \( C_A \) such that \( V(C) \cap X = A \). We also find necessary conditions for a graph in \( \mathcal{G}(n, m, \delta) \) to be super-pancyclic that are often sufficient for it. We conjecture that these conditions are equivalent for a graph in \( \mathcal{G}(n, m, \delta) \) to be super-pancyclic.

Joint work with Ruth Luo and Dara Zirlin

6.5 Mykhaylo Tyomkyn *When Ramsey met Brown, Erdős and Sós*

Brown, Erdős and Sós [10] conjectured that for every \( c > 0 \) and \( k \geq 3 \) any \( r \)-uniform hypergraph with \( n \) vertices and \( cn^2 \) edges contains \( k \) edges spanned by at most \( (r - 2)k + 3 \) vertices – an \((r - 2)k + 3, k\)-configuration, for short. It is well known that this question reduces to linear hypergraphs, that is hypergraphs with no two edges sharing more than one vertex. A complete linear \( r \)-graph (also known as \( r \)-Steiner System) is a linear hypergraph corresponding to an edge-decomposition of a complete graph \( K_n \) into copies of \( K_r \). We prove the following Ramsey version of the Brown-Erdős-Sós conjecture.

**Theorem 5.** For every integer \( c \) there exists \( r_0 = r_0(c) \) such that for every \( r \geq r_0 \) and integer \( k \geq 3 \) there exists \( n_0 = n_0(c, r, k) \) such that every \( c \)-colouring of a complete linear \( r \)-graph on \( n \geq n_0 \) vertices contains a monochromatic \((r - 2)k + 3, k\)-configuration.

In the important special case of \( c = 2 \) we show that \( r_0 \) can be chosen as small as 4.

**Theorem 6.** For any integers \( r \geq 4 \) and \( k \geq 3 \) there exists \( n_0 = n_0(r, k) \) such that every \( 2 \)-colouring of a complete linear \( r \)-graph on \( n \geq n_0 \) vertices contains a monochromatic \((r - 2)k + 3, k\)-configuration.

This is a joint work with Asaf Shapira.

6.6 Jacques Verstraete *Ordered graphs and hypergraphs: new results and open problems*

There has been substantial recent interest in extremal problems for ordered graphs, with the recent breakthrough of Marcus and Tardos on excluded permutations and the Stanley-Wilf and Furedi-Hajnal conjectures.

We give a survey of selected results, techniques and applications, as well as new results and techniques for ordered hypergraphs, with applications to ordered trees in graphs, tight paths in hypergraphs, and directed paths in eulerian digraphs, amongst others.

Joint work with Z. Furedi, T. Jiang, A. Kostochka, D. Mubayi

7 Design Theory

7.1 Adam Wagner *Completion and deficiency problems*

Given a partial Steiner triple system (STS) of order \( n \), what is the order of the smallest complete STS it can be embedded into? The study of this question goes back more than 40 years. In this talk we answer it for relatively sparse STSs, showing that given a partial STS of order \( n \) with at most \( r \leq 10n^2 \) triples, it can always be embedded into a complete STS of order \( n + O(\sqrt{n}) \), which is asymptotically optimal. We also obtain similar results for completions of Latin squares and other designs.

This suggests a new, natural class of questions, called *deficiency problems*. Given a global spanning property \( \mathcal{P} \) and a graph \( G \), we define the deficiency \( \text{defi}(G) \) of the graph \( G \) with respect to the property \( \mathcal{P} \) to be
the smallest positive integer $t$ such that the join $G \ast K_t$ has property $P$. To illustrate this concept we consider deficiency versions of some well-studied properties, such as having a $K_k$-decomposition, Hamiltonicity, having a triangle-factor and having a perfect matching in hypergraphs [35].

The main goal of this talk is to propose a systematic study of these problems; thus several future research directions are also given.

This is a joint work with Rajko Nenadov and Benny Sudakov.

8 Probability theory

8.1 Omer Angel Pairwise optimal coupling of multiple random variables

A coupling of a collection of random variables $(X_i)_{i \in I}$ is a set of variables $(X'_i)_{i \in I}$ on some common probability space with the given marginals, i.e. $X_i$ and $X'_i$ have the same law. We omit the primes when there is no risk of confusion. The total variation distance between two random variables $X$ and $Y$ is defined as

$$d_{TV}(X,Y) = \sup_A \{|\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|\},$$

where the supremum is over all (measurable) sets $A$. The fundamental, classical theorem relating the total variation distance to coupling is the following folklore theorem.

**Theorem 7.** For any two random variables $X$ and $Y$, there exists a coupling such that $\mathbb{P}(X \neq Y) = d_{TV}(X,Y)$. Moreover, this is the smallest possible value of $\mathbb{P}(X \neq Y)$ for any coupling.

When coupling more than two random variables, the total variation bound cannot be achieved simultaneously for all pairs even in very simple cases. For example, there are 3 random variables with pairwise total variation distance 1/2, but in any coupling some pair disagrees with probability at least 2/3. We prove a generalization of Theorem 7 with a slightly higher probability of disagreement, based on two different constructions.

**Theorem 8.** For any countable collection $S$ of real random variables, there exists a coupling such that, for any $X, Y \in S$,

$$\mathbb{P}(X \neq Y) \leq f(d_{TV}(X,Y)),$$

where $f(x) = \frac{2x}{1+x}$.

For rational values of $x$, proving that $f(x)$ is optimal is equivalent to a certain extremal combinatorial problem, which we can solve for some but not all values of $x$.

Joint work Yinon Spinka.

9 Expander graphs

9.1 Nati Linial Expander Graphs – Both Local and Global

Let $G = (V,E)$ be a finite graph. For $v \in V$ we denote by $G_v$ the subgraph of $G$ that is induced by $v$’s neighbor set. We say that $G$ is $(a,b)$-regular for $a > b > 0$ integers, if $G$ is $a$-regular and $G_v$ is $b$-regular for every $v \in V$. Recent advances in PCP theory call for the construction of infinitely many $(a,b)$-regular expander graphs $G$ that are expanders also locally. Namely, all the graphs $\{G_v|v \in V\}$ should be expanders as well. While random regular graphs are expanders with high probability, they almost surely fail to expand locally. Here we construct two families of $(a,b)$-regular graphs that expand both locally and globally. We also analyze the possible local and global spectral gaps of $(a,b)$-regular graphs. In addition, we examine our constructions vis-a-vis properties which are considered characteristic of high-dimensional expanders.

It is hard to overstate the significance of expander graphs in theoretical computer science and the impact their study has had on a number of mathematical areas. A particularly fascinating example of such an application is Dinur’s proof of the PCP Theorem. However, in recent advances in PCP theory more specialized expander graphs are required. If $v$ is a vertex in a graph $G$ we denote by $G_v$ the subgraph of $G$ that is induced
by \(v\)'s neighbors and call it the \textit{link} of \(v\) in \(G\). We seek large regular expanders \(G\) such that \(G_v\) is an expander for every \(v \in V(G)\).

One of the first discoveries in the study of expanders is that for every \(d \geq 3\) asymptotically almost every \(d\)-regular graph is a very good expander. However, it is easy to verify that almost every \(d\)-regular graph is very far from satisfying the above requirement, as \(G_v\) is typically an antie clique. So, here is our basic question: Given positive integers \(a > b\) do there exist arbitrarily large \((a,b)\)-expanders? Namely, \(a\)-regular expander graphs \(G\) such that every \(G_v\) is a \(b\)-regular expander. If so, how good can the expansion properties (edge expansion, spectral gap) of \(G\) and the graphs \(G_v\) be?

These investigations are closely related to the recently emerging field of high-dimensional expanders. Vertex-expansion, edge-expansion, spectral gaps and the speed of convergence of the simple random walk on the graph are key ingredients in the theory of expander graphs. While these parameters need not perfectly coincide, they mutually control each other quite tightly. In contrast, the high-dimensional theory suggests a number of inherently different ways to quantify expansion. Namely, the connections between these concepts are nowhere as tight as in the one-dimensional case of expander graphs. It is very suggestive to explore families of \((a,b)\)-expanders in light of this array of quantitative measures of high-dimensional expansion.

Let \(G\) be a graph and \(v \in V(G)\). The link of \(v\) denoted \(G_v\) is the subgraph of \(G\) that is induced by the vertex set \(\{u \in V \mid uv \in E\}\).

**Definition 9.** Let \(a > b \geq 0\) be integers. An \((a,b)\)-regular graph \(G\) is an \(a\)-regular graph, where for every vertex \(v \in V(G)\) the link \(G_v\) is \(b\)-regular.

It is natural to ask how large the spectral gaps can be in an \((a,b)\)-regular graph. We prove an optimal Alon-Boppana type bound which makes no reference to the graph's local expansion:

**Theorem 10.** The second eigenvalue of an \((a,b)\)-regular graph satisfies

\[
\lambda_2 \geq b + 2\sqrt{a - b - 1} - o_n(1).
\]

The bound is tight.

In the graphs that we construct to prove the tightness of the bound in Theorem 10, all the links are disconnected. So, the next obvious question is whether the same bound can be attained by graphs whose links are all expanders, or at least connected. The following theorem shows that the answer is negative, by describing some \textit{tradeoff} between local and global expansion.

**Theorem 11.** Consider an \((a,b)\)-regular graph each of whose links has edge expansion at least \(\delta > 0\). Then there exists some \(\epsilon = \epsilon(a,b,\delta) > 0\) such that the second eigenvalue of the graph satisfies:

\[
\lambda_2 \geq \left(b + 2\sqrt{a - b - 1}\right)(1 + \epsilon) - o_n(1).
\]

For fixed \(a\) and \(b\) with \(a \geq b^2 + O(b)\), \(\epsilon\) strictly increases with \(\delta\). For any other fixed values of \(a\) and \(b\), \(\epsilon\) increases for small enough \(\delta\).

The main new construction that we introduce here is the \textit{Polygraph}. It can be viewed as a family of new graph products which transform a high-girth regular expander into an \((a,b)\)-expander. To illustrate this idea, let \(q > p \geq 0\) be integers, let \(G\) be a graph with distance function \(\rho\) and girth larger then \(3p + 3q\). The vertex set of the \textit{polygraph} \(G_S\) is \(V(G)^3\) and \((x_1, x_2, x_3)\) is a neighbor of \((y_1, y_2, y_3)\) iff the multiset of three distances \(|\rho(x_i, y_i)| \mid i = 1, 2, 3\) coincides with the multiset \([p, q, p + q]\).

For illustration, here is a way of viewing the Polygraph corresponding to \(p = 0\) and \(q = 1\). Take three copies of a \(d\)-regular triangle-free graph \(G\) and have a token move on each of them. At every step two of the tokens move to a neighboring vertex and the third token stays put. Any configuration of tokens is a vertex of the graph and the above process defines its adjacency relation.

**Theorem 12.** Let \(q > p \geq 0\) be even integers. If \(G\) is connected, non-bipartite and its girth is bigger than \(3p + 3q\), then \(G_S\) is an \((a,b)\)-regular local \(\epsilon\)-spectral expander and global \(\epsilon'\)-spectral expander. Here \(a, b\) and \(\epsilon\) depend only on \(p\) and \(q\), while \(\epsilon'\) depends also on the spectral gap of \(G\).

Joint work with Michael Chapman and Yuval Peled.


10  Matroid theory

10.1  Alexey Pokrovskiy  Halfway to Rota’s basis conjecture

In 1989, Rota made the following conjecture. Given \( n \) bases \( B_1, \ldots, B_n \) in an \( n \)-dimensional vector space \( V \), one can always find \( n \) disjoint bases of \( V \), each containing exactly one element from each \( B_i \) (we call such bases rainbow bases). Rota’s basis conjecture remains open despite its apparent simplicity and the efforts of many researchers (for example, the conjecture was recently the subject of the collaborative Polymath project). In this talk, it was discussed how to find \((0.5 - o(1))n\) disjoint rainbow bases, improving the previously best known bound of \( n/\log n \).

**Theorem 13** (Bucić, Kwan, Pokrovskiy, and Sudakov, [13]). For any \( \varepsilon > 0 \), the following holds for sufficiently large \( n \). Given bases \( B_1, \ldots, B_n \) of a rank-\( n \) matroid, there are at least \( (1/2 - \varepsilon) n \) disjoint rainbow bases.

See [13] for the full details. In order to simplify the presentation, the talk focused on proving the following weaker result:

**Proposition**  In any family of \( n \)-edge trees \( T_1, \ldots, T_n \) on the same vertex set, it is possible to find \( \sqrt{n} \) edge-disjoint rainbow forests of size \( n - 1 \) (here a rainbow forest is one which contains at most one edge from each \( T_1, \ldots, T_n \)).

This proposition is easier than the full theorem from [13] in three different ways:

- It is about graphic matroids rather than general matroids. This doesn’t affect the proof strategy, but makes the matroids easier to visualise.
- If only finds rainbow forests of size \( n - 1 \) (rather than \( n - 1 \) like Rota’s Conjecture would predict). This is a convenient simplification because it means that every rainbow forest we consider misses some colour.
- It finds only \( \sqrt{n} \) rainbow forests. This is an illuminating because the full proof of Theorem 13 uses an iterative argument to find \((1/2 - \varepsilon) n\) disjoint rainbow bases. When one performs only one step of the iteration, then one naturally obtains \( \sqrt{n} \) rainbow objects.

Joint work with Bucić, Kwan, and Sudakov.

11  Other subjects

11.1  Matthew Kwan An algebraic inverse theorem for the quadratic Littlewood–Offord problem

Consider a quadratic polynomial \( f(\xi_1, \ldots, \xi_n) \) of independent Bernoulli random variables. What can be said about the concentration of \( f \) on any single value? This generalises the classical Littlewood–Offord problem, which asks the same question for linear polynomials. As in the linear case, it is known that the point probabilities of \( f \) can be as large as about \( 1/\sqrt{n} \), but still poorly understood is the “inverse” question of characterising the algebraic and arithmetic features \( f \) must have if it has point probabilities comparable to this bound (see [36]).

In this talk we discuss some results joint with Lisa Sauermann [26] of an algebraic flavour, showing that if \( f \) has point probabilities much larger than \( 1/n \) then it must be close to a quadratic form with low rank. We also give an application to Ramsey graphs, asymptotically answering a question of Kwan, Sudakov and Tran [27].

11.2  Bhargav Narayanan Disproportionate division

The cake-cutting problem, whose study was initiated by Banach and Steinhaus in 1949, is a classical measure partitioning problem concerned with the division of a ‘cake’, here the unit interval \([0, 1]\), amongst \( n \geq 2 \)
agents each with their own ‘utilities’, here non-atomic Borel probability measures $\mu_1, \mu_2, \ldots, \mu_n$ on $[0, 1]$. Given non-negative demands $\alpha_1, \alpha_2, \ldots, \alpha_n$ summing to 1, a disproportionate division for these demands is a partition $X_1 \cup X_2 \cdots \cup X_n$ of $[0, 1]$ with $\mu_i(X_i) \geq \alpha_i$ for all $1 \leq i \leq n$. Folklore arguments from algebraic topology show that a disproportionate division for $n$ agents with arbitrary demands may always be found with $O(n^2)$ cuts, and a more efficient rendition of this argument, recently discovered by Segal-Halevi, shows that in fact $O(n \log n)$ cuts always suffice. Our main result improves on these decades-old topological arguments as follows.

**Theorem 14.** For all $n \geq 2$, given non-atomic probability measures $\mu_1, \mu_2, \ldots, \mu_n$ on $[0, 1]$ and non-negative reals $\alpha_1, \alpha_2, \ldots, \alpha_n$ summing to 1, there exists a disproportionate division for these demands with at most $3n - 4$ cuts.

It is clear that $n - 1$ cuts are always necessary for $n$ agents, so this result is tight up to multiplicative constants. Our proof of Theorem 14 is combinatorial as opposed to topological; an attractive byproduct of this approach is that the proof is constructive, yielding an effective procedure for disproportionate division. While Theorem 14 determines the optimal number of cuts for disproportionate division with $n$ agents up to multiplicative constants, the problem of pinning down this extremal number still remains. Unlike with fair division, it turns out that $n - 1$ cuts do not always suffice; there is a construction demonstrating that $2n - 2$ cuts may be necessary in general. We suspect that this construction, not Theorem 14, reflects the truth, and that the tightness of this construction should follow from topological considerations: to this end, we present the following conjecture.

**Conjecture 15.** For any $n \geq 2$ non-atomic probability measures $\mu_1, \mu_2, \ldots, \mu_n$ on the unit circle $S^1$ and non-negative reals $\alpha_1, \alpha_2, \ldots, \alpha_n$ summing to 1, there exists a partition of the set $[n] = P \cup Q$ into two nonempty sets and a partition of the circle $S^1 = X \cup X^c$ into two intervals such that

$$\min_{i \in P} \mu_i(X) = \sum_{j \in P} \alpha_j \text{ and } \min_{i \in Q} \mu_i(X^c) = \sum_{j \in Q} \alpha_j.$$  

**References**


