Proceedings of the International Congress of Mathematicians
Hyderabad, India, 2010

Recent Developments in Extremal Combinatorics: Ramsey and Turán Type Problems

Benny Sudakov*

Abstract

Extremal combinatorics is one of the central branches of discrete mathematics and has experienced an impressive growth during the last few decades. It deals with the problem of determining or estimating the maximum or minimum possible size of a combinatorial structure which satisfies certain requirements. Often such problems are related to other areas including theoretical computer science, geometry, information theory, harmonic analysis and number theory. In this paper we discuss some recent advances in this subject, focusing on two topics which played an important role in the development of extremal combinatorics: Ramsey and Turán type questions for graphs and hypergraphs.

Mathematics Subject Classification (2010). 05C35, 05C65, 05D10, 05D40

Keywords. Extremal combinatorics, Ramsey theory, Turán problems, Probabilistic methods

1. Introduction

Discrete mathematics (or combinatorics) is a fundamental mathematical discipline which focuses on the study of discrete objects and their properties. Although it is probably as old as the human ability to count, the field experienced tremendous growth during the last fifty years and has matured into a thriving area with its own set of problems, approaches and methodologies. The development of powerful techniques, based on ideas from probability, algebra, harmonic analysis and topology, is one of the main reasons for the rapid growth

*Research supported in part by NSF CAREER award DMS-0812005 and by USA-Israeli BSF grant.

Department of Mathematics, UCLA, Los Angeles, CA 90095.
E-mail: bsudakov@math.ucla.edu.

of combinatorics. Such tools play an important organizing role in combinatorics, similar to the one that deep theorems of great generality play in more classical areas of mathematics.

Extremal combinatorics is one of the central branches of discrete mathematics. It deals with the problem of determining or estimating the maximum or minimum possible cardinality of a collection of finite objects (e.g., numbers, graphs, vectors, sets, etc.) satisfying certain restrictions. Often such problems appear naturally in other areas, and one can find applications of extremal combinatorics in theoretical computer science, geometry, information theory, analysis, and number theory. Extremal combinatorics has developed spectacularly in the last few decades, and two topics which played a very important role in its development are Ramsey theory and Turán type problems for graphs and hypergraphs.

The foundations of Ramsey theory rest on the following general phenomenon: every large object, chaotic as it may be, contains a sub-object that is guaranteed to be well structured, in a certain appropriately chosen sense. This phenomenon is truly ubiquitous and manifests itself in different mathematical areas, ranging from the most basic Pigeonhole principle to intricate statements from set theory. Extremal theory of graphs and hypergraphs considers problems such as the maximum possible number of edges in a triangle-free graph with a given number of vertices. The development of this subject was instrumental in turning Graph Theory into a modern, deep and versatile field.

Both areas use a variety of sophisticated methods and arguments (for example, algebraic and probabilistic considerations, geometric constructions, the stability approach and the regularity method) and there is a considerable overlap between them. Indeed, Ramsey theory studies which configurations one can find in every finite partition of the large structure. On the other hand, extremal graph theory deals with the inevitable occurrence of some specified configuration when the edge density of graph or hypergraph exceeds a certain threshold.

In this paper we survey recent progress on some classical Ramsey and Turán type problems, focusing on the basic ideas and connections to other fields. It is of course impossible to cover everything in such a short article, and therefore the choice of results we present is inevitably biased. Yet we hope to describe enough examples, problems and techniques from this fascinating subject to appeal to researchers not only in discrete mathematics but in other areas as well.

2. Ramsey Theory

Ramsey theory refers to a large body of deep results in mathematics whose underlying philosophy is captured succinctly by the statement that "Every large system contains a large well organized subsystem." This is an area in which a great variety of techniques from many branches of mathematics are used and whose results are important not only to combinatorics but also to logic, analysis, number theory, and geometry. Since the publication of the seminal paper of Ramsey [81] in 1930, this subject has experienced an impressive growth, and is currently among the most active areas in combinatorics.

The *Ramsey number* $r_k(s_1, s_2, \ldots, s_\ell)$ is the least integer N such that every ℓ-coloring of the unordered k-tuples of an N-element set contains a monochromatic set of size s_i in color i for some $1 \le i \le \ell$, where a set is called monochromatic if all k-tuples from this set have the same color. Ramsey's theorem states that these numbers exist for all values of the parameters. In the case of graphs (i.e., $k = 2$) it is customary to omit the index k and to write simply $r(s_1, \ldots, s_\ell)$.

Originally, Ramsey applied his result to a problem in logic, but his theorem has many additional applications. For example, the existence of the ℓ-colored Ramsey number $r(3, 3, \ldots, 3)$ can be used to deduce the classical theorem of Schur from 1916. Motivated by Fermat's last theorem, he proved that any ℓ-coloring of a sufficiently large initial segment of natural numbers contains a monochromatic solution of the equation $x + y = z$. Another application of this theorem to geometry was discovered by Erdős and Szekeres [48]. They showed that any sufficiently large set of points in the plane in general position (no 3 of which are collinear) contains the vertices of a convex n-gon. They deduced this result from Ramsey's theorem together with the simple fact that any 5 points in general position contain a convex 4-gon. Another early result of Ramsey theory is van der Waerden's theorem, which says that every finite coloring of the integers contains arbitrarily long arithmetic progressions. The celebrated density version of this theorem, proved by Szemerédi [98], has led to many deep and beautiful results in various areas of mathematics, including the recent spectacular result of Green and Tao [63] that there are arbitrarily long arithmetic progressions in the primes.

Determining or estimating Ramsey numbers is one of the central problems in combinatorics, see [62] for details. Erdős and Szekeres [48] proved a quantitative version of Ramsey's theorem showing that $r(s, n) \le \binom{n+s-2}{s-1}$. To prove this simple statement, one can fix a vertex and, depending on the number of its neighbors in colors 1 and 2, apply an induction to one of these two sets. In particular, for the diagonal case when $s = n$ it implies that $r(n, n) \le 2^{2n}$ for every positive integer n. The first exponential lower bound for this numbers was obtained by Erdős [32], who showed that $r(n, n) > 2^{n/2}$ for $n > 2$. His proof, which is one of the first applications of probabilistic methods in combinatorics is extremely short. The probability that a random 2-edge coloring of the complete graph K_N on $N = 2^{n/2}$ vertices contains a monochromatic set of size n is at most $\binom{N}{n} 2^{1-\binom{n}{2}} < 1$. Hence there is a coloring with the required properties. Although the proofs of both bounds for $r(n, n)$ are elementary, obtaining significantly better results appears to be notoriously difficult. Over the last sixty years, there have been several improvements on these estimates (most recently by Conlon, in [22]), but the constant factors in the above exponents remain the same. Improving these exponents is a very fundamental problem and will probably require novel techniques and ideas. Such techniques will surely have many applications to other combinatorial problems as well.

The probabilistic proof of Erdős [32], described above, leads to another important open problem which seems very difficult. Can one explicitly construct for some fixed $\epsilon > 0$ a 2-edge coloring of the complete graph on $N > (1+\epsilon)^n$ vertices with no monochromatic clique of size n? *Explicit* here means that there is an algorithm which produces the coloring in polynomial time in the number of its vertices. Despite a lot of efforts this question is still open. For many years the best known result was due to Frankl and Wilson [54] who gave an elegant explicit construction of such a coloring on $n^{c\frac{\log n}{\log\log n}}$ vertices for some fixed $c > 0$. (All logarithms in this paper are in base e unless otherwise stated.) Recently a new approach to this problem and its bipartite variant was proposed in [9, 10]. In particular, for any constant C the algorithm of Barak, Rao, Shaltiel and Wigderson efficiently constructs a 2-edge coloring of the complete graph on $n^{\log^C n}$ vertices with no monochromatic clique of size n.

Off-diagonal Ramsey numbers, i.e., $r(s,n)$ with $s \neq n$, have also been intensely studied. After several successive improvements, the asymptotic behavior of $r(3,n)$ was determined by Kim [69] and by Ajtai, Komlos and Szemerédi [1].

Theorem 2.1. *There are absolute constants c_1 and c_2 such that*

$$c_1 n^2/\log n \leq r(3,n) \leq c_2 n^2/\log n.$$

This is an important result, which gives an infinite family of Ramsey numbers that are known up to a constant factor. The upper bound of [1] is proved by analyzing a certain randomized greedy algorithm. The lower bound construction of [69] uses a powerful *semi-random* method, which generates it through many iterations, applying probabilistic reasoning at each step. The analysis of this construction is subtle and is based on large deviation inequalities.

For $s > 4$ we only have estimates for $r(s,n)$ which are far apart. From the results of [1, 92] it follows that

$$c_3\left(\frac{n}{\log n}\right)^{(s+1)/2} \leq r(s,n) \leq c_4\frac{n^{s-1}}{\log^{s-2} n}, \tag{1}$$

for some absolute constants $c_3, c_4 > 0$. Recently, by analyzing the asymptotic behavior of certain random graph processes, Bohman [12] gave a new proof of the lower bound for $r(3,n)$. Together with Keevash [13], they used this approach to improve the above lower bound for $r(s,n)$ by a factor of $\log^{1/(s-2)} n$.

2.1. Hypergraphs.

Although already for graph Ramsey numbers there are significant gaps between the lower and upper bounds, our knowledge of hypergraph Ramsey numbers ($k \geq 3$) is even weaker. Recall that $r_k(s,n)$ is the minimum N such that every red-blue coloring of the k-tuples of an N-element set contains a red set of size s or a blue set of size n. Erdős, Hajnal, and Rado [43] showed that there are positive constants c and c' such that

$$2^{cn^2} < r_3(n,n) < 2^{2^{c'n}}.$$

They also conjectured that $r_3(n,n) > 2^{2^{cn}}$ for some constant $c > 0$ and Erdős offered a \$500 reward for a proof. Similarly, for $k \geq 4$, there is a difference of one exponential between the known upper and lower bounds for $r_k(n,n)$, i.e.,

$$t_{k-1}(cn^2) \leq r_k(n,n) \leq t_k(c'n),$$

where the tower function $t_k(x)$ is defined by $t_1(x) = x$ and $t_{i+1}(x) = 2^{t_i(x)}$.

The study of 3-uniform hypergraphs is particularly important for our understanding of hypergraph Ramsey numbers. This is because of an ingenious construction called the stepping-up lemma due to Erdős and Hajnal (see, e.g., Chapter 4.7 in [62]). Their method allows one to construct lower bound colorings for uniformity $k+1$ from colorings for uniformity k, effectively gaining an extra exponential each time it is applied. Unfortunately, the smallest k for which it works is $k = 3$. Therefore, proving that $r_3(n,n)$ has doubly exponential growth will allow one to close the gap between the upper and lower bounds for $r_k(n,n)$ for all uniformities k. There is some evidence that the growth rate of $r_3(n,n)$ is closer to the upper bound, namely, that with four colors instead of two this is known to be true. Erdős and Hajnal (see, e.g., [62]) constructed a 4-coloring of the triples of a set of size $2^{2^{cn}}$ which does not contain a monochromatic subset of size n. This is sharp up to the constant factor c in the exponent. It also shows that the number of colors matters a lot in this problem and leads to the question of what happens in the intermediate case when we use three colors. In this case, Erdős and Hajnal have made some improvement on the lower bound 2^{cn^2} (see [42, 20]), showing that $r_3(n,n,n) \geq 2^{cn^2 \log^2 n}$. Recently, extending the above mentioned stepping-up lemma approach, the author, together with Conlon and Fox [25], gave a strong indication that $r_3(n,n,n)$ is probably also double-exponential.

Theorem 2.2. *There is a constant $c > 0$ such that*

$$r_3(n,n,n) \geq 2^{n^{c \log n}}.$$

A simple induction approach which was used to estimate $r(s,n)$ gives extremely poor bounds for off-diagonal hypergraph Ramsey numbers when $k \geq 3$. In 1952 Erdős and Rado [45] gave an intricate argument which shows how to bound the Ramsey numbers for uniformity k using estimates for uniformity $k-1$. They proved that

$$r_k(s,n) \leq 2^{\binom{r_{k-1}(s-1,n-1)}{k-1}}. \tag{2}$$

Together with the upper bound in (1) this gives, for fixed s, that

$$r_3(s,n) \leq 2^{\binom{r(s-1,n-1)}{2}} \leq 2^{cn^{2s-4}/\log^{2s-6} n}.$$

Progress on this problem was slow and for several decades this was the best known bound. In [25], the authors discovered an interesting connection between the problem of bounding $r_3(s,n)$ and a new game-theoretic parameter, which we describe next.

Consider the following game, played by two players, the builder and the painter: at step $i+1$ a new vertex v_{i+1} is revealed; then, for every existing vertex v_j, $j = 1, \cdots, i$, the builder decides, in order, whether to draw the edge $v_j v_{i+1}$; if he does expose such an edge, the painter has to color it either red or blue immediately. The *vertex on-line Ramsey number* $\tilde{r}(k,l)$ is then defined as the minimum number of edges that the builder has to draw in order to force the painter to create a red K_k or a blue K_l. It appears that one can bound the Ramsey number $r_3(s,n)$ roughly by exponential in $\tilde{r}(s-1,n-1)$ and also provide an upper bound on $\tilde{r}(s-1,n-1)$ which is much smaller than the best known estimate on $\binom{r(s-1,n-1)}{2}$. These facts together with some additional ideas were used in [25] to show the following result, which improves the exponent of the upper bound by a factor of $n^{s-2}/\text{polylog}\, n$.

Theorem 2.3. *For fixed $s \geq 4$ and sufficiently large n, there exists a constant $c > 0$ such that*

$$r_3(s,n) \leq 2^{cn^{s-2}\log n}.$$

A similar improvement for off-diagonal Ramsey numbers of higher uniformity follows from this result together with (2).

Clearly one should also ask, how accurate are these estimates? For the first nontrivial case when $s = 4$, this problem was first considered by Erdős and Hajnal [41] in 1972. Using the following clever construction they showed that $r_3(4,n)$ is exponential in n. Consider a random tournament with vertex set $[N] = \{1, \dots, N\}$. This is a complete graph on N vertices whose edges are oriented uniformly at random. Color the triples from $[N]$ red if they form a cyclic triangle and blue otherwise. Since it is well known and easy to show that every tournament on four vertices contains at most two cyclic triangles and a random tournament on N vertices with high probability does not contain a transitive subtournament of size $c' \log N$, the resulting coloring neither has a red set of size 4 nor a blue set of size $c' \log N$. In the same paper [41], Erdős and Hajnal conjectured that $\frac{\log r_3(4,n)}{n} \to \infty$. This was recently confirmed in [25], where the authors obtained a more general result which in particular implies that $r_3(4,n) \geq 2^{cn \log n}$. This should be compared with the above upper bound that $r_3(4,n) \leq 2^{cn^2 \log n}$.

2.2. Almost monochromatic subsets.

Despite the fact that Erdős [36, 20] believed $r_3(n,n)$ is closer to $2^{2^{cn}}$, he discovered together with Hajnal [42] the following interesting fact which they thought might indicate the opposite. They proved that there are $c, \epsilon > 0$ such that every 2-coloring of the triples of an N-element set contains a subset S of size $s > c(\log N)^{1/2}$ such that at least $(1/2+\epsilon)\binom{s}{3}$ triples of S have the same color. That is, this subset deviates from having density 1/2 in each color by at least some fixed positive constant. Erdős ([37], page 67) further remarks that he would begin to doubt that $r_3(n,n)$ is double-exponential in n if one could prove that any 2-coloring of the triples of an N-set contains some set of size $s = c(\epsilon)(\log N)^\delta$ for which at least $(1-\epsilon)\binom{s}{3}$ triples have the same color, where $\delta > 0$ is an absolute constant and $\epsilon > 0$ is arbitrary. Erdős and Hajnal proposed [42] that such a statement may even be true with $\delta = 1/2$. The following result in [26] shows that this is indeed the case.

Theorem 2.4. *For each $\epsilon > 0$ and ℓ, there is $c = c(\ell,\epsilon) > 0$ such that every ℓ-coloring of the triples of an N-element set contains a subset S of size $s = c\sqrt{\log N}$ such that at least $(1-\epsilon)\binom{s}{3}$ triples of S have the same color.*

A random ℓ-coloring of the triples of an N-element set in which every triple gets one of ℓ colors uniformly at random shows that this theorem is tight up to the constant factor c. Indeed, using a standard tail estimate for the binomial distribution, one can show that in this coloring, with high probability, every subset of size $\gg \sqrt{\log N}$ has a $1/\ell + o(1)$ fraction of its triples in each color.

The above theorem shows a significant difference between the discrepancy problem in graphs and that in hypergraphs. As we already mentioned in the previous section, Erdős and Hajnal constructed a 4-coloring of the triples of an N-element set which does not contain a monochromatic subset of size $c \log\log N$. Also, by Theorem 2.2, there is a 3-coloring of the triples which does not contain a monochromatic subset of size $2^{c\sqrt{\log\log N}}$. Thus, Theorem 2.4 demonstrates (at least for $\ell \geq 3$) that the maximum almost monochromatic subset that an ℓ-coloring of the triples must contain is much larger than the corresponding monochromatic subset. This is in a striking contrast with graphs, where these two quantities have the same order of magnitude, as demonstrated by a random ℓ-coloring of the edges of a complete graph.

It would be very interesting to extend Theorem 2.4 to uniformity $k \geq 4$. In [25] the authors proved that for all k, ℓ and $\epsilon > 0$ there is $\delta = \delta(k,\ell,\epsilon) > 0$ such that every ℓ-coloring of the k-tuples of an N-element set contains a subset of size $s = (\log N)^\delta$ which contains at least $(1-\epsilon)\binom{s}{k}$ k-tuples of the same color. Unfortunately, δ here depends on ϵ. On the other hand, this result probably holds even with $\delta = 1/(k-1)$ (which is the case for $k = 3$).

3. Graph Ramsey Theory

The most famous question in Ramsey Theory is probably that of estimating $r(n,n)$. Since this problem remains largely unsolved with very little progress over the last 60 years, the focus of the field has shifted to the study of general graphs. Given an arbitrary fixed graph G, the Ramsey number $r(G)$ is the smallest integer N such that any 2-edge coloring of the complete graph K_N contains a monochromatic copy of G. For the classical Ramsey numbers G itself is taken to be a complete graph K_n. When ℓ colors are used to color the edges of K_N instead of two, we will denote the corresponding value of N by $r(G;\ell)$. The original motivation to study Ramsey numbers of general graphs was the hope that it would eventually lead to methods that would give better estimates for $r(n,n)$. While this hope has not been realized, a beautiful subject has emerged with many fascinating problems and results. Graph Ramsey Theory, which started about 35 years ago, quickly became one of the most active areas of Ramsey theory. Here we discuss several problems which have played an important role in this development.

3.1. Linear Ramsey numbers. Among the most interesting questions about Ramsey numbers are the linear bounds for graphs with certain degree constraints. In 1975, Burr and Erdős [17] conjectured that, for each positive integer Δ, there is a constant $c(\Delta)$ such that every graph G with n vertices and maximum degree Δ satisfies $r(G) \leq c(\Delta)n$. This conjecture was proved by Chvatál, Rödl, Szemerédi, and Trotter [21]. Their proof is a beautiful illustration of the power of Szemerédi's celebrated regularity lemma (see, e.g., [70]). Remarkably, this means that for graphs of fixed maximum degree the Ramsey number only has a linear dependence on the number of vertices. Because the original method used the regularity lemma, it gave tower type bound on $c(\Delta)$. More precisely, $c(\Delta)$ was bounded by exponential tower of 2s with a height that is itself exponential in Δ. Since then, the problem of determining the correct order of magnitude of $c(\Delta)$ as a function of Δ has received considerable attention from various researchers.

The situation was remedied somewhat by Eaton, who proved, still using a variant of the regularity lemma, that the function $c(\Delta)$ can be taken to be of the form $2^{2^{c\Delta}}$ (here and later in this section c is some absolute constant). A novel approach of Graham, Rödl, and Rucinski [60] gave the first linear upper bound on Ramsey numbers of bounded degree graphs without using any form of the regularity lemma. Their proof implies that $c(\Delta) < 2^{c\Delta \log^2 \Delta}$. In [61], they also proved that there are bipartite graphs with n vertices and maximum degree Δ for which the Ramsey number is at least $2^{c'\Delta}n$. Recently, refining their approach further, together with Conlon and Fox [27] the author proved that

$$c(\Delta) < 2^{c\Delta \log \Delta},$$

which brings it a step closer to the lower bound.

The case of bipartite graphs with bounded degree was studied by Graham, Rödl, and Rucinski more thoroughly in [61], where they improved their upper bound, showing that $r(G) \leq 2^{c\Delta \log \Delta}n$ for every bipartite graph G with n vertices and maximum degree Δ. Using a totally different approach, Conlon [23] and, independently, Fox and Sudakov [51] have shown how to remove the $\log \Delta$ factor in the exponent, achieving an essentially best possible bound of $r(G) \leq 2^{c\Delta}n$ in the bipartite case. This gives strong evidence that in the general case $c(\Delta)$ should also be exponential in Δ. The bound proved in [51] has the following form (the estimate in [23] is slightly weaker).

Theorem 3.1. *If G is a bipartite graph with n vertices and maximum degree $\Delta \geq 1$, then*

$$r(G) \leq \Delta 2^{\Delta+5} n\,.$$

One family of bipartite graphs that has received particular attention are the d-cubes. The d-cube Q_d is the d-regular graph with 2^d vertices whose vertex set is $\{0,1\}^d$ and two vertices are adjacent if they differ in exactly one coordinate. More than 30 years ago, Burr and Erdős [17] conjectured that the Ramsey number $r(Q_d)$ is linear in the number of vertices of the d-cube, i.e., there exists an absolute constant $c > 0$ such that $r(Q_d) \leq c2^d$. Since then, several authors have improved the upper bound for $r(Q_d)$, but the problem is still open. Beck [11] proved that $r(Q_d) \leq 2^{cd^2}$. The bound of Graham et al. [61] shows that $r(Q_d) \leq 8(16d)^d$. Using ideas from [72], Shi [88] proved the first exponential bound $r(Q_d) \leq 2^{cd}$, with exponent $c = (1+o(1))\frac{3+\sqrt{5}}{2} \approx 2.618$. A very special case of Theorem 3.1, when $G = Q_d$, gives immediately that for every positive integer d,

$$r(Q_d) \leq d2^{2d+5},$$

which is roughly quadratic in the number of vertices of the d-cube.

Given the recent advances in developing the hypergraph regularity method it was natural to expect that linear bounds might also be provable for Ramsey numbers of bounded degree k-uniform hypergraphs. Such a result was indeed established for general k in [30] (extending two earlier proofs for $k = 3$). A short proof of this result, not using regularity and thus giving much better bounds was obtained in [24]. It is based on the approach from [23, 51] used to prove Theorem 3.1.

3.2. Sparse graphs.

A graph is *d-degenerate* if every subgraph of it has a vertex of degree at most d. This notion nicely captures the concept of sparse graphs as every t-vertex subgraph of a d-degenerate graph has at most td edges. (Indeed, remove from the subgraph a vertex of minimum degree, and repeat this process in the remaining subgraph.) Notice that graphs with maximum degree d are d-degenerate. On the other hand, it is easy to construct a d-degenerate graph on n vertices whose maximum degree is linear in n. One of the most famous open problems in Graph Ramsey Theory is the following conjecture of Burr and Erdős [17] from 1975.

Conjecture 3.2. *For each positive integer d, there is a constant $c(d)$ such that $r(G) \leq c(d)n$ for every d-degenerate graph G on n vertices.*

This difficult conjecture is a substantial generalization of the results on Ramsey numbers of bounded degree graphs from Section 3.1 and progress on this problem was made only recently.

Kostochka and Rödl [73] gave a polynomial upper bound on the Ramsey numbers of d-degenerate graphs. The first nearly linear bound for this conjecture was obtained, by Kostochka and the author, in [74]. They proved that

d-degenerate graphs on n vertices satisfy $r(G) \leq c_d n^{1+\epsilon}$ for any fixed $\epsilon > 0$. The following is the best current estimate, which appeared in [52].

Theorem 3.3. *For each positive integer d there is a constant c_d such that every d-degenerate graph G with order n satisfies $r(G) \leq 2^{c_d \sqrt{\log n}}\, n$.*

In the past two decades Conjecture 3.2 was also proved for some special families of d-degenerate graphs (see, e.g., [2, 18, 85]). For example, we know that planar graphs and more generally graphs which can be drawn on a surface of bounded genus have linear Ramsey numbers. One very large and natural family of d-degenerate graphs are sparse random graphs. The *random graph* $G_{n,p}$ is the probability space of labeled graphs on n vertices, where every edge appears independently with probability p. When $p = d/n$ it is easy to show using standard large deviation estimates for binomial distribution that with high probability $G_{n,p}$ is $O(d)$-degenerate. Hence it is natural to test the above conjecture on random graphs. This was done in [52], where it was proved that sparse random graphs do indeed have typically linear Ramsey numbers.

3.3. Maximizing the Ramsey number. Another related problem on Ramsey numbers of general graphs was posed in 1973 by Erdős and Graham. Among all graphs with m edges, they wanted to find a graph G with maximum Ramsey number. Since the results we mentioned so far clearly show that sparse graphs have slowly growing Ramsey numbers, one would probably like to make such a G as dense as possible. Indeed, Erdős and Graham [40] conjectured that among all the graphs with $m = \binom{n}{2}$ edges (and no isolated vertices), the complete graph on n vertices has the largest Ramsey number. This conjecture is very difficult and so far there has been no progress on this problem. Because of the lack of progress, in the early 80s Erdős [35] (see also [20]) asked whether one could at least show that the Ramsey number of any graph with m edges is not substantially larger than that of the complete graph with the same size. Since the number of vertices in a complete graph with m edges is a constant multiple of $\sqrt{m}$, Erdős conjectured that $r(G) \leq 2^{c\sqrt{m}}$ for every graph G with m edges and no isolated vertices. The authors of [3] showed that for all graphs with m edges $r(G) \leq 2^{c\sqrt{m}\log m}$ and also proved this conjecture in the special case when G is bipartite. Recently, Erdős' conjecture was established in full generality in [95].

Theorem 3.4. *If G is a graph on m edges without isolated vertices, then $r(G) \leq 2^{250\sqrt{m}}$.*

This theorem is best possible up to a constant factor in the exponent, since a complete graph with m edges has Ramsey number at least $2^{\sqrt{m/2}}$. Based on the results from Section 3.1, it seems plausible that the following strengthening of Conjecture 3.2 holds as well. For all d-degenerate graphs G on n vertices, $r(G) \leq 2^{cd}n$. Such a bound would be a far-reaching generalization of the estimates on Ramsey numbers of bounded-degree graphs and also of Theorem 3.4. Indeed, it is easy to check that every graph with m edges is $\sqrt{2m}$-degenerate.

3.4. Methods. The result of Chvatál et al. [21] which gave the first linear bound on Ramsey numbers of bounded degree graphs (see Section 3.1), was proved using the regularity lemma. This is a surprising and extremely powerful result proved by Szemerédi that has numerous applications in various areas including combinatorial number theory, computational complexity, and mainly extremal graph theory. The regularity lemma was an essential tool in the proof of the celebrated theorem of Szemerédi that any dense subset of integers contains long arithmetic progressions. The precise statement of the lemma is somewhat technical and can be found in [70] together with the description of several of its famous applications.

Roughly this lemma states that the vertices of every large enough graph can be partitioned into a finite number of parts such that the edges between almost all of the parts behave like a random graph. The strength of the regularity lemma is that it applies to every graph and provides a good approximation of its structure which enables one to extract a lot of information about it. It is also known that there is an efficient algorithm for finding such a regular partition. Although the regularity lemma is a great tool for proving qualitative statements, the quantitative bounds which one usually gets from such proofs are rather weak. This is because the number of parts M in the partition of the graph given by the regularity lemma may be very large, more precisely of tower type. Moreover, Gowers [57] constructed examples of graphs for which M has to grow that fast. Therefore, to obtain good quantitative estimates, one should typically use a different approach.

One such approach was proposed by Graham, Rödl, and Rucinski [60] (see also [50] for some extensions). They noticed that in some applications, instead of having tight control on the distribution of edges (which the regularity lemma certainly gives), it is enough to satisfy a bi-density condition, i.e., to have a lower bound on the density of edges between any two sufficiently large disjoint sets. Using this observation one can show that in every red-blue edge coloring of K_N, either the red color satisfies a certain bi-density condition or there is a large set in which the proportion of blue edges is very close to 1. Then, for example, one can find a blue copy of any bounded-degree graph in this almost blue set. On the other hand, this approach is highly specific to the 2-color case and it would be of considerable interest to make it work for k colors.

Another basic tool used to prove several results mentioned in Sections 3.1-3.3 as well as some other recent striking results in extremal combinatorics is a simple and yet surprisingly powerful lemma, whose proof is probabilistic. Early variants of this lemma, have been proved and applied by various researchers starting with Rödl, Gowers, Kostochka and Sudakov (see [72], [58], [94]).

The lemma asserts, roughly, that every graph with sufficiently many edges contains a large subset U in which every set of d vertices has many common neighbors. The proof uses a process that may be called a *dependent random choice* for finding the set U; U is simply the set of all common neighbors of an appropriately chosen random set R. Intuitively, it is clear that if some set of

d vertices has only few common neighbors, it is unlikely all the members of R will be chosen among these neighbors. Hence, we do not expect U to contain any such subset of d vertices.

The main idea of this approach is that in the course of a probabilistic proof, it is often better not to make the choices uniformly at random, but to try and make them depend on each other in a way tailored to the specific argument needed. While this sounds somewhat vague, this simple reasoning and its various extensions have already found many applications to extremal graph theory, additive combinatorics, Ramsey theory and combinatorial geometry. For more information about this technique and its applications we refer the interested reader to the recent survey [53].

4. Turán Numbers

Extremal problems are at the heart of graph theory. These problems were extensively studied during the last half century. One of the central questions from which extremal graph theory originated can be described as follows. Given a *forbidden graph* H, determine $\mathrm{ex}(n, H)$, the maximal number of edges in a graph on n vertices that does not contain a copy of H. This number is also called the *Turán number of* H. Instances of this problem appear naturally in discrete geometry, additive number theory, probability, analysis, computer science and coding theory. In this section we describe classical results in this area, mention several applications and report on some recent progress on the problem of determining $\mathrm{ex}(n, H)$ for bipartite graphs.

4.1. Classical results. How dense can a graph G on n vertices be if it contains no triangles? One way to obtain such a graph is to split the vertices into two nearly equal parts A and B and to connect every vertex in A with every vertex in B by an edge. This graph clearly has no triangles and is also very dense. Moreover, it is maximal triangle-free graph, since adding any other edge to G creates a triangle. But is it the densest triangle-free graph on n vertices? More than a hundred years ago Mantel [78] proved that this is indeed the case and therefore $\mathrm{ex}(n, K_3) = \lfloor n^2/4 \rfloor$. This, earliest extremal result in graph theory already has an interesting application, found by Katona [65].

Consider $v_1, \ldots, v_n$, vectors in $\mathbb{R}^d$ of length $|v_i| \geq 1$. How many pairs of these vectors have sum of length less than 1? Suppose we have v_i, v_j, v_k such that all three pairwise sums have length less than 1. Then an easy computation shows that

$$|v_i + v_j + v_k|^2 = |v_i + v_j|^2 + |v_i + v_k|^2 + |v_j + v_k|^2 - |v_i|^2 - |v_j|^2 - |v_k|^2 < 0\,.$$

This contradiction together with Mantel's theorem shows that the number of pairs i, j with $|v_i + v_j| < 1$ is at most $\lfloor n^2/4 \rfloor$. Suppose now we have two independent identical copies X and Y of some arbitrary distribution with values in $\mathbb{R}^d$. By sampling many independent copies of this distribution and using the above claim on the vectors in $\mathbb{R}^d$ one can prove the following general inequality

$$\Pr\big[\,|X+Y|\geq 1\,\big] \geq \frac{1}{2}\big(\Pr[\,|X|\geq 1\,]\big)^2.$$

The starting point of extremal graph theory is generally considered to be the following celebrated theorem of Turán [99]. Partition n vertices into r parts $V_1, \ldots, V_r$ of nearly equal size, i.e., $\big||V_i|-|V_j|\big| \leq 1$. Let the *Turán graph* $T_{n,r}$ be the complete r-partite graph obtained by putting the edges between all the pairs of vertices in different parts. In 1941 Turán proved that the largest n-vertex graph, not containing a clique K_{r+1} is precisely $T_{n,r}$. In addition, he posed the problem of determining $\mathrm{ex}(n,H)$ for general graphs and also for hypergraphs (see Section 6).

A priori one might think that the answer to Turán's problem would be messy and that to deal with every particular graph might require each time a new approach. The important and deep theorem of Erdős and Stone [47] together with an observation of Erdős and Simonovits [46] shows that this is not the case. Their very surprising result says that for most graphs there is a single parameter, the chromatic number, which determines the asymptotic behavior of $\mathrm{ex}(n,H)$. The *chromatic number* of a graph H is the minimal number of colors needed to color the vertices of H such that adjacent vertices get different colors. Erdős, Stone and Simonovits proved that for a fixed H and large n

$$\mathrm{ex}(n,H) = \left(1-\frac{1}{\chi(H)-1}\right)\frac{n^2}{2} + o(n^2)\,.$$

A moment's thought shows that this determines the asymptotics of $\mathrm{ex}(n,H)$ for all graphs H with chromatic number at least 3. For example, if H is a graph formed by the edges of the icosahedron, then it is easy to check that the chromatic number of H is 4 and therefore $\mathrm{ex}(n,H) = (1+o(1))n^2/3$.

4.2. Bipartite graphs.

As we already mentioned, the theorem of Erdős, Stone and Simonovits determines asymptotically $\mathrm{ex}(n,H)$ for all graphs with chromatic number at least 3. However, for bipartite graphs it only gives $\mathrm{ex}(n,H) = o(n^2)$. The determination of Turán numbers for bipartite graphs remains a challenging project with many open problems. In fact, even the order of magnitude of $\mathrm{ex}(n,H)$ is not known for quite simple bipartite graphs, such as the complete bipartite graph with four vertices in each part, the cycle of length eight, and the 3-cube graph. Here we describe some partial results obtained so far, which use a variety of techniques from different fields including probability, number theory and algebraic geometry.

Let $t \leq s$ be positive integers and let $K_{t,s}$ denote the complete bipartite graph with parts of size t and s. For every fixed t and $s \geq t$, Kővári, Sós and Turán [75] proved, more than 60 years ago, that

$$\mathrm{ex}(n,K_{t,s}) \leq \frac{1}{2}(s-1)^{1/t}n^{2-1/t} + \frac{1}{2}(t-1)n\,.$$

It is conjectured that the right hand side gives the correct order of magnitude of $\mathrm{ex}(n, K_{t,s})$. However, progress on this problem was slow and despite several results by various researchers this is known only for $s > (t-1)!$ (see [4] and its references). In particular, in the most interesting case $s = t$ the Turán number of $K_{4,4}$ is already unknown. All constructions for this problem are algebraic and the more recent ones require some tools from elementary algebraic geometry.

The Turán numbers for $K_{t,s}$ appear naturally in problems in other areas of mathematics. For example, in 1946, Erdős [31] asked to determine the maximum possible number of unit distances among n points on the plane. One might think that potentially the number of such distances may be even quadratic. Given such a set of n points, consider a graph whose vertices are the points and two of them are adjacent if the distance between them is one. Since on the plane for any two fixed points p and p' there are precisely two other points whose distance to both p, p' is one, the resulting graph has no $K_{2,3}$. Therefore, by the above result there are at most $O(n^{3/2})$ unit distances. Erdős conjectured that the number of such distances is always at most $n^{1+o(1)}$, but the best current bound for this problem, obtained in [93], is $O(n^{4/3})$.

Suppose we have a set of integers A such that $A + A = \{a + a' \mid a, a' \in A\}$ contains all numbers $1^2, 2^2, \ldots, n^2$. How small can the set A be? This is a special case of the question asked by Wooley [100] at the AIM conference on additive combinatorics in 2004. Clearly, A has size at least $\sqrt{n}$ but the truth is probably $n^{1-o(1)}$. It appears that Erdős and Newman [44] already considered this problem earlier and noticed that using extremal graph theory one can show that $|A| \geq n^{2/3-o(1)}$. Consider a graph whose vertices are elements of A and for every $1 \leq x \leq n$ choose some pair a, a' such that $x^2 = a + a'$ and connect them by an edge. Erdős and Newman use bounds on $\mathrm{ex}(n, K_{2,s})$ to conclude that if $|A| = n^{2/3-\epsilon}$ then this graph must contain two vertices a_1 and a_2 with at least n^δ common neighbors. Thus one can show that $a_1 - a_2$ can be written as a difference of two squares in n^δ different ways and therefore will have too many divisors, a contradiction.

Not much is known for Turán numbers of general bipartite graphs. Moreover, we do not even have a good guess what parameter of a bipartite graph might determine the order of growth of its Turán number. Some partial answers to this question were proposed by Erdős. Recall that a graph is t-degenerate if its every subgraph contains a vertex of degree at most t. In 1966 Erdős [33] (see also [20]) conjectured that every t-degenerate bipartite graph H satisfies $\mathrm{ex}(n, H) \leq O\big(n^{2-1/t}\big)$. Recently, progress on this conjecture was obtained in [3]. One of the results in this paper says that the conjecture holds for every bipartite graph H in which the degrees of all vertices in one part are at most t. This result, which can be also derived from an earlier result of Füredi [55], is a far reaching generalization of the above estimate of Kővári, Sós and Turán. It is tight for every fixed t as was shown, e.g., by constructions in [4]. Another result in [3] gives the first known estimate on the Turán numbers of degenerate bipartite graphs.

Theorem 4.1. *Let H be a bipartite t-degenerate graph on h vertices. Then for all $n \geq h$*

$$ex(n, H) \leq h^{1/2t} n^{2-\frac{1}{4t}} .$$

The proof of this theorem and also of the first result from [3] mentioned above is based on the dependent random choice approach, which we briefly discussed in Section 3.4.

4.3. Subgraph multiplicity.

Turán's theorem says that any graph with $m > \left(1 - \frac{1}{r}\right)\frac{n^2}{2}$ edges contains at least one copy of K_{r+1}. The question of how many such copies $f_r(m, n)$ must exist in an n-vertex graph with m edges received quite a lot of attention and has turned out to be notoriously difficult. When m is very close to the $\mathrm{ex}(n, K_{r+1})$ this function was computed by Erdős. Let $m = p\binom{n}{2}$, where the edge density p (the fraction of the pairs which are edges) is a fixed constant strictly greater than $1 - 1/r$. One very interesting open question is to determine the asymptotic behavior of $f_r(m, n)$ as a function of p only. Further results in this direction were obtained by Goodman, Lovász, Simonovits, Bollobás, and Fisher (for more details see [82, 80] and their references). Recently Razborov [82] and Nikiforov [80] resolved this problem for the cases $r = 2$ and $r = 3$, respectively. It appears that in these cases the solution corresponds to the complete $(t + 1)$-partite graph in which t parts are roughly equal and are larger than the remaining part, and the integer t is such that $p \in \left[1 - \frac{1}{t}, 1 - \frac{1}{t+1}\right]$.

For bipartite graphs the situation seems to be very different. The beautiful conjectures of Erdős and Simonovits [90] and of Sidorenko [89] suggest that for any bipartite H there is $\gamma(H) > 0$ such that the number of copies of H in any graph G on n vertices and edge density $p > n^{-\gamma(H)}$ is asymptotically at least the same as in the n-vertex random graph with edge density p. The original formulation of the conjecture by Sidorenko is in terms of graph homomorphisms. A homomorphism from a graph H to a graph G is a mapping $f : V(H) \to V(G)$ such that for each edge (u, v) of H, $(f(u), f(v))$ is an edge of G. Let $h_H(G)$ denote the number of homomorphisms from H to G. We also consider the normalized function $t_H(G) = h_H(G)/|G|^{|H|}$, which is the fraction of mappings $f : V(H) \to V(G)$ which are homomorphisms. Sidorenko's conjecture states that for every bipartite graph H with q edges and every graph G,

$$t_H(G) \geq t_{K_2}(G)^q.$$

This conjecture also has the following appealing analytical form. Let μ be the Lebesgue measure on $[0, 1]$ and let $h(x, y)$ be a bounded, symmetric, nonnegative and measurable function on $[0, 1]^2$. Let H be a bipartite graph with vertices $u_1, \ldots, u_t$ in the first part and vertices $v_1, \ldots, v_s$ in the second part. Denote by E the set of edges of H, i.e., all the pairs (i, j) such that u_i and v_j are adjacent, and let $|E| = q$.

Conjecture 4.2.

$$\int \prod_{(i,j)\in E} h(x_i, y_j) d\mu^{s+t} \geq \left(\int h d\mu^2\right)^q .$$

The expression on the left hand side of this inequality is quite common. Such integrals are called Feynman integrals in quantum field theory and they also appear in classical statistical mechanics. Unsurprisingly then, Sidorenko's conjecture has connections to a broad range of topics, such as matrix theory, Markov chains, graph limits and quasirandomness. So far this conjecture was established only in very special cases, e.g., for complete bipartite graphs, trees, even cycles (see [89]), and also for cubes [64].

Recently, Sidorenko's conjecture was proved for a new class of graphs. In [28], it was shown that the conjecture holds for every bipartite graph H which has a vertex adjacent to all the vertices in the other part. Using this result, one can easily deduce an approximate version of Sidorenko's conjecture for all graphs. For a connected bipartite graph H with parts V_1, V_2, define the bipartite graph $\bar{H}$ with parts V_1, V_2 such that $(v_1, v_2) \in V_1 \times V_2$ is an edge of $\bar{H}$ if and only if it is not an edge of H. Define the *width* of H to be the minimum degree of $\bar{H}$. If H is not connected, the width of H is the sum of the widths of the connected components of H. Note that the width of a connected bipartite graph is 0 if and only if it has a vertex that is complete to the other part. Moreover, the width of a bipartite graph with h vertices is always at most $h/2$.

Theorem 4.3. *If H is a bipartite graph with q edges and width w, then $t_H(G) \geq t_{K_2}(G)^{q+w}$ holds for every graph G.*

5. Generalizations

Turán's theorem, which determines the maximum number of edges in a K_{r+1}-free graph on n vertices, is probably the most famous result in extremal combinatorics and there are many interesting generalizations and extensions of this theorem. In this section we discuss several such results.

5.1. Local density. A generalization of Turán's theorem that takes into account edge distribution, or local density, was introduced by Erdős [34] in 1975. He asked the following question. Suppose that G is a K_{r+1}-free graph on n vertices in which every set of αn vertices spans at least βn^2 edges for some $0 \leq \alpha, \beta \leq 1$. How large can β be as a function of α? Erdős, Faudree, Rousseau and Schelp [39] studied this problem and conjectured that for α sufficiently close to 1 the Turán graph $T_{n,r}$ has the highest local density. They proved this for triangle-free graphs ($r = 2$) and the general case of this conjecture was established in [66]. It is easy to check that for $\alpha \geq \frac{r-1}{r}$ every subset of $T_{n,r}$ of size αn contains at least $\frac{r-1}{2r}(2\alpha - 1)n^2$ edges. The result in [66] says that if G

is a K_{r+1}-free graph on n vertices and $1 - \frac{1}{2r^2} \leq \alpha \leq 1$, then G contains a set of αn vertices spanning at most $\frac{r-1}{2r}(2\alpha - 1)n^2$ edges and equality holds only when G is a Turán graph.

For triangle-free graphs and general α it was conjectured in [39] that β is determined by a family of extremal triangle-free graphs. Besides the complete bipartite graph $T_{n,2}$ already mentioned, another important graph is $C_{n,5}$, which is obtained from a 5-cycle by replacing each vertex i by an independent set V_i of size $n/5$ (assuming for simplicity that n is divisible by 5), and each edge ij by a complete bipartite graph joining V_i and V_j (this operation is called a 'blow-up'). Erdős et al. conjectured that for α above 17/30 the Turán graph has the highest local density and for $1/2 \leq \alpha \leq 17/30$ the best graph is $C_{n,5}$. On the other hand, for $r \geq 3$ Chung and Graham [19] conjectured that the Turán graph has the best local density even for α as low as 1/2. When α is a small constant the situation is unclear and there are no natural conjectures.

The case $r = 2$ and $\alpha = 1/2$ is one of the favorite questions of Erdős that he returned to often and offered a \$250 prize for its solution. Here the conjecture is that any triangle-free graph on n vertices should contain a set of $n/2$ vertices that spans at most $n^2/50$ edges. This conjecture is one of several important questions in extremal graph theory where the optimal graph is suspected to be the blow-up of the 5-cycle $C_{n,5}$. So far these problems are completely open and we need new techniques to handle them.

Another question, that is similar in spirit, is to determine how many edges one may need to delete from a K_{r+1}-free graph on n vertices in order to make it bipartite. This is an instance of the well known Max-Cut problem, which asks for the largest bipartite subgraph of a given graph G. This problem has been the subject of extensive research both from the algorithmic perspective in computer science and the extremal perspective in combinatorics.

A long-standing conjecture of Erdős [34] says that one needs to delete at most $n^2/25$ edges from a triangle-free graph to make it bipartite, and $C_{n,5}$ shows that this estimate would be the best possible. This problem is still open and the best known bound is $(1/18 - \epsilon)n^2$ for some constant $\epsilon > 0$, obtained by Erdős, Faudree, Pach and Spencer [38]. Erdős also conjectured that for K_4-free graphs on n vertices the answer for this problem is at most $n^2/9$. This was recently proved in [96].

Theorem 5.1. *Every K_4-free graph on n vertices can be made bipartite by deleting at most $n^2/9$ edges. Moreover, the only extremal graph which requires deletion of so many edges is the Turán graph $T_{n,3}$.*

It is also plausible to conjecture that, for all $r > 3$, the K_{r+1}-free n-vertex graph that requires the most edge deletions in order to make it bipartite is the Turán graph $T_{n,r}$.

It was observed in [76] that for regular graphs, a bound for the local density problem implies a related bound for the problem of making the graph bipartite. Indeed, suppose n is even, G is a d-regular graph on n vertices and S is a set of $n/2$ vertices. Then $dn/2 = \sum_{s\in S} d(s) = 2e(S) + e(S,\bar{S})$ and $dn/2 = \sum_{s\notin S} d(s) = 2e(\bar{S}) + e(S,\bar{S})$. This implies that $e(S) = e(\bar{S})$, i.e., S and $\bar{S}$ span the same number of edges. Deleting the $2e(S)$ edges within S and $\bar{S}$ makes the graph bipartite. Thus, for example, if in a regular triangle-free graph G one can find a set S with $|S| = n/2$ which spans at most $n^2/50$ edges, then G can be made bipartite by deleting at most $n^2/25$ edges. This relation, together with Theorem 5.1, gives some evidence that indeed for $r \geq 3$ the Turán graph should have the best local density for all $0 \leq \alpha \leq 1$.

5.2. Graphs with large minimum degree. Clearly, for any graph G, the largest K_{r+1}-free subgraph of G has at least as many edges as does the largest r-partite subgraph. For which graphs do we have an equality? This question was raised by Erdős [35], who noted that by Turán's theorem there is an equality for the complete graph K_n. In [7] and [16] it was shown that the equality holds with high probability for sufficiently dense random graphs. Recently, a general criteria implying equality was obtained in [5], where it is proved that a minimum degree condition is sufficient. Given a fixed graph H and a graph G let $e_r(G)$ and $e_H(G)$ denote the number of edges in the largest r-partite and the largest H-free subgraphs of G, respectively. The following theorem shows that both Turán's and Erdős-Stone-Simonovits' theorems holds not only for K_n but also for any graph with large minimum degree.

Theorem 5.2. *Let H be a graph with chromatic number $r+1 \geq 3$. Then there are constants $\gamma = \gamma(H) > 0$ and $\mu = \mu(H) > 0$ such that if G is a graph on n vertices with minimum degree at least $(1-\mu)n$, then*

$$e_r(G) \leq e_H(G) \leq e_r(G) + O(n^{2-\gamma})\,.$$

Moreover, if $H = K_{r+1}$ then $e_H(G) = e_r(G)$.

The assertion of this theorem for the special case when H is a triangle is proved in [14] and in a stronger form in [8].

As well as being interesting in its own right, this theorem was motivated by the following question in computer science. Given some *property* $\mathcal{P}$ and a graph G, it is a fundamental computational problem to find the smallest number of edge deletions and additions needed to turn G into a graph satisfying this property. We denote this quantity by $E_{\mathcal{P}}(G)$. Specific instances of graph modification problems arise naturally in several fields, including molecular biology and numerical algebra. A graph property is *monotone* if it is closed under removal of vertices and edges. Note that, when trying to turn a graph into one satisfying a monotone property, we will only need to use edge deletions and therefore in these cases the problem is called an edge-deletion problem. Two examples of interesting monotone properties are k-colorability and the property of not containing a copy of a fixed graph H. It appears that using combinatorial methods it is possible to give a nearly complete answer to the question of how accurately one can approximate (up to additive error) the solution of the edge-deletion problem for monotone properties.

For any fixed $\epsilon > 0$ and any monotone property $\mathcal{P}$ there is a deterministic algorithm, obtained in [5], which, given a graph G on n vertices, approximates $E_{\mathcal{P}}(G)$ within an additive error of ϵn^2 (i.e., it computes a number X such that $|X - E_{\mathcal{P}}(G)| \leq \epsilon n^2$). Moreover, the running time of the algorithm is linear in the size of the graph. This algorithm uses a strengthening of Szemerédi's regularity lemma which implies that every graph G can be approximated by a small (fixed size) weighted graph W, so that $E_{\mathcal{P}}(G)$ is an approximate solution of a related problem on W. Since W has a fixed size, we can now resort to a brute force solution. Given the above, a natural question is for which monotone properties one can obtain better additive approximations of $E_{\mathcal{P}}$. Another result in [5] essentially resolves this problem by giving a precise characterization of the monotone graph properties for which such approximations exist.

On the one hand, if there is a bipartite graph that does not satisfy property $\mathcal{P}$, then there is a $\delta > 0$ for which it is possible to approximate $E_{\mathcal{P}}$ within an additive error of $n^{2-\delta}$ in polynomial time. On the other hand, if all bipartite graphs satisfy $\mathcal{P}$, then for any $\delta > 0$ it is NP-hard to approximate distance to $\mathcal{P}$ within an additive error of $n^{2-\delta}$. The proof of this result, among the other tools, uses Theorem 5.2 together with spectral techniques. Interestingly, prior to [5], it was not even known that computing $E_{\mathcal{P}}$ *precisely* for most properties satisfied by all bipartite graphs (e.g., being triangle-free) is NP-hard. It thus answers (in a strong form) a question of Yannakakis, who asked in 1981 if it is possible to find a large and natural family of graph properties for which computing $E_{\mathcal{P}}$ is NP-hard.

5.3. Spectral Turán theorem. Given an arbitrary graph G, consider a partition of its vertices into r parts which maximizes the number of edges between the parts. Then the degree of each vertex within its own part is at most $1/r$-times its degree in G, since otherwise we can move this vertex to some other part and increase the total number of edges connecting different parts. This simple construction shows that the largest r-partite and hence also largest K_{r+1}-free subgraph of G has at least a $\frac{r-1}{r}$-fraction of its edges. We say that a graph G (or rather a family of graphs) is *r-Turán* if this trivial lower bound is essentially an upper bound as well, i.e., the largest K_{r+1}-free subgraph of G has at most $(1 + o(1))\frac{r-1}{r}|E(G)|$ edges. Note that Turán's theorem says that this holds when G is a complete graph on n vertices. Thus it is very natural to ask, which other graphs are r-Turán?

It has been shown that for any fixed r, there exists $p(r, n)$ such that for all $p \gg p(r, n)$ with high probability the random graph $G_{n,p}$ is r-Turán. The value of p for which this result holds was improved several times by various researches. Recently, resolving the longstanding conjecture, Conlon and Gowers [29] and independently Schacht [87] established the optimal value of $p(r, n) = n^{-2/(r+2)}$. These results about random graphs do not yet provide a deterministic sufficient condition for a graph to be r-Turán. However, they suggest that one should look at graphs whose edges are distributed sufficiently evenly. It turns out that under certain circumstances, such an edge distribution can be guaranteed by a simple assumption about the spectrum of the graph.

For a graph G, let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the eigenvalues of its adjacency matrix. The quantity $\lambda(G) = \max\{\lambda_2, -\lambda_n\}$ is called the *second eigenvalue* of G. A graph $G = (V, E)$ is called an (n, d, λ)*-graph* if it is d-regular, has n vertices and the second eigenvalue of G is at most λ. It is well known (see [6, 77] for more details) that if λ is much smaller than the degree d, then G has certain random-like properties. Thus, λ could serve as some kind of "measure of randomness" in G. The following recent result from [97] shows that Turán's theorem holds asymptotically for graphs with small second eigenvalue.

Theorem 5.3. *Let $r \geq 2$ be an integer and let $G = (V, E)$ be an (n, d, λ)-graph. If $d^r/n^{r-1} \gg \lambda$ then the largest K_{r+1}-free subgraph of G has at most $(1 + o(1))\frac{r-1}{r}|E(G)|$ edges*

This result generalizes Turán's theorem, since the second eigenvalue of the complete graph K_n is 1 and thus it satisfies the above condition. Theorem 5.3 is also part of the fast-growing comprehensive study of graph theoretical properties of (n, d, λ)-graphs, which has recently attracted lots of attention both in combinatorics and theoretical computer science. For a recent survey about these fascinating graphs and their properties, we refer the interested reader to [77].

6. Turán-type Problems for Hypergraphs

Given a k-uniform hypergraph H, the Turán number $\mathrm{ex}(n, H)$ is the maximum number of edges in a k-uniform hypergraph on n vertices that does not contain a copy of H. Determining these numbers is one of the main challenges in extremal combinatorics. For ordinary graphs (the case $k = 2$), a rich theory has been developed, whose highlights we described in Section 4. In 1941, Turán also posed the question of finding $\mathrm{ex}(n, K_s^{(k)})$ for complete k-uniform hypergraphs with $s > k > 2$ vertices, but to this day not one single instance of this problem has been solved. It seems very hard even to determine the *Turán density*, which is defined as $\pi(H) = \lim_{n\to\infty} \mathrm{ex}(n, H)/\binom{n}{k}$.

The most famous problem in this area is the conjecture of Turán that $\mathrm{ex}(n, K_4^{(3)})$ is given by the following construction, which we denote by T_n. Partition n vertices into 3 sets V_0, V_1, V_2 of equal size. Consider all triples which either intersect all these sets or contain two vertices in V_i and one in $V_{i+1 (\mathrm{mod}\ 3)}$. This hypergraph has density 5/9 and every 4 vertices span at most 3 edges. In memory of Turán, Erdős offered \$1000 for proving that $\pi(K_4^{(3)}) = 5/9$. Despite several results giving rather close estimates for the Turán density of $K_4^{(3)}$, this problem remains open. One of the main difficulties is that Turán's conjecture, if it is true, has exponentially many non-isomorphic extremal configurations (see [71]).

Recently the problem of finding the numbers $\mathrm{ex}(n, H)$ got a lot of attention and these numbers were determined for various hypergraphs. One such example is the Fano plane $PG_2(2)$, which is the projective plane over the field with 2 elements. It is the unique 3-uniform hypergraph with 7 vertices and 7 edges, in which every pair of vertices is contained in a unique triple and triples corresponds to the lines in the projective plane. A hypergraph is *2-colorable* if its vertices can be labeled as red or blue so that no edge is monochromatic. It is easy to check that the Fano plane is not 2-colorable, and therefore any 2-colorable hypergraph cannot contain the Fano plane. Partition an n-element set into two almost equal parts, and take all the triples that intersect both of them. This is clearly the largest 2-colorable 3-uniform hypergraph on n vertices. In 1976 Sós conjectured that this construction gives the exact value of $\mathrm{ex}\big(n, PG_2(2)\big)$. This was proved independently in [67] and [56], where it was also shown that the extremal construction, which we described above, is unique.

The strategy of the proof is first to obtain an approximate structure theorem, and then to show that any imperfection in the structure leads to a suboptimal configuration. This is the so-called "stability approach" which was first introduced for graphs by Simonovits. Following the above two papers, this approach has become a standard tool for attacking extremal problems for hypergraphs as well, and was used successfully to determine several hypergraph Turán numbers.

Let $\mathcal{C}_r^{(2k)}$ be the $2k$-uniform hypergraph obtained by letting $P_1, \cdots, P_r$ be pairwise disjoint sets of size k and taking as edges all sets $P_i \cup P_j$ with $i \neq j$. This can be thought of as the 'k-expansion' of the complete graph K_r: each vertex has been replaced with a set of size k. The Turán problem for $\mathcal{C}_r^{(2k)}$ was first considered by Frankl and Sidorenko, as a possible generalization of Turán's theorem for graphs. Using a clever reduction of this problem to the case of graphs they showed that the Turán density of $\mathcal{C}_r^{(2k)}$ is at most $\frac{r-2}{r-1}$. Frankl and Sidorenko also gave a matching lower bound construction, which was essentially algebraic but existed only when $r = 2^a + 1$. In [68], among other results, it was shown that, surprisingly, when r is not of the form $2^a + 1$ then the Turán density of $\mathcal{C}_r^{(4)}$ is strictly smaller than $\frac{r-2}{r-1}$. Interestingly, this result, showing that certain constructions do not exist, also uses a stability argument. By studying the properties of a $\mathcal{C}_r^{(4)}$-free hypergraph with density close to $\frac{r-2}{r-1}$ the authors show that it gives rise to an edge coloring of the complete graph K_{r-1} with special properties. Next they show that for such an edge-coloring there is a natural $GF(2)$ vector space structure on the colors. Of course, such a space has cardinality 2^a, for some integer a, so one gets a contradiction unless $r = 2^a + 1$.

It is interesting to note that T_n, the conjectured extremal example for $K_4^{(3)}$, also does not contain 4 vertices which span a single edge. Thus, there is a 3-uniform hypergraph with edge density 5/9 in which every 4 vertices span either zero or two edges. In [83], Razborov showed that 5/9 is the maximum possible density for such a hypergraph. Combining his result with the stability approach (described above) Pikhurko proved that the unique extremal configuration for this problem is T_n. Razborov's proof uses the formalism of flag algebras, which, roughly speaking, allows one to computerize the search for inequalities which should be satisfied by various statistics of the extremal hypergraph. Then the "right inequalities" can be proved using Cauchy-Schwarz type arguments. This approach works for various other extremal problems as well. For example, one can use it to improve the best known bounds for Turán's original conjecture (see [83]).

6.1. Hypergraphs and arithmetic progressions. Extremal problems for hypergraphs have many connections to other areas of mathematics. Here we describe one striking application of hypergraphs to number theory.

An old question of Brown, Erdős and Sós asks to determine the maximum number of edges in the k-uniform hypergraph which has no s edges whose union has at most t vertices. This is a very difficult question which is solved only for few specific values of parameters. One such special case is the so-called $(6,3)$-problem. Here, one wants to maximize the number of edges in a 3-uniform hypergraph such that every 6 vertices span at most 2 edges. In 1976, Ruzsa and Szemerédi [86] proved that such a hypergraph can have only $o(n^2)$ edges. Surprisingly, this purely combinatorial result has a tight connection with number theory. Using it one can give a short proof of the well-known theorem of Roth that every $A \subset [n]$ of size ϵn (for constant ϵ and large n) contains a 3-term arithmetic progression. Indeed, consider a 3-uniform hypergraph whose vertex set is the disjoint union of $[n], [2n]$ and $[3n]$ and whose edges are all the triples $x, x+a, x+2a$ with $x \in [n]$ and $a \in A$. This hypergraph has $O(n)$ vertices, $n|A|$ edges and, one can check that every 3-term arithmetic progression in A corresponds to 6 vertices spanning at least 3 edges and vice versa.

The $(6,3)$-theorem of Ruzsa and Szemerédi is closely related to the triangle removal lemma, which says that for every ϵ there is a δ such that every graph on n vertices with at most δn^3 triangles can be made triangle-free by removing ϵn^2 edges. The original proof of both results used the regularity lemma and therefore gave a very poor dependence of δ on ϵ. Very recently, this result was substantially improved by Fox [49]. Still, the dependence of δ on ϵ in [49] is of tower-type and compared with the Fourier-analytical approach it gives much weaker bounds for the number-theoretic applications.

A remarkable extension of the triangle removal lemma to hypergraphs was obtained by Gowers [59] and independently by Nagle, Rödl, Schacht and Skokan [84, 79]. They proved that if a k-uniform hypergraph on n vertices has at most δn^{k+1} copies of the complete hypergraph $K^{(k)}_{k+1}$, then all these copies can be destroyed by removing ϵn^k edges. This result was obtained by developing a new, very useful and important tool: the hypergraph analogue of the regularity lemma. The hypergraph removal lemma can be used to give a short proof of Szemerédi's theorem that dense subsets of integers contain long arithmetic progressions (see [91]).

7. Conclusion

We mentioned several specific problems of extremal combinatorics throughout this paper. Many of them are of a fundamental nature, and we believe that any progress on these questions will require the development of new techniques which will have wide applicability. We also gave examples of connections between extremal combinatorics and other areas of mathematics. In the future it is safe to predict that the number of such examples will only grow. Combinatorics will employ more and more advanced tools from algebra, topology, analysis and geometry and, on the other hand, there will be more applications of purely combinatorial techniques to non-combinatorial problems. One spectacular instance of such an interaction is a series of recent results on approximate subgroups and expansion properties of linear groups which combine combinatorics, number theory, algebra and model theory (see, e.g., [15] and its references).

The open problems which we mentioned, as well as many more additional ones which we skipped due to the lack of space, will provide interesting challenges for future research in extremal combinatorics. These challenges, the fundamental nature of the area and its tight connection with other mathematical disciplines will ensure that in the future extremal combinatorics will continue to play an essential role in the development of mathematics.

Acknowledgments

I would like to thank Conlon, Fox, Keevash, Krivelevich and Markman for carefully reading this manuscript and many valuable comments, which greatly improved the presentation.

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi, A note on Ramsey numbers, *J. Combinatorial Theory, Ser. A* **29** (1980), 354–360.

[2] N. Alon, Subdivided graphs have linear Ramsey numbers, *J. Graph Theory* **18** (1994), 343–347.

[3] N. Alon, M. Krivelevich, B. Sudakov, Turán numbers of bipartite graphs and related Ramsey-type questions, *Combin. Probab. Comput.* **12** (2003), 477–494.

[4] N. Alon, L. Rónyai and T. Szabó, Norm-graphs: variations and applications, *J. Combinatorial Theory, Ser. B* **76** (1999), 280–290.

[5] N. Alon, A. Shapira and B. Sudakov, Additive approximation for edge-deletion problems, *Annals of Mathematics* **170** (2009), 371–411.

[6] N. Alon and J. Spencer, **The probabilistic method**, 3^{rd} Ed., Wiley, New York, 2008.

[7] L. Babai, M. Simonovits and J. Spencer, Extremal subgraphs of random graphs, *J. Graph Theory* **14** (1990), 599–622.

[8] J. Balogh, P. Keevash and B. Sudakov, On the minimal degree implying equality of the largest triangle-free and bipartite subgraphs, *J. Combinatorial Theory Ser. B* **96** (2006), 919–932.

[9] B. Barak, G. Kindler, R. Shaltiel, B. Sudakov and A. Wigderson, Simulating Independence: New Constructions of Condensers, Ramsey Graphs, Dispersers and Extractors, *Proc. of the* 37^{th} *Symposium on Theory of Computing (STOC)*, ACM (2005), 1–10.

[10] B. Barak, A. Rao, R. Shaltiel and A. Wigderson, 2-source dispersers for subpolynomial entropy and Ramsey graphs beating the Frankl-Wilson construction, *Proc.* 38^{th} *Symposium on Theory of Computing (STOC)*, ACM (2006), 671–680.

[11] J. Beck, An upper bound for diagonal Ramsey numbers, *Studia Sci. Math. Hungar* **18** (1983), 401–406.

[12] T. Bohman, The Triangle-Free Process, *Advances in Mathematics* **221** (2009), 1653–1677.

[13] T. Bohman and P. Keevash, The early evolution of the H-free process, preprint.

[14] J. Bondy, J. Shen, S. Thomassé and C. Thomassen, Density conditions implying triangles in k-partite graphs, *Combinatorica* **26** (2006), 121–131.

[15] E. Breuillard, B. Green and T. Tao, Linear Approximate Groups, preprint.

[16] G. Brightwell, K. Panagiotou and A. Steger, On extremal subgraphs of random graphs, *Proc. of the* 18^{th} *Symposium on Discrete Algorithms (SODA)*, ACM-SIAM (2007), 477–485.

[17] S. A. Burr and P. Erdős, On the magnitude of generalized Ramsey numbers for graphs, in: *Infinite and Finite Sets I*, 10, Colloq. Math. Soc. Janos Bolyai, North-Holland, Amsterdam, 1975, 214–240.

[18] G. Chen and R. H. Schelp, Graphs with linearly bounded Ramsey numbers, *J. Combin. Theory Ser. B* **57** (1993), 138–149.

[19] F. Chung and R. Graham, On graphs not containing prescribed induced subgraphs, in: *A tribute to Paul Erdős*, Cambridge Univ. Press, Cambridge, 1990, 111–120.

[20] F. Chung and R. Graham, **Erdős on Graphs. His Legacy of Unsolved Problems**, A K Peters, Ltd., Wellesley, MA, 1998.

[21] V. Chvátal, V. Rödl, E. Szemerédi, and W. T. Trotter, Jr., The Ramsey number of a graph with bounded maximum degree, *J. Combin. Theory Ser. B* **34** (1983), 239–243.

[22] D. Conlon, A new upper bound for diagonal Ramsey numbers, *Annals of Mathematics* **170** (2009), 941–960.

[23] D. Conlon, Hypergraph packing and sparse bipartite Ramsey numbers, *Combin. Probab. Comput.,* **18** (2009), 913–923.

[24] D. Conlon, J. Fox and B. Sudakov, Ramsey numbers of sparse hypergraphs, *Random Structures Algorithms* **35** (2009), 1–14.

[25] D. Conlon, J. Fox and B. Sudakov, Hypergraph Ramsey numbers, *J. Amer. Math. Soc.* **23** (2010), 247–266.

[26] D. Conlon, J. Fox and B. Sudakov, Large almost monochromatic subsets in hypergraphs, *Israel Journal of Mathematics*, to appear.

[27] D. Conlon, J. Fox and B. Sudakov, On two problems in graph Ramsey theory, submitted.

[28] D. Conlon, J. Fox and B. Sudakov, An approximate version of Sidorenko's conjecture, preprint.

[29] D. Conlon and T. Gowers, Combinatorial theorems relative to a random set, preprint.

[30] O. Cooley, N. Fountoulakis, D. Kühn and D. Osthus, Embeddings and Ramsey numbers of sparse k-uniform hypergraphs, *Combinatorica* **28** (2009), 263–297.

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

[32] P. Erdős, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* **53** (1947), 292–294.

[33] P. Erdős, Some recent results on extremal problems in graph theory, in: em Theory of Graphs (Rome, 1966), Gordon and Breach, New York, (1967), 117-123.

[34] P. Erdős, Problems and results in graph theory and combinatorial analysis. Proceedings of the Fifth British Combinatorial Conference (Univ. Aberdeen, Aberdeen, 1975), pp. 169–192. *Congressus Numerantium*, No. XV, Utilitas Math., Winnipeg, Man., 1976.

[35] P. Erdős, On some problems in graph theory, combinatorial analysis and combinatorial number theory, in: *Graph theory and combinatorics (Cambridge, 1983)*, Academic Press, London, 1984, 1–17.

[36] P. Erdős, Problems and results on graphs and hypergraphs: similarities and differences, in *Mathematics of Ramsey theory,* Algorithms Combin., Vol. 5 (J. Nešetřil and V. Rödl, eds.) 12–28. Berlin: Springer-Verlag, 1990.

[37] P. Erdős, Problems and results in discrete mathematics, *Discrete Math.* **136** (1994), 53–73.

[38] P. Erdős, R. Faudree, J. Pach and J. Spencer, How to make a graph bipartite, *J. Combin. Theory Ser. B* **45** (1988), 86–98.

[39] P. Erdős, R. Faudree, C. Rousseau and R. Schelp, A local density condition for triangles, *Discrete Math.* **127** (1994), 153–161.

[40] P. Erdős and R. Graham, On partition theorems for finite graphs, in *Infinite and finite sets (Colloq., Keszthely, 1973)*, Vol. I; Colloq. Math. Soc. János Bolyai, Vol. 10, North-Holland, Amsterdam, 1975, 515–527.

[41] P. Erdős and A. Hajnal, On Ramsey like theorems, Problems and results, Combinatorics (Proc. Conf. Combinatorial Math., Math. Inst., Oxford, 1972) , pp. 123–140, Inst. Math. Appl., Southend-on-Sea, 1972.

[42] P. Erdős and A. Hajnal, Ramsey-type theorems, *Discrete Appl. Math.* **25** (1989), 37–52.

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

[44] P. Erdős and D.J. Newman, Bases for sets of integers, *J. Number Theory* **9** (1977), 420–425.

[45] P. Erdős and R. Rado, Combinatorial theorems on classifications of subsets of a given set, *Proc. London Math. Soc.* **3** (1952), 417–439.

[46] P. Erdős and M. Simonovits, A limit theorem in graph theory, em Studia Sci. Math. Hungar **1** (1966), 51–57.

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

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

[49] J. Fox, A new proof of the triangle removal lemma, preprint.

[50] J. Fox and B. Sudakov, Induced Ramsey-type theorems, *Advances in Mathematics* **219** (2008), 1771–1800.

[51] J. Fox and B. Sudakov, Density theorems for bipartite graphs and related Ramsey-type results, *Combinatorica* **29** (2009), 153-196.

[52] J. Fox and B. Sudakov, Two remarks on the Burr-Erdős conjecture, *European J. Combinatorics,* **30** (2009), 1630–1645.

[53] J. Fox and B. Sudakov, Dependent Random Choice, *Random Structures and Algorithms*, to appear.

[54] P. Frankl and R. Wilson, Intersection theorems with geometric consequences, *Combinatorica* **1** (1981), 357–368.

[55] Z. Füredi, On a Turán type problem of Erdős, *Combinatorica* **11** (1991), 75–79.

[56] Z. Füredi and M. Simonovits, Triple systems not containing a Fano configuration *Combin. Probab. Comput.* **14** (2005), 467–484.

[57] W.T. Gowers, Lower bounds of tower type for Szemerédi's uniformity lemma, *Geom. Funct. Anal.* **7** (1997), 322–337.

[58] W.T. Gowers, A new proof of Szemerédi's theorem for arithmetic progressions of length four, *Geom. Funct. Anal.* **8** (1998), 529–551.

[59] W.T. Gowers, Hypergraph regularity and the multidimensional Szemerdi theorem, *Ann. of Math.* **166** (2007), 897–946.

[60] R. Graham, V. Rödl, and A. Ruciński, On graphs with linear Ramsey numbers, *J. Graph Theory* **35** (2000), 176–192.

[61] R. Graham, V. Rödl, and A. Ruciński, On bipartite graphs with linear Ramsey numbers, *Combinatorica* **21** (2001), 199–209.

[62] R. L. Graham, B. L. Rothschild, and J. H. Spencer, **Ramsey theory**, 2nd edition, John Wiley & Sons (1980).

[63] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions, *Annals of Math.* **167** (2008), 481–547.

[64] H. Hatami, Graph norms and Sidorenko's conjecture, *Israel Journal of Mathematics*, to appear.

[65] G. Katona, Graphs, vectors and inequalities in probability theory (in Hungarian), *Mat. Lapok* **20** (1969), 123–127.

[66] P.Keevash and B. Sudakov, Local density in graphs with forbidden subgraphs, *Combinatorics, Probability and Computing* **12** (2003), 139–153.

[67] P.Keevash and B. Sudakov, The Turán number of the Fano Plane, *Combinatorica* **25** (2005), 561–574.

[68] P.Keevash and B. Sudakov, On a hypergraph Turán problem of Frankl, *Combinatorica* **25** (2005), 673–706.

[69] J. H. Kim, The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$, *Random Structures and Algorithms* **7** (1995), 173–207.

[70] J. Komlós and M. Simonovits, Szemerédi's regularity lemma and its applications in graph theory, in: *Combinatorics, Paul Erdős is eighty, Vol. 2* (Keszthely, 1993), 295–352, Bolyai Soc. Math. Stud., 2, Jnos Bolyai Math. Soc., Budapest, 1996.

[71] A. Kostochka, A class of constructions for Turán's $(3,4)$-problem, *Combinatorica* **2** (1982), 187–192.

[72] A. Kostochka and V. Rödl, On graphs with small Ramsey numbers, *J. Graph Theory* **37** (2001), 198–204.

[73] A. Kostochka and V. Rödl, On graphs with small Ramsey numbers II, *Combinatorica* **24** (2004), 389–401.

[74] A. Kostochka and B. Sudakov, On Ramsey numbers of sparse graphs, *Combin. Probab. Comput.* **12** (2003), 627–641.

[75] T. Kővári, V.T. Sós and P. Turán, On a problem of K. Zarankiewicz,*Colloquium Math.* **3** (1954), 50–57.

[76] M. Krivelevich, On the edge distribution in triangle-free graphs, *J. Combin. Theory Ser. B* **63** (1995), 245-260.

[77] M. Krivelevich and B. Sudakov, Pseudo-random graphs, in: *More Sets, Graphs and Numbers*, Bolyai Society Mathematical Studies 15, Springer, 2006, 199–262.

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

[79] B. Nagle, V. Rödl and M. Schacht, The counting lemma for regular k-uniform hypergraphs, *Random Structures and Algorithms* **28** (2006), 113–179.

[80] V. Nikiforov, The number of cliques in graphs of given order and size, *Transactions of AMS*, to appear.

[81] F.P. Ramsey, On a problem of formal logic, *Proc. London Math. Soc. Ser. 2* **30** (1930), 264–286.

[82] A. Razborov, On the minimal density of triangles in graphs, *Combin. Probab. Comput.* **17** (2008), 603–618.

[83] A. Razborov, On 3-hypergraphs with forbidden 4-vertex configurations, preprint.

[84] V. Rödl and J. Skokan, Regularity lemma for k-uniform hypergraphs, *Random Structures and Algorithms* **25** (2004), 1–42.

[85] V. Rödl and R. Thomas, Arrangeability and clique subdivisions, in: *The mathematics of Paul Erdős, II*, Algorithms Combin. 14, Springer, Berlin, 1997, 236–239.

[86] I. Ruzsa and E. Szemerédi, Triples systems with no six points carrying three triangles, in: *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976)*, Vol. II, Colloq. Math. Soc. Bolyai 18, North-Holland, Amsterdam, 1978, 939–945.

[87] M. Schacht, Extremal results for random discrete structures, preprint.

[88] L. Shi, Cube Ramsey numbers are polynomial, *Random Structures & Algorithms*, **19** (2001), 99–101.

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

[90] M. Simonovits, Extremal graph problems, degenerate extremal problems and super-saturated graphs, in: *Progress in graph theory* (A. Bondy ed.), Academic, New York, 1984, 419–437.

[91] J. Solymosi, Note on a generalization of Roth's theorem, in: *Discrete and computational geometry*, Algorithms Combin. 25, Springer, Berlin, 2003, 825–827.

[92] J. Spencer, Asymptotic lower bounds for Ramsey functions, *Discrete Math.* **20** (1977/78), 69–76.

[93] J. Spencer, E. Szemerédi and W. Trotter, Unit distances in the Euclidean plane, in: *Graph theory and combinatorics (Cambridge, 1983)*, Academic Press, London, 1984, 293–303.

[94] B. Sudakov, Few remarks on the Ramsey-Turan-type problems, *J. Combinatorial Theory Ser. B* **88** (2003), 99–106.

[95] B. Sudakov, A conjecture of Erdos on graph Ramsey numbers, submitted.

[96] B. Sudakov, Making a K_4-free graph bipartite, *Combinatorica* **27** (2007), 509–518.

[97] B. Sudakov, T. Szabo and V. Vu, A generalization of Turan's theorem, *J. Graph Theory* **49** (2005), 187–195.

[98] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arith.* **27** (1975), 199–245.

[99] P. Turán, On an extremal problem in graph theory (in Hungarian), *Mat. Fiz. Lapok* **48**, (1941) 436–452.

[100] T. Wooley, Problem 2.8, *Problem presented at the workshop on Recent Trends in Additive Combinatorics*, AIM, Palo Alto, 2004, http://www.aimath.org/WWN/additivecomb/additivecomb.pdf.