Combinatorics, Probability and Computing (2014) 23, 8–28. © Cambridge University Press 2013
doi:10.1017/S0963548313000448

Short Proofs of Some Extremal Results

DAVID CONLON[1†], JACOB FOX[2‡] and BENNY SUDAKOV[3§]

[1]Mathematical Institute, Oxford OX1 3LB, UK
(e-mail: david.conlon@maths.ox.ac.uk)

[2]Department of Mathematics, MIT, Cambridge, MA 02139-4307, USA
(e-mail: fox@math.mit.edu)

[3]Department of Mathematics, UCLA, Los Angeles, CA 90095, USA
(e-mail: bsudakov@math.ucla.edu)

Received 6 December 2012; revised 9 September 2013; first published online 4 November 2013

We prove several results from different areas of extremal combinatorics, giving complete or partial solutions to a number of open problems. These results, coming from areas such as extremal graph theory, Ramsey theory and additive combinatorics, have been collected together because in each case the relevant proofs are quite short.

2010 *Mathematics subject classification*: Primary 05C35
Secondary 05C55, 11B30

1. Introduction

We study several questions from extremal combinatorics, a broad area of discrete mathematics which deals with the problem of maximizing or minimizing the cardinality of a collection of finite objects satisfying a certain property. The problems we consider come mainly from the areas of extremal graph theory, Ramsey theory and additive combinatorics. In each case, we give a complete or partial solution to an open problem posed by researchers in the area.

While each of the results in this paper is interesting in its own right, the proofs are all quite short. Accordingly, in the spirit of Alon's 'Problems and results in extremal combinatorics' papers [3, 4], we have chosen to combine them. We describe the results in

† Research supported by a Royal Society University Research Fellowship.
‡ Research supported by a Simons Fellowship and NSF grant DMS-1069197.
§ Research supported in part by NSF grant DMS-1101185, by AFOSR MURI grant FA9550-10-1-0569 and by a USA–Israel BSF grant.

brief below. For full details on each topic we refer the reader to the relevant section, each of which is self-contained and may be read separately from all the others.

In Section 2 we improve a result of Alon [3] on the size of the largest induced forest in a bipartite graph of given average degree. In Section 3, we prove a conjecture of Balister, Lehel and Schelp [7] on Ramsey saturated graphs. We study the relationship between degeneracy and online Ramsey games in Section 4, addressing a question raised by Grytczuk, Hałuszczak and Kierstead [26]. In Section 5, we improve a recent result of Dudek and Mubayi [17] on generalized Ramsey numbers for hypergraphs. We prove a conjecture of Cavers and Verstraëte [11] in Section 6 by showing that any graph on n vertices whose complement has $o(n^2)$ edges has a clique partition using $o(n^2)$ cliques. In Section 7, we improve a result of Hegyvári [29] on the size of the largest Hilbert cube that may be found in a dense subset of the integers.

All logarithms are base 2 unless otherwise stated. For the sake of clarity of presentation, we systematically omit floor and ceiling signs whenever they are not crucial. We also do not make any serious attempt to optimize absolute constants in our statements and proofs.

2. Induced forests in sparse bipartite graphs

Every bipartite graph trivially has an independent set with at least half of its vertices. Under what conditions can we find a considerably larger sparse set? A *forest* is a graph without cycles. Akiyama and Watanabe [1] and, independently, Albertson and Haas [2] conjectured that every planar bipartite graph on n vertices contains an induced forest on at least $5n/8$ vertices. Motivated by this conjecture, Alon [3] considered induced forests in sparse bipartite graphs, showing that every bipartite graph on n vertices with average degree at most $d \geqslant 1$ contains an induced forest on at least $(\frac{1}{2} + e^{-bd^2})n$ vertices, for some absolute constant $b > 0$. On the other hand, there exist bipartite graphs on n vertices with average degree at most $d \geqslant 1$ that contain no induced forest on at least $(\frac{1}{2} + e^{-b'\sqrt{d}})n$ vertices. Alon remarks that it would be interesting to close the gap between the lower and upper bounds for this problem. We improve the lower bound here to $(\frac{1}{2} + d^{-bd})n$ for $d \geqslant 2$.

In particular, since the average degree of any planar bipartite graph is less than 4, there is an absolute positive constant ϵ such that every planar bipartite graph on n vertices contains an induced forest on at least $(1/2 + \epsilon)n$ edges. This gives some non-trivial result on the question raised in [1, 2]. More recently, it was shown in [32] (see also [41]) that every triangle-free (and hence every bipartite) planar graph on n vertices contains an induced forest on at least $71n/128$ vertices.

As in Alon's proof, we show that every sparse bipartite graph contains a large induced subgraph whose connected components are stars.

Theorem 2.1. *Let d be a positive integer. Every bipartite graph G on n vertices with average degree at most d contains an induced subgraph on at least $(\frac{1}{2} + \delta)n$ vertices with $\delta = (2^7 d^2)^{-4d}$ whose connected components are stars.*

Proof. Suppose, for contradiction, that the theorem is false. Let X and Y denote a bipartition of G into independent sets with $|X| \leqslant |Y|$. We may assume $\frac{1}{2}n \leqslant |Y| < (\frac{1}{2}+\delta)n$, as Y is an independent set and hence induces a star-forest. We have $|X| \geqslant (1/2-\delta)n \geqslant n/4$.

We will construct a sequence $Y_0 \subset Y_1 \subset \ldots \subset Y_{4d}$ of nested subsets of Y. Let $Y_0 = \emptyset$. Once Y_{i-1} has been defined, let $\delta_i = \delta + |Y_{i-1}|/n$, $d_i = 1/(2^7 d\delta_i)$ and Y_i consist of those vertices in Y with degree at least d_i. Note that δ_i increases, d_i decreases and Y_i grows as i increases. As G has at most $dn/2$ edges and every vertex in Y_i has degree at least d_i, we have $|Y_i|d_i \leqslant dn/2$ and hence $|Y_i|/n \leqslant d/(2d_i)$. We therefore have $\delta_1 = \delta$ and

$$\delta_{i+1} \leqslant \delta + d/(2d_i) = \delta + (d/2)(2^7 d\delta_i) = \delta + 2^6 d^2 \delta_i \leqslant 2^7 d^2 \delta_i.$$

Hence, for $i \geqslant 1$, by induction on i we have $\delta_i \leqslant (2^7 d^2)^{i-1}\delta$ and $d_i \geqslant 1/((2^7 d^2)^i \delta)$.

Let e_i denote the number of edges containing a vertex in $Y_i \setminus Y_{i-1}$. In the claim below we will prove that $e_i > n/8$ for all $i \geqslant 1$. This will complete the proof by contradiction, as then the number of edges of G is at most $dn/2$ and at least

$$\sum_{i=1}^{4d} e_i > 4d \cdot n/8 = dn/2.$$

Claim 2.2. *For $i \geqslant 1$, $e_i > n/8$.*

Indeed, suppose $e_i \leqslant n/8$. Let $X_i \subset X$ consist of those vertices not adjacent to any vertex in $Y_i \setminus Y_{i-1}$. We have $|X_i| \geqslant |X| - e_i \geqslant n/8$. Let $X_i' \subset X_i$ consist of those vertices of degree at most $8d$. We have $|X_i'| \geqslant |X_i|/2 \geqslant n/16$ as otherwise $e(G) > |X_i \setminus X_i'|8d > (n/16)8d = dn/2$, a contradiction. Pick out vertices $x_1, \ldots, x_t$ from X_i' greedily as follows. We pick x_j arbitrarily from the remaining vertices, and delete from X_i' all vertices which share a neighbour with x_j not in Y_{i-1}. Note that x_j has degree at most $8d$ and all its neighbours not in Y_{i-1} are also not in Y_i (by the definition of X_i) and so have degree less than $d_i \geqslant 1$. Hence, there are at most $8d(d_i - 1)$ vertices in X_i' which share a neighbour with x_j not in Y_{i-1}. We get $t \geqslant |X_i'|/(8dd_i) \geqslant n/(2^7 dd_i) = \delta_i n$ and the induced subgraph of G with vertex set $\{x_1, \ldots, x_t\} \cup (Y \setminus Y_{i-1})$ is a star-forest with at least $\delta_i n + |Y| - |Y_{i-1}| = \delta n + |Y| \geqslant (\frac{1}{2}+\delta)n$ vertices, a contradiction. This verifies the claim and completes the proof of the theorem. □

A much better bound may be proved for regular graphs. Indeed, it is shown in [6] that every d-regular bipartite graph on n vertices contains an induced forest with at least $(\frac{1}{2} + \frac{1}{2(d-1)^2})n$ vertices. Moreover, this result is sharp in its dependence on d. It seems to us that the bound given in Theorem 2.1 should also be close to best possible. It would, for example, be very interesting to improve Alon's upper bound to show that there exist bipartite graphs on n vertices with average degree at most $d \geqslant 1$ that contain no induced forest on at least $(\frac{1}{2} + e^{-b'd})n$ vertices.

3. Ramsey saturated graphs

The *Ramsey number* $r(G)$ of a graph G is the smallest natural number N such that every two-colouring of the edges of the complete graph K_N contains a monochromatic copy of G. The fact that these numbers exist was first proved by Ramsey [40].

Following Balister, Lehel and Schelp [7], we say that a graph G on n vertices is *Ramsey unsaturated* if there exists an edge $e \in E(K_n)\backslash E(G)$ such that $r(G+e) = r(G)$. However, if $r(G+e) > r(G)$ for all edges $e \in E(K_n)\backslash E(G)$, we say that G is *Ramsey saturated.*

There are many open questions about Ramsey saturated and unsaturated graphs. For example, it is not even known whether $K_n - e$ is saturated, that is, whether $r(K_n) > r(K_n - e)$, for any $n \geqslant 7$, though Balister, Lehel and Schelp conjecture that this should be the case.

One result proved by Balister, Lehel and Schelp [7] is that there are at least $\lfloor (n-2)/2 \rfloor$ non-isomorphic Ramsey saturated graphs on n vertices. Moreover, they conjectured that there should exist $c > 0$ and $\epsilon > 0$ for which there are at least $cn^{1+\epsilon}$ non-isomorphic Ramsey saturated graphs on n vertices. Here we prove this conjecture in a strong form, as follows.

Theorem 3.1. *There exists $c > 0$ such that there are at least 2^{cn^2} non-isomorphic Ramsey saturated graphs on n vertices.*

The proof of this theorem is a straightforward combination of two results from graph Ramsey theory. The first is a standard lower bound for the Ramsey number of a graph with n vertices and m edges which follows from the probabilistic method.

Lemma 3.2. *For any graph G with n vertices and m edges, the Ramsey number $r(G)$ satisfies*

$$r(G) > 2^{\frac{m}{n}-1}.$$

Proof. Let $N = 2^{\frac{m}{n}-1}$ and colour the edges of K_N at random, each edge being red or blue with probability $\frac{1}{2}$. Let X be the random variable counting the number of monochromatic copies of G. Then

$$\mathbb{E}[X] \leqslant 2^{1-m}N^n = 2^{1-m}(2^{\frac{m}{n}-1})^n = 2^{1-n} < 1.$$

It follows that there must be some colouring of K_N which does not contain a monochromatic copy of G. □

The second lemma we require is an upper bound for the Ramsey number of graphs with n vertices and m edges. The following result [13, 23] is sufficient for our purposes, though other results [15, 46] could also be used instead.

Lemma 3.3. *For any bipartite graph G with n vertices and maximum degree Δ, the Ramsey number $r(G)$ satisfies*

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

□

Proof of Theorem 3.1. We will show, for n sufficiently large, that there are at least $2^{n^2/80}$ non-isomorphic Ramsey saturated graphs. Consider a random labelled bipartite graph G_1 between two sets A and B, each of size $n/2$, containing $\frac{1}{5}\left(\frac{n}{2}\right)^2 = \frac{n^2}{20}$ edges. By a standard large deviation inequality for the hypergeometric distribution, for n sufficiently large, G_1 has maximum degree at most $n/8$ (and positive minimum degree) with probability at least $1/2$. That is, the number of such graphs is at least $\frac{1}{2}\binom{n^2/4}{n^2/20}$. For any such graph, Lemma 3.3 tells us that

$$r(G_1) \leqslant 2^{\frac{n}{8}+2}n^2 \leqslant 2^{\frac{n}{6}}$$

for n sufficiently large. On the other hand, by Lemma 3.2, any graph G_2 on $A \cup B$ with $m = \frac{n^2}{5}$ edges satisfies

$$r(G_2) > 2^{\frac{n}{5}-1} > 2^{\frac{n}{6}}.$$

It follows that for any G_1 for which $r(G_1) \leqslant 2^{\frac{n}{6}}$ there is a graph G with at most $\frac{n^2}{5}$ edges such that $G_1 \subseteq G$ and G is Ramsey saturated. Otherwise, starting from G_1 and adding edges which do not increase the Ramsey number one by one, we could find a sequence of graphs $G_1 \subset G_1 \cup \{e\} \subset \cdots \subset G_2$ up to a graph G_2 with $\frac{n^2}{5}$ edges such that $r(G_1) = r(G_1 \cup \{e\}) = \cdots = r(G_2)$, contradicting our estimate for the Ramsey number of graphs G_2 with $\frac{n^2}{5}$ edges.

Since any graph G with at most $\frac{n^2}{5}$ edges contains at most $\binom{n^2/5}{n^2/20}$ labelled subgraphs of size $n^2/20$, we see that the number of labelled Ramsey saturated graphs is at least

$$\begin{aligned}\frac{\frac{1}{2}\binom{n^2/4}{n^2/20}}{\binom{n^2/5}{n^2/20}} &= \frac{1}{2}\frac{\frac{n^2}{4}\left(\frac{n^2}{4}-1\right)\cdots\left(\frac{n^2}{4}-\frac{n^2}{20}+1\right)}{\frac{n^2}{5}\left(\frac{n^2}{5}-1\right)\cdots\left(\frac{n^2}{5}-\frac{n^2}{20}+1\right)} \\ &\geqslant \frac{1}{2}\left(\frac{5}{4}\right)^{n^2/20} \geqslant 2^{n^2/70},\end{aligned}$$

for n sufficiently large. Dividing through by $n!$ tells us that the number of non-isomorphic Ramsey saturated graphs is, again for n sufficiently large, at least $2^{n^2/80}$, as required. □

We note that this proof easily extends to more than two colours by using the appropriate analogues of Lemmas 3.2 and 3.3. We omit the details.

Balister, Lehel and Schelp [7] also conjecture that almost all graphs should be Ramsey unsaturated. While they prove that this is the case for paths and cycles of length at least five (a result which was recently extended by Skokan and Stein [43]), we feel that the truth probably lies in the other direction, that is, that almost all graphs should be Ramsey saturated. However, it would still be very interesting to find further classes of Ramsey unsaturated graphs.

4. Degeneracy and online Ramsey theory

Online Ramsey games were first introduced by Beck [8] and, independently, by Friedgut, Kohayakawa, Rödl, Ruciński and Tetali [24] (see also [35]). There are two players, Builder

and Painter, playing on a board consisting of an infinite, independent set of vertices. At each step, Builder exposes an edge and, as each edge appears, Painter decides whether to colour it red or blue. Builder's aim is to force Painter to draw a monochromatic copy of a fixed graph G.

These games have been studied from multiple perspectives. One variant asks for the least number of edges $\tilde{r}(G)$ which Builder needs to force Painter to draw a monochromatic G (see [14] and its references). Another asks how long the game lasts if the game is played on a board with n vertices and Builder chooses the edges at random (see [36, 37] and their references).

We will consider another variant, first introduced by Grytczuk, Hałuszczak and Kierstead [26]. Suppose that we have a positive integer-valued increasing graph parameter such as chromatic number, degeneracy, treewidth, thickness or genus. The question asked in [26] is whether, for each of these properties, there exists a function $f : \mathbb{N} \to \mathbb{N}$ such that Builder can force Painter to draw a monochromatic copy of any graph G with parameter k while himself only drawing graphs with parameter $f(k)$. To give a concrete example, it was proved in [26] (and extended in [31] to any number of colours) that Builder may force Painter to draw a monochromatic copy of any graph G with chromatic number k while only drawing graphs of chromatic number k himself.

Here we prove a similar result where the chosen parameter is degeneracy, partially addressing one of the questions raised by Grytczuk, Hałuszczak and Kierstead [26]. A graph is *d-degenerate* if every subgraph of it has a vertex of degree at most d. Equivalently, a graph G is d-degenerate if there is an ordering of the vertices of G, say $u_1, u_2, \ldots, u_n$, such that for each $1 \leqslant i \leqslant n$ the vertex u_i has at most d neighbours u_j with $j < i$. The *degeneracy* of G is the smallest d for which G is d-degenerate. We will consider the q-colour online Ramsey game, where Painter has a choice of q colours at each step, proving the following result.

Theorem 4.1. *In the q-colour online Ramsey game, Builder may force Painter to draw a monochromatic copy of any d-degenerate graph while only drawing a $(qd-(q-1))$-degenerate graph.*

We note that the $d = 1$ case of this theorem was already proved by Grytczuk, Hałuszczak and Kierstead [26]. Since 1-degenerate graphs are forests, the theorem in this case states that, for any fixed number of colours, Painter may force Builder to create a monochromatic copy of any forest while only building a forest. To prove Theorem 4.1 in the general case, we will need to use the hypergraph version of Ramsey's theorem [21, 40].

Lemma 4.2. *For all natural numbers k, ℓ and q with $\ell \geqslant k$, there exists an integer n such that if the edges of the complete k-uniform hypergraph on n vertices are q-coloured, then there is a monochromatic copy of $K_\ell^{(k)}$.* □

The smallest such number n is known as the *Ramsey number* $r_k(\ell; q)$.

Proof of Theorem 4.1. Suppose that G is a d-degenerate graph with n vertices. We will show that while only drawing a $(qd-(q-1))$-degenerate graph it is possible for Builder to force Painter to construct a sequence of subsets $V_0 \supset V_1 \supset \cdots \supset V_t$ and a sequence of colours $C_1, \ldots, C_t$, where $t = q(n-1)-(q-1)$, such that $|V_t| \geqslant d$ and the following holds. For every d-set D in V_i there is a vertex v_D in V_{i-1} such that every vertex $w \in D$ is connected to v_D in colour C_i. By the choice of t, the pigeonhole principle implies that there is a subsequence $V_0 = V_{i_0} \supset V_{i_1} \supset \cdots \supset V_{i_{n-1}}$ such that $C_{i_1} = \cdots = C_{i_{n-1}}$. Without loss of generality, we may assume that this colour is red.

It is straightforward to show that the constructed graph contains a red copy of G. Let $u_1, \ldots, u_n$ be a d-degenerate ordering of the vertices of G, that is, such that u_i has at most d neighbours u_j with $j < i$. We will construct an embedding $f : V(G) \to V$ of the vertices of G into the red graph constructed above by induction, embedding the vertex u_p into the set $V_{i_{n-p}}$. We begin by mapping u_1 to any vertex in $V_{i_{n-1}}$. Suppose now that $u_1, \ldots, u_p$ have been embedded and we wish to embed u_{p+1}. We know that u_{p+1} has $f \leqslant d$ neighbours $u_{a_1}, \ldots, u_{a_f}$ with $a_j \leqslant p$. Since the images of each of these vertices under the embedding lie in $V_{i_{n-p}}$, we see, by taking

$$D = \{f(u_{a_1}), \ldots, f(u_{a_f})\},$$

that there is a vertex $w \in V_{i_{n-(p+1)}}$ such that the edge between w and $f(u_{a_j})$ is red for all $1 \leqslant j \leqslant f$. Taking $f(u_{p+1}) = w$ completes the induction.

It remains to construct the subsets $V_0 \supset V_1 \supset \cdots \supset V_t$. Let $n_t = d$ and for each $1 \leqslant i \leqslant t$ let

$$m_{t-i} = r_s(qd(n_{t-i+1}+1); q^s) \quad \text{and } \mathrm{n}_{t-i} = m_{t-i} + \binom{m_{t-i}}{s},$$

where $s = qd-(q-1)$. We begin by taking an independent set of size n_0 for V_0. Suppose now that Builder has forced Painter to construct $V_0 \supset V_1 \supset \cdots \supset V_{i-1}$ and that V_{i-1} is an independent set of size n_{i-1}. We will show how Builder may force Painter to construct an independent set $V_i \subset V_{i-1}$ of size n_i such that for every d-set D in V_i there is a vertex v_D in $V_{i-1} \backslash V_i$ such that every vertex $w \in D$ is connected to v_D in a fixed colour C_i.

Suppose, therefore, that V_{i-1} is an independent set of size n_{i-1}. By the choice of n_{i-1}, we may partition V_{i-1} into two pieces W_i and $V_{i-1} \backslash W_i$ of sizes m_{i-1} and $\binom{m_{i-1}}{s}$, respectively. For each s-set S in W_i, Builder now chooses a unique vertex v_S in $V_{i-1} \backslash W_i$ (which is possible by the choice of size for $V_{i-1} \backslash W_i$) and joins every $w \in S$ to v_S. Moreover, v_S will have no other neighbours in V_{i-1}. Note that W_i is an independent set and every vertex in $V_{i-1} \backslash W_i$ has degree exactly $s = qd-(q-1)$.

We now consider the complete s-uniform hypergraph on W_i. Suppose that the vertices of W_i have been ordered. For any edge $e = \{w_1, \ldots, w_s\}$ of this graph with $w_1 < \cdots < w_s$, we assign it the colour $(\chi(w_1 v_e), \ldots, \chi(w_s v_e))$, where v_e is the unique vertex in $V_{i-1} \backslash W_i$ joined to each of $w_1, \ldots, w_s$ and $\chi(w_i v_e)$ is the colour assigned by Painter to the edge between w_i and v_e. This gives a q^s-colouring of the edges of the complete s-uniform hypergraph on W_i. By the choice of m_{i-1}, this set must contain a monochromatic complete subgraph of order $qd(n_i+1)$; we call its vertex set M_i.

For any edge $e = \{w_1, \ldots, w_s\}$ in M_i with $w_1 < \cdots < w_s$, we know that $\chi(w_j v_e) = \chi_j$ for a fixed sequence of colours $\chi_1, \ldots, \chi_s$. By the choice of s, we know that there must be a colour C_i and a set of indices $j_1, \ldots, j_d$ such that $\chi(w_{j_k} v_e) = C_i$ for all edges e and all $1 \leqslant k \leqslant d$. Suppose now that the vertices in M_i are $w'_1 < \cdots < w'_\ell$, where $\ell = qd(n_i + 1)$. Consider the subset of M_i containing the vertices $w'_{qd}, w'_{2qd}, w'_{3qd}, \ldots, w'_{n_i qd}$. Then, for any $1 \leqslant a_1 < \cdots < a_d \leqslant n_i$, there exists an edge $e = \{w_1, \ldots, w_s\}$ of the complete s-uniform hypergraph on M_i such that $w_{j_k} = w'_{a_k qd}$. This follows since the vertices in the subsequence are a distance qd apart and we may place up to $qd - 1 \geqslant s$ dummy vertices between any pair (and before and after the first and the last terms in the sequence). Therefore, $\chi(w'_{a_k qd} v_e) = C_i$ for all $1 \leqslant k \leqslant d$. The set $V_i = \{w'_{qd}, w'_{2qd}, w'_{3qd}, \ldots, w'_{n_i qd}\}$ has the required property that for every d-set D in V_i there is a vertex v_D in $V_{i-1} \backslash V_i$ such that every vertex $w \in D$ is connected to v_D in colour C_i.

To show that the graph constructed by Builder has a $(qd - (q-1))$-degenerate ordering, we note that each vertex in $V_{i-1} \backslash W_i$ has no neighbours in $V_{i-1} \backslash W_i$ and degree $qd - (q-1)$ in W_i. Moreover, there are no edges between V_i and $W_i \backslash V_i$. If, therefore, we choose our ordering so that for all $i = 1, 2, \ldots, t$, the set of vertices in V_i comes before those in $W_i \backslash V_i$ and the set of vertices in $W_i \backslash V_i$ comes before those in $V_{i-1} \backslash W_i$, we will have an s-degenerate ordering. This completes the proof. □

In the other direction, it is obvious that Painter must draw a graph with degeneracy at least d. For $d = 1$, a matching upper bound is given by Theorem 4.1 and was first proved in [26]. It would be interesting to decide whether this simple lower bound is always sharp. It would also be interesting to know whether a similar theorem holds when degeneracy is replaced by maximum degree. This question, first studied in [10], appears difficult.

5. Generalized Ramsey numbers for hypergraphs

Given natural numbers s and t with $s < t$, the *Erdős–Rogers function* $f_{s,t}(n)$ is defined as the size of the largest K_s-free subset that may be found in any K_t-free graph on n vertices. This function generalizes the usual Ramsey function, since determining $f_{2,t}(n)$ is the problem of determining the size of the largest independent set which is guaranteed in any K_t-free graph on n vertices.

Since its introduction [20, 22] and particularly in recent years [5, 19, 33, 34, 44, 45], this function has been studied quite intensively. Much of this work has focused on the case where $t = s + 1$, culminating in the result [17, 18, 53] that there are constants c_1 and c_2 depending only on s such that

$$c_1 \left(\frac{n \log n}{\log \log n} \right)^{1/2} \leqslant f_{s,s+1}(n) \leqslant c_2 n^{1/2} (\log n)^{4s^2}.$$

The analogous function for hypergraphs was recently studied by Dudek and Mubayi [17]. For $s < t$, let $f^{(k)}_{s,t}(n)$ be given by

$$f^{(k)}_{s,t}(n) = \min\{\max\{|W| : W \subseteq V(\mathcal{G}) \quad \text{and } \mathcal{G}[W] \text{ contains no } K^{(k)}_s\}\},$$

where the minimum is taken over all $K_t^{(k)}$-free k-uniform hypergraphs $\mathcal{G}$ on n vertices. Dudek and Mubayi proved that, for $k = 3$ and $3 \leqslant s < t$, this function satisfies

$$f^{(2)}_{s-1,t-1}(\lfloor\sqrt{\log n}\rfloor) \leqslant f^{(3)}_{s,t}(n) \leqslant c_s \log n.$$

In particular, for $t = s + 1$, this gives constants c_1 and c_2 depending only on s such that

$$c_1(\log n)^{1/4}\left(\frac{\log\log n}{\log\log\log n}\right)^{1/2} \leqslant f^{(3)}_{s,s+1}(n) \leqslant c_2 \log n.$$

Here we make an improvement to the lower bound, using ideas on hypergraph Ramsey numbers developed by the authors in [16].

We will need the following lemma, due to Shearer [42], which gives an estimate for the size of the largest independent set in a sparse K_s-free graph.

Lemma 5.1. *There exists a constant c_s such that if G is a K_s-free graph on n vertices with average degree at most d then G contains an independent set of size at least*

$$c_s \frac{\log d}{d \log\log d} n. \qquad \square$$

The main result of this section is now as follows.

Theorem 5.2. *For any natural number $s \geqslant 3$, there exists a constant c such that*

$$f^{(3)}_{s,s+1}(n) \geqslant c\left(\frac{\log n}{\log\log\log n}\right)^{1/3}.$$

Proof. Let $\mathcal{G}$ be a $K^{(3)}_{s+1}$-free 3-uniform hypergraph on n vertices. Let

$$p = c\left(\frac{\log n}{\log\log\log n}\right)^{1/3},$$

where c is a constant to be determined later. Our aim will be to show that $\mathcal{G}$ contains a $K^{(3)}_s$-free subgraph of size at least p.

We will construct, by induction, a sequence of vertices $v_1, v_2, \ldots, v_\ell$ and a non-empty set V_ℓ such that, for any $1 \leqslant i < j \leqslant \ell$, all triples $\{v_i, v_j, v_k\}$ with $j < k \leqslant \ell$ and all triples $\{v_i, v_j, w\}$ with $w \in V_\ell$ are either all edges of $\mathcal{G}$ or all not edges of $\mathcal{G}$. At each step, we consider the auxiliary graph G_ℓ on vertex set $\{v_1, \ldots, v_\ell\}$ formed by connecting v_i and v_j with $i < j$ if and only if all triples $\{v_i, v_j, v_k\}$ with $j < k \leqslant \ell$ and all triples $\{v_i, v_j, w\}$ with $w \in V_\ell$ are edges of $\mathcal{G}$. We note that G_ℓ must be K_s-free. Otherwise, if $\{v_{i_1}, \ldots, v_{i_s}\}$ are the vertices of a K_s, we may take any vertex w in V_ℓ to form a $K^{(3)}_{s+1}$ in $\mathcal{G}$, namely, $\{v_{i_1}, \ldots, v_{i_s}, w\}$.

If at any point G_ℓ contains a vertex u with at least p neighbours, we stop the process. Since G_ℓ is K_s-free, it follows that the neighbourhood U of u in G_ℓ is K_{s-1}-free. The set U must also be $K^{(3)}_s$-free in $\mathcal{G}$. Indeed, suppose otherwise and that $\{v_{i_1}, \ldots, v_{i_s}\}$ with $i_1 < \cdots < i_s$ is a $K^{(3)}_s$ with vertices from U. Then, by the construction of G_ℓ, the set

$\{v_{i_1}, \ldots, v_{i_{s-1}}\}$ must be a K_{s-1} in U. We may therefore assume that the maximum degree in each G_ℓ is at most p.

To begin our induction, we fix v_1 and let $V_1 = V(\mathcal{G})\backslash\{v_1\}$. Suppose now that we have constructed the sequence $v_1, v_2, \ldots, v_\ell$ and a set V_ℓ satisfying the required conditions and we wish to find a vertex $v_{\ell+1}$ and a set $V_{\ell+1}$. We let $v_{\ell+1}$ be any vertex from the set V_ℓ. Let $V_{\ell,0} = V_\ell\backslash\{v_{\ell+1}\}$. We will construct a sequence of subsets $V_{\ell,0} \supset V_{\ell,1} \supset \cdots \supset V_{\ell,\ell}$ such that all triples $\{v_i, v_{\ell+1}, w\}$ with $1 \leqslant i \leqslant j$ and $w \in V_{\ell,j}$ are either all edges of $\mathcal{G}$ or all not edges of $\mathcal{G}$, depending only on the value of i. (Note that since each $V_{\ell,j} \subset V_\ell$ it follows that all triples $\{v_i, v_j, w\}$ with $1 \leqslant i < j \leqslant \ell$ and $w \in V_{\ell,j}$ are either all edges of $\mathcal{G}$ or all not edges of $\mathcal{G}$, depending only on the values of i and j.)

Suppose then that $V_{\ell,j}$ has been constructed in an appropriate fashion. To construct $V_{\ell,j+1}$, we consider the neighbourhood of the vertices v_{j+1} and $v_{\ell+1}$ in $V_{\ell,j}$. If this neighbourhood has size at least $\alpha|V_{\ell,j}|$, we let $V_{\ell,j+1}$ be this neighbourhood. Otherwise, we let $V_{\ell,j+1}$ be the complement of this neighbourhood in $V_{\ell,j}$. Note that in this case $|V_{\ell,j+1}| \geqslant (1-\alpha)|V_{\ell,j}|$. To finish the construction of $V_{\ell+1}$, we let $V_{\ell+1} = V_{\ell,\ell}$. It is easy to check that it satisfies the required conditions.

We halt the process when $\ell = m$. Recall that the maximum degree of G_m is at most p. Since G_m is also K_s-free, Lemma 5.1 implies that there is a constant c_s such that the graph contains an independent set of size at least

$$c_s \frac{\log p}{p \log\log p} m = p,$$

by choosing

$$m = \frac{1}{c_s} \frac{\log\log p}{\log p} p^2.$$

This in turn implies an independent set of size p in $\mathcal{G}$, completing the proof provided only that $|V_m| \geqslant 1$. To verify this, note that if in G_i the vertex v_i has degree d_i then

$$|V_i| \geqslant \alpha^{d_i}(1-\alpha)^{i-1-d_i}(|V_{i-1}| - 1) \geqslant \alpha^{d_i}(1-\alpha)^{i-1-d_i}|V_{i-1}| - 1.$$

Telescoping over all $i = 2, \ldots, m$, it follows that, for $\alpha \leqslant \frac{1}{2}$,

$$\begin{aligned} |V_m| &\geqslant \alpha^{\sum_{i=2}^m d_i}(1-\alpha)^{\binom{m}{2} - \sum_{i=2}^m d_i} n - m \\ &\geqslant \alpha^{\frac{pm}{2}}(1-\alpha)^{\binom{m}{2} - \frac{pm}{2}} n - m. \end{aligned}$$

The second inequality follows by noting that $\sum_{i=2}^m d_i = e(G_m) \leqslant \frac{pm}{2}$ and observing that the function $\alpha^t(1-\alpha)^{\binom{m}{2}-t}$ is decreasing in t for $\alpha \leqslant \frac{1}{2}$. Therefore, taking $\alpha = \frac{p}{m}\log\left(\frac{m}{p}\right)$ (note that $\alpha \leqslant \frac{1}{2}$ for n sufficiently large) and using that $2^{-2x} \leqslant 1 - x$ for $0 \leqslant x \leqslant \frac{1}{2}$,

$$|V_m| \geqslant \alpha^{\frac{pm}{2}}(1-\alpha)^{\frac{m^2}{2}} n - m \geqslant (p/m)^{\frac{pm}{2}} 2^{-\alpha m^2} n - m = 2^{-\frac{3}{2}pm\log(m/p)} n - m.$$

For n sufficiently large depending on c_s, $\log(m/p) \leqslant \log p$ and, therefore,

$$\frac{3}{2} pm \log(m/p) \leqslant \frac{3}{2c_s} p^3 \frac{\log\log p}{\log p} \log p = \frac{3}{2c_s} p^3 \log\log p.$$

Hence, since

$$p = c\left(\frac{\log n}{\log\log\log n}\right)^{1/3},$$

for c sufficiently small we have

$$\frac{3}{2}pm\log(m/p) \leqslant \frac{1}{2}\log n.$$

Hence

$$|V_n| \geqslant \sqrt{n} - m \geqslant 1,$$

completing the proof. □

This result easily extends to higher uniformities to give that $f^{(k)}_{s,s+1}(n) \geqslant (\log_{(k-2)} n)^{1/3-o(1)}$. Here $\log_{(0)} x = x$ and $\log_{(i+1)} x = \log(\log_{(i)} x)$. This improves an analogous result of Dudek and Mubayi [17] with a 1/4 in the exponent, but remains far from their upper bound $f^{(k)}_{s,s+1}(n) \leqslant c_{s,k}(\log n)^{1/(k-2)}$. It is an interesting open question to close the gap between the upper and lower bounds.

6. Clique partitions of very dense graphs

A *clique partition* of a graph G is a collection of complete subgraphs of G that partition the edge set of G. The *clique partition number* $\mathrm{cp}(G)$ is the smallest number of cliques in a clique partition of G. Despite receiving considerable attention over the last 60 years, this interesting graph parameter is still not well understood.

One class of graphs for which this parameter has been studied quite extensively is when the graph G is the complement $\overline{F}$ of a sparse graph F. Orlin [39] was the first to ask about the asymptotics of the clique partition number for the complement of a perfect matching. If F is a perfect matching on n vertices, Wallis [50] showed that $\mathrm{cp}(\overline{F}) = o(n^{1+\epsilon})$ for any $\epsilon > 0$. This was later improved by Gregory, McGuinness and Wallis [25] to $\mathrm{cp}(\overline{F}) = O(n\log\log n)$. Wallis [51] later showed that the same bound holds if F is a Hamiltonian path on n vertices. For any forest F on n vertices, Cavers and Verstraëte [11] proved that $\mathrm{cp}(\overline{F}) = O(n\log n)$. They also conjectured that there are forests F on n vertices such that $\mathrm{cp}(\overline{F})$ grows superlinearly in n. Here we extend their proof to give the following improvement.

Theorem 6.1. *If F is a forest on n vertices, then* $\mathrm{cp}(\overline{F}) = O(n\log\log n)$.

A *Steiner (n,k)-system* is a family of k-element subsets of an n-element set such that each pair of elements appears in exactly one of the subsets. One can view a Steiner (n,k)-system as a clique partition of K_n into cliques of size k. It is an open problem to determine for which pairs (n,k) a Steiner (n,k)-system exists. Necessary conditions for the existence of a Steiner (n,k)-system are $n \equiv 1 \bmod k-1$ and $n(n-1) \equiv 0 \bmod k(k-1)$. Wilson [52] showed that for each k there exists $n(k)$ such that the necessary conditions for

the existence of a Steiner (n,k)-system are sufficient provided that $n \geqslant n(k)$. An explicit upper bound on $n(k)$ that is triple exponential in $k^2 \log k$ was proved by Chang [12].

For sparse graphs F, Cavers and Verstraëte [11] proved an upper bound on the clique partition number $\mathrm{cp}(\overline{F})$ that is conditional on the existence of certain Steiner systems. They prove that if F is a graph on n vertices with maximum degree $\Delta = o(n/\log^4 n)$ and there is a Steiner (n,k)-system with

$$k = \left\lfloor \left(\frac{n}{2\Delta} \right)^{1/2} \right\rfloor,$$

then $\mathrm{cp}(\overline{F}) = O\big(n^{3/2}\Delta^{1/2}\log^2 n\big)$. They also conjectured that if the maximum degree of F is $o(n)$ then $\mathrm{cp}(\overline{F}) = o(n^2)$. This conjecture follows from the next theorem, which gives an unconditional improvement on their bound.

Theorem 6.2. *If F has n vertices and $m \geqslant \sqrt{n}$ edges, then* $\mathrm{cp}(\overline{F}) = O\big((mn)^{2/3}\big)$.

We have the following immediate corollary, showing that we can partition the complement of a sparse graph into a few cliques.

Corollary 6.3. *If F has n vertices and $o(n^2)$ edges, then* $\mathrm{cp}(\overline{F}) = o(n^2)$. □

To prove Theorem 6.2, we will first prove a useful lemma saying that, for $2 \leqslant k < n$, we can partition the complete graph K_n into $O(\max((n/k)^2, n))$ cliques of order at most k. We begin with some elementary observations. For graphs G and H, let $G \cup H$ denote the disjoint union of G and H.

Lemma 6.4. *For s,t,n with $s+t \leqslant n$, we have* $\mathrm{cp}(K_n \setminus (K_s \cup K_t)) \geqslant st$.

Proof. Indeed, in a clique-partition of $K_n \setminus (K_s \cup K_t)$, each of the st edges between the independent set of size s and the independent set of size t must be in different cliques. □

Lemma 6.5. *If G has n vertices and all but $t \geqslant 1$ vertices of G have degree $n-1$, then* $\mathrm{cp}(G) < tn$.

Proof. The clique partition of G consisting of the clique on the $n-t$ vertices of degree $n-1$ and a K_2 for each remaining edge uses at most $1 + t(n-2) < tn$ cliques. □

In order to prove our main auxiliary lemma, we need to know a little about the particular class of Steiner systems known as projective planes. A *projective plane* consists of a set of points and a set of lines and a relation between points and lines called incidence having the following properties.

- For any two distinct points, there is exactly one line incident to both of them,
- For any two distinct lines, there is exactly one point incident to both of them,
- There are four points such that no line is incident with more than two of them.

The first condition says that any two points determine a line, the second condition says that any two lines intersect in one point and the last condition excludes some degenerate cases. It can be shown that a projective plane has the same number of lines as it has points. Any finite projective plane has q^2+q+1 points for some positive integer q and we denote such a projective plane by P_q. Each line is incident with $q+1$ points and each point is incident with $q+1$ lines. Therefore, the lines of P_q form a Steiner $(q^2+q+1, q+1)$-system. It is known that if q is a prime power then there is a projective plane on q^2+q+1 points. It is a famous open problem to determine if there exist finite projective planes of other orders.

Lemma 6.6. *Let $k \geqslant 2$ and n be positive integers and let $f(n,k)$ denote the minimum number of cliques each on at most k vertices needed to clique partition K_n. If $n \leqslant k$, trivially $f(n,k) = 1$. If $n > k$, then*

$$f(n,k) = \Theta\big(\max\big((n/k)^2, n\big)\big).$$

Proof. We first prove the lower bound. The pigeonhole principle implies the lower bound $f(n,k) \geqslant \binom{n}{2}/\binom{k}{2} \geqslant (n/k)^2$, as we need to cover $\binom{n}{2}$ edges by cliques each with at most $\binom{k}{2}$ edges. Now suppose $\sqrt{n} \leqslant k < n$ and we have a partition of the edges of K_n into cliques each with at most k vertices. Let $s \geqslant t$ be the sizes of the two largest cliques used in the partition, so $t \geqslant 2$. If $s \leqslant 2\sqrt{n}$, then the number of cliques used is at least $f(n,s) \geqslant n^2/s^2 \geqslant n/4$. So suppose $s > 2\sqrt{n}$. By Lemma 6.4, there are at least $(s-1)(t-1)$ remaining cliques in the partition (the cliques may intersect in one vertex). So suppose $(s-1)(t-1) < n/4$. Then $st < n$. Since all cliques besides the largest have at most t vertices, the number of cliques used is at least

$$\begin{aligned} 1 + \frac{\binom{n}{2} - \binom{s}{2}}{\binom{t}{2}} &= 1 + \frac{(n-s)(n+s-1)}{t(t-1)} > 1 + \frac{(n-s)(n+s-1)}{(n/s)((n/s)-1)} \\ &= 1 + \frac{s^2(n+s-1)}{n} \geqslant 1 + s^2 > n, \end{aligned}$$

as $\sqrt{n} \leqslant s$ and $t < n/s$. We have thus proved that $f(n,k) \geqslant \max((n/k)^2, n/4)$.

We now turn to proving the upper bound. We will prove by induction on n that the bound $f(n,k) \leqslant \max(200(n/k)^2, 4n)$ holds. We may assume that $k > 20$, as otherwise we can clique partition K_n into $\binom{n}{2} \leqslant 200(n/k)^2$ edges. We will use Bertrand's postulate that there is a prime between x and $2x$ for every $x \geqslant 1$. Much better estimates are known on the distribution of the primes, but this is sufficient for our purposes.

If $k \geqslant 2\sqrt{n}$, let q be the smallest prime such that $q^2+q+1 \geqslant n$. Bertrand's postulate implies that $q^2+q+1 < 4n$, and hence $q < k$. Consider a projective plane P_q on q^2+q+1 points and edge-partition the complete graph on these points into cliques of size $q+1 \leqslant k$ given by the lines of P_q. There are $q^2+q+1 < 4n$ such cliques and, restricting to n of these points, we get in this case $f(n,k) \leqslant q^2+q+1 < 4n$.

If $k < 2\sqrt{n}$, let q be the smallest prime which is at least $n/k + k$. Bertrand's postulate implies that $q \leqslant 2(n/k+k) < 2(n/k + 2\sqrt{n}) < 10n/k$. The Turán graph $T_{n,k}$ is the complete k-partite graph on n vertices with parts of size as equal as possible. We will find a partition

of the edges of $T_{n,k}$ into cliques of size at most k using at most $q^2+q-k+1$ cliques. Let $S_1,\ldots,S_k$ denote the k independent sets of $T_{n,k}$. Let $L_1,L_2,\ldots,L_k$ denote k lines of the projective plane P_q. View each S_i as a subset of points of the line L_i such that the points of $S_1,\ldots,S_k$ do not include any of the intersection points of $L_1,\ldots,L_k$. We can do this since $|S_i| \leqslant \lceil n/k \rceil$, $|L_i| = q+1 \geqslant n/k+k+1$, each pair of lines have one intersection point, and so each L_i has precisely $k-1$ intersection points in total with all the other L_j. Every pair of points in $S_1 \cup \ldots \cup S_k$ other than those inside one of the S_i is contained in exactly one of the $q^2+q-k+1$ lines of P_q other than $L_1,\ldots,L_k$. Since each of these lines intersects L_i and hence S_i in at most one point, these $q^2+q-k+1$ lines give an edge-partition of $T_{n,k}$ into $q^2+q-k+1$ cliques each with at most k vertices. This gives the bound

$$\begin{aligned} f(n,k) &\leqslant q^2+q-k+1+kf(\lceil n/k \rceil,k) \leqslant 110(n/k)^2+k\max(200(\lceil n/k \rceil)^2/k^2, 4\lceil n/k \rceil) \\ &\leqslant 110(n/k)^2+800n^2/k^3+8n \leqslant 200(n/k)^2, \end{aligned}$$

where we use $20<k<2\sqrt{n}$ and that, for each k, $f(n,k)$ is a monotonically increasing function of n. This latter fact follows by noting that restricting a clique partition of a clique to a subclique is a clique partition of the subclique. This completes the proof by induction on n. □

One can easily modify the above proof, using the prime number theorem instead of Bertrand's postulate, to show that

$$f(n,k) = (1+o(1))\frac{n^2}{k(k-1)}$$

as long as $k = o(\sqrt{n})$.

Proof of Theorem 6.2. Let F be a graph with n vertices and $m \geqslant \sqrt{n}$ edges and

$$k = (n^2/m)^{1/3} \leqslant \sqrt{n}.$$

By Lemma 6.6, we can partition the complete graph on n vertices into $N = O\big((n/k)^2\big)$ cliques $Q_1,\ldots,Q_N$, each of order at most k. For each clique Q_i let t_i be the number of edges of F it contains. Each Q_i that contains no edge of F will be used in the clique partition of $\overline{F}$. If $t_i \geqslant 1$, we use Lemma 6.5 to partition the induced subgraph of $\overline{F}$ on the vertex set of Q_i into $2t_i|Q_i| \leqslant 2t_ik$ cliques. Thus, we get that $\overline{F}$ can be partitioned into at most $N+2\sum_i t_ik = N+2mk = O\big((mn)^{2/3}\big)$ cliques, which completes the proof of Theorem 6.2. □

Note that adding an edge to a graph can increase the clique partition number by at most 1. It follows that for any forest F, if T is a spanning tree containing F then $\mathrm{cp}(\overline{F}) \leqslant \mathrm{cp}(\overline{T})+n-1$. So, to prove Theorem 6.1, it is sufficient to prove it for trees. Let $g(n)$ denote the maximum of $\mathrm{cp}(\overline{T})$ over all trees T on $n \geqslant 2$ vertices. We will prove by induction that

$$g(n) \leqslant 100n(1+\log\log n),$$

which verifies Theorem 6.1. This inequality clearly holds for $n \leqslant 200$ as $g(n) \leqslant \binom{n}{2} \leqslant 100n$ in this case. So suppose $n > 200$.

A *tree partition* of a graph G is a collection $\{T_1, \ldots, T_r\}$ of subtrees of G such that each edge of G is in exactly one tree and, for all $i \neq j$, T_i and T_j share at most one vertex in common. We will use the following simple lemma from [11].

Lemma 6.7. *Let T be a tree on n vertices and $2 \leqslant k \leqslant n$. Then there exists a tree partition $\{T_1, \ldots, T_r\}$ of T into at most $2n/k$ trees such that the number of vertices of each T_i is between $k/3$ and k.* □

Proof of Theorem 6.1. Let $k = \sqrt{n}$ and apply the above lemma to find a tree partition of T into $r \leqslant 2n/k = 2\sqrt{n}$ trees, each with at most $k = \sqrt{n}$ vertices. Order these trees $\{T_1, \ldots, T_r\}$ so that for $i \geqslant 2$ the union of $\bigcup_{j=1}^{i-1} V(T_j)$ is connected. Then there is exactly one vertex v_i of T_i that is contained in $\bigcup_{j=1}^{i-1} V(T_j)$. By Bertrand's postulate, there is a prime q satisfying $3\sqrt{n} \leqslant q \leqslant 6\sqrt{n}$. We will show that there is a one-to-one mapping of the vertices of T into the points of the projective plane P_q on $q^2 + q + 1$ points such that each T_i is contained in a line L_i and no vertices of T other than those in T_i map to a point in L_i.

Let L_1 be an arbitrary line in the projective plane P_q on $q^2 + q + 1$ vertices. Arbitrarily embed the vertices of T_1 into the points of L_1 with the vertex v_2 identified with some point w_2. This can be done since $|T_1| \leqslant k < q + 1 = |L_1|$. Suppose we have already embedded $L_1, \ldots, L_{i-1}$ and let w_i denote the image of the vertex v_i which is in at least one line L_j with $j < i$. Pick an arbitrary line L_i through w_i with $L_i \neq L_j$ for $j < i$. Since there are $q + 1 > r > i - 1$ lines through each point, and, in particular, through w_i, we can indeed pick such a line L_i. Arbitrarily embed the remaining vertices $V(T_i) \setminus \{v_i\}$ of T_i amongst the points of L_i not in any L_j with $j < i$. Since any two lines intersect at exactly one point, L_i has $q + 1 - (i - 1) \geqslant q + 2 - r > k \geqslant |T_i|$ points not in any L_j with $j < i$. Thus we can indeed embed these remaining vertices. This demonstrates that we can find the desired mapping of the vertices of T into the points of P_q.

We next construct a clique partition of $\overline{T}$. For each line of the projective plane not containing an edge of T, we use the corresponding clique restricted to the vertices of T. For each line of the projective plane containing an edge of T, this line contains the vertices of T_i and no other vertex of T, so we use at most $g(|T_i|)$ cliques to partition the edges of $\overline{T_i}$. Note that $\sum_{i=1}^{r} |T_i| = |T| + r - 1 \leqslant n + 2n/k - 1 < n + 2\sqrt{n}$. Using this estimate, together with the inequalities $|T_i| \leqslant \sqrt{n}$, $n > 200$ and the induction hypothesis, we get

$$\begin{aligned} g(n) &\leqslant q^2 + q + 1 + \sum_{i=1}^{r} g(|T_i|) \leqslant 45n + 100(n + 2\sqrt{n})(1 + \log\log\sqrt{n}) \\ &= 45n + 100(n + 2\sqrt{n})\log\log n \leqslant 100n(1 + \log\log n), \end{aligned}$$

as required. □

7. Hilbert cubes in dense sets

A *Hilbert cube* or an affine cube is a set $H \subset \mathbb{N}$ of the form

$$H = H(x_0, x_1, \ldots, x_d) = \left\{ x_0 + \sum_{i \in I} x_i : I \subseteq [d] \right\},$$

where x_0 is a non-negative integer and $x_1, \ldots, x_d$ are positive integers. We will always assume here that the generators $x_1, \ldots, x_d$ are all distinct. We refer to the index d as the dimension.

One of the earliest results in Ramsey theory is a theorem of Hilbert [30] stating that if n is sufficiently large then any colouring of the set $[n]$ with a fixed number of colours, say r, must contain a monochromatic Hilbert cube of dimension d. The smallest such n we denote by $h(d, r)$. The best known upper bound for this function is

$$h(d, r) \leqslant (2r)^{2^{d-1}}.$$

As noted in [9], a double exponential upper bound already follows from Hilbert's original argument. The result we have quoted is a slight strengthening noted by Gunderson, Rödl and Sidorenko [27] which follows from a stronger density statement.

This density version [47] states that for any natural number d and $\delta > 0$ there exists an n_0 such that if $n \geqslant n_0$, any subset of $[n]$ containing δn elements contains a Hilbert cube of dimension d. Quantitative versions of this lemma imply that any subset of $[n]$ of fixed positive density δ contains a Hilbert cube of dimension at least $c' \log\log n$, where c' depends only on δ. This in turn gives a double-exponential upper bound for the original colouring problem.

On the other hand, Hegyvári [29] gave a lower bound by proving that with high probability a random subset of $[n]$ of small but fixed positive density δ does not contain Hilbert cubes of dimension $c\sqrt{\log n \log\log n}$, where c depends only on δ. Here we improve this result as follows.

Theorem 7.1. *For any $0 < \delta < 1$, there exists $c > 0$ such that with high probability a random subset of the set $[n]$, where each element is chosen independently with probability δ, does not contain Hilbert cubes of dimension $c\sqrt{\log n}$.*

By another result of Hegyvári [29], this theorem is sharp up to the constant c. That is, with high probability, dense random subsets of $[n]$ contain Hilbert cubes of dimension $c'\sqrt{\log n}$ for some c' depending on the density.

Let X be a set with elements $1 \leqslant x_1 < x_2 < \cdots < x_d$ and write

$$\Sigma(X) = \left\{ \sum_{i \in I} x_i : I \subseteq [d] \right\},$$

where the sum over the empty set is taken to be zero. Note that a Hilbert cube is just a translation of an appropriate $\Sigma(X)$. The following basic estimates for $|\Sigma(X)|$ will be useful to us. For a proof, see [28].

Lemma 7.2. *For any set X with d elements,*

$$\binom{d+1}{2} + 1 \leqslant |\Sigma(X)| \leqslant 2^d.$$

The main new ingredient that we use is the following inverse Littlewood–Offord theorem due to Tao and Vu [48] (see also [38] and [49] for improved versions). This says that if $|\Sigma(X)|$ is small then X must be highly structured. A *generalized arithmetic progression* (or GAP) is a subset Q of $\mathbb{Z}$ of the form

$$Q = \{x_0 + a_1 x_1 + \cdots + a_r x_r : 1 \leqslant a_i \leqslant A_i\}.$$

We refer to r as the rank of the GAP Q and the product $A_1 \ldots A_r$ as the volume of Q. Rephrased in our terms, the inverse Littlewood–Offord theorem states that if $|\Sigma(X)|$ is small then almost all of X must be contained in a GAP Q of low rank.

Lemma 7.3. *For every $C > 0$ and $0 < \epsilon < 1$, there exist positive constants r and C' such that if X is a multiset with d elements and $|\Sigma(X)| \leqslant d^C$ then there is a GAP Q of dimension r and volume at most $d^{C'}$ such that all but at most $d^{1-\epsilon}$ elements of X are in Q.* □

The following lemma is the key step in our proof.

Lemma 7.4. *For $s \leqslant \log d$, the number of d-sets $X \subset [n]$ with $|\Sigma(X)| \leqslant 2^s d^2$ is at most*

$$n^{O(s)} d^{O(d)}.$$

Proof. Let $C = 3$, $\epsilon = 1/2$ and let r and C' be the constants given by Lemma 7.3. If a d-set X satisfies $|\Sigma(X)| \leqslant 2^s d^2 \leqslant d^3$, then, by Lemma 7.3, it has a subset $X' \subset X$ which lies in a GAP Q of dimension r and volume at most $d^{C'}$ such that $|X \setminus X'| \leqslant d^{1-\epsilon} = d^{1/2}$.

The number of possible choices for Q is $d^{O(1)} n^{O(1)}$ since it is of constant dimension. For each such Q, the number of subsets of Q of size at most d is at most $\binom{|Q|}{\leqslant d} \leqslant |Q|^d = d^{O(d)}$. Thus, the number of choices of X' is at most $d^{O(d)} n^{O(1)}$.

Let $m = |X \setminus X'|$ and order the elements of $X \setminus X'$ as $x_1 < \cdots < x_m$. Consider the nested sequence $X_0 \subset X_1 \subset \cdots \subset X_m$ of sets where $X_0 = X'$ and $X_{i+1} = X_i \cup \{x_{i+1}\}$, so $X_m = X$. If

$$|\Sigma(X_{i+1})| < 2|\Sigma(X_i)|,$$

then x_{i+1} is in the difference set $\Sigma(X_i) - \Sigma(X_i)$ and there are at most

$$|\Sigma(X_i)|^2 \leqslant |\Sigma(X)|^2 \leqslant 2^{2s} d^4 \leqslant d^6$$

such choices for x_{i+1}. Otherwise,

$$|\Sigma(X_{i+1})| = 2|\Sigma(X_i)|,$$

and there are at most n choices for x_{i+1}.

Note that, by Lemma 7.2, we have

$$|\Sigma(X')| \geqslant \binom{|X'|+1}{2} + 1 \geqslant d^2/4.$$

Suppose there are t elements x_{i+1} where $|\Sigma(X_{i+1})| = 2|\Sigma(X_i)|$. We have $t \leqslant s+2$ as

$$|\Sigma(X_0)| = |\Sigma(X')| \geqslant d^2/4 \quad \text{and} \quad 2^s d^2 \geqslant |\Sigma(X)| = |\Sigma(X_m)| \geqslant 2^t |\Sigma(X_0)|.$$

Thus, after selecting the at most $s+2$ indices i for which $|\Sigma(X_{i+1})| = 2|\Sigma(X_i)|$, we have that the number of possible d-sets $X \subset [n]$ such that $|\Sigma(X)| \leqslant 2^s d^2$ is at most

$$d^{O(d)} n^{O(1)} \binom{m}{\leqslant s+2} n^{s+2} d^{6m} = d^{O(d)} n^{O(s)}.$$

This completes the proof of the lemma. □

Proof of Theorem 7.1. Let A be the random set formed by choosing each element independently with probability δ and let E be the event that A contains a Hilbert cube of dimension $d = c\sqrt{\log n}$. We wish to show that the probability of the event E tends to zero with n.

Let m_t denote the number of d-sets $X \subset [n]$ with $|\Sigma(X)| = t$. We have

$$\begin{aligned}
\mathbb{P}[E] &\leqslant n \sum_X \delta^{|\Sigma(X)|} \\
&= n \sum_t m_t \delta^t, \\
&= n \sum_{t<d^3} m_t \delta^t + n \sum_{t \geqslant d^3} m_t \delta^t.
\end{aligned}$$

Note that the extra factor of n comes from summing over all possible first coordinates x_0 for the Hilbert cube. The total number of choices for X is at most n^d by choosing the basis elements arbitrarily. Hence, the second term is at most $n \cdot n^d \cdot \delta^{d^3} = o(1)$, where we use $d = c\sqrt{\log n}$ and c is a sufficiently large constant depending on δ.

We next bound the first term above. Letting $a_s = 2^{s-3} d^2$ and

$$n_s = \sum_{t=a_s}^{a_{s+1}-1} m_t,$$

we use Lemma 7.4, which gives $n_s \leqslant n^{O(s)} d^{O(d)}$. Hence, using a dyadic partition of the sum, the first term is at most

$$n \sum_{s=1}^{2+\log d} n_s \delta^{2^{s-3} d^2} \leqslant n \sum_{s=1}^{2+\log d} n^{O(s)} d^{O(d)} \delta^{2^{s-3} d^2} = o(1),$$

where the last estimate uses $d = c\sqrt{\log n}$ and c is a sufficiently large constant depending on δ to ensure that each of the terms in the sum is $o(n^{-2})$.

It therefore follows that $\mathbb{P}[E]$ tends to zero as n tends to infinity, completing the proof. □

By considering a random r-colouring of $[n]$ and carefully checking the dependence of the constant factor on δ in the above proof, we have the following corollary.

Corollary 7.5. *There exists a constant c such that*

$$h(d,r) \geqslant r^{cd^2}. \qquad \square$$

We will not be able to improve this further unless we can improve the bound on the van der Waerden number $W(k,r)$, the smallest number which guarantees a k-term AP in any r-colouring of $[W(k,r)]$. Indeed, the lower bound on $W(k,2)$ is exponential in k and, since a d^2-term AP contains a d-cube, $W(d^2,r) \geqslant h(d,r)$. It seems plausible that $W(d^2,2)$ and hence $h(d,2)$ grow as exponentials in d^2.

References

[1] Akiyama, J. and Watanabe, M. (1987) Maximum induced forests of planar graphs. *Graphs Combin.* **3** 201–202.

[2] Albertson, M. and Haas, R. (1998) A problem raised at the DIMACS Graph Coloring Week.

[3] Alon, N. (2003) Problems and results in extremal combinatorics I. *Discrete Math.* **273** 31–53.

[4] Alon, N. (2008) Problems and results in extremal combinatorics II. *Discrete Math.* **308** 4460–4472.

[5] Alon, N. and Krivelevich, M. (1997) Constructive bounds for a Ramsey-type problem. *Graphs Combin.* **13** 217–225.

[6] Alon, N., Mubayi, D. and Thomas, R. (2001) Large induced forests in sparse graphs. *J. Graph Theory* **38** 113–123.

[7] Balister, P., Lehel, J. and Schelp, R. H. (2006) Ramsey unsaturated and saturated graphs. *J. Graph Theory* **51** 22–32.

[8] Beck, J. (1993) Achievement games and the probabilistic method. In *Combinatorics: Paul Erdős is Eighty*, Vol. 1, Bolyai Mathematical Society, pp. 51–78.

[9] Brown, T. C., Chung, F. R. K., Erdős, P. and Graham, R. L. (1985) Quantitative forms of a theorem of Hilbert. *J. Combin. Theory Ser. A* **38** 210–216.

[10] Butterfield, J., Grauman, T., Kinnersley, W. B., Milans, K. G., Stocker, C. and West, D. B. (2011) On-line Ramsey theory for bounded degree graphs. *Electron. J. Combin.* **18** #136.

[11] Cavers, M. and Verstraëte, J. (2008) Clique partitions of complements of forests and bounded degree graphs. *Discrete Math.* **308** 2011–2017.

[12] Chang, Y. (1996) A bound for Wilson's theorem III. *J. Combin. Des.* **4** 83–93.

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

[14] Conlon, D. (2009/10) On-line Ramsey numbers. *SIAM J. Discrete Math.* **23** 1954–1963.

[15] Conlon, D. (2013) The Ramsey number of dense graphs. *Bull. London Math. Soc.* **45** 483–496.

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

[17] Dudek, A. and Mubayi, D. On generalized Ramsey numbers for 3-uniform hypergraphs. *J. Graph Theory*. Preprint.

[18] Dudek, A., Retter, T. and Rödl, V. On generalized Ramsey numbers of Erdős and Rogers. Preprint.

[19] Dudek, A. and Rödl, V. (2011) On K_s-free subgraphs in K_{s+k}-free graphs and vertex Folkman numbers. *Combinatorica* **31** 39–53.

[20] Erdős, P. and Gallai, T. (1961) On the minimal number of vertices representing the edges of a graph. *Publ. Math. Inst. Hungar. Acad. Sci.* **6** 181–202.
[21] Erdős, P. and Rado, R. (1952) Combinatorial theorems on classifications of subsets of a given set. *Proc. London Math. Soc.* **3** 417–439.
[22] Erdős, P. and Rogers, C. A. (1962) The construction of certain graphs. *Canad. J. Math.* **14** 702–707.
[23] Fox, J. and Sudakov, B. (2009) Density theorems for bipartite graphs and related Ramsey-type results. *Combinatorica* **29** 153–196.
[24] Friedgut, E., Kohayakawa, Y., Rödl, V., Ruciński, A. and Tetali, P. (2003) Ramsey games against a one-armed bandit. *Combin. Probab. Comput.* **12** 515–545.
[25] Gregory, D. A., McGuinness, S. and Wallis, W. (1986) Clique partitions of the cocktail party graph. *Discrete Math.* **59** 267–273.
[26] Grytczuk, J. A., Hałuszczak, M. and Kierstead, H. A. (2004) On-line Ramsey theory. *Electron. J. Combin.* **11** #57.
[27] Gunderson, D. S., Rödl, V. and Sidorenko, A. (1999) Extremal problems for sets forming Boolean algebras and complete partite hypergraphs. *J. Combin. Theory Ser. A* **88** 342–367.
[28] Hegyvári, N. (1996) On representation problems in the additive number theory. *Acta Math. Hungar.* **72** 35–44.
[29] Hegyvári, N. (1999) On the dimension of the Hilbert cubes. *J. Number Theory* **77** 326–330.
[30] Hilbert, D. (1892) Über die Irreduzibilität ganzer rationaler Funktionen mit ganzzahligen Koeffizienten. *J. Reine Angew. Math.* **110** 104–129.
[31] Kierstead, H. A. and Konjevod, G. (2009) Coloring number and on-line Ramsey theory for graphs and hypergraphs. *Combinatorica* **29** 49–64.
[32] Kowalik, Ł, Lužar, B. and Škrekovski, R. (2010) An improved bound on the largest induced forests for triangle-free planar graphs. *Discrete Math. Theor. Comput. Sci.* **12** 87–100.
[33] Krivelevich, M. (1994) K^s-free graphs without large K^r-free subgraphs. *Combin. Probab. Comput.* **3** 349–354.
[34] Krivelevich, M. (1995) Bounding Ramsey numbers through large deviation inequalities. *Random Struct. Alg.* **7** 145–155.
[35] Kurek, A. and Ruciński, A. (2005) Two variants of the size Ramsey number. *Discuss. Math. Graph Theory* **25** 141–149.
[36] Marciniszyn, M., Spöhel, R. and Steger, A. (2009) Online Ramsey games in random graphs. *Combin. Probab. Comput.* **18** 271–300.
[37] Marciniszyn, M., Spöhel, R. and Steger, A. (2009) Upper bounds for online Ramsey games in random graphs. *Combin. Probab. Comput.* **18** 259–270.
[38] Nguyen, H. and Vu, V. (2011) Optimal inverse Littlewood–Offord theorems. *Adv. Math.* **226** 5298–5319.
[39] Orlin, J. (1977) Contentment in graph theory. *Indag. Math.* **39** 406–424.
[40] Ramsey, F. P. (1930) On a problem of formal logic. *Proc. London Math. Soc. Ser. 2* **30** 264–286.
[41] Salavatipour, M. R. (2006) Large induced forests in triangle-free planar graphs. *Graphs Combin.* **22** 113–126.
[42] Shearer, J. B. (1995) On the independence number of sparse graphs. *Random Struct. Alg.* **7** 269–271.
[43] Skokan, J. and Stein, M. Cycles are strongly Ramsey-unsaturated. *Combin. Probab. Comput.*, to appear.
[44] Sudakov, B. (2005) A new lower bound for a Ramsey-type problem. *Combinatorica* **25** 487–498.
[45] Sudakov, B. (2005) Large K_r-free subgraphs in K_s-free graphs and some other Ramsey-type problems. *Random Struct. Alg.* **26** 253–265.
[46] Sudakov, B. (2011) A conjecture of Erdős on graph Ramsey numbers. *Adv. Math.* **227** 601–609.
[47] Szemerédi, E. (1969) On sets of integers containing no four elements in arithmetic progression. *Acta Math. Acad. Sci. Hungar.* **20** 199–245.

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

[49] Tao, T. and Vu, V. (2010) A sharp inverse Littlewood–Offord theorem. *Random Struct. Alg.* **37** 525–539.

[50] Wallis, W. D. (1985) Clique partitions of the complement of a one-factor. *Congr. Numer.* **46** 317–319.

[51] Wallis, W. D. (1990) Finite planes and clique partitions. *Contemp. Math.* **111** 279–285.

[52] Wilson, R. M. (1974) Constructions and uses of pairwise balanced designs. In *Combinatorics: Proc. NATO Advanced Study Inst., Breukelen, 1974*, Part 1: *Theory of Designs, Finite Geometry and Coding Theory*, No. 55 of Math. Centre Tracts, Math Centrum, Amsterdam, pp. 18–41.

[53] Wolfovitz, G. K_4-free graphs without large induced triangle-free subgraphs. *Combinatorica*, to appear.