Proof of a conjecture on induced subgraphs of Ramsey graphs

Matthew Kwan * Benny Sudakov†

Abstract

An n-vertex graph is called *C-Ramsey* if it has no clique or independent set of size $C \log n$. All known constructions of Ramsey graphs involve randomness in an essential way, and there is an ongoing line of research towards showing that in fact all Ramsey graphs must obey certain "richness" properties characteristic of random graphs. More than 25 years ago, Erdős, Faudree and Sós conjectured that in any C-Ramsey graph there are $\Omega(n^{5/2})$ induced subgraphs, no pair of which have the same numbers of vertices and edges. Improving on earlier results of Alon, Balogh, Kostochka and Samotij, in this paper we prove this conjecture.

1 Introduction

An induced subgraph of a graph is said to be *homogeneous* if it is a clique or independent set. A classical result in Ramsey theory, proved in 1935 by Erdős and Szekeres [18], is that every n-vertex graph has a homogeneous subgraph with at least $\frac{1}{2}\log_2 n$ vertices. On the other hand, Erdős [14] famously used the probabilistic method to prove that, for all n, there exists an n-vertex graph with no homogeneous subgraph on $2\log_2 n$ vertices. Despite significant effort (see for example [24, 8, 12, 11]), there are no non-probabilistic constructions of graphs with comparably small homogeneous sets.

For some fixed C, say an n-vertex graph is *C-Ramsey* if it has no homogeneous subgraph of size $C\log_2 n$. It is widely believed that C-Ramsey graphs must in some sense resemble random graphs, and this belief has been supported by a number of theorems showing that certain "richness" properties characteristic of random graphs hold for all C-Ramsey graphs. The first result of this type was due to Erdős and Szemerédi [19], who showed that C-Ramsey graphs have density bounded away from 0 and 1. This basic result was the foundation for a large amount of further research; over the years many conjectures have been proposed and resolved as our understanding of Ramsey graphs has improved. Improving a result of Erdős and Hajnal [16], Prömel and Rödl [32] proved that for every constant C there is $c > 0$ such that every n-vertex C-Ramsey graph contains every possible graph on $c\log_2 n$ vertices as an induced subgraph. Shelah [33] proved that every n-vertex C-Ramsey graph contains $2^{\Omega(n)}$ non-isomorphic induced subgraphs. Answering a question of Erdős, Faudree and Sós [20, 21], Bukh and Sudakov [10] showed that every n-vertex C-Ramsey graph has an induced subgraph with $\Omega(\sqrt{n})$ different degrees.

Despite this progress, there are several problems that have remained open for quite some time. Two of them deal with the variation in the numbers of edges and vertices of induced subgraphs of Ramsey graphs. For a graph G, let

$$\Phi(G) = \{e(H) : H \text{ is an induced subgraph of } G\}.$$

*Department of Mathematics, Stanford University, Stanford, CA 94305. Email: mattkwan@stanford.edu. This research was done while the author was working at ETH Zurich, and is supported in part by SNSF project 178493.

†Department of Mathematics, ETH, 8092 Zürich, Switzerland. Email: benjamin.sudakov@math.ethz.ch. Research supported in part by SNSF grant 200021-175573.

Erdős and McKay [20, 21] conjectured that for any C there is $\delta > 0$ such that for every n-vertex C-Ramsey graph G, the set $\Phi(G)$ contains the interval $\{0, \dots, \delta n^2\}$. Progress on this conjecture has come from two directions. First, Alon, Krivelevich and Sudakov [5] proved a weaker result with n^δ in place of δn^2. Second, recently Narayanan, Sahasrabudhe and Tomon [30] proposed a natural relaxation of the Erdős–McKay conjecture, that $\Phi(G)$ contains at least $\Omega(n^2)$ values (not necessarily forming an interval). They showed that $|\Phi(G)| = n^{2-o(1)}$, and in [27] we proved their conjecture that Ramsey graphs induce subgraphs of quadratically many sizes.

Next, for a graph G let

$$\Psi(G) = \{(v(H), e(H)) : H \text{ is an induced subgraph of } G\}.$$

Strengthening a conjecture of Alon and Bollobás, it was conjectured by Erdős, Faudree and Sós that for any fixed C and any n-vertex C-Ramsey graph G, we have $|\Psi(G)| = \Omega(n^{5/2})$. This problem appeared in several of Erdős' problem papers [20, 15, 21]. Of course, since $|\Psi(G)| \ge |\Phi(G)|$, our result in [27] implies that $|\Psi(G)| = \Omega(n^2)$, which was also proved much earlier by Alon and Kostochka [4]. Until now, the best progress on the Erdős–Faudree–Sós conjecture was due to Alon, Balogh, Kostochka and Samotij [1], who proved it with the exponent 2.369 in place of 5/2.

In this paper we establish the Erdős–Faudree–Sós conjecture, combining ideas from many of the aforementioned papers.

Theorem 1.1. *For any fixed $C > 0$, there is $\gamma > 0$ such that every n-vertex C-Ramsey graph G has $|\Psi(G)| = \gamma n^{5/2}$.*

As mentioned in [15], we remark that the order of magnitude $n^{5/2}$ is best-possible. This can be seen by considering a random graph $\mathbb{G}(n, 1/2)$ where each edge is present independently with probability 1/2 (it is well known that this is an $O(1)$-Ramsey graph with probability $1 - o(1)$). Briefly, one can use a concentration inequality to show that with probability $1 - o(2^n)$, the number of edges in any fixed vertex subset of $\mathbb{G}(n, 1/2)$ lies in an interval of length $O(n^{3/2})$, and by the union bound it follows that with probability $1 - o(1)$, for each $0 \le \ell \le n$ there are at most $O(n^{3/2})$ different numbers of edges among ℓ-vertex induced subgraphs. This proves that $|\Psi(\mathbb{G}(n, 1/2))| = O(n^{5/2})$ with probability $1 - o(1)$. See also [4, Section 4] for further discussion of $\Psi(\mathbb{G}(n, 1/2))$.

The rest of the paper is organised as follows. In Section 2 we give a very high-level outline of the basic ideas of our proof and briefly compare it to previous work. In Section 3 we collect a number of basic tools which will be useful for our proof (some of which are standard, and some of which are new), and in Section 4 we present the technical details of our proof. Finally, in Section 5 we discuss some potential further directions of research.

1.1 Notation and basic definitions

We use standard asymptotic notation throughout. For functions $f = f(n)$ and $g = g(n)$ we write $f = O(g)$ to mean that there is a constant C such that $|f| \le C|g|$, we write $f = \Omega(g)$ to mean there is a constant $c > 0$ such that $f \ge c|g|$ for sufficiently large n, we write $f = \Theta(g)$ to mean that $f = O(g)$ and $f = \Omega(g)$, and we write $f = o(g)$ or $g = \omega(f)$ to mean that $f/g \to 0$ as $n \to \infty$. All asymptotics are as $n \to \infty$ unless stated otherwise. Floor and ceiling symbols will be systematically omitted where they are not crucial.

For two multisets A and B, let $A \triangle B$ be the set of elements which have different multiplicities in A and B (so if A and B are ordinary sets, then $A \triangle B$ is the ordinary symmetric difference $(A \setminus B) \cup (B \setminus A)$). For a set A, we denote by $\binom{A}{k}$ the set of all k-subsets of elements of A.

We also use standard graph theoretic notation throughout. In particular, in a graph, $e(A)$ is the number of edges which are contained inside a vertex subset A, and $e(A, B)$ is the number of edges between two disjoint vertex subsets A and B. For a vertex v and a set of vertices A, we denote the set of neighbours of v in A by $N_A(v) = N(v) \cap A$ and we denote the degree of v into A by $d_A(v) = |N_A(v)|$.

We also make some less standard graph theoretic definitions that will be convenient for the proof. For a set of vertices $\boldsymbol{v} = \{v_1, \dots, v_k\}$, let $N(\boldsymbol{v})$ (respectively $N_U(\boldsymbol{v})$) be the multiset union of $N(v_1), \dots, N(v_k)$ (respectively, of $N_U(v_1), \dots, N_U(v_k)$. Let $d(v) = d(v_1) + \dots + d(v_k)$ (respectively $d_U(\boldsymbol{v}) = d_U(v_1) + \dots + d_U(v_k)$) be the size of $N(\boldsymbol{v})$ (respectively, of $N_U(\boldsymbol{v})$), accounting for multiplicity.

Finally, we remark that we will often use variable names of the form n_A to denote the size of a set A. (This is really only a convention, not a definition; we will often introduce n_A before the set A has actually been defined).

2 Discussion and main ideas of the proof

According to Erdős [20], at the time the problem was proposed, he and Sós had already proved the weaker bound that $|\Psi(G)| = \Omega(n^{3/2})$ for $O(1)$-Ramsey graphs. In fact, there are at least two reasonably simple ways to prove this weak bound, and both are instructive for our proof. To describe these, we define

$$\Psi(\ell, G) = \{e(H) : H \text{ is an } \ell\text{-vertex induced subgraph of } G\}.$$

To prove that $|\Psi(G)| = \Omega(n^{3/2})$, it suffices to prove that $|\Psi(\ell, G)| = \Omega(\sqrt{n})$ for each of $\Omega(n)$ different choices of ℓ.

One way to do this, described by Alon and Kostochka [4], is to use a discrepancy theorem and a switching argument. Erdős, Goldberg, Pach and Spencer [22] proved that in any n-vertex graph G with density bounded away from 0 and 1, and any $\alpha \in (0, 1)$ bounded away from 0 and 1, there are two induced subgraphs $G[W^-]$ and $G[W^+]$, with $|W^-| = |W^+| = \alpha n$, such that $e(W^+) - e(W^-) = \Omega(n^{3/2})$. Recalling the Erdős–Szemerédi theorem that $O(1)$-Ramsey graphs have density bounded away from 0 and 1, we can find such W^- and W^+ in any n-vertex $O(1)$-Ramsey graph G. One can then obtain a sequence of induced subgraphs $G[W_0], \dots, G[W_{\alpha n}]$ by starting with $W_0 = W^-$ and switching vertices one-by-one from W^- into W^+. Formally, fix an ordering $w_1^-, \dots, w_{\alpha n}^-$ of W^- and an ordering $w_1^+, \dots, w_{\alpha n}^+$ of W^+ and let

$$W_i = \{w_1^-, \dots, w_{\alpha n - i}^-\} \cup \{w_1^+, \dots, w_i^+\}.$$

Then, we have $|e(G[W_i]) - e(G[W_{i-1}])| = \left|d_{W_i}(w_i^+) - d_{W_{i-1}}(w_{\alpha n - i + 1}^-)\right| \le \alpha n$, so as $e(G[W_i])$ varies over an interval of length $\Omega(n^{3/2})$, it must attain $\Omega(\sqrt{n})$ different values. This proves $|\Psi(\alpha n, G)| = \Omega(\sqrt{n})$, and we can apply this fact for $\Omega(n)$ different choices of $\alpha = \ell/n$, proving that $|\Psi(G)| = \Omega(n^{3/2})$. We remark that this basic approach was refined by Alon and Kostochka [4] and by Alon, Balogh, Kostochka and Samotij [1], to prove stronger bounds.

A second completely different way to prove that $|\Psi(G)| = \Omega(n^{3/2})$, due to Bukh and Sudakov [10, Proposition 3.1], is to make use of the fact that Ramsey graphs have induced subgraphs with many distinct degrees. Specifically, what Bukh and Sudakov proved was that in any $O(1)$-Ramsey graph, there is an induced subgraph with $\Omega(n)$ vertices which is *diverse* in the sense that most pairs of vertices have very different neighbourhoods (to be precise, the symmetric difference of their neighbourhoods has size $\Omega(n)$). In an n'-vertex diverse graph (with $n' = \Omega(n)$), consider a random subset U of $\alpha n'$ vertices (with $\alpha \in (0,1)$ bounded away from 0 and 1). By the diversity assumption, for most pairs of vertices u, v their degrees $d_U(u)$, $d_U(v)$ into U are not too strongly correlated, and the probability they are exactly equal turns out to be $O(1/\sqrt{n})$. (A simple intuitive reason for this probability is that $d_U(u) - d_U(v)$ is approximately normally distributed with standard deviation $\Theta(\sqrt{n})$). A simple linearity-of-expectation argument then shows that there is an outcome of $G[U]$ with $\Omega(\sqrt{n})$ different degrees. Finally, given an $\alpha n'$-vertex graph with $\Omega(\sqrt{n})$ different degrees, we can obtain $(\alpha n' - 1)$-vertex graphs with $\Omega(\sqrt{n})$ different numbers of edges, simply by choosing different vertices to delete. This proves that $|\Psi(\alpha n' - 1, G)| = \Omega(\sqrt{n})$, and again applying this fact for $\Omega(n)$ different choices of $\alpha = \ell/n'$, it follows that $|\Psi(G)| = \Omega\big(n^{3/2}\big)$.

Observe that both the approaches described above seem to be somewhat complementary. The discrepancy/switching argument, in its most basic form, gives $\Omega(\sqrt{n})$ different values of $e(G[U])$ that are distributed fairly evenly over a range of length $\Omega\big(n^{3/2}\big)$. On the other hand, the diversity/anticoncentration argument gives $\Omega(\sqrt{n})$ values of $e(G[U])$ contained in an interval of length $O(n)$. It is natural to try to combine both types of arguments to obtain better bounds.

In fact, recent developments bounding $|\Phi(G)|$ due to Narayanan, Sahasrabudhe and Tomon [30], and ourselves [27], make this idea seem even more promising. In [30], the authors made the simple observation (using the pigeonhole principle) that in any n-vertex graph G, there is a set A of $\sqrt{n}$ vertices with degrees lying in an interval of length $\sqrt{n}$. If G is diverse, and U is a random vertex set of linear size, then the degrees $d_U(x)$, for $x \in A$, are likely to take $n^{1/2-o(1)}$ different values, very tightly packed in an interval of length $O(\sqrt{n})$. By augmenting U with different combinations of vertices in A, we can obtain subgraphs of many different sizes, all lying in a fixed interval of length $O(n)$. Adapting these ideas to our context, and using the further refinements in [27], one can prove that we can actually obtain $\Omega(n)$ values of $e(G[U \cup Y])$ among subsets $Y \subseteq A$ of a certain fixed size, tightly packed in an interval of length $O(n)$.

So, as a rough plan to prove Theorem 1.1, one might start with vertex subsets W^-, W^+ of fixed size $\ell = \Theta(n)$ such that $e(W^+) - e(W^-) = \Omega\big(n^{3/2}\big)$, provided by a discrepancy theorem. We would then switch between W^- and W^+ to obtain subsets $W_1, \dots, W_t$ such that among the $e(W_i)$ there are $\Omega(\sqrt{n})$ different values $e(W_{i_1}), e(W_{i_2}), \dots$ each separated by a distance of $\Omega(n)$. One might then hope to somehow use diversity and anticoncentration to show that each such W_{i_j} has an "augmenting set" A_j such that $e\big(W_{i_j} \cup Y\big)$ takes $\Omega(n)$ different values as Y varies over subsets of A_j with some fixed size $f(n)$. We would moreover hope that for each j, the augmented values $e\big(W_{i_j} \cup Y\big)$ fall in a specific interval of length $O(n)$ that does not intersect the corresponding interval for any other j. This would prove that $|\Psi(\ell + f(n), G)| = \Omega\big(n^{3/2}\big)$, and this fact could be applied for $\Omega(n)$ different choices of ℓ to prove that $|\Psi(G)| = \Omega\big(n^{5/2}\big)$.

There are several serious challenges associated with this kind of approach. First, we need some way to introduce a random set U of linear size in order to use anticoncentration for our augmenting sets. We have very little control over the number of edges in such a random set (this number has variance $\Theta\big(n^3\big)$), so it seems we must use the same random set for each W_i, and apply our switching argument *after* our random set has been exposed. However, it seems that doing this would introduce new complications: the anticoncentration probabilities we are interested in are of order $O(1/\sqrt{n})$, which is not small enough to apply the union bound over all i, given a single source of randomness. (It does not suffice to prove things for *most* i, because the subsequence (i_j) arising from the switching argument comprises a negligible fraction of all i).

Our approach is to first prepare vertex sets U^0, W^-, W^+, each of a certain linear size, such that

$$\big(e\big(W^+\big) + \alpha e\big(W^+, U^0\big)\big) - \big(e\big(W^-\big) + \alpha e\big(W^-, U^0\big)\big) = \Omega\Big(n^{3/2}\Big),$$

for some $\alpha \in (0,1)$. Then, as above, we switch between W^- and W^+ to obtain a sequence of sets W_i, and identify a well-separated subsequence of $\Omega(\sqrt{n})$ sets W_{i_j} such that

$$\big(e\big(W_{i_j}\big) + \alpha e\big(W_{i_j}, U^0\big)\big) - \big(e\big(W_{i_{j-1}}\big) + \alpha e\big(W_{i_{j-1}}, U^0\big)\big) = \Omega(n)$$

for each j. Only then do we choose a random subset $U \subseteq U^0$ of size $\alpha\big|U^0\big|$, which we may use for anticoncentration. By construction, the $e\big(W_{i_j} \cup U\big)$ are well-separated in expectation, and the added randomness does not too severely disturb the increments $e(W_i \cup U) - e(W_{i-1} \cup U)$. Because we do not have any real control over the spacing of the i_j, we must additionally carefully compensate for the buildup of deviations caused by "large gaps" between the i_j.

Of course, before we even expose the random set U we need to decide which vertices should be in the augmenting sets A_j. Recall that we would like to be able to use anticoncentration to obtain $\Omega(n)$ subgraph sizes of the form $e\big(W_{i_j} \cup U \cup Y\big)$, for $Y \subseteq A_j$ of a fixed size. Provided that we have been carefully maintaining appropriate diversity properties through the construction, the only real requirement for this is that the A_j are sufficiently large (of size at least $\Omega(\sqrt{n})$). However, ensuring that the different A_j do not "interfere" with each other is a much more delicate task. With the pigeonhole principle, for each j we can show that there are $\sqrt{n}$ vertices v such that each $d_{W_{i_j}}(v) + \alpha d_{U^0}(v)$ is contained in an interval I_j of length $\sqrt{n}$, and we might hope to use such a set of vertices as our augmenting set A_j. However, the pigeonhole principle gives us no guarantee of "consistency" between different j, and it might happen that the intervals I_j jump around in such a way that there is a lot of overlap between the augmented values $e\big(U \cup W_{i_j} \cup Y\big)$ for different j. It seems to be quite difficult to carefully choose the A_j in such a way that the I_j are well-behaved.

Instead, we sidestep this issue, with the insight that it is not actually necessary for all the vertices in A_j to have similar degrees into $W_{i_j} \cup U$; it suffices that A_j has a large hypergraph matching $M_j \subseteq \binom{A_j}{k}$, such that the *sums* $d_{W_{i_j} \cup U}(\boldsymbol{v}) = d_{W_{i_j} \cup U}(v_1) + \cdots + d_{W_{i_j} \cup U}(v_k)$ are similar for each $\boldsymbol{v} = \{v_1, \ldots, v_k\} \in M_j$. We may treat the edges of M_j as we would treat single vertices, forming our augmented values $e\big(U \cup W_{i_j} \cup \bigcup_{\boldsymbol{v} \in Z} \boldsymbol{v}\big)$ from subsets $Z \subseteq M_j$.

Being able to use a hypergraph matching instead of a set of vertices affords us a lot of flexibility. For fixed $K \in \mathbb{N}$, there are $\Omega\big(n^K\big)$ sets of K vertices, and by the pigeonhole principle, $\Omega\big(n^{K-3}\big)$ of these K-sets $\boldsymbol{v}$ have exactly the same values of $d_{W^-}(\boldsymbol{v})$, the same values of $d_{W^+}(\boldsymbol{v})$ and the same values of $d_{U^0}(\boldsymbol{v})$. If we obtain the W_i by switching *randomly* between W^- and W^+, then we can show that the degrees $d_{W_i}(\boldsymbol{v})$ are concentrated around a certain convex combination of $d_{W^-}(\boldsymbol{v})$ and $d_{W^+}(\boldsymbol{v})$. In this way we can produce a collection of K-sets $\boldsymbol{v}$ such that the degrees $d_{W_{i_j} \cup U}(\boldsymbol{v})$ are quite well-behaved.

Of course, these K-sets are not disjoint, but for large K we may apply a weak form of the *sunflower lemma* of Erdős and Rado, to produce a hypergraph matching $M \subseteq \binom{V}{k}$ (with $k \leq K$) of almost linear size, which has similarly well-behaved degrees. With this as a starting point, it becomes feasible to use the pigeonhole principle to obtain appropriate sub-matchings $M'_j \subseteq M$, and modulo a lot of technical details we are able to more or less implement the plan described above. To summarise, for each of $\sqrt{n}$ choices of j we use anticoncentration and a generalised notion of diversity to produce $\Omega(n)$ values of $e\big(U \cup W_{i_j} \cup \bigcup_{\boldsymbol{v} \in Z} \boldsymbol{v}\big)$ among $Z \subseteq M'_j$ of a certain fixed size, in such a way that there is little overlap between the values for different j. This gives us $\Omega\big(n^{3/2}\big)$ subgraphs with the same number of vertices and different numbers of edges, and varying the size of U allows us to prove that $|\Psi(G)| = \Omega\big(n^{5/2}\big)$, as desired.

3 Basic tools

3.1 Diverse neighbourhoods in Ramsey graphs

In [10] Bukh and Sudakov introduced the notion of *diversity*: an n-vertex graph is said to be diverse if $|N(x)\triangle N(y)| = \Omega(n)$ for most pairs of distinct vertices x, y. We will need a slightly stronger notion than diversity, which we introduced in [27]. Say an n-vertex graph is (δ,ε)-*rich* if for any vertex subset W with $|W| \geq \delta n$, at most $n^{1/5}$ vertices v have $|N(v)\cap W| < \varepsilon|W|$ or $\left|\overline{N(v)}\cap W\right| < \varepsilon|W|$. Note that a graph which is (δ,ε)-rich is also (δ',ε)-rich, if $\delta' > \delta$. We remark that a slightly different definition of richness appeared in the published version of this paper, which was not quite suitable for our application. We thank Mantas Baksys and Xuanang Chen for bringing this to our attention. The next lemma appears as [27, Lemma 4], showing that Ramsey graphs contain large rich induced subgraphs.

Lemma 3.1. *For any $C, \delta > 0$, there exist $\varepsilon = \varepsilon(C) > 0$ and $c = c(C,\delta) > 0$ and $n_0 = n_0(\delta)$ such that if $n \geq n_0$ then every n-vertex C-Ramsey graph contains a (δ,ε)-rich induced subgraph on at least cn vertices.*

In [27], the reason we introduced (δ,ε)-richness was to derive a type of diversity for pairs of vertices. Here we will need a type of diversity for larger sets of vertices. (recall from Section 1.1 the non-standard multiset definitions of $N(\boldsymbol{x}), N(\boldsymbol{y})$ and $N(\boldsymbol{x})\triangle N(\boldsymbol{y})$).

Lemma 3.2. *Fix $k \in \mathbb{N}$ and let G be a (δ,ε)-rich graph on an n-vertex set V. Then, for each $\boldsymbol{x} \in \binom{V}{k}$ with $\left|\bigcap_{v\in\boldsymbol{x}} N(v)\right| \geq \delta n$, one cannot find a collection of $n^{1/5}$ vertex subsets $\boldsymbol{y} \in \binom{V}{k}$ (disjoint from $\boldsymbol{x}$ and each other) such that $|N(\boldsymbol{x})\triangle N(\boldsymbol{y})| < \delta\varepsilon n$.*

Proof. Let $W = \bigcap_{v\in\boldsymbol{x}} N(v)$. Suppose the statement of the lemma were false, and such a collection Y of vertex subsets existed. Then, for each $y \in \boldsymbol{y}$, for $\boldsymbol{y} \in Y$, we would have $\left|\overline{N(y)}\cap W\right| \leq |N(\boldsymbol{x})\triangle N(\boldsymbol{y})| < \varepsilon|W|$, and the set of all such y would contradict (δ,ε)-richness. $\square$

Lemma 3.2 only applies to $\boldsymbol{x} \in \binom{V}{k}$ such that $\bigcap_{v\in\boldsymbol{x}} N(v)$ is large. In order to apply it, we next show that in a rich graph, $\bigcap_{v\in\boldsymbol{x}} N(v)$ is large for almost all $\boldsymbol{x} \in \binom{V}{k}$.

Lemma 3.3. *Fix $k \in \mathbb{N}$ and let G be a (δ,ε)-rich graph on an n-vertex set V, for $\delta \leq \varepsilon^{k-1}$. Then there are at most $n^{k-1+1/5}$ subsets $\boldsymbol{v} \in \binom{V}{k}$ such that $\left|\bigcap_{v\in\boldsymbol{v}} N(v)\right| < \varepsilon^k n$.*

Proof. We will prove by induction that there are at most $qn^{q-1+1/5}$ "bad" ordered q-tuples $\boldsymbol{v} \in V^q$ such that $\left|\bigcap_{v\in\boldsymbol{v}} N(v)\right| < \varepsilon^q n$, for all $1 \leq q \leq k$. This will prove that there are at most $kn^{k-1+1/5}/k! \leq n^{k-1+1/5}$ subsets $\boldsymbol{v} \in \binom{V}{k}$ such that $\left|\bigcap_{v\in\boldsymbol{v}} N(v)\right| < \varepsilon^k n$.

First note that the base case $q = 1$ follows directly from (δ,ε)-richness, with $W = V$. Then, assume for induction that our desired bound holds for $q-1$; we will prove it for q. First, there are at most $(q-1)n^{q-1+1/5}$ bad q-tuples obtained by appending a vertex to a bad $(q-1)$-tuple. Then, for each $(q-1)$-tuple $\boldsymbol{v}$ which is not bad (meaning $\left|\bigcap_{v\in\boldsymbol{v}} N(v)\right| \geq \varepsilon^{q-1}n$), by (δ,ε)-richness there are at most $n^{1/5}$ vertices w with $\left|N(w)\cap\bigcap_{v\in\boldsymbol{v}} N(v)\right| < \varepsilon\left|\bigcap_{v\in\boldsymbol{v}} N(v)\right|$, meaning that there are at most $n^{q-1+1/5}$ bad-q-tuples that can be obtained by appending a vertex to a not-bad $(q-1)$-tuple, and at most $qn^{q-1+1/5}$ bad q-tuples total. $\square$

3.2 Tools from extremal (hyper)graph theory

We will make frequent use of Turán's theorem to find large independent sets in various auxiliary graphs. The following form of the theorem appears, for example, in [6].

Proposition 3.4. *Every n-vertex graph G contains an independent set of size at least*

$$\sum_{v\in V(G)} \frac{1}{d(v)+1} \geq \frac{n^2}{\sum_{v\in V(G)}(d(v)+1)} = \Omega\left(\min\left\{n, \frac{n^2}{e(G)}\right\}\right).$$

Next, a *sunflower* in a hypergraph is a subgraph in which every pair of edges has the same intersection (this common intersection is called the *kernel*, and removing the kernel from each edge gives the *petals*). We will need the following weak form of the Erdős–Rado sunflower lemma [17], which one can easily prove by induction on the uniformity of a hypergraph.

Lemma 3.5. *Fix $k \in \mathbb{N}$ and let H be a k-uniform hypergraph with m edges. Then H contains an $\Omega\big(m^{1/k}\big)$-edge sunflower.*

3.3 Probabilistic tools

We will need concentration and anticoncentration inequalities for random variables arising from random subsets of given sizes. Say a random variable X is of (n,p,b)-*hypergeometric* type if it can be expressed in the form $X = \sum_{i\in I} a_i$, where $a_1,\dots,a_n \in \mathbb{R}$ are fixed, $|a_i| \leq b$ for each i, and I is a uniformly random subset of $\{1,\dots,n\}$ of size pn. The following concentration lemma follows directly from [25, Corollary 2.2].

Lemma 3.6. *Suppose X is of (n,p,b)-hypergeometric type. Then, for any $t \in \mathbb{R}$,*

$$\Pr(|X - \mathbb{E}X| \geq t) = \exp\left(-\Omega\left(\frac{t^2}{nb^2\min\{p, 1-p\}}\right)\right).$$

Next, say that X as above (of (n,p,b)-hypergeometric type) is of $(n,p,b,r)^*$-hypergeometric type if moreover $|a_i| \geq 1/b$ for at least r indices i (that is, many a_i are bounded away from zero as well as being bounded in size). The following central limit theorem directly follows from a classical quantitative central limit theorem first proved by Bikelis [9] (see also [26]).

Lemma 3.7. *Fix $b > 0$ and suppose X is of $(n,p,b,n/b)^*$-hypergeometric type, with $|\mathbb{E}X| \leq n/(2b^2)$. Let F be the distribution function of $(X-\mathbb{E}X)/\sqrt{\operatorname{Var} X}$ and let G be the standard Gaussian distribution function. Then for all $z \in \mathbb{R}$,*

$$|F(z) - G(z)| = O\left(\frac{1}{\sqrt{p(1-p)n}}\right).$$

We only need Lemma 3.7 for anticoncentration, so we state a simple corollary for later use.

Lemma 3.8. *Suppose X is of $(n,p,O(1),\Omega(n))^*$-hypergeometric type. Then, for any $-\sqrt{n} < x < \sqrt{n}$,*

$$\Pr(X = x) = O\left(\frac{1}{\sqrt{p(1-p)n}}\right).$$

Proof. If say $|\mathbb{E}X| \le n^{2/3}$ then the desired result follows from Lemma 3.7. Otherwise, the desired result follows from Lemma 3.6, since with probability $1 - e^{-\Omega(n^{1/3})}$, X does not even fall in the interval between $-\sqrt{n}$ and $\sqrt{n}$. $\square$

We also make the following simple observation, which will be convenient to show that various discrepancy properties we are able to establish will persist with positive probability through certain kinds of random sampling. If a random variable is of $(n, 1/2, b)$-hypergeometric type for some n and b, say it is of $(1/2)$-hypergeometric type.

Lemma 3.9. *Suppose X is of $(1/2)$-hypergeometric type. Then, $X - \mathbb{E}X$ has the same distribution as $\mathbb{E}X - X$, and in particular, $X \ge \mathbb{E}X$ with probability at least $1/2$.*

Proof. Suppose that $X = \sum_{i\in I} a_i$, and let $X = \sum_{i\notin I} a_i$. Since I is a random subset of exactly half the indices $\{1, \dots, n\}$, it has the same distribution as its complement $\overline{I}$, so X has the same distribution as X'. But observe that

$$\big(X + X'\big)/2 = \sum_{i=1}^{n} a_i/2 = \mathbb{E}X = \mathbb{E}X',$$

so $\mathbb{E}X - X = X' - \mathbb{E}X'$. $\square$

We remark that Lemma 3.6 (respectively Lemma 3.9) trivially remains true when the relevant random variables are translated by a fixed constant. We will therefore frequently abuse notation and say that translations of random variables of (n, p, b)-hypergeometric type (respectively $(1/2)$-hypergeometric type) are themselves of (n, p, b)-hypergeometric type (respectively $(1/2)$-hypergeometric type).

Throughout the proof we will also frequently use Markov's inequality; the statement and proof can be found, for example, in [6].

3.4 Switching analysis

In this subsection we collect some simple lemmas that will be useful for tracking how certain parameters change as we gradually switch from one vertex subset to another. First, we show that if we move between two distant values, and most of the incremental steps are not too extreme, then there are many intermediate steps with "well-separated" values.

Lemma 3.10. *Consider a sequence $p_0, \dots, p_\tau$ with $p_\tau - p_0 \ge \lambda$. Let $\Delta_i = p_i - p_{i-1}$ and suppose that for some ρ we have*

$$\sum_{i:\Delta_i > \rho} \Delta_i \le \kappa.$$

Then, for any $\sigma \le \rho$ there is an increasing subsequence $0 = i_1, \dots, i_s = \tau$, with $s \ge \lambda/(\rho + \sigma) - \kappa/\rho$, such that $p_{i_j} - p_{i_{j-1}} \ge \sigma$ for all $1 \le j \le s$.

Proof. We view p_i as the position of a "particle" at "time" i. In the interval from p_0 to p_τ , consider $\lambda/(\rho + \sigma)$ sub-intervals of length ρ separated by a distance of at least σ, with the first sub-interval containing p_0 and the last containing p_τ. We say a sub-interval I is "further" than a sub-interval I' if I is closer to p_τ than I'.

Let $i_1 = 0$ and let I_1 be the sub-interval containing p_0. For $j > 1$ let $i_j > i_{j-1}$ be the first time i that p_i is in a sub-interval further than I_{j-1} and let I_j be this sub-interval. This process terminates when there is no sub-interval further than I_j (let $s = j$ for this value of j, and redefine $i_s = \tau$). Observe that at most κ/ρ intervals were skipped, so $s \ge \lambda/(\rho + \sigma) - \kappa/\rho$. $\square$

Next, the following lemma shows that if an ensemble of values move slowly in a bounded region, then at least one value "follows the crowd" for quite a long time.

Lemma 3.11. *Consider an interval $I \subseteq \mathbb{Z}$ with $|I| = \lambda$, and consider a "time horizon" $\tau \in \mathbb{N}$. Consider a set of "particles" R, and for each $a \in R$ let $p_i(a) \in I$ represent the "position" of a at time i, in such a way that $|p_i(a) - p_{i-1}(a)| \leq \rho$ for each $0 < i \leq \tau$ (that is, the particles move with "speed" at most ρ). For $\sigma, \mu > 0$, say a particle a is* lonely *at time $0 \leq i \leq \tau$ if*

$$|\{b \in R : |p_i(b) - p_i(a)| \leq \sigma\}| < \mu.$$

(That is, a particle is lonely if there are few other particles close to it). Now, if $\tau \leq |R|\sigma^2/(8\mu\rho\lambda)$ then there is a particle a which is never lonely.

Proof. Say a particle a is *crowded* if

$$|\{b \in R : |p_i(b) - p_i(a)| \leq \sigma/2\}| \geq \mu.$$

At any time i, fewer than $2\mu\lambda/\sigma$ particles are not crowded. To see this, divide I into $2\lambda/\sigma$ sub-intervals of length $\sigma/2$. If a sub-interval contains at least μ particles then all particles in that sub-interval are crowded.

Now, if a particle is crowded for every time $j\sigma/(4\rho)$ (among $j \in \mathbb{N}$ with $j \leq 4\rho\tau/\sigma$), then it is never lonely. To see this, observe that it takes at least $\sigma/(4\rho)$ time steps for a crowded particle to become lonely. This is because the separation between that particle and the particles within distance $\sigma/2$ must increase by $\sigma/2$, and if two particles are moving away from each other their separation increases by at most 2ρ per time step. By this fact and the preceding paragraph, there are fewer than $(4\rho\tau/\sigma)(2\mu\lambda/\sigma)$ particles that are ever lonely, and if $8\rho\tau\mu\lambda/\sigma^2 \leq |R|$ then there is a particle that is never lonely. □

4 Proof of Theorem 1.1

As in Section 2, define

$$\Psi(\ell, G) = \{e(H) : H \text{ is an } \ell\text{-vertex induced subgraph of } G\}.$$

As discussed in Section 2, in the previous bounds on $|\Psi(G)|$ in [10, 4, 1], the approach was to show that $\Psi(\ell, G)$ is large for each of $\Omega(n)$ specific choices of ℓ. In this paper it will be convenient to have slightly more flexibility: we show that $\Psi(\ell', G)$ is large for $\Omega(n)$ different choices of ℓ', but we do not specify precisely which choices they are. We will prove the following lemma, which suffices to prove Theorem 1.1.

Lemma 4.1. *For any fixed C, there is $c > 0$ such that the following holds. For any n-vertex C-Ramsey graph G, there are $f, h \in \mathbb{N}$ such that for any $cn \leq \ell \leq 2cn$, either $|\Psi((\ell - f) + h, G)| = \Omega\big(n^{3/2}\big)$ or $|\Psi(2(\ell - f) + h, G)| = \Omega\big(n^{3/2}\big)$.*

Proof of Theorem 1.1 given Lemma 4.1. We have

$$|\Psi(G)| \geq \frac{1}{2}\sum_{\ell=cn}^{2cn}(|\Psi((\ell - f) + h, G)| + |\Psi(2(\ell - f) + h, G)|) = \Omega\Big(n^{5/2}\Big). \qquad \square$$

The first ingredient for the proof of Lemma 4.1 will be the following lemma asserting the existence of a collection of vertex sets with certain discrepancy, regularity and diversity properties.

Lemma 4.2. *For any fixed C, there are $K \in \mathbb{N}$ and $c > 0$ such that the following holds. For any n-vertex C-Ramsey graph G, any $\alpha = \alpha(n) \geq 1/2$ and any $cn \leq \ell \leq 2cn$, there are disjoint vertex sets W^-, W^+, U^0, A, and a k-uniform hypergraph perfect matching $M \subseteq \binom{A}{k}$ of A for some $k \leq K$, satisfying the following properties.*

1. *$|W^-| = |W^+| = cn$, $|A| = \Omega(n^{3/4})$, and either $|U^0| = \ell$ or $|U^0| = 2\ell$;*
2. *$\left(e(W^+) + \alpha e(U^0, W^+)\right) - \left(e(W^-) + \alpha e(U^0, W^-)\right) = \Omega(n^{3/2})$;*
3. *there are $d_{W^-}, d_{W^+}, d_{U^0} \in \mathbb{N}$ such that $d_{W^-}(\boldsymbol{v}) = d_{W^-}$, $d_{W^+}(\boldsymbol{v}) = d_{W^+}$ and $d_{U^0}(\boldsymbol{v}) = d_{U^0}$ for all $\boldsymbol{v} \in M$;*
4. *for each $\{\boldsymbol{x}, \boldsymbol{y}\} \in \binom{M}{2}$ we have $|N_{U^0}(\boldsymbol{x}) \triangle N_{U^0}(\boldsymbol{y})| = \Omega(n)$.*

(Here, the implied constants in all asymptotic notation depend on C but not α).

We will prove Lemma 4.2 in Section 4.1. We remark that our proof can be easily modified to give $|A| = \Omega(n^{1-\eta})$ for any $\eta > 0$, and all that we actually need for the proof of Lemma 4.1 is that $|A| = \Omega(n^{1/2+\eta})$ for some $\eta > 0$. The choice of the exponent $3/4$ is merely for concreteness.

The next ingredient is the following lemma, showing that with positive probability we can augment a random set of vertices in many different ways to get induced subgraphs with many different numbers of edges.

Lemma 4.3. *Consider any $n_D = n_D(n) \in \mathbb{N}$ with $n_D = \omega(\log n)$, and suppose in a graph G we have disjoint vertex subsets W, A, U^0 and a hypergraph perfect matching $M \subseteq \binom{A}{k}$ for some $k = O(1)$, satisfying the following properties.*

1. *$|U^0| \geq 3n_D$, and $|M| = \Omega(\sqrt{n_D})$;*
2. *$|N_{U^0}(\boldsymbol{x}) \triangle N_{U^0}(\boldsymbol{y})| = \Omega(|U^0|)$ for each $\{\boldsymbol{x}, \boldsymbol{y}\} \in \binom{M}{2}$;*
3. *there are $d_W, d_{U^0} \in \mathbb{N}$ such that $d_{U^0}(\boldsymbol{v}) = d_{U^0}$ and $d_W(\boldsymbol{v}) = d_W + o(\sqrt{n_D})$ for all $\boldsymbol{v} \in M$.*

Then, there are $B = O(1)$ and $\delta = \Omega(1)$ (depending on the implied constants in the above asymptotic notation, but not depending on n_D) such that the following holds. Consider any $n_Z \leq \delta\sqrt{n_D}$, let D be a uniformly random subset of n_D elements of U^0, let $U = U^0 \setminus D$ and define α to satisfy $n_D = (1-\alpha)|U^0|$. With probability at least $1/4$,

$$\left|\left\{ e\left(W \cup U \cup \bigcup_{\boldsymbol{z} \in Z} \boldsymbol{z}\right) :\ Z \subseteq M,\ |Z| = n_Z,\ \left| e\left(U, \bigcup_{\boldsymbol{z} \in Z} \boldsymbol{z}\right) - \alpha n_Z d_{U^0} \right| \leq B n_D \right\}\right| = \Omega(n_Z \sqrt{n_D}).$$

We will prove Lemma 4.3 in Section 4.2, using some ideas from [27, 30]. To interpret its conclusion in words, it says that one can obtain $\Omega(n_Z\sqrt{n_D})$ induced subgraphs with different numbers of edges, by augmenting $W \cup U$ with different subsets $Z \subseteq M$ of size n_Z. Moreover, this is still true if we restrict our attention to those subsets Z such that there are about the expected number of edges $\alpha n_Z d_{U^0}$ between U and Z.

Finally, we show how to combine Lemma 4.2 and Lemma 4.3 to prove Lemma 4.1.

Proof of Lemma 4.1. Apply Lemma 4.2 with $\alpha = (\ell - c'n)/\ell$, for some small c' (depending on c) that will be chosen later to satisfy certain inequalities. Until we finally determine the value of c', the constants implied by all asymptotic notation in this section will be independent of c' (that is, if say

$f \le c'n$, we may write $f = O(c'n)$ but not $f = O(n)$). Choose n_D to satisfy $n_D = (1-\alpha)|U^0|$ (so $n_D = c'n$ or $n_D = 2c'n$, and in particular $n_D \le 2c'n$). Let $n_W = cn$ and consider uniformly random orderings $w_1^-, \dots, w_{n_W}^-$ of W^- and $w_1^+, \dots, w_{n_W}^+$ of W^+. For $0 \le i \le n_W$ let

$$W_i^- = \left\{w_1^-, \dots, w_{n_W - i}^-\right\}, \quad W_i^+ = \left\{w_1^+, \dots, w_i^+\right\}, \quad W_i = W_i^- \cup W_i^+.$$

This means each individual W_i^- (respectively W_i^+) is a uniformly random subset of $n_W - i$ elements of W^- (respectively, i elements of W^+). Define

$$d_{W_i^-} = \frac{n_W - i}{n_W} d_{W^-}, \quad d_{W_i^+} = \frac{i}{n_W} d_{W^+}, \quad d_{W_i} = d_{W_i^-} + d_{W_i^+}.$$

Now, for each $0 \le i \le n_W$ and $\boldsymbol{v} \in M$, the random variable $d_{W_i^-}(\boldsymbol{v})$ (respectively $d_{W_i^+}(\boldsymbol{v})$) is of $(n_W, p, O(1))$-hypergeometric type, for $p = (n_W - i)/n_W$ (respectively, for $p = i/n_W$), and has mean $d_{W_i^-}$ (respectively, mean $d_{W_i^+}$). By Lemma 3.6 (with $t = \sqrt{n}\log n$) and the union bound, we can fix an outcome of the orderings $w_1^-, \dots, w_{n_W}^-$ and $w_1^+, \dots, w_{n_W}^+$ such that $|d_{W_i}(\boldsymbol{v}) - d_{W_i}| \le \sqrt{n}\log n$ for each $0 \le i \le n_W$ and $\boldsymbol{v} \in M$. (Note that if $p = o(1)$ the estimate in Lemma 3.6 only becomes stronger).

This concentration would suffice to prove an approximate version of Theorem 1.1, that $|\Psi(G)| = n^{5/2}/\log^{O(1)} n$ (simply using the single matching M to augment each W_i). However, in order to obtain an exact result we need to eliminate the logarithmic factor in the estimate for $|d_{W_i}(\boldsymbol{v}) - d_{W_i}|$. We have the freedom to do this because M (coming from Lemma 4.2, of size $\Omega(n^{3/4})$) is much larger than the necessary size $\Theta(\sqrt{n})$ of our "augmenting sets" (as outlined in Section 2). In fact, we could use the pigeonhole principle to easily show that for each i there is a subset $M_i \subseteq M$ of size $\Theta(\sqrt{n})$ such that the degrees $d_{W_i}(\boldsymbol{v})$, for $\boldsymbol{v} \in M_i$, are contained in a tiny interval of length only $O(n^{1/4}\log n)$, centered at some point d_i. But because we require consistency between the i (in particular, we do not want the d_i to vary too much), things are a bit more delicate, and we will apply Lemma 3.11.

Claim 4.4. *There are $d_0, \dots, d_{n_W} \in \mathbb{N}$ and $M_1, \dots, M_{n_W} \subseteq M$ such that the following hold.*

(i) Each $|M_i| \ge \sqrt{n}$;

(ii) For each $0 \le i \le n_W$ and each $\boldsymbol{v} \in M_i$, we have $|d_{W_i}(\boldsymbol{v}) - d_i| = o(\sqrt{n})$;

(iii) For each $0 < i \le n_W$ we have $|d_i - d_{i-1}| = O(\sqrt{n}\log n)$, and actually $|d_i - d_{i-1}| = o(\sqrt{n})$ for all but $O\big(n^{1/4}\log^3 n\big)$ indices i.

Proof. The indices i will represent points in time. Let $\mu = \sqrt{n}$, let $\sigma = \sqrt{n}/\log n = o(\sqrt{n})$, let $\lambda = 2\sqrt{n}\log n$ and let I be the interval of integers between $-\lambda/2$ and $\lambda/2$. Let $R = M$ (recalling that $|M| = \Omega\big(n^{3/4}\big)$) and for each $0 \le i \le n_W$ and $\boldsymbol{v} \in R$ let $p_i(\boldsymbol{v}) = d_{W_i}(\boldsymbol{v}) - d_{W_i} \in I$.

Note that each $\boldsymbol{v} \in M$ has size k, so for each $0 < i \le n_W$, we have $\left|d_{\boldsymbol{v}}\left(w_i^+\right) - d_{\boldsymbol{v}}\left(w_{n_W - i + 1}^-\right)\right| \le k$ and therefore $\left|d_{W_i}(\boldsymbol{v}) - d_{W_{i-1}}(\boldsymbol{v})\right| \le k$. Also, we can compute

$$\left|d_{W_i} - d_{W_{i-1}}\right| = \left|\frac{d_{W^+} - d_{W^-}}{n_W}\right| \le k. \tag{1}$$

So, with $\rho = 2k$, we have $|p_i(\boldsymbol{v}) - p_{i-1}(\boldsymbol{v})| \le \rho$. Divide the range of "times" between 0 and n_W into n_W/τ sub-ranges of lengths $\tau = |R|\sigma^2/(8\rho\mu\lambda) = O(n^{3/4}/\log^3 n)$. For each such sub-range T, by Lemma 3.11 there is some $\boldsymbol{v}_T \in R$ which is never lonely in that range; fix such a $\boldsymbol{v}_T$ and for each $i \in T$ let $d_i = d_{W_i}(\boldsymbol{v}_T)$. For each $0 \le i \le n_W$ let $M_i \subseteq M$ be a set of μ elements $\boldsymbol{v} \in M$ satisfying $|d_{W_i}(\boldsymbol{v}) - d_i| \le \sigma$, which exists by the definition of loneliness. Recalling (1), observe that $|d_i - d_{i-1}| \le \lambda = O(\sqrt{n}\log n)$ for all $0 < i \le n_W$. Moreover, for all i except the $n_W/\tau = O\big(n^{1/4}\log^3 n\big)$ times where there is a "transition" between sub-ranges, there is $\boldsymbol{v}$ such that $|d_i - d_{i-1}| = \big|d_{W_i}(\boldsymbol{v}) - d_{W_{i-1}}(\boldsymbol{v})\big| \le k = o(\sqrt{n})$. $\square$

Next, (more or less) as described in Section 2, we identify a subsequence of indices i leading to subgraph sizes that are "well-separated" in a certain sense. Let $e_i = e(W_i) + \alpha e\big(U^0, W_i\big) + n_Z d_i$, where $n_Z = \delta\sqrt{c'n}/k \le \delta\sqrt{n_D}$ for some small $\delta = \delta(C) > 0$ (not depending on c') to be determined. The precise significance of these quantities e_i will become clear later, but the rough idea (as sketched in Section 2) is that we will eventually want to consider subgraphs consisting of some W_i, a random α-proportion of the elements of U^0, and n_Z vertices of M_i. Note that each d_i was defined to be equal to some $d_{W_i}(\boldsymbol{v}) \le kn$, so $|d_{n_W} - d_0| \le kn = O(n)$, and recall that $n_Z \le \sqrt{c'n}$. So, for small c', by property 2 of Lemma 4.2, we have

$$e_{n_W} - e_0 = \big(e\big(W^+\big) + \alpha e\big(U^0, W^+\big)\big) - \big(e\big(W^-\big) + \alpha e\big(U^0, W^-\big)\big) + n_Z(d_{n_W} - d_0) = \Omega\Big(n^{3/2}\Big).$$

Now, for $0 < i \le n_W$ let $\Delta_i = e_i - e_{i-1}$. Observe that

$$\begin{aligned}&\big|\big(e(W_i) + \alpha e\big(U^0, W_i\big)\big) - \big(e(W_{i-1}) + \alpha e\big(U^0, W_{i-1}\big)\big)\big| \\ &\quad = \Big|\Big(d_{W_i}\Big(w_i^+\Big) + \alpha d_{U^0}\Big(w_i^+\Big)\Big) - \Big(d_{W_{i-1}}\Big(w_{n_W-i+1}^-\Big) + \alpha d_{U^0}\Big(w_{n_W-i+1}^-\Big)\Big)\Big| \le (1+\alpha)n \le 2n,\end{aligned}$$

so the only way to have $\Delta_i > 3n$ is if $|d_i - d_{i-1}| = \Omega(\sqrt{n})$. By (iii) of Claim 4.4,

$$\sum_{i:|\Delta_i|>3n} |\Delta_i| = O\Big(n_Z\Big(n^{1/4}\log^3 n\Big)\big(\sqrt{n}\log n\big)\Big) = o\Big(n^{3/2}\Big).$$

By Lemma 3.10 (with $\tau = n_W = \Theta(n)$, $\lambda = \Omega(n^{3/2})$, $\rho = 3n$, $\kappa = o(n^{3/2})$ and $\sigma = n$) there is an increasing subsequence of indices $0 = i_1, \ldots, i_t = n_W$, with $t = \Omega(\sqrt{n})$, such that $e_{i_j} - e_{i_{j-1}} \ge n$ for each $1 < j \le t$.

Now, let D be a uniformly random subset of n_D elements of U^0, and let $U = U^0 \setminus D$. For a collection Z of vertex sets we write $V_Z = \bigcup_{\boldsymbol{z}\in Z} \boldsymbol{z}$, and for each $0 \le i \le n_W$ and some B to be determined, define

$$\Psi_i = \{e(W_i \cup U \cup V_Z) :\ Z \subseteq M_i,\ |Z| = n_Z,\ |e(U, V_Z) - \alpha n_Z d_{U^0}| \le B n_D\}.$$

Now the significance of the quantities e_i should be more clear: we expect the values in Ψ_i to be about $e(U) + e_i + \alpha n_Z d_{U^0}$, so the idea is that the separation we have established between the e_{i_j} should translate to the Ψ_{i_j} not interfering too much with each other.

Note that we can apply Lemma 4.3 to determine $\delta = \Omega(1)$ and $B = O(1)$ such that for each $0 \le i \le n_W$, $|\Psi_i| = \Omega\big(n_Z\sqrt{n_D}\big) = \Omega(n_D)$ with probability at least $1/4$. Indeed, the first condition of Lemma 4.3 follows from (i) in Claim 4.4 and a sufficiently small choice of c', the second condition follows from property 4 of Lemma 4.2, and the third condition follows from (ii) in Claim 4.4 and property 3 of Lemma 4.2. We will next show that there is an outcome of U for which many Ψ_{i_j} are large, and in addition the cumulative deviations introduced by the randomness of U do not too severely affect the separation we established so far. To this end, for each $0 < i \le n_W$ define

$$g_i = \Big(d_U\Big(w_i^+\Big) - d_U\Big(w_{n_W-i+1}^-\Big)\Big) - \Big(\alpha d_{U^0}\Big(w_i^+\Big) - \alpha d_{U^0}\Big(w_{n_W-i+1}^-\Big)\Big).$$

Basically, $|g_i|$ measures the deviation of the separation $e(U,W_i) - e(U,W_{i-1})$ from its expected value $\alpha e(U^0,W_i) - \alpha e(U^0,W_{i-1})$. We will control the cumulative deviation $\sum_{i=1}^{n_W}|g_i|$; the absolute deviations $\left|e(U,W_i) - \alpha e(U^0,W_i)\right|$ are unfortunately too large to control directly.

Claim 4.5. *The following hold together with positive probability.*

(i) There is a subset $\mathcal{J}$ of $(0.1)t$ indices j for which $\left|\Psi_{i_j}\right| = \Omega(n_D)$ (that is, a positive proportion of Ψ_{i_j} are large);

(ii) $\sum_{i=1}^{n_W}|g_i| \le O\big(n\sqrt{n_D}\big)$.

Proof. First we show that (i) holds with probability at least $1/6$. As discussed above, for each $1 \le j \le t$, by Lemma 4.3 we have $\left|\Psi_{i_j}\right| = \Omega(n_D)$ with probability at least $1/4$. Let $\overline{\mathcal{J}}$ be the set of j for which this fails, so $\mathbb{E}\left|\overline{\mathcal{J}}\right| \le 3t/4$ and by Markov's inequality, $\left|\overline{\mathcal{J}}\right| \le (0.9)t$ with probability at least $1/6$.

Next we show that (ii) holds with probability at least 0.9, meaning that we can use the union bound to show that (i) and (ii) hold simultaneously with positive probability. For this, note that for each $0 < i \le n_W$, g_i is of $\big(\left|U^0\right|, n_D/\left|U^0\right|, O(1)\big)$-hypergeometric type and has mean zero (because $\mathbb{E}d_U(w) = \alpha d_{U^0}(w)$ for any $w \in W$), so by Lemma 3.6 we have

$$\Pr(|g_i| \ge r) \le e^{-\Omega\left(r^2/n_D\right)}.$$

For a sufficiently large constant Q we have

$$\mathbb{E}|g_i| = \sum_{r=1}^{\infty} \Pr(|g_i| \ge r) \le Q\sqrt{n_D} + \sum_{r=Q\sqrt{n_D}}^{\infty} e^{-\Omega\left(r^2/n_D\right)} \le 2Q\sqrt{n_D},$$

$$\mathbb{E}\sum_{i=1}^{n_W}|g_i| \le 2Qn_W\sqrt{n_D},$$

and by Markov's inequality $\sum_{i=1}^{n_W}|g_i| \le 20Qn_W\sqrt{n_D} = O\big(n\sqrt{n_D}\big)$ with probability at least 0.9. □

Fix an outcome of U such that the above properties hold.

We now take a moment to summarise the situation so far. We have $c'n \le n_D \le 2c'n$ and $n_Z = \Theta(\sqrt{n_D})$ for some small constant c' (and the constants in all asymptotic notation are independent of c'). With $e_i = e(W_i) + \alpha e\big(U^0, W_i\big) + n_Z d_i$, we have a subsequence of indices $0 = i_1, \dots, i_t = n_W$, for $t = \Omega(\sqrt{n})$, such that $e_{i_j} - e_{i_{j-1}} \ge n$ for each $1 < j \le t$. We also have matchings M_i such that the degrees $d_{W_i}(\boldsymbol{v})$, for $0 \le i \le n_W$ and $\boldsymbol{v} \in M_i$, are very tightly controlled (to be precise, Claim 4.4 (ii) says that $|d_{W_i}(\boldsymbol{v}) - d_i| = o(\sqrt{n})$). Moreover, Claim 4.5 shows that many $\left|\Psi_{i_j}\right|$ are large (specifically, $\left|\Psi_{i_j}\right| = \Omega(n_D)$ for $\Omega(\sqrt{n})$ different j), and the cumulative deviation $\sum_{i=1}^{n_W}|g_i| \le O\big(n\sqrt{n_D}\big)$ caused by dropping to a random subset $U = U^0 \setminus D$ is not too severe. We next show that many of the Ψ_{i_j} are disjoint, which essentially completes the proof of Theorem 1.1.

Claim 4.6. *For sufficiently small c', there is a subset $\mathcal{J}' \subseteq \mathcal{J}$ of $\Omega\big(n^{1/2}\big)$ indices j among which each Ψ_{i_j} is disjoint.*

Proof. For $1 \le j < t$, let $\Sigma_j = n(j-1) - \sum_{i=1}^{i_j}|g_i|$, so that $\Sigma_1 = 0$ and $\Sigma_t \ge \Big(1 - O\Big(\sqrt{c'}\Big)\Big)tn$, by (ii) in Claim 4.5. The significance of these quantities is that we have established the separation $e_{i_j} - e_{i_{j-1}} \ge n$, but this may be offset by the buildup of deviations $|g_i|$. That is, each increment $\Sigma_j - \Sigma_{j-1} = n - \sum_{i=i_{j-1}+1}^{i_j}|g_i|$ is a lower bound on the separation between $e\big(W_{i_{j-1}}\big) + e\big(U, W_{i_{j-1}}\big) +$

$n_Z d_{i_{j-1}}$ and $e(W_{i_j}) + e(U, W_{i_j}) + n_Z d_{i_j}$, which approximates the separation between the values in $\Psi_{i_{j-1}}$ and the values in Ψ_{i_j}.

By Lemma 3.10 (with $\tau = t = \Omega(n)$, $\lambda = \Sigma_t = (1 - O(c'))tn$, $\sigma = (0.01)t$, $\rho = n$ and $\kappa = 0$), for small c' we can find an increasing sequence $j_1^0, \dots, j_{s^0}^0$, for $s^0 \geq ((1 - O(c'))/1.01)t \geq (0.95)t$, such that $\Sigma_{j_q^0} - \Sigma_{j_{q-1}^0} \geq (0.01)n$ for each $1 < q \leq s^0$. By (i) in Claim 4.5, deleting the indices not in $\mathcal{J}$ gives an increasing sequence $j_1, \dots, j_s$, with $s \geq (0.05)t$, also satisfying $\Sigma_{j_q} - \Sigma_{j_{q-1}} \geq (0.01)n$ for each $1 < q \leq s$.

To avoid too many layered subscripts, for $1 \leq q \leq s$ define $W_q' = W_{i_{j_q}}$, $d_q' = d_{i_{j_q}}$, $M_q' = M_{i_{j_q}}$, $e_q' = e_{i_{j_q}}$, $i_q' = i_{j_q}$. Also, for $1 < q \leq s$ define $\Gamma_q = \sum_{i=i_{q-1}'+1}^{i_q'} |g_i|$.

Our goal is now to show that quantities of the form $e(W_q' \cup U \cup V_Z)$ arising from the definition of $|\Psi_{i_{j_q}}|$ are well-separated for different q. This will basically follow from the fact that $\Sigma_{j_q} - \Sigma_{j_{q-1}} = \Omega(n)$, our control over the $d_{W_i}(\boldsymbol{v})$ for $\boldsymbol{v} \in M_i$, and the definition of the $\Psi_{i_q'}$.

First, for each $1 < q \leq s$ observe that

$$\begin{aligned}
e(W_q', U) - e(W_{q-1}', U) &= \sum_{i=i_{q-1}'+1}^{i_q'} \Big(d_U\big(w_i^+\big) - d_U\big(w_{n_W - i+1}^-\big)\Big) \\
&\geq \alpha \sum_{i=i_{q-1}'+1}^{i_q'} \Big(d_{U^0}\big(w_i^+\big) - d_{U^0}\big(w_{n_W - i+1}^-\big)\Big) - \Gamma_q \\
&= \alpha e(W_q', U^0) - \alpha e(W_{q-1}', U^0) - \Gamma_q.
\end{aligned}$$

Next, for $Z \subseteq M_q'$ and $Z' \subseteq M_{q-1}'$ satisfying $|Z| = |Z'| = n_Z$ and

$$|e(U, V_Z) - \alpha n_Z d_{U^0}|,\ |e(U, V_{Z'}) - \alpha n_Z d_{U^0}| \leq B n_D = O(n_D),$$

we also have

$$\begin{aligned}
&e(W_q' \cup U, V_Z) - e(W_{q-1}' \cup U, V_{Z'}) + e(V_Z) - e(V_{Z'}) \\
&\quad = \sum_{\boldsymbol{v} \in Z} d_{W_q'}(\boldsymbol{v}) - \sum_{\boldsymbol{v} \in Z'} d_{W_{q-1}'}(\boldsymbol{v}) + O(n_D) \\
&\quad = \sum_{\boldsymbol{v} \in Z} \big(d_q' + o(\sqrt{n_D})\big) - \sum_{\boldsymbol{v} \in Z'} \big(d_{q-1}' + o(\sqrt{n_D})\big) + O(n_D) \\
&\quad = n_Z d_q' - n_Z d_{q-1}' + O(n_D).
\end{aligned}$$

Recall that $\Sigma_{j_q} - \Sigma_{j_{q-1}} = \Omega(n)$ and $n_D \leq 2c'n$. For small c', it follows that

$$\begin{aligned}
&e(W_q' \cup U \cup V_Z) - e(W_{q-1}' \cup U \cup V_{Z'}) \\
&\quad = e(W_q') - e(W_{q-1}') + e(W_q', U) - e(W_{q-1}', U) \\
&\qquad\quad + e(W_q' \cup U, V_Z) - e(W_{q-1}' \cup U, V_{Z'}) + e(V_Z) - e(V_{Z'}) \\
&\quad \geq \big(e(W_q') + \alpha e(W_q', U^0) + n_Z d_q'\big) - \big(e(W_{q-1}') + \alpha e(W_{q-1}', U^0) + n_Z d_{q-1}'\big) - \Gamma_q - O(n_D) \\
&\quad = e_q' - e_{q-1}' - \Gamma_q - O(n_D) \\
&\quad \geq (j_q - j_{q-1})n - \Gamma_q - O(n_D) = \Sigma_{j_q} - \Sigma_{j_{q-1}} - O(n_D) = \Omega(n) > 0.
\end{aligned}$$

We conclude that the minimum value in $\Psi_{i_q'}$ is greater than the maximum value in $\Psi_{i_{q-1}'}$. Since this is true for all $1 < q \leq s$, it follows that each $\Psi_{i_{j_q}}$ is disjoint, so we may take $\mathcal{J}' = \{j_1, \dots, j_s\}$. $\square$

Finally, let $f = c'n$ and $h = n_W + kn_Z = cn + \delta\sqrt{c'n}$. For $1 \le i \le n_W$ observe that if $Z \subseteq M_i$ satisfies $|Z| = n_Z$ then $W_i \cup U \cup V_Z$ has exactly $\left|U^0\right| - n_D + h$ vertices, and this number is equal to $(\ell - f) + h$ or $2(\ell - f) + h$. Therefore, for $\ell' = (\ell - f) + h$ or $\ell' = 2(\ell - f) + h$, we have

$$\left|\Psi(\ell', G)\right| \ge \sum_{j \in \mathcal{J}'} \left|\Psi_{i_j}\right| = \Omega\Big(c'n^{3/2}\Big). \qquad \square$$

4.1 Proof of Lemma 4.2

As outlined in Section 2, we will first construct W^-, W^+ and U^0 satisfying properties 1 and 2, and we will then use richness (Lemma 3.1) and the sunflower lemma (Lemma 3.5) to construct M satisfying properties 3 and 4. We remark that it would be possible to use an existing discrepancy theorem (for example, a theorem in [22], as mentioned in Section 2) to construct sets W^-, W^+ and U^0 satisfying property 2, using only the fact that G has density bounded away from 0 and 1. However, since we are already using Lemma 3.1 for property 4, it is convenient to instead use richness and anticoncentration.

So, consider $\varepsilon = \varepsilon(C)$ from Lemma 3.1, note that we can assume $\varepsilon < 1/8$, and let $\delta = \varepsilon^K$ for some large absolute constant K which we will determine later. Let $G[V']$ be a (δ, ε)-rich induced subgraph of G, with $n' := |V'| \ge 15cn$ vertices, which exists for small $c > 0$ by Lemma 3.1. We will only work inside V', so all degrees and neighbourhoods should be interpreted as being restricted to V'.

First, let U^1 be a uniformly random subset of V' with size $2\ell \le 4cn$. Let $H \subseteq \binom{V'}{2}$ be the auxiliary graph with an edge $\{x, y\} \in \binom{V'}{2}$ whenever $d_{U^1}(x) = d_{U^1}(y)$. We show that with positive probability, the diversity of neighbourhoods in $G[V']$ is maintained for neighbourhoods in U^1, and simultaneously H is quite sparse, which implies that there is a lot of variation between degrees into U^1 (this will be the starting point from which we obtain our discrepancy for property 2).

Claim 4.7. *The following hold together with positive probability.*

(i) For each $k \le K$ and $\boldsymbol{x}, \boldsymbol{y} \in \binom{V}{k}$ with $|N(\boldsymbol{x}) \triangle N(\boldsymbol{y})| \ge \varepsilon^K n'$, we have $|N_{U^1}(\boldsymbol{x}) \triangle N_{U^1}(\boldsymbol{y})| \ge \varepsilon^K \ell$;

(ii) there is a set W of at least $7cn$ vertices such that $d_H(x) = O(\sqrt{n})$ for each $x \in W$.

Proof. We will show that (i) and (ii) each hold with probability greater than 1/2. The proofs will be quite routine, using the concentration and anticoncentration theorems in Section 3.3.

For (i), observe that for each $\boldsymbol{x}, \boldsymbol{y} \in \binom{V}{k}$, $|N_{U^1}(\boldsymbol{x}) \triangle N_{U^1}(\boldsymbol{y})| = \left|(N(\boldsymbol{x}) \triangle N(\boldsymbol{y})) \cap U^1\right|$ is of $(n', 2\ell/n', 1)$-hypergeometric type, and apply Lemma 3.6 and the union bound. (Recall from Section 1.1 the nonstandard multiset definition of $A \triangle B$).

For (ii), note that each $d_{U^1}(x) - d_{U^1}(y)$ is of $(n', 2\ell/n', 1, |N(x) \triangle N(y)|)^*$-hypergeometric type, so if $|N(x) \triangle N(y)| = \Omega(n)$ then by Lemma 3.8, $\Pr(d_{U^1}(x) = d_{U^1}(y)) = O(1/\sqrt{n})$. By Lemma 3.3 (taking $k = 1$), there are at most $n^{1/5}$ vertices $x \in V'$ with $N(x) < \varepsilon n'$, and by Lemma 3.2, for every other vertex $x \in V'$ there are at most $n^{1/5}$ vertices $y \ne x$ with $|N(x) \triangle N(y)| < \varepsilon^2 n'$. For each $x \in V'$ of the latter type, we have $\mathbb{E} d_H(x) = O\big(n^{1/5} + \sqrt{n}\big) = O(\sqrt{n})$, so by Markov's inequality, $d_H(x) = O(\sqrt{n})$ (for a sufficiently large constant implied by the big-oh notation) with probability at least 7/8. Let W be the set of all $x \in V'$ for which this holds, so that $\mathbb{E}|V' \setminus W| \le n'/8 + n^{1/5} < n'/4$. Therefore, $|W| \ge n'/2 \ge 7cn$ with probability greater than 1/2. $\square$

Fix an outcome of U^1 satisfying both the properties in the above claim, and note that $\left|W \setminus U^1\right| \ge 3cn$. Order the vertices $x \in W \setminus U^1$ by their values of $d_{U^1}(x)$ (breaking ties arbitrarily), let W^1

contain the first cn vertices in this ordering and let W^2 contain the last cn. By (ii) in Claim 4.7, for the (at least cn) vertices x between W^1 and W^2 in this ordering, we have $d_H(x) = O(\sqrt{n})$, so there are at least $\Omega(\sqrt{n})$ values of $d_{U^1}(x)$, and

$$\min_{x\in W^2} d_{U^1}(x) - \max_{x\in W^1} d_{U^1}(x) = \Omega\big(\sqrt{n}\big).$$

Recalling that $\alpha \geq 1/2$, this implies that

$$\alpha e\big(W^2, U^1\big) - \alpha e\big(W^1, U^1\big) = \Omega\Big(n^{3/2}\Big).$$

Now, if

$$\big(e\big(W^2\big) + \alpha e\big(W^2, U^1\big)\big) - \big(e\big(W^1\big) + \alpha e\big(W^1, U^1\big)\big) \geq \big(\alpha e\big(W^2, U^1\big) - \alpha e\big(W^1, U^1\big)\big)/4$$

then let $W^- = W^1$ and $W^+ = W^2$ and $U^0 = U^1$; property 2 is satisfied. Otherwise, there must be a large discrepancy between $e\big(W^1\big)$ and $e\big(W^2\big)$. To be specific, we must have

$$\big(e\big(W^1\big) + \alpha e\big(W^1, U^1\big)/2\big) - \big(e\big(W^2\big) + \alpha e\big(W^2, U^1\big)/2\big) \geq \big(\alpha e\big(W^2, U^1\big) - \alpha e\big(W^1, U^1\big)\big)/4. \tag{2}$$

In this case, let U^0 be a random subset of $\ell = \big|U^0\big|/2$ elements of U^1, let $W^- = W^2$ and let $W^+ = W^1$. Then

$$\begin{aligned}&\big(e\big(W^+\big) + \alpha e\big(W^+, U^0\big)\big) - \big(e\big(W^-\big) + \alpha e\big(W^-, U^0\big)\big)\\ &\quad= \big(e\big(W^+\big) - e\big(W^-\big)\big) + \alpha \sum_{u\in U^0} \big(d_{W^+}(u) - d_{W^+}(u)\big)\end{aligned}$$

is of (1/2)-hypergeometric type and has mean $\Omega\big(n^{3/2}\big)$, given by (2). So, by Lemma 3.9, this random value is $\Omega\big(n^{3/2}\big)$ with probability at least 1/2. Also, for each $k \leq K$ and $\boldsymbol{x}, \boldsymbol{y} \in \binom{V}{k}$ with $|N(\boldsymbol{x}) \triangle N(\boldsymbol{y})| \geq \varepsilon^K n'$, the random variable $|N_{U^0}(\boldsymbol{x}) \triangle N_{U^0}(\boldsymbol{y})|$ is of $(\Omega(n), 1/2, 1)$-hypergeometric type with mean $\Omega(n)$, so by Lemma 3.6 and the union bound, with probability $1 - o(1)$ we have $|N_{U^0}(\boldsymbol{x}) \triangle N_{U^0}(\boldsymbol{y})| = \Omega(n)$ for all such $k, \boldsymbol{x}, \boldsymbol{y}$. So, we can fix an outcome of U^0 satisfying both of these properties.

In either of the above two cases, property 2 is satisfied and $|N_{U^0}(\boldsymbol{x}) \triangle N_{U^0}(\boldsymbol{y})| = \Omega(n)$ for each $\boldsymbol{x}, \boldsymbol{y} \in \binom{V}{k}$ with $|N(\boldsymbol{x}) \triangle N(\boldsymbol{y})| \geq \varepsilon^K n'$. We also have $|U^0| = \ell$ or $|U^0| = 2\ell$, satisfying property 1. Now, fix some $\Omega(n)$-vertex subset A^0 disjoint from U^1 and W, and let $M^0 \subseteq \binom{A^0}{K}$ contain every $\boldsymbol{v} \in \binom{A^0}{K}$ with $\big|\bigcap_{v\in \boldsymbol{v}} N(v)\big| \geq \varepsilon^K n'$. By Lemma 3.3, we have $\big|M^0\big| = \Omega\big(n^K\big)$.

Observe that there are only $(kn + 1)^3$ possible values of the tuples $(d_{W^+}(\boldsymbol{x}), d_{W^+}(\boldsymbol{x}), d_{U^0}(\boldsymbol{x}))$, so by the pigeonhole principle there are $d'_{W^-}, d'_{W^+}, d'_{U^0} \in \mathbb{N}$, and a collection $M^1 \subseteq M^0$ of size $\Omega\big(n^{K-3}\big)$, such that for each $\boldsymbol{x} \in M^1$ we have $d_{W^-}(\boldsymbol{x}) = d'_{W^-}$, $d_{W^+}(\boldsymbol{x}) = d'_{W^+}$ and $d_{U^0}(\boldsymbol{x}) = d'_{U^0}$. For sufficiently large K, by Lemma 3.5, M^1 has a sunflower with $\Omega\big(n^{(K-3)/K}\big) = \Omega\big(n^{3/4+1/5}\big)$ petals; take M^2 as this set of petals, and let k be the common size of these petals. Let $\boldsymbol{v}$ be the kernel of the sunflower, and let $d_{W^-} = d'_{W^-} - d_{W^-}(\boldsymbol{v})$, $d_{W^+} = d'_{W^+} - d_{W^+}(\boldsymbol{v})$ and $d_{U^0} = d'_{U^0} - d_{U^0}(\boldsymbol{v})$, so for $\boldsymbol{x} \in M^2$ we have $d_{W^-}(\boldsymbol{x}) = d_{W^-}$, $d_{W^+}(\boldsymbol{x}) = d_{W^+}$ and $d_{U^0}(\boldsymbol{x}) = d_{U^0}$.

Finally, consider the auxiliary graph $F \subseteq \binom{M^2}{2}$ which has an edge $\{\boldsymbol{x}, \boldsymbol{y}\} \in \binom{M^2}{2}$ whenever $|N(\boldsymbol{x}) \triangle N(\boldsymbol{y})| < \varepsilon^K n'$. By Lemma 3.2, the degrees in F are at most $n^{1/5}$ so by Proposition 3.4 (Turán's theorem) there is $M \subseteq M^2$ with $|M| = \Omega\big(n^{3/4}\big)$ such that $|N(\boldsymbol{x}) \triangle N(\boldsymbol{y})| \geq \varepsilon^K n'$, and therefore $|N_{U^0}(\boldsymbol{x}) \triangle N_{U^0}(\boldsymbol{y})| = \Omega(n)$, for all pairs $\{\boldsymbol{x}, \boldsymbol{y}\} \in \binom{M}{2}$.

4.2 Proof of Lemma 4.3

As in the deduction of Theorem 1.1 in Section 4, for a collection Z of vertex sets let $V_Z = \bigcup_{\boldsymbol{z}\in Z} \boldsymbol{z}$.

Our proof of Lemma 4.3 will be quite similar to the proof of the main theorem in [27]. Roughly speaking, we will first expose a random superset D^1 of D (we may view this as "partially exposing" the random subset D). Using this randomness for anticoncentration, we will construct sub-matchings $S^-, S^+ \subseteq M$ of size $\Omega(\sqrt{n_D})$, such that all the degrees from elements of S^+ into D^1 are higher by $\sqrt{n_D}$ than the degrees from S^- into D^1. Starting with any $S_0 \subseteq S^+$ of some size $n_Z - 1$, we can therefore obtain n_Z subsets $S_0, \dots, S_{n_Z-1}$ such that the values $e(W \cup U^0 \cup V_{S_i})$ are separated by a distance of $\Omega(\sqrt{n_D})$, simply by switching elements of M one-by-one from S^+ into S^-. Then, we fully expose the random set D (therefore exposing $U = U^0 \setminus D$), and show that the values $e(W \cup U \cup V_{S_i})$ remain fairly well-separated. We use this further randomness, and anticoncentration, to show that for most i, there is a set X_i of $\sqrt{n_D}$ elements of M which have different degrees into $W \cup U \cup V_{S_i}$, still concentrated in a known interval of length $O(\sqrt{n_D})$. This will prove that there are $\Omega(n_Z\sqrt{n_D})$ values $e(W \cup U \cup V_{S_i} \cup \boldsymbol{z})$, for $\boldsymbol{x} \in X_i$. (So, our sets Z in the lemma statement are of the form $S_i \cup \{\boldsymbol{x}\}$, for $\boldsymbol{x} \in X_i$). The additional requirement that there are about the expected number of edges between U and Z will follow from our proof basically for free.

We now proceed with this plan to prove Lemma 4.3. Arbitrarily split M into two subsets S^0 and X^0 each of size $\Omega(\sqrt{n_D})$. Let D^1 be a uniformly random subset of U^0 of size $2n_D$, so that we may realise the desired distribution of D as a uniformly random subset of D^1 of size n_D. We will first observe some regularity and discrepancy properties that hold with probability at least 3/4 with respect to the random choice of D^1. Let $H \subseteq \binom{S^0}{2}$ be the auxiliary random graph (depending on D^1) with an edge $\{\boldsymbol{x}, \boldsymbol{y}\} \in \binom{S^0}{2}$ if $d_{D^1}(\boldsymbol{x}) = d_{D^1}(\boldsymbol{y})$. Also, let $d_D = (1-\alpha)d_{U^0}$, recalling from the statement of Lemma 4.3 that $1 - \alpha = n_D/|U^0|$.

Claim 4.8. *The following hold together with probability at least* $3/4$.

(i) $|N_{D^1}(\boldsymbol{x}) \triangle N_{D^1}(\boldsymbol{y})| = \Omega(n_D)$ *for each* $\{\boldsymbol{x}, \boldsymbol{y}\} \in \binom{X^0}{2}$*;*

(ii) there are $X \subseteq X^0$ *and* $S^1 \subseteq S^0$*, each with size* $\Omega(\sqrt{n_D})$*, such that* $d_{D^1}(\boldsymbol{x}) = 2d_D + O(\sqrt{n_D})$ *for each* $\boldsymbol{x} \in X \cup S^1$*;*

(iii) H *has* $O(\sqrt{n_D})$ *edges.*

Proof. We will prove that each of (i)-(iii) individually hold with high probability, then apply the union bound. The proofs will be rather routine, using the concentration and anticoncentration theorems in Section 3.3 in a similar way to the proof of Claim 4.7.

For (i), observe that for each $\{\boldsymbol{x}, \boldsymbol{y}\} \in \binom{X^0}{2}$, the random variable

$$|N_{D^1}(\boldsymbol{x}) \triangle N_{D^1}(\boldsymbol{y})| = \left|N_{U^0}(\boldsymbol{x}) \triangle N_{U^0}(\boldsymbol{y}) \cap D^1\right|$$

is of $(|U^0|, 2(1-\alpha), 1)$-hypergeometric type with mean $\Omega(n_D)$, so by the second assumption of this lemma, Lemma 3.6 and the union bound, (i) holds with probability $1 - |X^0|^2 e^{-\Omega(n_D)} = 1 - o(1)$.

We next show that (ii) holds with probability at least 7/8. For each $\boldsymbol{x} \in X^0$, the random variable $d_{D^1}(\boldsymbol{x})$ is of $(|U^0|, 2(1-\alpha), k)$-hypergeometric type, so by Lemma 3.6 (with t a large multiple of $\sqrt{n_D}$), with probability at least 31/32 we have $d_{D^1}(\boldsymbol{x}) = \mathbb{E}d_{D^1}(\boldsymbol{x}) + O(\sqrt{n_D}) = 2d_D + O(\sqrt{n_D})$. Therefore the expected number of $\boldsymbol{x} \in X^0$ failing to satisfy this bound is at most $|X^0|/32$, and the probability more than $|X^0|/2$ fail to satisfy this bound is at most 1/16. If this does not occur, we can find an appropriate subset $X \subseteq X^0$ of size $|X^0|/2$. A very similar argument shows that an appropriate subset $S^1 \subseteq S^0$ with size $\left|S^0\right|/2$ exists with probability at least $15/16$, and by the union bound we can simultaneously find suitable X, S^1 with probability at least $7/8$.

Finally, we show that (iii) holds with probability at least $15/16$. This will suffice to apply the union bound over parts (i)-(iii). Note that the random variable $d_{D^1}(\boldsymbol{x}) - d_{D^1}(\boldsymbol{y})$ is of $\left(|U^0|, 2(1-\alpha), O(1), |N_{U^0}(\boldsymbol{x}) \triangle N_{U^0}(\boldsymbol{y})|\right)^*$-hypergeometric type. Recalling the second assumption of this lemma that $|N_{U^0}(\boldsymbol{x}) \triangle N_{U^0}(\boldsymbol{y})| = \Omega\left(\left|U^0\right|\right)$, we may apply Lemma 3.8 to see that for any $\{\boldsymbol{x}, \boldsymbol{y}\} \in \binom{S^0}{2}$, the probability $\{\boldsymbol{x}, \boldsymbol{y}\}$ is an edge in H is $O\left(1/\sqrt{(1-\alpha)|U^0|}\right) = O\left(1/\sqrt{n_D}\right)$, and the expected number of edges is $O\left(\sqrt{n_D}\right)$. The desired result then follows from Markov's inequality. $\square$

Condition on an outcome of D^1 satisfying all the above properties (we will treat D^1 as fixed for the remainder of the proof). By Proposition 3.4, the graph H (which has $|S^0| = \Omega(\sqrt{n_D})$ vertices) has an independent set S^2 of size $\Omega\left(\sqrt{n_D}\right)$, meaning that the values of $d_{D^1}(\boldsymbol{x})$, for $\boldsymbol{x} \in S^2$, are all different. Now, let $n_S = n_Z - 1$, and note that for small δ we have $n_S < \delta\sqrt{n_D} \le \left|S^2\right|/3$. Order the vertices $\boldsymbol{x} \in S^2$ by their values of $d_{D^1}(\boldsymbol{x})$, let S^- contain the first n_S elements of this ordering and let S^+ contain the last n_S elements. By construction, we have

$$\min_{\boldsymbol{x} \in S^+} d_{D^1}(\boldsymbol{x}) - \max_{\boldsymbol{x} \in S^-} d_{D^1}(\boldsymbol{x}) = \Theta(\sqrt{n_D}). \tag{3}$$

(Here and from now on, the constants implied by all asymptotic notation are independent of δ).

Now, fix orderings $\boldsymbol{v}_1^-, \ldots, \boldsymbol{v}_{n_S}^-$ of S^- and $\boldsymbol{v}_1^+, \ldots, \boldsymbol{v}_{n_S}^+$ of S^+. For $0 \le i \le n_S$, define

$$S_i = \left\{\boldsymbol{v}_1^-, \ldots, \boldsymbol{v}_i^-\right\} \cup \left\{\boldsymbol{v}_1^+, \ldots, \boldsymbol{v}_{n_S - i}^+\right\},$$

let $U_i = W \cup U \cup V_{S_i}$, and let $e_i = e(V_{S_i}) + e(V_{S_i}, W \cup U) = e(U_i) - e(W \cup U)$. For $0 < i \le n_S$ define

$$\begin{aligned}
\Delta_i &= e_i - e_{i-1} \\
&= e(V_{S_i}, W \cup U) - e\left(V_{S_{i-1}}, W \cup U\right) + e(V_{S_i}) - e\left(V_{S_{i-1}}\right) \\
&= d_{W \cup U}\left(\boldsymbol{v}_i^-\right) - d_{W \cup U}\left(\boldsymbol{v}_{n_S - i+1}^+\right) + e(V_{S_i}) - e\left(V_{S_{i-1}}\right) \\
&= \left(d_{W \cup U^0}\left(\boldsymbol{v}_i^-\right) - d_D\left(\boldsymbol{v}_i^-\right)\right) - \left(d_{W \cup U^0}\left(\boldsymbol{v}_{n_S-i+1}^+\right) - d_D\left(\boldsymbol{v}_{n_S-i+1}^+\right)\right) + e(V_{S_i}) - e\left(V_{S_{i-1}}\right). \quad (4)
\end{aligned}$$

Next we observe that with probability at least $1/3$, our discrepancy properties are to some extent maintained, while for many i we can find many vertices in X with distinct degrees into U_i. Recall that D is a random subset of half the elements of $|D^1|$.

Claim 4.9. *There are $\gamma_1, \gamma_3 = \Omega(1)$ and $Q_2, Q_4 = O(1)$ such that the following hold together with probability at least $1/3$.*

(i) there is a set $\mathcal{I}_1$ of $(1 - \gamma_1/(8Q_2))n_S$ indices i such that for each $i \in \mathcal{I}_1$, we have $e(D, V_{S_i}) = n_S d_D + O(n_D)$;

(ii) There is a set $\mathcal{I}_2$ of $(1 - \gamma_1/(8Q_2))n_S$ indices i, each with a set $X_i \subseteq X$ of size $2\gamma_3|X|$, such that the $d_{U_i}(\boldsymbol{x})$, for $\boldsymbol{x} \in X_i$, are distinct;

(iii) there is a set X^* of size $(1 - \gamma_3)|X|$ such that for each $\boldsymbol{x} \in X^*$ we have $|d_D - d_D(\boldsymbol{x})| \le Q_4\sqrt{n_D}$;

(iv) $e_{n_S} - e_0 \ge 3\gamma_1 n_S \sqrt{n_D}$;

(v) $\sum_{i:|\Delta_i|\ge Q_2\sqrt{n_D}}|\Delta_i| \le \gamma_1 n_S\sqrt{n_D}$.

Proof. We will prove that each part holds with probability at least 0.99, except (iv), which holds with probability at least 1/2. The values of $\gamma_1, Q_2, \gamma_3, Q_4$ will be determined in order, and will depend on each other.

For (iv), recalling (4) we observe

$$\begin{aligned}\mathbb{E}\,\Delta_i &= \mathbb{E}[e_i - e_{i-1}]\\ &= \Big(d_{W\cup U^0}\Big(\boldsymbol{v}_i^-\Big) - d_{D^1}\Big(\boldsymbol{v}_i^-\Big)/2\Big) - \Big(d_{W\cup U^0}\Big(\boldsymbol{v}_{n_S-i+1}^+\Big) - d_{D^1}\Big(\boldsymbol{v}_{n_S-i+1}^+\Big)/2\Big) - O(n_S).\end{aligned}$$

Recall from the third assumption of this lemma that $d_{W\cup U^0}(\boldsymbol{v}) = d_{U^0} + d_W + o\big(\sqrt{n_D}\big)$ for all $\boldsymbol{v}\in M$, and recall from (3) that the degrees from S^+ into D^1 are larger by $\Theta\big(\sqrt{n_D}\big)$ than the degrees from S^- into D^1. Also, recall that $n_S < \delta\sqrt{n_D}$. For small δ it follows that

$$\mathbb{E}[e_i - e_{i-1}] = \Theta(\sqrt{n_D}) - o(\sqrt{n_D}) - O(n_S) = \Theta(\sqrt{n_D}).$$

So, $\mathbb{E}[e_{n_S} - e_0] = \Theta\big(n_S\sqrt{n_D}\big)$. Since $e_{n_S} - e_0$ is of (1/2)-hypergeometric type, we may apply Lemma 3.9 to show that for small γ_1 it is at least as large as its expectation $\Omega\big(n_S\sqrt{n_D}\big) \ge 3\gamma_1 n_S\sqrt{n_D}$, with probability at least 1/2.

For (v), observe that for each $0 < i \le n_S$, the random variable Δ_i is of $(2n_D, 1/2, k)$-hypergeometric type, because it is a translation of the random variable $d_D\Big(\boldsymbol{v}_{n_S-i+1}^+\Big) - d_D\Big(\boldsymbol{v}_i^-\Big)$. We have just computed that $\mathbb{E}\,\Delta_i = O\big(\sqrt{n_D}\big)$, so by Lemma 3.6 we therefore have $\Pr(|\Delta_i| \ge t) = \exp\big(-\Omega\big(t^2/n_D\big)\big)$. Now, for any nonnegative integer random variable ξ, we have $\mathbb{E}\xi = \sum_{t=1}^{\infty}\Pr(\xi \ge t)$, so

$$\begin{aligned}\mathbb{E}\Big[|\Delta_i|\mathbb{1}_{|\Delta_i|\ge Q_2\sqrt{n_D}}\Big] &= \sum_{t=1}^{\infty}\Pr\Big(|\Delta_i|\mathbb{1}_{|\Delta_i|\ge Q_2\sqrt{n_D}} \ge t\Big)\\ &= Q_2\sqrt{n_D}\Pr(|\Delta_i| \ge Q_2\sqrt{n_D}) + \sum_{t=Q_2\sqrt{n_D}}^{\infty}\Pr(|\Delta_i| \ge t)\\ &= Q_2\sqrt{n_D}e^{-\Omega\left(Q_2^2\right)} + \sum_{t=Q_2\sqrt{n_D}}^{\infty}\exp\big(-\Omega\big(t^2/n_D\big)\big) = e^{-\Omega\left(Q_2^2\right)}\sqrt{n_D}.\end{aligned}$$

For sufficiently large Q_2, this is at most $(\gamma_1/100)\sqrt{n_D}$, so

$$\mathbb{E}\sum_{i:|\Delta_i|\ge Q_2\sqrt{n_D}}|\Delta_i| \le (\gamma_1/100)n_S\sqrt{n_D}$$

and (v) holds with probability at least 0.99 by Markov's inequality.

For (i), recall from (ii) of Claim 4.8 that each $\boldsymbol{x}\in S^1$ has degree $2d_D + O\big(\sqrt{n_D}\big)$ into D^1. Therefore, for each $0 \le i \le n_S$, $e(D, V_{S_i})$ is of $\big(2n_D, 1/2, O\big(\sqrt{n_D}\big)\big)$-hypergeometric type, and has mean $n_S d_D + O\big(n_S\sqrt{n}_D\big) = n_S d_D + O(n_D)$. So, applying Lemma 3.6 with t a large multiple of n_D, we have $e(D, V_Z) = n_S d_D + O(n_D)$ with probability at least $1 - \gamma_1/(800Q_2)$. The expected number of indices i for which this fails is $(\gamma_1/(800Q_2))n_S$, so by Markov's inequality, the probability it fails for more than $(\gamma_1/(8Q_2))n_S$ indices i is at most 0.99.

Next we consider (ii). For each i and each $\{\boldsymbol{x}, \boldsymbol{y}\}\in\binom{X}{2}$, let

$$d_i = \Big(d_W(\boldsymbol{x}) + d_{U^0}(\boldsymbol{x}) + d_{V_{S_i}}(\boldsymbol{x})\Big) - \Big(d_W(\boldsymbol{y}) + d_{U^0}(\boldsymbol{y}) + d_{V_{S_i}}(\boldsymbol{y})\Big) = o(\sqrt{n_D}) + O(n_Z),$$

so $|d_i| \le \sqrt{n_D}$ for small δ. Then, observe that the random variable

$$d_{U_i}(\boldsymbol{x}) - d_{U_i}(\boldsymbol{y}) - d_i = d_D(\boldsymbol{y}) - d_D(\boldsymbol{x})$$

is of $(2n_D, 1/2, O(1), |N_{D^1}(\boldsymbol{x})\triangle N_{D^1}(\boldsymbol{y})|)^*$-hypergeometric type. So, by part (i) of Claim 4.8 and Lemma 3.8, $\Pr(d_{U_i}(\boldsymbol{x}) = d_{U_i}(\boldsymbol{y})) = O\big(1/\sqrt{n_D}\big)$. Let H_i be the graph of pairs $\{\boldsymbol{x}, \boldsymbol{y}\} \in \binom{X}{2}$ satisfying $d_{U_i}(\boldsymbol{x}) = d_{U_i}(\boldsymbol{y})$, so we have $\mathbb{E}e(H_i) = O\big(\sqrt{n_D}\big)$. By Markov's inequality, with probability at least $1-\gamma_1/(800Q_2)$ we have $e(H_i) = O\big(\sqrt{n_D}\big)$, in which case by Proposition 3.4, H_i has an independent set X_i of size $2\gamma_3\sqrt{n}$, for some $\gamma_3 > 0$. The expected proportion of indices i for which this fails to occur is $\gamma_1/(800Q_2)$, and by Markov's inequality again, with probability at least 0.99 it fails for only a $\gamma_1/(8Q_2)$ proportion.

Finally we consider (iii). For each $\boldsymbol{x} \in X$, $d_D(\boldsymbol{x})$ is of $(2n_D,\ 1/2,\ O(1))$-hypergeometric type, and by (ii) in Claim 4.8, it has mean $d_D + O(\sqrt{n_D})$. Therefore, by Lemma 3.6, with large enough Q_4, we have $|d_D - d_D(\boldsymbol{x})| \le Q_4\sqrt{n_D}$ with probability at least $1-\gamma_3/100$, and by Markov's inequality the probability this fails for more than $\gamma_3|X|$ vertices is at most 0.99. □

Now it is a relatively simple matter to put everything together to prove Lemma 4.3. Fix $\gamma_1, Q_2, \gamma_3, Q_4$ and U such that all parts of the above claim are satisfied. By (iii), for any $0 \le i \le n_S$, any $\boldsymbol{x} \in X^*$, and small δ, we have

$$|d_{U_i}(\boldsymbol{x}) - (\alpha d_U + d_W)| \le d_{V_{S_i}}(\boldsymbol{x}) + Q_4\sqrt{n_D} + o(\sqrt{n_D}) = O(n_S) + Q_4\sqrt{n_D} < 2Q_4\sqrt{n_D}. \quad (5)$$

By Lemma 3.10 (with $\lambda = 3\gamma_1 n_S\sqrt{n_D}$, $\rho = Q_2\sqrt{n_D}$, $\kappa = \gamma_1 n_S\sqrt{n_D}$ and $\sigma = \sqrt{n_D}$) and parts (iv) and (v) of the above claim, for large enough Q_2 there is an increasing subsequence $i_1, \dots, i_t$, with $t \ge \gamma_1 n_S/(2Q_2)$, such that $e_{i-1} - e_i \ge \sqrt{n_D}$ for each $1 < i \le t$. Delete all indices not in $\mathcal{I}_1 \cap \mathcal{I}_2$ (there are at most $\gamma_1 n_S/(4Q_2)$ such) to obtain a subsubsequence $i'_1, \dots, i'_s$ with $s \ge \gamma_1 n_S/(4Q_2)$. Let $\mathcal{I}$ contain every $4Q_4$th element of this subsubsequence, so that $|\mathcal{I}| = \Theta(n_S) = \Theta(n_Z)$ and

$$|e_i - e_{i'}| = |e(U_i) - e(U_{i'})| \ge 4Q_4\sqrt{n_D}$$

for every pair of distinct indices $i, i' \in \mathcal{I}$. Recalling (5), this means that for different $i \in \mathcal{I}$, there is no overlap between the sets of values $\{e(U_i) + d_{U_i}(\boldsymbol{x}) : \boldsymbol{x} \in X^*\}$. By the definition of X_i in (ii) of Claim 4.9, this means that for each of the $\Theta\big(n_Z\sqrt{n_D}\big)$ choices of $i \in \mathcal{I}$ and $\boldsymbol{x} \in X_i \cap X^*$, the values $e\big(W \cup U \cup V_{S_i\cup\{\boldsymbol{x}\}}\big) = e(U_i) + d_{U_i}(\boldsymbol{x})$ are in fact distinct. It remains to show that the $e\big(U, V_{S_i\cup\{\boldsymbol{x}\}}\big)$ are close to their expectations $\alpha n_Z d_{U^0}$. We have $e\big(U^0, V_{S_i\cup\{\boldsymbol{x}\}}\big) = n_Z d_{U^0}$, $d_D = (1-\alpha)d_{U^0}$ and $n_S = n_Z - 1$, so by (i) and (iii) in Claim 4.9, for sufficiently large B,

$$\begin{aligned}\big|e\big(U, V_{S_i\cup\{\boldsymbol{x}\}}\big) - \alpha n_Z d_{U^0}\big| &= \big|e\big(U^0, V_{S_i\cup\{\boldsymbol{x}\}}\big) - e(D, V_{S_i}) - d_D(\boldsymbol{x}) - n_Z d_{U^0} + n_S d_D + d_D\big| \\ &\le O(n_D + \sqrt{n_D}) \le Bn_D.\end{aligned}$$

We have proved that the statements of Claims 4.8 and 4.9 hold together with probability at least $(3/4)(1/3) = 1/4$, in which case the desired conclusion holds.

5 Concluding remarks

We have proved the Erdős–Faudree–Sós conjecture that for any fixed C, if G is an n-vertex graph with no homogeneous subgraph on $C\log n$ vertices, then G contains $\Omega\big(n^{5/2}\big)$ induced subgraphs, no pair of which have the same numbers of vertices and edges. We feel that this area is still a long way from maturity, and there is much more room for further research towards understanding the structure of C-Ramsey graphs. We hope that such research will inform future work on explicit constructions of Ramsey graphs.

Regarding specific open questions, of course the Erdős–McKay conjecture remains an intriguing problem. We would also like to draw attention to the subject of subgraphs with many different degrees: as mentioned in the introduction, answering a different conjecture of Erdős, Faudree and Sós [20, 21], Bukh and Sudakov [10] proved that C-Ramsey graphs have induced subgraphs with $\Omega(\sqrt{n})$ different degrees. However, in random graphs one can actually find induced subgraphs with $\Omega(n^{2/3})$ distinct degrees (this was proved in an unpublished paper of Conlon, Morris, Samotij and Saxton [13]), and it is not clear whether such an improved bound also holds for C-Ramsey graphs.

Additionally, observe that the main result of this paper can be rephrased as the fact that in an $O(1)$-Ramsey graph, for most choices of ℓ, there are many possibilities for the number of edges in a subset of ℓ vertices. We believe a natural next step would be to study statistical properties of the number of edges in a *random* set of ℓ vertices. For example, is this random variable anticoncentrated? For general graphs this question was first studied by Alon, Hefetz, Krivelevich and Tyomkyn [3] (see [28, 23, 29] for further work). Regarding Ramsey graphs, as we recently proposed in a paper with Tuan Tran [28], could it be true that in any $O(1)$-Ramsey graph G, if A is a uniformly random set of $n/2$ vertices, then $\Pr(e(G[A]) = x) = O(1/n)$ for all x? In [27] we also formulated a version of this question for random subsets where the presence of each vertex is chosen independently, which may be more tractable.

Finally, we believe an interesting further direction of research would be to consider regimes where larger homogeneous subgraphs are forbidden (see [2, 7, 5, 31] for some examples of theorems of this type). In [27] we proposed the conjecture that $|\Phi(G)| = \Omega(e(G))$ for graphs G which have no homogeneous subgraph on $n/4$ vertices; we do not know a good counterpart of this conjecture for $|\Psi(G)|$, but it seems likely that some nontrivial bound should hold.

Acknowledgment. The authors would like to thank the referee for their careful reading of the manuscript and their valuable comments. We would also like to thank Mantas Baksys and Xuanang Chen for carefully reading the paper and finding an oversight in the proof (related to the definition of richness in Section 3.1).

References

[1] N. Alon, J. Balogh, A. Kostochka, and W. Samotij, *Sizes of induced subgraphs of Ramsey graphs*, Combin. Probab. Comput. **18** (2009), no. 4, 459–476.

[2] N. Alon and B. Bollobás, *Graphs with a small number of distinct induced subgraphs*, Discrete Math. **75** (1989), no. 1-3, 23–30, Graph theory and combinatorics (Cambridge, 1988).

[3] N. Alon, D. Hefetz, M. Krivelevich and M. Tyomkyn, *Edge-statistics on large graphs*, arXiv preprint arXiv:1805.06848 (2018).

[4] N. Alon and A. V. Kostochka, *Induced subgraphs with distinct sizes*, Random Structures Algorithms **34** (2009), no. 1, 45–53.

[5] N. Alon, M. Krivelevich, and B. Sudakov, *Induced subgraphs of prescribed size*, J. Graph Theory **43** (2003), no. 4, 239–251.

[6] N. Alon and J. H. Spencer, *The probabilistic method*, fourth ed., Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., Hoboken, NJ, 2016.

[7] M. Axenovich and J. Balogh, *Graphs having small number of sizes on induced k-subgraphs*, SIAM J. Discrete Math. **21** (2007), no. 1, 264–272.

[8] B. Barak, A. Rao, R. Shaltiel, and A. Wigderson, *2-source dispersers for $n^{o(1)}$ entropy, and Ramsey graphs beating the Frankl-Wilson construction*, Ann. of Math. (2) **176** (2012), no. 3, 1483–1543.

[9] A. Bikelis, *The estimation of the remainder term in the central limit theorem for samples taken from finite sets*, Studia Sci. Math. Hungar. **4** (1969), 345–354.

[10] B. Bukh and B. Sudakov, *Induced subgraphs of Ramsey graphs with many distinct degrees*, J. Combin. Theory Ser. B **97** (2007), no. 4, 612–619.

[11] E. Chattopadhyay and D. Zuckerman, *Explicit two-source extractors and resilient functions*, STOC'16—Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, ACM, New York, 2016, pp. 670–683.

[12] G. Cohen, *Two-source dispersers for polylogarithmic entropy and improved Ramsey graphs*, STOC'16—Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, ACM, New York, 2016, pp. 278–284.

[13] D. Conlon, R. Morris, W. Samotij, and D. Saxton, *The number of distinct degrees in an induced subgraph of a random graph*, personal communication.

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

[15] P. Erdős, *On some of my favourite problems in various branches of combinatorics*, Fourth Czechoslovakian Symposium on Combinatorics, Graphs and Complexity (Prachatice, 1990), Ann. Discrete Math., vol. 51, North-Holland, Amsterdam, 1992, pp. 69–79.

[16] P. Erdős and A. Hajnal, *On spanned subgraphs of graphs*, Contributions to graph theory and its applications (Internat. Colloq., Oberhof, 1977) (German), Tech. Hochschule Ilmenau, Ilmenau, 1977, pp. 80–96.

[17] P. Erdős and R. Rado, *Intersection theorems for systems of sets*, J. London Math. Soc. **35** (1960), 85–90.

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

[19] P. Erdős and A. Szemerédi, *On a Ramsey type theorem*, Period. Math. Hungar. **2** (1972), 295–299, Collection of articles dedicated to the memory of Alfréd Rényi, I.

[20] P. Erdős, *Some of my favourite problems in various branches of combinatorics*, Matematiche (Catania) **47** (1992), no. 2, 231–240 (1993), Combinatorics 92 (Catania, 1992).

[21] P. Erdős, *Some recent problems and results in graph theory*, Discrete Math. **164** (1997), no. 1-3, 81–85, The Second Krakow Conference on Graph Theory (Zgorzelisko, 1994).

[22] P. Erdős, M. Goldberg, J. Pach, and J. Spencer, *Cutting a graph into two dissimilar halves*, J. Graph Theory **12** (1988), no. 1, 121–131.

[23] J. Fox and L. Sauermann, *A completion of the proof of the Edge-statistics Conjecture*, arXiv preprint arXiv:1809.01352 (2018).

[24] P. Frankl and R. M. Wilson, *Intersection theorems with geometric consequences*, Combinatorica **1** (1981), no. 4, 357–368.

[25] C. Greenhill, M. Isaev, M. Kwan, and B. D. McKay, *The average number of spanning trees in sparse graphs with given degrees*, European J. Combin. **63** (2017), 6–25.

[26] T. Höglund, *Sampling from a finite population: a remainder term estimate*, Scand. J. Statist. **5** (1978), no. 1, 69–71.

[27] M. Kwan and B. Sudakov, *Ramsey graphs induce subgraphs of quadratically many sizes*, Int. Math. Res. Not. IMRN, to appear, arXiv preprint arXiv:1711.02937 (2017).

[28] M. Kwan and B. Sudakov and T. Tran, *Anticoncentration for subgraph statistics*, arXiv preprint arXiv:1807.05202 (2018).

[29] A. Martinsson, F. Mousset, A. Noever and M. Trujić, *The edge-statistics conjecture for $\ell \ll k^{6/5}$*, arXiv preprint arXiv:1809.02576 (2018).

[30] B. Narayanan, J. Sahasrabudhe, and I. Tomon, *Ramsey graphs induce subgraphs of many different sizes*, Combinatorica, to appear, arXiv preprint arXiv:1609.01705 (2016).

[31] B. Narayanan and I. Tomon, *Induced subgraphs with many distinct degrees*, Combin. Probab. Comput. (2017), 1–14.

[32] H. J. Prömel and V. Rödl, *Non-Ramsey graphs are $c \log n$-universal*, J. Combin. Theory Ser. A **88** (1999), no. 2, 379–384.

[33] S. Shelah, *Erdős and Rényi conjecture*, J. Combin. Theory Ser. A **82** (1998), no. 2, 179–185.