Forum of Mathematics, Pi (2023), **Vol. 11**:e19 1–13
doi:10.1017/fmp.2023.19

CAMBRIDGE
UNIVERSITY PRESS

RESEARCH ARTICLE

Resolution of the Erdős–Sauer problem on regular subgraphs

Oliver Janzer [1] and Benny Sudakov [2]

[1]Trinity College, University of Cambridge, Trinity Street, CB2 1TQ Cambridge, United Kingdom; E-mail: oj224@cam.ac.uk.
[2]Department of Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland;
E-mail: benjamin.sudakov@math.ethz.ch.

Received: 2 November 2022; **Accepted:** 29 June 2023

2020 Mathematics Subject Classification: *Primary* – 05C35; *Secondary* – 05C07

Abstract

In this paper, we completely resolve the well-known problem of Erdős and Sauer from 1975 which asks for the maximum number of edges an n-vertex graph can have without containing a k-regular subgraph, for some fixed integer $k \geq 3$. We prove that any n-vertex graph with average degree at least $C_k \log\log n$ contains a k-regular subgraph. This matches the lower bound of Pyber, Rödl and Szemerédi and substantially improves an old result of Pyber, who showed that average degree at least $C_k \log n$ is enough.

Our method can also be used to settle asymptotically a problem raised by Erdős and Simonovits in 1970 on almost regular subgraphs of sparse graphs and to make progress on the well-known question of Thomassen from 1983 on finding subgraphs with large girth and large average degree.

Contents

1. Introduction

The problem of finding regular subgraphs in graphs has a very long history. Note that finding a 1-regular subgraph is the same as finding a matching. One of the oldest results of graph theory is Petersen's theorem from 1891 [24], which states that every cubic, bridgeless graph contains a perfect matching. The celebrated Hall's theorem [17] gives a necessary and sufficient condition for a bipartite graph to have a perfect matching, while Tutte's theorem [29] gives such a condition for an arbitrary graph. Later, Tutte [30] found a necessary and sufficient condition for a graph to contain a k-regular spanning subgraph and, more generally, for it to contain an f-factor – that is, a spanning subgraph in which each vertex v has degree $f(v)$.

The problem of finding general, not necessarily spanning, regular subgraphs was also extensively studied. In 1975, Erdős and Sauer [14] asked the following natural extremal question. Given a positive integer k, what is the maximum number of edges that an n-vertex graph can have if it does not contain a k-regular subgraph? The problem also appeared in Bollobás's book on extremal graph theory [7] and in the book of Bondy and Murty [9]. Later, Erdős [16] mentioned this as one of his favourite unsolved problems. Let us write $f_k(n)$ for the smallest number of edges which guarantees a k-regular subgraph. Trivially $f_2(n) = n$, but already for $k = 3$, the answer is unclear. Erdős and Sauer observed that $f_3(n) = O(n^{8/5})$ follows from the known upper bound for the Turán number of the cube [15] and suggested that $f_k(n) \le n^{1+\varepsilon}$ for any fixed $\varepsilon > 0$ and sufficiently large n.

A very influential algebraic technique to find regular subgraphs was developed in the early 1980s by Alon, Friedland and Kalai. In [4], motivated by a conjecture of Berge and Sauer, they showed that any 4-regular multigraph plus an edge contains a 3-regular subgraph. Alon, Friedland and Kalai [5] also proved several results which state, roughly speaking, that nearly regular graphs have regular subgraphs with not too small degree. One of these results, together with an ingenious argument, was used by Pyber [25] in 1985 to show that indeed $f_k(n) \le n^{1+\varepsilon}$. More precisely, Pyber proved that any n-vertex graph with average degree at least $C_k \log n$ contains a k-regular subgraph (here and in the rest of the paper logarithms are to the base two).

At this point, the best lower bound was due to Chvátal [14], who had shown that $f_3(2n + 3) > 6n$. This was greatly improved by Pyber, Rödl and Szemerédi [26], who found a remarkable construction of graphs with a superlinear number of edges that do not contain k-regular subgraphs.

Theorem 1.1 (Pyber–Rödl–Szemerédi [26]). *There is some absolute constant $c > 0$ such that for every n there exists an n-vertex graph with at least $cn \log\log n$ edges which does not contain a k-regular subgraph for any $k \ge 3$.*

Furthermore, Pyber, Rödl and Szemerédi generalized Pyber's result to show that for any positive integer k, there is a constant $C = C(k)$ such that any graph with maximum degree Δ and average degree at least $C \log \Delta$ contains a k-regular subgraph.

Despite substantial interest from many researchers, the above bounds have not been improved in the last 30 years. At the same time, several variants of the original problem have been considered. Bollobás, Kim and Verstraëte [8] studied the threshold for a random graph to have a k-regular subgraph. Rödl and Wysocka [27] investigated the largest r for which every n-vertex graph with at least γn^2 edges has an r-regular subgraph. Many papers have been written on the existence of regular subgraphs in hypergraphs; see, for example, [23, 11, 20, 19, 18].

In this paper, we prove the following result, which completely resolves the problem of Erdős and Sauer.

Theorem 1.2. *For any positive integer k, there is a constant $C = C(k)$ such that any graph with maximum degree $\Delta \ge 3$ and average degree at least $C \log\log \Delta$ contains a k-regular subgraph. In particular, any n-vertex graph with average degree at least $C \log\log n$ contains such a subgraph.*

Our results also make progress on two other old and well-known problems. The first one, due to Erdős and Simonovits from 1970 [15], concerns the question of how dense an almost-regular subgraph there must exist in a graph with n vertices and $n \log n$ edges. An almost-regular graph here is one in which the maximum degree and the minimum degree differ by at most a constant factor. Our results resolve this question asymptotically, showing that one can find an almost-regular subgraph with $m = \omega(1)$ vertices and at least $m(\log m)^{1/2-o(1)}$ edges, which is tight by a result of Alon [1]. Since the full discussion of this topic takes considerable space, we postpone it to the concluding remarks (Section 6).

Finally, let us discuss the other application of our main result. In 1983, Thomassen [28] conjectured that, for every $t, g \in \mathbb{N}$, there exists some d such that any graph with average degree at least d contains a subgraph with average degree at least t and girth at least g. Kühn and Osthus [21] proved this for $g \le 6$ (see also [22] for an improved bound on d), but the general case is wide open. It is easy to see that for every $t, g \in \mathbb{N}$, if d is sufficiently large, then any d-regular graph has a subgraph with average degree at least t and girth at least g. Hence, our Theorem 1.2 implies that any graph G with average degree at least $C(t, g) \log\log \Delta(G)$ contains a subgraph with average degree at least t and girth at least g. This improves the best bound towards Thomassen's conjecture, due to Dellamonica and Rödl [12], which states that any graph G with average degree at least $\alpha(t,g)(\log\log\Delta(G))^{\beta(t,g)}$ contains a subgraph with average degree at least t and girth at least g.

The rest of the paper is organized as follows. In the next section, we give a sketch of the proof of Theorem 1.2. In Section 3, we prove some simple preliminary lemmas. The key lemma is proved in Section 4. The proof of Theorem 1.2 is then completed in Section 5. We give some concluding remarks about large almost-regular subgraphs in Section 6.

Notation. We write $e(G)$ for the number of edges in a graph G. For a set $S \subset V(G)$, $G[S]$ stands for the induced subgraph of G on vertex set S. For a vertex $u \in V(G)$, $N_G(u)$ denotes the neighbourhood of u in G and we write $d_G(u)$ for the degree of u. Given vertices u and u', $d_G(u,u')$ stands for the number of common neighbours of u and u' in G. We write $\Delta(G)$ for the maximum degree of G.

2. An overview of the proof

Although our proof is short, we think that it may be useful to give a sketch of the main ideas. We start by briefly discussing the lower bound construction of Pyber, Rödl and Szemerédi since it partly motivates our argument. Their (random) graph can be described roughly as follows. The vertex set is $A \cup B$, where $|B| = n$, A is the disjoint union of sets $A(j)$ for $\frac{1}{4}\log\log n \le j \le \frac{1}{2}\log\log n$ with $|A(j)| = n/2^{2^j}$, and for each j and $v \in B$, v has a unique neighbour in $A(j)$, chosen uniformly at random. Note that a typical vertex in $A(j)$ has degree about 2^{2^j}.

In a general bipartite graph, we obtain a similar partitioning of part A according to the degrees and show, very roughly speaking, that if there is some j such that each $v \in B$ has at least a large constant number of neighbours in $A(j)$ (unlike in the above construction where each $v \in B$ has only one neighbour there), then G has a k-regular subgraph.

More precisely, let r be a large constant (that can depend on k), and let G be a bipartite graph with parts A and B such that for some positive integers $s, t \ge r$, every $v \in B$ has degree r, the average degree of a vertex in A is at least 2^s, the maximum degree of a vertex in A is at most 2^t and $t \le (1 + \frac{1}{r-1})s$. We remark that it is possible (and not too hard) to find such a subgraph in any graph with maximum degree Δ and average degree at least $100r^2\log\log\Delta$. For the sake of simplicity, let us also suppose that G is C_4-free, although this assumption can be significantly relaxed – it suffices to assume that the codegrees in G are not extremely large. We shall now argue that G has a k-regular subgraph.

Our key lemma (stated and proved in Section 4) provides a subgraph which, although is not necessarily regular, has much better regularity properties than G. More precisely, it asserts the existence of subsets $A'' \subset A$, $B'' \subset B$ and positive integers $s', t' \ge r$ such that in the graph $G[A'' \cup B'']$ all the previous conditions are satisfied (with A'', B'', s', t' in place of A, B, s, t), and, additionally, $t' - s' \le \log(40(t-s)r^2)$. Hence, as long as $t - s > 3\log r$, $t' - s'$ is smaller than $t - s$. Iteratively applying this lemma, eventually we obtain subsets $A^* \subset A$, $B^* \subset B$ and positive integers $s^*, t^* \ge r$ such that in the graph $G[A^* \cup B^*]$, every $v \in B^*$ has degree r, the average degree of a vertex in A^* is at least 2^{s^*}, the maximum degree of a vertex in A^* is at most 2^{t^*} and $t^* - s^* \le 3\log r$. The last property means that the maximum degree is at most $2^{3\log r}$ times the average degree in A^*. Taking a random subset $\hat{B} \subset B^*$ of size $|A^*|$ and deleting the vertices in $G[A^* \cup \hat{B}]$ whose degree is much larger than expected, we obtain a subgraph with average degree about r and maximum degree at most about $r2^{3\log r} = r^4$. We can then use a result of Pyber, Rödl and Szemerédi (see Theorem 3.8) to find a k-regular subgraph in this graph.

We shall now sketch the proof of our key lemma. Let $A_{s+1} = \{u \in A : d_G(u) \le 2^{s+1}\}$, and for each $s+2 \le i \le t$, let $A_i = \{u \in A : 2^{i-1} < d_G(u) \le 2^i\}$. Observe that the sets $A_{s+1}, \dots, A_t$ partition A. For each $u \in A$, let $\alpha(u)$ be the unique i with $u \in A_i$. For $v \in B$, let $\beta(v) = \sum_{u \in N_G(v)} \alpha(u)$. Note that for every $v \in B$, we have $(s+1)r \le \beta(v) \le tr$, so there are at most $(t-s)r$ possible values for $\beta(v)$. Hence, by the pigeon hole principle, there are some $\gamma \ge (s+1)r$ and a subset $\tilde{B} \subset B$ of size at least $\frac{|B|}{(t-s)r}$ such that $\beta(v) = \gamma$ for all $v \in \tilde{B}$.

Let A' be a random subset of A where each $u \in A$ is kept independently with probability $2^{\alpha(u)-t}$. Let $B' = \{v \in \tilde{B} : N_G(v) \subset A'\}$. Let us see why we expect $G[A' \cup B']$ to be nearly biregular. Firstly, every

https://doi.org/10.1017/fmp.2023.19 Published online by Cambridge University Press

$v \in B'$ has degree precisely r in $G[A' \cup B']$. Now, let $u \in A$. We claim that, conditional on $u \in A'$, $|N_G(u) \cap B'|$ is distributed as a binomial random variable $\mathrm{Bin}\big(|N_G(u) \cap \tilde{B}|, 2^{\gamma-\alpha(u)-(r-1)t}\big)$. Indeed, conditional on $u \in A'$, each $v \in N_G(u) \cap \tilde{B}$ belongs to B' with probability $\prod_{w \in N_G(v)\setminus\{u\}} 2^{\alpha(w)-t} = 2^{\beta(v)-\alpha(u)-(r-1)t} = 2^{\gamma-\alpha(u)-(r-1)t}$ and these events are independent for all v since $N_G(v) \setminus \{u\}$ are pairwise disjoint as G is C_4-free. Thus, if $u \in A'$, then the degree of u in $G[A' \cup B']$ is concentrated around $|N_G(u) \cap \tilde{B}|2^{\gamma-\alpha(u)-(r-1)t}$. Since $|N_G(u) \cap \tilde{B}| \le d_G(u) \le 2^{\alpha(u)}$, we see that it is very unlikely that the degree of u in $G[A' \cup B']$ is much larger than $2^{\gamma-(r-1)t}$. On the other hand, $\tilde{B}$ has size similar to B, so on average $|N_G(u) \cap \tilde{B}|$ is not much smaller than $d_G(u)$, which is in turn at least $2^{\alpha(u)-1}$ (unless $\alpha(u) = s+1$). Hence, it is not hard to see that the average degree in $G[A' \cup B']$ of a vertex in A' is expected to be not much smaller than $2^{\gamma-(r-1)t}$. A small loss arises from the fact that $\tilde{B}$ is potentially smaller than B by a factor of $(t-s)r$; this partly explains why $t'-s'$ can be as large as $\log(40(t-s)r^2)$. We remark that the condition $t \le (1+\frac{1}{r-1})s$ ensures that $\gamma-(r-1)t \ge sr-(r-1)t+r \ge r$, so the expected average degree of the vertices in A' is not too small.

Now, we take A'' to be the set of those vertices in A' whose degree in $G[A' \cup B']$ is not much larger than expected (i.e., not much larger than $2^{\gamma-(r-1)t}$) and set $B'' = \{v \in B' : N_G(v) \subset A''\}$. By the strong concentration of the degrees, $G[A'' \cup B'']$ has almost as many edges as $G[A' \cup B']$, so the average degree of a vertex in A'' is not much smaller than it was in A'. Hence, the desired conditions about A'' and B'' are satisfied (for suitable $s' \approx \gamma-(r-1)t$ and $t' \approx \gamma-(r-1)t$).

3. Preliminaries

As we have mentioned in the outline of the proof, one of the conditions in our key lemma is that no two vertices in the graph have very large codegree. We prove that with the loss of a constant factor in the number of edges, we can pass to a subgraph satisfying this property. For this, we can assume that our graphs are $K_{k,k}$-free (or else, they contain a k-regular subgraph).

Lemma 3.1. *Let k be a positive integer, and let H be a $K_{k,k}$-free bipartite graph with parts S and T. Assume that $d_H(u) \ge k|T|^{1-1/k}$ holds for every $u \in S$. Then $e(H) \le k|T|$.*

Proof. Suppose that $e(H) > k|T|$. Then, by Jensen's inequality,

$$\sum_{v \in T} \binom{d_H(v)}{k} \ge |T|\binom{e(H)/|T|}{k} \ge |T|\left(\frac{e(H)}{k|T|}\right)^k.$$

By the degree condition we have $e(H) \ge k|S||T|^{1-1/k}$. Then $\frac{e(H)}{k|T|} \ge |S||T|^{-1/k}$ and therefore, $\sum_{v \in T}\binom{d_H(v)}{k} \ge |S|^k \ge k\binom{|S|}{k}$. This implies that there is some $R \subset S$ of size k whose common neighbourhood in T has size at least k. Then H contains $K_{k,k}$, which is a contradiction. □

The proof of the next lemma is somewhat similar to that of Lemma 2.3 from [22].

Lemma 3.2. *Let k be a positive integer. Let G be a $K_{k,k}$-free bipartite graph with parts A and B, and assume that $d_G(u) \le m$ holds for all $u \in A$. Then G has a spanning subgraph G' such that $e(G') \ge e(G)/(k+1)$ and $d_{G'}(u,u') \le km^{1-1/k}$ for any two distinct $u,u' \in A$.*

Proof. Let $A = \{u_1, u_2, \dots, u_n\}$, and let $G_0 = G$. Let us define spanning subgraphs $G_1, \dots, G_n$ of G recursively as follows. Having defined G_{i-1} for some $1 \le i \le n$, let $T = N_{G_{i-1}}(u_i)$, let $S = \{u_j : j > i \text{ and } d_{G_{i-1}}(u_i,u_j) > km^{1-1/k}\}$ and let $H = G_{i-1}[S \cup T]$. Since H is a subgraph of G, it is $K_{k,k}$-free. Moreover, for any $u_j \in S$, $d_H(u_j) = d_{G_{i-1}}(u_i,u_j) > km^{1-1/k} \ge k|T|^{1-1/k}$. Hence, by Lemma 3.1, $e(H) \le k|T|$. Let G_i be the subgraph of G_{i-1} obtained by deleting the edges in H. Then $e(G_{i-1}) - e(G_i) = e(H) \le kd_{G_{i-1}}(u_i)$.

https://doi.org/10.1017/fmp.2023.19 Published online by Cambridge University Press

Observe that for any $i < j$, we have $d_{G_n}(u_i, u_j) \le km^{1-1/k}$. Moreover, note that for any $1 \le i \le n$, $d_{G_n}(u_i) = d_{G_{i-1}}(u_i)$, so

$$e(G_n) = \sum_{i=1}^{n} d_{G_n}(u_i) = \sum_{i=1}^{n} d_{G_{i-1}}(u_i) \ge \sum_{i=1}^{n} (e(G_{i-1}) - e(G_i))/k = (e(G_0) - e(G_n))/k.$$

Thus, $e(G_n) \ge e(G_0)/(k+1) = e(G)/(k+1)$, which means that $G' = G_n$ satisfies the conditions described in the lemma. □

We now define two notions of 'near-regularity' that will be used in our proofs.

Definition 3.3. A graph G is called *K-almost-regular* if the maximum degree of G is at most K times the minimum degree of G.

Definition 3.4. We say that a bipartite graph G is *(L, d)-almost-biregular* if the following holds. G has parts A and B, where $d_G(v) = d$ for every $v \in B$, and, writing $D = e(G)/|A|$, we have $D \ge d$ (equivalently $|A| \le |B|$) and $d_G(u) \le LD$ for every $u \in A$.

Our key lemma will provide an (L, d)-almost-biregular subgraph where L is fairly small. The next lemma allows us to find a large almost-regular subgraph in it.

Lemma 3.5. *Let G be an (L, δ)-almost-biregular graph for some $L \ge \delta \ge 2$. Then G has a 64-almost-regular subgraph with average degree at least $\frac{\delta}{16 \log L}$.*

The proof of this lemma uses the following result, which is a slight variant of Lemma 2.7 from [26]. It states that as long as L is subexponential in the average degree, we can pass to a subgraph with large average degree and constant L.

Lemma 3.6. *Let G be an (L, δ)-almost-biregular graph. Suppose that $L\delta \le 2^{\lfloor \delta/(d-1) \rfloor}$ and $d \le \delta$. Then G has a $(4, d)$-almost-biregular subgraph.*

Since the proof of this result is almost identical to that of Lemma 2.7 from [26], it is omitted here.

The next simple lemma shows that in an almost-biregular graph we can find an almost-regular subgraph with similar average degree.

Lemma 3.7. *Let $L, d \ge 1$, and let G be an (L, d)-almost-biregular graph. Then G has a nonempty subgraph G' with average degree at least $d/2$ and maximum degree at most $4Ld$.*

Proof. Choose A, B and D according to Definition 3.4. Define B' to be a random subset of B where each $v \in B$ is kept independently with probability $\frac{d}{D}$. Let G' be the subgraph of $G[A \cup B']$ obtained by deleting all edges uv with $|N_G(u) \cap B'| \ge 4Ld$. Clearly, $\Delta(G') \le \max(4Ld, d) = 4Ld$.

Let $X = e(G[A \cup B'])$, and let Y be the number of edges $uv \in E(G[A \cup B'])$ with $|N_G(u) \cap B'| \ge 4Ld$. Now,

$$\mathbb{E}[X] = \sum_{uv \in E(G)} \mathbb{P}(v \in B') = e(G)\frac{d}{D}$$

and

$$\begin{aligned}\mathbb{E}[Y] &= \sum_{uv \in E(G)} \mathbb{P}(v \in B' \text{ and } |N_G(u) \cap B'| \ge 4Ld) \\ &= \sum_{uv \in E(G)} \mathbb{P}(v \in B')\mathbb{P}\left(|N_G(u) \cap B'| \ge 4Ld \mid v \in B'\right).\end{aligned}$$

For any $uv \in E(G)$, we have

$$\mathbb{E}[|N_G(u) \cap B'| \mid v \in B'] = 1 + (d_G(u) - 1)\frac{d}{D} \le 1 + LD\frac{d}{D} = 1 + Ld \le 2Ld,$$

https://doi.org/10.1017/fmp.2023.19 Published online by Cambridge University Press

so it follows by Markov's inequality that

$$\mathbb{P}\left(|N_G(u) \cap B'| \geq 4Ld \mid v \in B'\right) \leq 1/2.$$

Hence,

$$\mathbb{E}[Y] \leq \sum_{uv \in E(G)} \mathbb{P}(v \in B')/2 = e(G)\frac{d}{2D}.$$

Thus, using $|A|D = |B|d = e(G)$,

$$\begin{aligned}\mathbb{E}[X - Y - (|A| + |B'|)d/4] &\geq e(G)\frac{d}{D} - e(G)\frac{d}{2D} - |A|d/4 - |B|\frac{d}{D}d/4 \\ &= e(G)\frac{d}{D} - e(G)\frac{d}{2D} - e(G)\frac{d}{4D} - e(G)\frac{d}{4D} = 0.\end{aligned}$$

In particular, there is an outcome for which $X - Y \geq (|A| + |B'|)d/4$. Since $e(G') = X - Y$, this means that G' has average degree at least $d/2$. □

Lemma 3.5 can now be proven using the last two lemmas.

Proof of Lemma 3.5. Let $d = \lceil \frac{\delta}{4\log L} \rceil$. Now, $\delta/(d-1) \geq 4\log L$, so $\lfloor \delta/(d-1) \rfloor \geq 2\log L$. Hence, $2^{\lfloor \delta/(d-1) \rfloor} \geq L^2 \geq L\delta$. By Lemma 3.6, G has a $(4, d)$-almost-biregular subgraph G'. By Lemma 3.7, G' has a subgraph with average degree at least $d/2$ and maximum degree at most $16d$. Repeatedly discarding vertices of degree less than $d/4$, we end up with a nonempty subgraph G'' with minimum degree at least $d/4$ and maximum degree at most $16d$. Clearly, G'' is 64-almost-regular. Moreover, the average degree of G'' is at least $d/4 \geq \frac{\delta}{16\log L}$. □

In order to find a k-regular subgraph in an almost-regular graph with sufficiently large average degree, one can use the result of Pyber, Rödl and Szemerédi.

Theorem 3.8 (Pyber–Rödl–Szemerédi [26]). *For any positive integer k, there is a constant $C = C(k)$ such that any graph with maximum degree Δ and average degree at least $C\log\Delta$ contains a k-regular subgraph.*

4. The key lemma

In this section, we prove the following lemma, which is the main ingredient in our proof. Given a bipartite graph with parts A and B in which B is regular, the maximum degree in A is at most about $1+\varepsilon$ power of the average degree in A and the codegrees are not extremely large, the lemma provides an induced subgraph with much better regularity properties. In the statement of the lemma, we hide the condition bounding the maximum degree in terms of the average degree of A in the upper bound on the codegrees: For the codegree condition to hold, we must have $t \leq (1 + \frac{1}{r-1})s$.

Lemma 4.1. *Let r, s, t be positive integers such that $s < t$. Let G be a bipartite graph with parts A and B such that $d_G(v) = r$ for every $v \in B$, $d_G(u) \leq 2^t$ for every $u \in A$, $d_G(u, u') \leq 2^{rs-(r-1)t}$ for any two distinct $u, u' \in A$ and $e(G) \geq 2^s|A|$.*

Then there are subsets $A'' \subset A$ and $B'' \subset B$ such that $N_G(v) \subset A''$ for every $v \in B''$ and, writing $G' = G[A'' \cup B'']$ and $d' = e(G')/|A''|$, we have $d' \geq \frac{2^{rs-(r-1)t}}{10(t-s)r}$ and $d_{G'}(u) \leq 40(t-s)r^2 d'$ for all $u \in A''$.

Proof. Let $A_{s+1} = \{u \in A : d_G(u) \leq 2^{s+1}\}$ and for each $s+2 \leq i \leq t$, let $A_i = \{u \in A : 2^{i-1} < d_G(u) \leq 2^i\}$. Observe that the sets $A_{s+1}, \dots, A_t$ partition A. For each $u \in A$, let $\alpha(u)$ be the unique i with $u \in A_i$. For $v \in B$, let

https://doi.org/10.1017/fmp.2023.19 Published online by Cambridge University Press

$$\beta(v) = \sum_{u \in N_G(v)} \alpha(u).$$

Note that for every $v \in B$, we have $(s+1)r \leq \beta(v) \leq tr$, so there are at most $(t-s)r$ possible values for $\beta(v)$. Hence, by the pigeonhole principle, there are some $\gamma \geq (s+1)r$ and a subset $\tilde{B} \subset B$ of size at least $\frac{|B|}{(t-s)r}$ such that $\beta(v) = \gamma$ for all $v \in \tilde{B}$.

Let A' be a random subset of A where each $u \in A$ is kept independently with probability $2^{\alpha(u)-t}$. Let $B' = \{v \in \tilde{B} : N_G(v) \subset A'\}$. Let A'' be the subset of A' consisting of those vertices u with $|N_G(u) \cap B'| \leq 4r2^{\gamma-(r-1)t}$, and let $B'' = \{v \in B' : N_G(v) \subset A''\}$. Let $G' = G[A'' \cup B'']$. Then $d_{G'}(u) \leq 4r2^{\gamma-(r-1)t}$ for all $u \in A''$.

The following claim shows that we expect $E(G[A' \cup B']) \setminus E(G')$ to be small.

Claim. *For any $uv \in E(G[A \cup \tilde{B}])$ with $u \in A$ and $v \in \tilde{B}$,*

$$\mathbb{P}(u \in A', v \in B' \text{ and } |N_G(u) \cap B'| \geq 4r2^{\gamma-(r-1)t}) \leq \mathbb{P}(v \in B')/(2r).$$

Proof of Claim. Note that

$$\mathbb{E}[|N_G(u) \cap B'| \mid v \in B'] = \sum_{w \in N_G(u) \cap \tilde{B}} \mathbb{P}(w \in B' \mid v \in B')$$
$$= \sum_{\substack{w \in N_G(u) \cap \tilde{B}: \\ N_G(w) \cap N_G(v) = \{u\}}} \mathbb{P}(w \in B' \mid v \in B') + \sum_{\substack{w \in N_G(u) \cap \tilde{B}: \\ N_G(w) \cap N_G(v) \neq \{u\}}} \mathbb{P}(w \in B' \mid v \in B'). \tag{4.1}$$

We now bound the two sums separately. If some $w \in N_G(u) \cap \tilde{B}$ satisfies $N_G(w) \cap N_G(v) = \{u\}$, then

$$\mathbb{P}(w \in B' \mid v \in B') = \mathbb{P}(w \in B' \mid u \in A') = \prod_{z \in N_G(w) \setminus \{u\}} \mathbb{P}(z \in A') = \prod_{z \in N_G(w) \setminus \{u\}} 2^{\alpha(z)-t}$$
$$= 2^{\sum_{z \in N_G(w) \setminus \{u\}} \alpha(z) - (r-1)t} = 2^{\beta(w) - \alpha(u) - (r-1)t} = 2^{\gamma - \alpha(u) - (r-1)t}.$$

Since $d_G(u) \leq 2^{\alpha(u)}$, we have that

$$\sum_{\substack{w \in N_G(u) \cap \tilde{B}: \\ N_G(w) \cap N_G(v) = \{u\}}} \mathbb{P}(w \in B' \mid v \in B') \leq d_G(u) 2^{\gamma - \alpha(u) - (r-1)t} \leq 2^{\gamma-(r-1)t}.$$

As for the second sum, the number of $w \in N_G(u) \cap \tilde{B}$ with $N_G(w) \cap N_G(v) \neq \{u\}$ is at most $r2^{rs-(r-1)t}$. Indeed, v has r neighbours in G, so there are at most r ways to choose a vertex $u' \in N_G(w) \cap N_G(v) \setminus \{u\}$, and, by assumption, any such u' has at most $2^{rs-(r-1)t}$ common neighbours with u. Now, using $\gamma \geq (s+1)r$, we have $2^{\gamma-(r-1)t} \geq 2^r 2^{rs-(r-1)t} \geq r2^{rs-(r-1)t}$. This implies that

$$\sum_{\substack{w \in N_G(u) \cap \tilde{B}: \\ N_G(w) \cap N_G(v) \neq \{u\}}} \mathbb{P}(w \in B' \mid v \in B') \leq r2^{rs-(r-1)t} \leq 2^{\gamma-(r-1)t}.$$

Thus, by (4.1),

$$\mathbb{E}[|N_G(u) \cap B'| \mid v \in B'] \leq 2^{\gamma-(r-1)t+1}.$$

This implies, by Markov's inequality, that

$$\mathbb{P}(|N_G(u) \cap B'| \geq 4r2^{\gamma-(r-1)t} \mid v \in B') \leq 1/(2r).$$

Thus,

$$\begin{aligned}&\mathbb{P}(u \in A', v \in B' \text{ and } |N_G(u) \cap B'| \geq 4r2^{\gamma-(r-1)t})\\&= \mathbb{P}(v \in B' \text{ and } |N_G(u) \cap B'| \geq 4r2^{\gamma-(r-1)t})\\&= \mathbb{P}(v \in B')\mathbb{P}(|N_G(u) \cap B'| \geq 4r2^{\gamma-(r-1)t} \mid v \in B') \leq \mathbb{P}(v \in B')/(2r),\end{aligned}$$

completing the proof of the claim. □

For any $uv \in E(G[A \cup \tilde{B}])$,

$$\mathbb{P}(u \in A', v \in B') = \mathbb{P}(v \in B') = \prod_{w \in N_G(v)} \mathbb{P}(w \in A') = \prod_{w \in N_G(v)} 2^{\alpha(w)-t} = 2^{\gamma-rt}. \tag{4.2}$$

Let $X = e(G[A' \cup B'])$, and let Y be the number of edges $uv \in E(G[A' \cup B'])$ with $|N_G(u) \cap B'| \geq 4r2^{\gamma-(r-1)t}$. Note that $e(G') \geq X - rY$. Indeed, $E(G')$ is obtained from $E(G[A' \cup B'])$ by deleting all edges incident to vertices $v \in B'$ which have a neighbour u with $|N_G(u) \cap B'| > 4r2^{\gamma-(r-1)t}$. Clearly, there are at most Y such vertices $v \in B'$, so at most rY edges are deleted. By equation (4.2) and the claim, we have

$$\begin{aligned}&\mathbb{E}[X - rY]\\&= \sum_{uv \in E(G[A \cup \tilde{B}])} \left(\mathbb{P}(u \in A', v \in B') - r\mathbb{P}(u \in A', v \in B' \text{ and } |N_G(u) \cap B'| \geq 4r2^{\gamma-(r-1)t})\right)\\&\geq \sum_{uv \in E(G[A \cup \tilde{B}])} (\mathbb{P}(v \in B') - \mathbb{P}(v \in B')/2) = e(G[A \cup \tilde{B}])2^{\gamma-rt-1}\\&= |\tilde{B}|r2^{\gamma-rt-1} \geq \frac{|B|}{(t-s)r} r2^{\gamma-rt-1} = \frac{e(G)}{(t-s)r} 2^{\gamma-rt-1}.\end{aligned}$$

Note that, for every $u \in A$, we have $2^{\alpha(u)} \leq 2^{s+1} + 2d_G(u)$. Indeed, this is trivial if $\alpha(u) = s + 1$, and else $d_G(u) \geq 2^{\alpha(u)-1}$. Also recall that $e(G) \geq 2^s|A|$. Therefore,

$$\begin{aligned}\mathbb{E}[|A'|] &= \sum_{u \in A} \mathbb{P}(u \in A') = \sum_{u \in A} 2^{\alpha(u)-t} \leq 2^{-t} \sum_{u \in A} (2^{s+1} + 2d_G(u))\\&= 2^{-t}(|A|2^{s+1} + 2e(G)) \leq 2^{-t}4e(G).\end{aligned}$$

Hence,

$$\mathbb{E}\left[X - rY - |A'|\frac{2^{\gamma-(r-1)t}}{10(t-s)r}\right] > 0.$$

It follows that there exists an outcome for which $X - rY - |A'|\frac{2^{\gamma-(r-1)t}}{10(t-s)r} > 0$. Then $e(G') \geq X - rY \geq |A'|\frac{2^{\gamma-(r-1)t}}{10(t-s)r}$. Hence, $d' = e(G')/|A''|$ satisfies $d' \geq e(G')/|A'| \geq \frac{2^{\gamma-(r-1)t}}{10(t-s)r} \geq \frac{2^{rs-(r-1)t}}{10(t-s)r}$ and $d_{G'}(u) \leq 4r2^{\gamma-(r-1)t} \leq 40(t-s)r^2d'$ for all $u \in A''$, completing the proof. □

The next lemma is obtained by iterative applications of Lemma 4.1.

Lemma 4.2. *Let r, s, t be positive integers such that r is sufficiently large and $s < t$. Let G be a bipartite graph with parts A and B such that $d_G(v) = r$ for every $v \in B$, $d_G(u) \leq 2^t$ for every $u \in A$, $d_G(u, u') \leq 2^{2rs-(2r-1)t}$ for any two distinct $u, u' \in A$ and $e(G) \geq 2^s|A|$.*

Then there exist positive integers $s^* \geq 2rs - (2r-1)t$, $t^* > s^*$ and sets $A^* \subset A$, $B^* \subset B$ such that $N_G(v) \subset A^*$ for every $v \in B^*$, $t^* - s^* \leq 5\log r$ and, writing $G^* = G[A^* \cup B^*]$ and $d^* = e(G^*)/|A^*|$, we have $d^* \geq 2^{s^*}$ and $d_{G^*}(u) \leq 2^{t^*}$ for all $u \in A^*$.

Proof. We prove the lemma by induction on $t-s$ (with r fixed). If $t-s \le 5\log r$, then we can take $s^*=s$, $t^*=t$, $A^*=A$ and $B^*=B$. Assume now that $t-s>5\log r$. Note that $2rs-(2r-1)t \le rs-(r-1)t$, so for any two distinct $u,u' \in A$, we have $d_G(u,u') \le 2^{rs-(r-1)t}$. By Lemma 4.1, there are subsets $A' \subset A$ and $B' \subset B$ such that $N_G(v) \subset A'$ for every $v \in B'$ and, writing $G'=G[A' \cup B']$ and $d'=e(G')/|A'|$, we have $d' \ge \frac{2^{rs-(r-1)t}}{10(t-s)r}$ and $d_{G'}(u) \le 40(t-s)r^2 d'$ for all $u \in A'$. By the codegree condition, $2rs-(2r-1)t \ge 0$, so $rs-(r-1)t \ge (t-s)r$ and hence $d' \ge 2$. Let $s'=\lfloor \log d' \rfloor$, and let $t'=\lceil \log(40(t-s)r^2 d') \rceil$. Then $s'<t'$, $d_{G'}(v)=r$ for every $v \in B'$, $d_{G'}(u) \le 2^{t'}$ for every $u \in A'$ and $e(G') \ge 2^{s'}|A'|$. Moreover, using $d' \ge \frac{2^{rs-(r-1)t}}{10(t-s)r}$, we have $t' \ge rs-(r-1)t$. Hence,

$$\begin{aligned} 2rs'-(2r-1)t' &= t'-2r(t'-s') \ge t'-2r(\log(40(t-s)r^2)+2) \\ &\ge rs-(r-1)t-2r\log(160(t-s)r^2) \\ &= 2rs-(2r-1)t+r(t-s)-2r\log(160(t-s)r^2) \\ &\ge 2rs-(2r-1)t, \end{aligned}$$

where the last inequality uses that $t-s \ge 5\log r$ and that r is sufficiently large. Hence, for any distinct $u,u' \in A'$, we have $d_{G'}(u,u') \le d_G(u,u') \le 2^{2rs-(2r-1)t} \le 2^{2rs'-(2r-1)t'}$. Finally, $t'-s' \le \log(40(t-s)r^2)+2 < t-s$.

Thus, by the induction hypothesis, there exist positive integers $s^* \ge 2rs'-(2r-1)t'$, $t^*>s^*$ and sets $A^* \subset A'$, $B^* \subset B'$ such that $N_{G'}(v) \subset A^*$ for every $v \in B^*$, $t^*-s^* \le 5\log r$ and, writing $G^*=G'[A^* \cup B^*]=G[A^* \cup B^*]$ and $d^*=e(G^*)/|A^*|$, we have $d^* \ge 2^{s^*}$ and $d_{G^*}(u) \le 2^{t^*}$ for all $u \in A^*$. Since $N_G(v)=N_{G'}(v)$ for every $v \in B'$, we have $N_G(v) \subset A^*$ for every $v \in B^*$. As $2rs'-(2r-1)t' \ge 2rs-(2r-1)t$, it follows that $s^* \ge 2rs-(2r-1)t$, so s^*, t^*, A^* and B^* are suitable for the conclusion of the lemma. □

5. Completing the proof

The next lemma is the upshot of what we have proved so far.

Lemma 5.1. *Let r,s,t be positive integers such that r is sufficiently large and $s<t$. Let G be a bipartite graph with parts A and B such that $d_G(v)=r$ for every $v \in B$, $d_G(u) \le 2^t$ for every $u \in A$, $d_G(u,u') \le 2^{2rs-(2r-1)t-r}$ for any two distinct $u,u' \in A$ and $e(G) \ge 2^s|A|$.*

Then G contains a 64-almost-regular subgraph with average degree at least $\frac{r}{80\log r}$.

Proof. By the assumption on the codegrees, we have $2rs-(2r-1)t-r \ge 0$, so $2rs-(2r-1)t \ge r$. Hence, by Lemma 4.2, G has a $(2^{5\log r}, r)$-almost-biregular subgraph G^* (the condition $2rs-(2r-1)t \ge r$ ensures that $s^* \ge r$ and so $d^* \ge r$). Then, by Lemma 3.5 with $L=2^{5\log r}$ and $\delta=r$, G^* has a 64-almost-regular subgraph with average degree at least $\frac{r}{80\log r}$. □

The next result is very similar to Lemma 5.1 – the main difference is that the codegree condition is no longer present. This can be achieved using Lemma 3.2.

Lemma 5.2. *Let k,r,s,t be positive integers such that r is sufficiently large, $s<t$ and $s \ge t(1-\frac{1}{6r})$. Let G be a $K_{k,k}$-free bipartite graph with parts A and B such that $d_G(v)=r$ for every $v \in B$, $d_G(u) \le 2^t$ for every $u \in A$ and $e(G) \ge 4(k+1)^2 2^s|A|$.*

Then G contains a 64-almost-regular subgraph with average degree at least $\frac{r}{160(k+1)\log r}$.

Proof. We may assume that $r \ge k\log r$, otherwise the conclusion of the lemma is trivial. By Lemma 3.2, G has a spanning subgraph G' such that $e(G') \ge e(G)/(k+1)$ and $d_{G'}(u,u') \le k2^{(1-1/k)t}$ for any two distinct $u,u' \in A$. Since $s \le t-1$ and $s \ge t(1-\frac{1}{6r})$, we have $t \ge 6r$. Hence, $2^{\frac{t}{2k}} \ge 2^{3r/k} \ge 2^{\log r}=r \ge k$. Thus, $k2^{(1-1/k)t} \le 2^{(1-\frac{1}{2k})t}$, so $d_{G'}(u,u') \le 2^{(1-\frac{1}{2k})t}$ for any two distinct $u,u' \in A$. Let $\tilde{B}$ be the subset of B consisting of those vertices v which satisfy $d_{G'}(v) \ge r/(2k+2)$. Now, the number of edges in G' between A and $B \setminus \tilde{B}$ is at most $|B|r/(2k+2)=e(G)/(2k+2) \le e(G')/2$, so there are at least

$e(G')/2$ edges between A and $\tilde{B}$. Let $r' = \lceil r/(2k+2)\rceil$ and let G'' be a spanning subgraph of $G'[A \cup \tilde{B}]$ obtained by keeping precisely r' edges from each $v \in \tilde{B}$. Clearly, $e(G'') \geq e(G'[A \cup \tilde{B}])/(2k+2) \geq e(G')/(4k+4) \geq e(G)/(4(k+1)^2) \geq 2^s|A|$. Note that

$$2r's - (2r'-1)t - r' = t - r'(2t-2s+1) \geq t - \frac{r}{k}(2t-2s+1) \geq t - \frac{3r}{k}(t-s)$$
$$\geq t - \frac{3r}{k} \cdot \frac{t}{6r} = (1 - \frac{1}{2k})t,$$

so $d_{G''}(u,u') \leq 2^{2r's-(2r'-1)t-r'}$ holds for any distinct $u, u' \in A$. Thus, we can apply Lemma 5.1 with G'' in place of G and r' in place of r to get a 64-almost-regular subgraph with average degree at least $\frac{r'}{80 \log r'} \geq \frac{r}{160(k+1)\log r}$. □

We are now in a position to prove our main result. In the proof and later in the paper, we omit floor and ceiling signs whenever they are not crucial.

Theorem 5.3. *Let k, r and Δ be positive integers such that r is sufficiently large. Let G be a $K_{k,k}$-free graph with maximum degree at most Δ and average degree at least $80r^2 \log\log\Delta$. Then G has a 64-almost-regular subgraph with average degree at least $\frac{r}{160(k+1)\log r}$.*

Proof. Since r is sufficiently large, the degree conditions imply that Δ is also sufficiently large. Moreover, we may assume that $r \geq k$, otherwise the conclusion of the lemma is trivial. Note that G has a bipartite subgraph G' with average degree at least $40r^2 \log\log\Delta$ and G' has a nonempty subgraph G'' with minimum degree at least $20r^2 \log\log\Delta$ (this subgraph can be obtained by repeatedly discarding vertices of degree less than $20r^2 \log\log\Delta$). Let A and B be the parts of G'' such that $|A| \leq |B|$. Let H be a spanning subgraph of G'' such that $d_H(v) = 20r^2 \log\log\Delta$ for every $v \in B$.

Let $t_0 = r(\log r)(\log\log\Delta)^{1/2}$ and let ℓ be the smallest nonnegative integer such that $t_0/(1-\frac{1}{10r})^{\ell} \geq \log\Delta$. Note that $(1-\frac{1}{10r})^{10r\log\log\Delta} \leq \exp(-\log\log\Delta) \leq 1/\log\Delta$, so $\ell \leq 10r\log\log\Delta$. For $1 \leq i \leq \ell$, let $t_i = t_0/(1-\frac{1}{10r})^i$. Clearly, $t_\ell \geq \log\Delta$.

Let $A_0 = \{u \in A : d_H(u) \leq 2^{t_0}\}$ and for $1 \leq i \leq \ell$, let $A_i = \{u \in A : 2^{t_{i-1}} < d_H(u) \leq 2^{t_i}\}$. Clearly, these sets partition A. Hence, for every $v \in B$, either v has at least $d_H(v)/2 = 10r^2 \log\log\Delta$ neighbours (in the graph H) in A_0 or $\ell > 0$ and there is some $1 \leq i \leq \ell$ such that v has at least $d_H(v)/(2\ell) \geq r$ neighbours in A_i.

Therefore, at least one of the following two cases must occur.

Case 1. There are at least $|B|/2$ vertices $v \in B$ which have at least $10r^2 \log\log\Delta$ neighbours in A_0.

Case 2. There exist some $1 \leq i \leq \ell$ and at least $|B|/(2\ell)$ vertices $v \in B$ which have at least r neighbours in A_i.

In Case 1, let $B' \subset B$ be a set of size at least $|B|/2$ such that every $v \in B'$ has at least $10r^2 \log\log\Delta$ neighbours in A_0. For technical reasons, let us take a random subset $A_0' \subset A_0$ of size $|A_0|/3$. With positive probability, there is a set $B'' \subset B'$ of at least $2|B'|/3$ vertices which all have at least $r^2 \log\log\Delta$ neighbours in A_0'. Let H' be a spanning subgraph of $H[A_0' \cup B'']$ obtained by keeping precisely $r^2 \log\log\Delta$ edges from each vertex in B''. Now, note that $|A_0'| = |A_0|/3 \leq |A|/3 \leq |B|/3 \leq 2|B'|/3 \leq |B''|$. Moreover, $d_{H'}(u) \leq d_H(u) \leq 2^{t_0}$ for every $u \in A_0'$. This implies that H' is (L,d)-almost biregular for $L = 2^{t_0} = 2^{r(\log r)(\log\log\Delta)^{1/2}}$ and $d = r^2 \log\log\Delta$. Hence, by Lemma 3.5, H' has a 64-almost-regular subgraph with average degree at least $\frac{r^2\log\log\Delta}{16r(\log r)(\log\log\Delta)^{1/2}} \geq \frac{r}{160(k+1)\log r}$.

In Case 2, let us choose some $1 \leq i \leq \ell$ and a set $B' \subset B$ of size at least $|B|/(2\ell)$ such that for every $v \in B'$, v has at least r neighbours in A_i. Let H' be a subgraph of $H[A_i \cup B']$ obtained by keeping precisely r edges incident to each $v \in B'$. We will apply Lemma 5.2 for this graph. Let $t = t_i$ and let $s = t(1-\frac{1}{6r})$. Note that for any $u \in A_i$, $d_{H'}(u) \leq d_H(u) \leq 2^t$. Let $C = 4(r+1)^2$. Using that $|B| \cdot 20r^2 \log\log\Delta = e(H) \geq |A_i| 2^{t_{i-1}}$, we get

https://doi.org/10.1017/fmp.2023.19 Published online by Cambridge University Press

$$e(H') = |B'|r \geq |B|r/(2\ell) = \frac{e(H)}{40\ell r \log\log\Delta} \geq \frac{e(H)}{400r^2(\log\log\Delta)^2} \geq \frac{C|A_i|2^{t_{i-1}}}{400Cr^2(\log\log\Delta)^2}$$
$$= C|A_i|2^{t_{i-1}-\log(400Cr^2(\log\log\Delta)^2)}.$$

Note that

$$t_{i-1} - \log(400Cr^2(\log\log\Delta)^2) = t\left(1 - \frac{1}{10r}\right) - \log(400Cr^2(\log\log\Delta)^2) \geq t\left(1 - \frac{1}{6r}\right) = s,$$

where the inequality follows from $t/r \geq t_0/r = (\log r)(\log\log\Delta)^{1/2}$ and since Δ is sufficiently large. Hence, $e(H') \geq C|A_i|2^s \geq 4(k+1)^2|A_i|2^s$.

Thus, we can apply Lemma 5.2 to the graph H' and get a 64-almost-regular subgraph with average degree at least $\frac{r}{160(k+1)\log r}$. □

Theorem 1.2 now follows easily.

Proof of Theorem 1.2. Let r be sufficiently large in terms of k, and let $C = 80r^2$. Let G be a graph with maximum degree Δ and average degree at least $C\log\log\Delta$. If G contains $K_{k,k}$ as a subgraph, then it has a k-regular subgraph. Else, by Theorem 5.3, G has a 64-almost-regular subgraph G' with average degree at least $\frac{r}{160(k+1)\log r}$. Since r is sufficiently large in terms of k, Theorem 3.8 implies that G' has a k-regular subgraph. □

6. Concluding remarks

Motivated by the study of the Turán number of the cube, in 1970, Erdős and Simonovits [15] proved the following result.

Theorem 6.1 (Erdős–Simonovits [15]). *For every $\alpha > 0$ there exist some $K = K(\alpha)$ and $n_0 = n_0(\alpha)$ such that any graph with $n \geq n_0$ vertices and at least $n^{1+\alpha}$ edges contains a K-almost-regular subgraph with m vertices and at least $\frac{2}{5}m^{1+\alpha}$ edges for some $m \geq n^{\alpha(1-\alpha)/(1+\alpha)}$.*

This result has since become one of the most widely used tools for Turán type problems. Its extreme usefulness comes from the fact that it allows us to replace a general host graph by an almost-regular one at negligible cost. Usually extremal problems are much easier to deal with when the host graph is almost-regular.

In their paper, Erdős and Simonovits asked whether a similar ‘regularization’ is possible for sparser graphs. More precisely, they asked whether there exist absolute constants $\varepsilon, K > 0$ such that any n-vertex graph with at least $n\log n$ edges contains a K-almost-regular subgraph with m vertices and at least $\varepsilon m\log m$ edges, where $m \to \infty$ as $n \to \infty$. Using a variant of the construction of Pyber, Rödl and Szemerédi from Theorem 1.1, Alon [1] gave a negative answer to this question as follows.

Theorem 6.2 (Alon [1]). *For every $K > 0$ and $n > 10^6$, there is an n-vertex graph with at least $n\log n$ edges in which any K-almost-regular m-vertex subgraph has at most $72m\sqrt{\log m} + 18\log(64K) + 324$ edges.*

Thus, we cannot necessarily pass to an almost-regular m-vertex subgraph with much more than $m\sqrt{\log m}$ edges. This naturally leads to the question what density we can actually guarantee in an almost-regular subgraph of a graph with $n\log n$ edges. Our Theorem 5.3 essentially answers this question. Indeed, it implies that we can always find an almost-regular m-vertex subgraph with nearly $m\sqrt{\log m}$ edges, showing that Theorem 6.2 is tight up to $\log\log$ factors.

Theorem 6.3. *For any positive integer m_0, there exist some $n_0 = n_0(m_0)$ and $\varepsilon = \varepsilon(m_0) > 0$ such that any graph with $n \geq n_0$ vertices and at least $n\log n$ edges has a 64-almost-regular subgraph with $m \geq m_0$ vertices and at least $\varepsilon m\sqrt{\log m}/(\log\log m)^{3/2}$ edges.*

Proof. Let n_0 be sufficiently large in terms of m_0, and let G be a graph with $n \geq n_0$ vertices and at least $n \log n$ edges. Let $r = \frac{\sqrt{\log n}}{10\sqrt{\log\log n}}$, and let Δ be the maximum degree of G. Since $\Delta \leq n$, the average degree of G is at least $\log n \geq 100r^2 \log\log \Delta$. If K_{m_0,m_0} is a subgraph of G, then we are done. Else, by Theorem 5.3, G has a 64-almost-regular subgraph G' with average degree at least $\frac{r}{160(m_0+1)\log r}$. This is at least $2\varepsilon\sqrt{\log n}/(\log\log n)^{3/2}$ for some $\varepsilon > 0$ that depends only on m_0. Now, if m is the number of vertices in G', then $m \geq m_0$ (since n is sufficiently large) and G' has at least $\varepsilon m\sqrt{\log n}/(\log\log n)^{3/2} \geq \varepsilon m\sqrt{\log m}/(\log\log m)^{3/2}$ edges, as desired. □

In certain extremal problems one wants to pass to almost-regular subgraphs whose average degree is large not just compared to the number of vertices in the subgraph but also compared to the number of vertices in the original graph (which we call n). Using a variant of Pyber's argument, it was shown in [6] and [10] that passing to such a subgraph is possible with the loss of only a $\log n$ factor in the average degree. Also, by considering a variant of the construction of Pyber, Rödl and Szemerédi, it was shown in [10] that this $\log n$ loss is necessary for graphs with average degree at least n^{ε}. However, when the graph has at most $n \log n$ edges, then this $\log n$ loss means that the above result becomes trivial. Our methods apply in this sparse regime and give the following.

Theorem 6.4. *There is an absolute constant $\varepsilon > 0$ such that any n-vertex graph with average degree $d \geq 2\log\log n$ has a 64-almost-regular subgraph with average degree at least $\varepsilon\frac{(d/\log\log n)^{1/4}}{(\log(d/\log\log n))^{1/2}}$. Furthermore, for any positive integer t there is some $\varepsilon_t > 0$ such that any $K_{t,t}$-free n-vertex graph with average degree $d \geq 2\log\log n$ has a 64-almost-regular subgraph with average degree at least $\varepsilon_t\frac{(d/\log\log n)^{1/2}}{\log(d/\log\log n)}$.*

Proof. We only prove the first assertion – the second one has a very similar proof.

Let G be an n-vertex graph with average degree d. We may assume that $d/\log\log n$ is sufficiently large, otherwise the statement is trivial. Let ε be a sufficiently small positive absolute constant, and let $k = \lceil \varepsilon\frac{(d/\log\log n)^{1/4}}{(\log(d/\log\log n))^{1/2}} \rceil$. We want to show that G has a 64-almost-regular subgraph with average degree at least k. Let $r = (\frac{d}{80\log\log n})^{1/2}$. If G contains $K_{k,k}$ as a subgraph, then we are done; else, by Theorem 5.3, G has a 64-almost-regular subgraph with average degree at least $\frac{r}{160(k+1)\log r}$. But for sufficiently small ε, we have $\frac{r}{160(k+1)\log r} \geq k$, so the proof is complete. □

Since the problem of finding regular subgraphs arises naturally in many combinatorial settings, it seems very likely that our results will have further applications. We conclude the paper by discussing two such applications.

- Our results can be used to answer a recent question posed by Alon et al. in [6]. Motivated by applications to neural networks, in [6] the authors defined the notion of an *α-multitasker* graph and asked (see the discussion after Theorem 1.3 in the full version of their paper [3]) if there exists an n-vertex α-multitasker with average degree $\Theta(\log n)$ for some $\alpha > 0$, independent of n. Our Theorem 6.4 (or Theorem 1.2) can be used to show that there is no such multitasker with average degree $\omega(\log\log n)$. This is tight as Alon et al. showed that there are α-multitaskers with average degree $\Theta(\log\log n)$ and constant α.
- The second application concerns the notion of spectral degeneracy, introduced by Dvořák and Mohar [13]. The spectral radius of a graph G is the largest eigenvalue of the adjacency matrix of G and is denoted by $\rho(G)$. We say that G is *spectrally d-degenerate* if for every subgraph H of G, we have $\rho(H) \leq \sqrt{d\Delta(H)}$ and call the smallest such d the *spectral degeneracy of* G. It was observed in [13] that every d-degenerate graph is spectrally $4d$-degenerate. Moreover, they showed the rough converse that any spectrally d-degenerate graph with maximum degree $D \geq 2d$ is $4d\log(D/d)$-degenerate. In addition, they asked whether there exists a function f such that any spectrally d-degenerate graph is $f(d)$-degenerate. This was answered negatively by Alon [2], who constructed, for every M, a spectrally 50-degenerate graph which is not M-degenerate. Analyzing his construction more carefully, one can

https://doi.org/10.1017/fmp.2023.19 Published online by Cambridge University Press

see that in fact there is a positive constant c such that for every n there is a spectrally 50-degenerate n-vertex graph with degeneracy at least $c \log \log n$. Our results can be used to show that this is best possible, that is, that for every d there is a constant $C = C(d)$ such that any spectrally d-degenerate n-vertex graph has degeneracy at most $C \log \log n$. This follows from our Theorem 1.2 and the simple fact that the spectral degeneracy of a d-regular graph is precisely d.

Acknowledgments. We are grateful to the referees for their careful reading of the manuscript and their helpful comments.

Competing interests. The authors have no competing interest to declare.

Financial support. The research of the first author was supported by an ETH Zürich Postdoctoral Fellowship (grant number 20-1 FEL-35). The research of the second author was supported in part by an SNSF grant (grant number 200021_196965).

References

[1] N. Alon, 'Problems and results in extremal combinatorics—II', *Discrete Mathematics* **308**(19) (2008), 4460–4472, 2008.
[2] N. Alon, 'A note on degenerate and spectrally degenerate graphs', *Journal of Graph Theory* **72**(1) (2013), 1–6, 2013.
[3] N. Alon, J. D. Cohen, B. Dey, T. Griffiths, S. Musslick, K. Ozcimder, D. Reichman, I. Shinkar and T. Wagner, 'A graph-theoretic approach to multitasking', Preprint, 2016, arXiv:1611.02400.
[4] N. Alon, S. Friedland and G. Kalai, 'Every 4-regular graph plus an edge contains a 3-regular subgraph', *Journal of Combinatorial Theory, Series B* **37** (1984), 92–93.
[5] N. Alon, S. Friedland and G. Kalai, 'Regular subgraphs of almost regular graphs', *Journal of Combinatorial Theory, Series B* **37**(1) (1984), 79–91.
[6] N. Alon, D. Reichman, I. Shinkar, T. Wagner, S. Musslick, J. D. Cohen, T. Griffiths, B. Dey and K. Ozcimder, 'A graph-theoretic approach to multitasking', *Advances in Neural Information Processing Systems* **30** (2017).
[7] B. Bollobás, *Extremal Graph Theory* (Academic Press, 1978).
[8] B. Bollobás, J. H. Kim and J. Verstraëte, 'Regular subgraphs of random graphs', *Random Structures & Algorithms* **29**(1) (2006), 1–13.
[9] J. A. Bondy and U. S. R. Murty, *Graph Theory with Applications*, vol. 290 (Macmillan, London, 1976).
[10] M. Bucić, M. Kwan, A. Pokrovskiy, B. Sudakov, T. Tran and A. Z. Wagner, 'Nearly-linear monotone paths in edge-ordered graphs', *Israel Journal of Mathematics* **238**(2) (2020), 663–685.
[11] D. Dellamonica, P. Haxell, T. Łuczak, D. Mubayi, B. Nagle, Y. Person, V. Rödl, M. Schacht and J. Verstraëte, 'On even-degree subgraphs of linear hypergraphs', *Combinatorics, Probability and Computing* **21**(1–2) (2012), 113–127.
[12] D. Dellamonica Jr and V. Rödl, 'A note on Thomassen's conjecture', *Journal of Combinatorial Theory, Series B* **101**(6) (2011), 509–515.
[13] Z. Dvořák and B. Mohar, 'Spectrally degenerate graphs: Hereditary case', *Journal of Combinatorial Theory, Series B* **102**(5) (2012), 1099–1109.
[14] P. Erdős, 'Some recent progress on extremal problems in graph theory', *Congr. Numer.* **14** (1975), 3–14.
[15] P. Erdős and M. Simonovits, 'Some extremal problems in graph theory', in *Combinatorial Theory and Its Applications, I* (North Holland, Amsterdam, 1970), 377–390.
[16] P. Erdős, 'On the combinatorial problems which I would most like to see solved', *Combinatorica* **1**(1) (1981), 25–42.
[17] P. Hall, 'On representatives of subsets', *J. London Math. Soc.* **10** (1935), 26–30.
[18] J. Han and J. Kim, 'Two-regular subgraphs of odd-uniform hypergraphs', *Journal of Combinatorial Theory, Series B* **128** (2018), 175–191.
[19] J. Kim, 'Regular subgraphs of uniform hypergraphs', *Journal of Combinatorial Theory, Series B* **119** (2016), 214–236.
[20] J. Kim and A. V. Kostochka, 'Maximum hypergraphs without regular subgraphs', *Discuss. Math. Graph Theory* **34**(1) (2014), 151–166.
[21] D. Kühn and D. Osthus, 'Every graph of sufficiently large average degree contains a C_4-free subgraph of large average degree', *Combinatorica* **24**(1) (2004), 155–162.
[22] R. Montgomery, A. Pokrovskiy and B. Sudakov, 'C_4-free subgraphs with large average degree', *Israel Journal of Mathematics* **246**(1) (2021), 55–71.
[23] D. Mubayi and J. Verstraëte, 'Two-regular subgraphs of hypergraphs', *Journal of Combinatorial Theory, Series B* **99**(3) (2009), 643–655.
[24] J. Petersen, 'Die theorie der regulären graphs', *Acta Mathematica* **15**(1) (1891), 193–220.
[25] L. Pyber, 'Regular subgraphs of dense graphs', *Combinatorica* **5**(4) (1985), 347–349.
[26] L. Pyber, V. Rödl and E. Szemerédi, 'Dense graphs without 3-regular subgraphs. *Journal of Combinatorial Theory, Series B* **63**(1) (1995), 41–54.
[27] V. Rödl and B. Wysocka, 'Note on regular subgraphs', *Journal of Graph Theory* **24**(2) (1997), 139–154.
[28] C. Thomassen, 'Girth in graphs', *Journal of Combinatorial Theory, Series B* **35**(2) (1983), 129–141.
[29] W. T. Tutte, 'The factorization of linear graphs', *Journal of the London Mathematical Society* **1**(2) (1947), 107–111.
[30] W. T. Tutte, 'The factors of graphs', *Canadian Journal of Mathematics* **4** (1952), 314–328.