INDUCED C_4-FREE SUBGRAPHS WITH LARGE AVERAGE DEGREE

XIYING DU, ANTÓNIO GIRÃO, ZACH HUNTER, ROSE MCCARTY, AND ALEX SCOTT

ABSTRACT. We prove that there exists a constant C so that, for all $s, k \in \mathbb{N}$, if G has average degree at least k^{Cs^3} and does not contain $K_{s,s}$ as a subgraph then it contains an induced subgraph which is C_4-free and has average degree at least k. It was known that some function of s and k suffices, but this is the first explicit bound. We give several applications of this result, including short and streamlined proofs of the following two corollaries.

We show that there exists a constant C so that, for all $s, k \in \mathbb{N}$, if G has average degree at least k^{Cs^3} and does not contain $K_{s,s}$ as a subgraph then it contains an induced subdivision of K_k. This is the first quantitative improvement on a well-known theorem of Kühn and Osthus; their proof gives a bound that is triply exponential in both k and s.

We also show that for any hereditary degree-bounded class $\mathcal{F}$, there exists a constant $C = C_{\mathcal{F}}$ so that C^{s^3} is a degree-bounding function for $\mathcal{F}$. This is the first bound of any type on the rate of growth of such functions. It is open whether there is always a polynomial degree-bounding function.

1. INTRODUCTION

A longstanding conjecture from 1983 due to Thomassen [29] states that for all $g, k \geqslant 2$, there exists $f(g, k)$ such that every graph G with average degree at least $f(g, k)$ contains a subgraph with girth at least g and average degree at least k. It is a standard exercise to show that every graph has a bipartite subgraph with at least half of its edges. So the first nontrivial open case is when $g = 5$ and we wish to find a C_4-free subgraph in a bipartite graph G. (A graph is *H-free*, for some graph H, if it has no subgraph isomorphic to H.) This case of $g = 5$ was resolved in a remarkable paper by Kühn and Osthus [18] in 2004. However, since then no further progress has been made, and the conjecture remains wide open for all $g \geqslant 7$.

A straightforward probabilistic argument shows that Thomassen's conjecture holds for every "almost-regular" graph G. Therefore a natural strategy is to pass to an almost-regular subgraph of G which preserves some of its average degree. Unfortunately, this strategy is bound to fail: Pyber, Rödl, and Szemerédi [24] proved that there exist n-vertex graphs with average degree $\Omega(\log \log n)$ which do not contain a k-regular subgraph for any $k \geqslant 3$. We remark that, very recently, Janzer and Sudakov [15] proved via an ingenious argument that $\Omega(\log \log n)$ is indeed the correct barrier. (Formally, they proved that for each k there exists a constant c_k so that every n-vertex graph with average degree at least $c_k \log \log n$ has a k-regular subgraph.) This fully resolved the Erdős-Sauer problem from [8].

Since Kühn and Osthus [18] first resolved the case of $g = 6$, two additional proofs of their theorem have been discovered. The first proof, which is due to Dellamonica, Koubek, Martin, and Rödl [6], uses a surprising result of Füredi [11] on hypergraphs. The second proof, which is due to Montgomery, Pokrovskiy, and Sudakov [23], is more recent and gives the best bounds currently known. In particular, they proved that there exists a constant C so that for all $k \in \mathbb{N}$, every graph with average degree at least k^{Ck^2} contains a

Date: November 14, 2023.

AS and AG were supported by EPSRC grant EP/V007327/1. RM was supported by NSF grant DMS-2202961.

subgraph which is C_4-free and has average degree at least k. The same paper also gives a lower bound of $k^{3-o(1)}$, and it is a very interesting problem to determine whether a polynomial in k suffices.

In 2021, the fourth author of this paper proved that it is possible to ask for a much stronger property [22]: for all $s, k \in \mathbb{N}$, there exists an integer $f(s,k)$ so that every $K_{s,s}$-free bipartite graph of average degree at least $f(s,k)$ contains an *induced* C_4-free subgraph of average degree at least k. The paper does not determine an explicit bound; however, with some unwrapping, we believe that the proof yields a function which is triply exponential in s, for fixed k. Moreover, the assumption about bipartiteness can be removed, at the cost of another exponent, using a theorem of Kwan, Letzter, Sudakov, and Tran [20]. Their theorem says that for all $s, k \in \mathbb{N}$, there exists an integer $g(s,k)$ so that every K_s-free graph with average degree at least $g(s,k)$ contains an *induced* bipartite subgraph with average degree at least k. (More precisely, they proved that there exists a constant C so that the function $g(s,k) = \exp\big(C^{s \log s} \cdot k \log k\big)$ suffices.)

In our main theorem, we prove that this stronger induced theorem for $K_{s,s}$-free graphs holds with bounds that essentially match the singly exponential bounds of [23] from the non-induced setting.

Theorem 1.1. *There exists a constant C so that for all $s, k \in \mathbb{N}$, every $K_{s,s}$-free graph with average degree at least k^{Cs^3} contains an induced subgraph which is C_4-free and has average degree at least k.*

First, we note that the assumption of $K_{s,s}$-freeness is essential, as any C_4-free induced subgraph of $K_{s,s}$ has average degree less than 2 (thus by taking $G = K_{s,s}$ for larger and larger s, one could make the average degree of G arbitrarily large without obtaining induced C_4-free subgraphs with average degree at least 2). Moreover, we remark that if a graph G contains a $K_{4k^2,4k^2}$-subgraph, then it contains a bipartite C_4-free subgraph with average degree at least k; see [23, Lemma 2.2]. (In fact, we can just take this subgraph to be the incidence graph of the projective plane of order q, where q is the smallest prime which is at least $k-1$.) Therefore we obtain the following as a simple corollary of Theorem 1.1.

Corollary 1.2. *There exists a constant C so that for all $k \in \mathbb{N}$, every graph with average degree at least k^{Ck^6} contains a bipartite C_4-free subgraph with average degree at least k.*

Corollary 1.2 nearly recovers the result of [23], only with the exponent 'k^2' replaced by 'k^6'.

We also prove some corresponding lower bounds for Theorem 1.1. So, let $f_{\text{Ind}}(s,k)$ denote the least integer D such that if G is a $K_{s,s}$-free graph with average degree at least D, then G contains an *induced* C_4-free subgraph with average degree at least k. Theorem 1.1 establishes that $f_{\text{Ind}}(s,k) \leqslant k^{Cs^3}$ for some absolute constant C. By considering random graphs, one can obtain some lower bounds for this quantity as well. Specifically, in Section 6 we prove the following.

Theorem 1.3. *The following bounds hold.*

(a) There exists a constant $c > 0$ so that for all $k \geqslant 2$, we have that $f_{\text{Ind}}(k,k) \geqslant k^{ck}$.
(b) For each fixed $k \geqslant 2$, we have that $f_{\text{Ind}}(s,k) \geqslant s^{(1/4-o(1))(k^2-3k-2)}$.
(c) For each fixed $s \geqslant 2$, we have that $f_{\text{Ind}}(s,k) \geqslant k^{(1/2-o(1))s-1}$.

1.1. **Further applications.** Now we discuss two additional corollaries of Theorem 1.1 which are motivated by χ-boundedness. A family of graphs $\mathcal{F}$ is *hereditary* if for every $G \in \mathcal{F}$, every induced subgraph of G is also in $\mathcal{F}$. A hereditary family of graphs $\mathcal{F}$ is χ*-bounded* if there exists a function $f : \mathbb{N} \to \mathbb{N}$ such that for every $G \in \mathcal{F}$, we have $\chi(G) \leqslant f(\omega(G))$, where we write $\chi(G)$ for the chromatic number of G and $\omega(G)$ for the clique number of G. These classes have been widely studied, following an influential paper of Gyárfás [13] raised a number of well-known conjectures (many of these have now been solved: see, for example, [5, 26]; and [27] for a survey).

A similar notion to χ-boundedness, but for average degree instead of chromatic number, has recently been receiving attention. Intuitively, a class is χ-bounded if large chromatic number forces the existence of large cliques; and it is "degree-bounded" if large average degree forces large balanced bicliques. Formally, we say that a hereditary family $\mathcal{F}$ is *degree-bounded* if there exists a function $g : \mathbb{N} \to \mathbb{N}$ such that for every $G \in \mathcal{F}$, we have $d(G) \leqslant g(\tau(G))$, where we write $d(G)$ for the average degree of G and $\tau(G)$ for the *biclique number* of G, which is the largest integer s such that G contains a (not necessarily induced) copy of $K_{s,s}$. Any such function g is called a *degree-bounding function* for the class.

Strengthening an unpublished theorem of Hajnal and Rödl (see [14]), Kierstead and Penrice [16] showed that for every tree T, the class of graphs without an induced copy of T is degree-bounded. Recently, Scott, Seymour and Spirkl [28] established a quantitative strengthening of the result from [16], by showing that the class of T-induced-free graphs is polynomially degree bounded (that is, the function g can be a polynomial in s, for fixed T). This improved upon another recent theorem of Bonamy, Bousquet, Pilipczuk, Rzążewski, Thomassé and Walczak [2], who proved the same result when T is a path.

A rephrasing of Theorem 1.1 immediately gives the following corollary, which says that every hereditary degree-bounded class is essentially exponentially bounded.

Corollary 1.4. *For any hereditary degree-bounded class of graphs $\mathcal{F}$, there exists a constant $C = C_{\mathcal{F}}$ so that C^{s^3} is a degree-bounding function for $\mathcal{F}$.*

This corollary is surprising since the analogous statement for χ-boundedness is false: Briański, Davies, and Walczak [3] recently constructed hereditary χ-bounded classes where the optimal χ-bounding function grows arbitrarily quickly. This disproved the popular conjecture of Esperet [10] that every hereditary χ-bounded class is polynomially χ-bounded. It could still be true, however, that every hereditary degree-bounded class is polynomially degree-bounded. We note that, informally, Theorem 1.1 also tells us that the value of $g(1)$ can be used to "efficiently control" the rate of growth of g. Again the analogous statement for χ-boundedness is very false: Carbonero, Hompe, Moore, and Spirkl [4] recently proved that there exist graphs of arbitrarily large chromatic number where every triangle-free induced subgraph has chromatic number at most 4 (for more general results see [12]).

Our final corollary obtains improved bounds for classes with a forbidden induced subdivision. A *subdivision* of a graph H is any graph which can be obtained from H by replacing the edges uv of H by internally-disjoint paths between u and v. A beautiful result of Kühn and Osthus [18] states that for every graph H, the class of graphs which do not contain an induced subdivision of H is degree-bounded.

Theorem 1.5 (Kühn and Osthus)**.** *For every graph H and integer s, there is an integer $p(s, H)$ such that every $K_{s,s}$-free graph G with average degree at least $p(s, H)$ contains an induced subdivision of H.*

Their bounds for $p(s, H)$ are roughly triply exponential in s, for fixed H. A conjecture raised by Bonamy et al. [2, Conjecture 33] claims that actually $p(s, H)$ could be taken to be a polynomial in s. In Section 7, we give a short and streamlined proof that a single exponential is enough.

Corollary 1.6. *There exists a constant C so that for all $k, s \in \mathbb{N}$, every $K_{s,s}$-free graph G with average degree at least k^{Cs^3} contains an induced subdivision of K_k.*

Our proof will use Theorem 5.3 (which is a technical strengthening of our main theorem, Theorem 1.1) and a theorem about (non-induced) subdivisions which was proved independently by Bollobás and Thomason [1] and Kómlos and Szemerédi [17].

The rest of this paper is organized as follows: after giving some notation in the next section, we prove some preliminary lemmas in sections 3 and 4. Theorem 1.1 is proved in Section 5, Theorem 1.3 in Section 6, and Corollary 1.6 in Section 7. We conclude with some discussion in section 8.

2. Notation

We mostly use standard notation, and we consider all graphs to be finite, simple, and loopless. Let G be a graph. We denote the vertex set of G by $V(G)$ and the edge set of G by $E(G)$. We write $|G| \coloneqq |V(G)|$ and $e(G) \coloneqq |E(G)|$. We write $d(G)$ for the average degree of G (so that $d(G) = 2e(G)/|G|$), $\delta(G)$ for the minimum degree of G, and $\Delta(G)$ for the maximum degree of G. Given a vertex v of G, we write $d_G(v)$ for the degree of v and $N_G(v)$ for the neighbourhood of v.

As usual, we say that a graph H is an *induced subgraph* of G if there is an injective map $f : V(H) \to V(G)$ so that $f(xy) \in E(G)$ if and only if $xy \in E(H)$. We say that G is d*-degenerate* if every non-empty subgraph G' of G contains a vertex of degree at most d. We say that G is H*-free* if G does not contain H as a subgraph. We say that G is H*-induced-free* if G does not contain H as an induced subgraph.

Finally, given two disjoint sets of vertices $A, B \subset V(G)$, we write $G[A, B]$ for the induced bipartite graph with parts A and B. That is, the vertex set of $G[A, B]$ is $A \cup B$, and the edge set of $G[A, B]$ is $\{xy \in E(G) : x \in A, y \in B\}$. We view bipartite graphs as being equipped with a fixed bipartition; typically we denote a bipartite graph G with sides A and B and edge set E by $G = (A, B, E)$.

3. Preliminary lemmas

Recall from the introduction that a straightforward probabilistic argument shows that Thomassen's Conjecture holds for graphs which are almost regular; this is done by including edges independently at random and deleting one edge from each short cycle. It turns out that we can similarly use the Kővári-Sós-Turán Theorem [21] to prove Theorem 1.1 for graphs which are almost regular, by including vertices independently at random and deleting one vertex from each short cycle. We will take care of this step in Lemma 3.3.

For now, we recall the classical result of Kővári, Sós, and Turán. (We note that this can be formulated either for graphs or for bipartite graphs. There is a standard reduction to the bipartite case: for a $K_{s,s}$-free graph G, we create a new graph with two vertices u_1 and u_2 for each vertex u of G, where u_i and v_j are adjacent in the new graph if $i \neq j$ and uv is an edge of G.)

Theorem 3.1 ([21])**.** *For all $s \geqslant 2$ and $n \geqslant 1$, every $K_{s,s}$-free graph on n vertices has $O((s-1)^{1/s}n^{2-1/s}+sn))$ edges. Thus, there exists an absolute constant C such that for all $s \geqslant 2$ and $n \geqslant s^2$, every $K_{s,s}$-free graph on n vertices has at most $Cn^{2-1/s}$ edges.*

Since we will be able to take care of the nearly regular case easily, it is helpful to have configurations which guarantee a nearly-regular induced subgraph of large average degree. Our first lemma says that it is enough to find a bipartite graph of large average degree where each side *individually* is nearly regular (but the average degrees of the two sides can be quite different from each other). This lemma is due to Janzer and Sudakov [15, Lemma 3.7], but we restate the proof since we will need to state a slightly stronger conclusion. We say that a bipartite graph $\Gamma = (A, B, E)$ is L*-almost-biregular* if $d_\Gamma(a) \leqslant L|E|/|A|$ for all $a \in A$, and

$d_\Gamma(b) \leqslant L|E|/|B|$ for all $b \in B$. Thus, informally, each vertex in A has degree at most a factor of L larger than the average degree of all vertices in A, and likewise for B.

Lemma 3.2. *Each L-almost-biregular graph Γ has an induced subgraph Γ' with $d(\Gamma') \geqslant d(\Gamma)/4$ and $\Delta(\Gamma') \leqslant 24Ld(\Gamma')$.*

Proof. Suppose $\Gamma = (A, B, E)$. If $E = \varnothing$, there is nothing to prove (we may take $\Gamma' = \Gamma$). Otherwise, as both A and B have vertices of nonzero degree, we have $L\frac{|E|}{|A|}, L\frac{|E|}{|B|} \geqslant 1$.

We may assume $|A| \leqslant |B|$. Let $p = |A|/|B|$. We take a random subset $B' \subset B$ by adding each $b \in B$ independently with probability p. Let $A' \subset A$ be the set of $a \in A$ such that $|N_\Gamma(v) \cap B'| \leqslant 1 + 2p(d_\Gamma(v) - 1)$. We shall take $\Gamma' = \Gamma[A', B']$, and show this works with positive probability.

By construction, we have $\Delta(\Gamma') \leqslant 1 + 2L\frac{|E|}{|B|} \leqslant 3L\frac{|E|}{|B|}$. Indeed, each vertex $a \in A'$ has degree at most $1 + 2pL|E|/|A| \leqslant 3L|E|/|B|$, and each vertex $b \in B'$ has degree at most $d_\Gamma(b) \leqslant L|E|/|B|$.

Consider any $e = ab \in E$, and let $U = (N_\Gamma(a) \setminus \{b\}) \cap B'$. Applying Markov's inequality, we see that

$$\mathbb{P}(a \in A' | b \in B') = \mathbb{P}(|U| \leqslant 2\mathbb{E}[|U|]) > 1/2.$$

Then $\mathbb{E}[|A'| + |B'|] \leqslant |A| + \mathbb{E}[|B'|] = 2|A|$. and

$$\mathbb{E}[e(\Gamma')] = \sum_{e=ab\in E} \mathbb{P}(b \in B')\mathbb{P}(a \in A' | b \in B') > p|E|/2.$$

Thus we have

$$\mathbb{E}[4e(\Gamma') - \frac{|E|}{|B|}(|A'| + |B'|)] > 4(p|E|/2) - \frac{|E|}{|B|}2|A| = 0.$$

Consequently, we can choose A', B' such that the expectation is positive. Then Γ' is non-empty and, dividing by $|\Gamma'| = |A'| + |B'|$, we get

$$d(\Gamma') \geqslant 2\frac{|E|}{4|B|} \geqslant d(\Gamma)/4.$$

Since $\Delta(\Gamma') \leqslant 3L|E|/|B|$,

$$12Le(\Gamma') \geqslant \Delta(\Gamma')(|A'| + |B'|)$$

which implies that $\Delta(\Gamma') \leqslant 24Ld(\Gamma')$ as desired. □

We can now use Theorem 3.1 to deal with the case of nearly regular graphs, as discussed.

Lemma 3.3. *There exists an absolute constant $\kappa > 0$ so that the following holds for all $s \geqslant 2$, $\delta \in (0, 1/10)$ and $d \geqslant \kappa^{s/\delta}$.*

Let G be a $K_{s,s}$-free graph with $\Delta(G) \leqslant d$ and average degree $d(G) \geqslant d^{1-\delta/s}$. Then, G contains an induced $\{C_3, C_4\}$-free subgraph H with average degree at least $d^{(1-10\delta)/5s}$ and $\Delta(H) \leqslant d(H)^{1/(1-10\delta)}$.

Remark. We will need the upper bound on $\Delta(H)$ to derive a more technical version of Theorem 1.1 (namely Theorem 5.3), which will be useful in Section 7 and future applications.

Proof. Let $\mathcal{C}_4$ be the set of 4-cycles C in G. Let $\mathcal{S}_4$ be the set of pairs $(e, C) \in E(G) \times \mathcal{C}_4$ such that e contains exactly one vertex of C (and so $e \cup E(C)$ spans 5 vertices).

By Kővári-Sós-Turán (Theorem 3.1), for each edge xy there are at most $O(d^{2-1/s})$ edges between $N(x)$ and $N(y)$ and so at most $O(d^{2-1/s})$ 4-cycles passing through xy. As there are at most nd edges, we have $|\mathcal{C}_4| = O(nd^{3-1/s})$. Thus $|\mathcal{S}_4| \leqslant 4d|\mathcal{C}_4| = O(nd^{4-1/s})$.

Similarly, let $\mathcal{C}_3$ be the set of 3-cycles $C \subset G$ and $\mathcal{S}_3$ be the set of pairs $(e, C) \in E(G) \times \mathcal{C}$ such that e contains exactly one vertex of C. By Kővári-Sós-Turán, for each vertex x there are at most $O(d^{2-1/s})$ edges in $N(x)$, and so each vertex belongs to at most $O(d^{2-1/s})$ triangles. We conclude that $|\mathcal{C}_3| = O(nd^{2-1/s})$, and $|\mathcal{S}_3| \leqslant 3d|\mathcal{C}_3| = O(nd^{3-1/s})$.

We will take $p = d^{(1/5s)-1}$. Let U be a random subset of $V(G)$ with each vertex independently included with probability p. Let $\mathcal{C}' \subset \mathcal{C}_3 \cup \mathcal{C}_4$ be the set of C such that $V(C) \subset U$, and define

$$U' := \bigcup_{C \in \mathcal{C}'} V(C)$$

and

$$U^* := \{u \in U : |N_G(u) \cap U| \geqslant 1 + 4pd_G(u)\}.$$

We shall consider $H := G[U - U' - U^*]$, which clearly lacks 3-cycles and 4-cycles, and also has $\Delta(H) \leqslant 1 + 4pd \leqslant 5d^{1/5s}$.

Let

$$\begin{aligned} X &= e(G[U]), \\ X^* &= \sum_{u \in U^*} |N_G(u) \cap U|, \\ Y &= |\mathcal{C}'|, \end{aligned}$$

and

$$Z = \#\{(e, C) \in \mathcal{S}_3 \cup \mathcal{S}_4 : e \in E(G[U]) \text{ and } C \in \mathcal{C}'\}.$$

We observe that $e(H) \geqslant X - X^* - 6Y - Z$ (indeed, for each $u \in U^*$, we lose at most $|N_G(u) \cap U|$ edges by deleting u; and for each $C \in \mathcal{C}'$, and any set $U \subset V(G)$, we have that $e(G[U]) - e(G[U - V(C)]) \leqslant \binom{|V(C)|}{2} + e(V(C), U \setminus V(C))$ (here the first term counts internal edges, and the second term counts the rest).

Now

$$\mathbb{E}[X] = p^2 e(G) \geqslant \frac{1}{2} n d^{1-\delta/s} p^2 = \frac{1}{2} npd^{(1/5-\delta)/s}.$$

Applying Markov's inequality (as in Lemma 3.2), we get

$$\begin{aligned} \mathbb{E}[X^*] &= \sum_{u \in V(G)} \sum_{v \in N_G(u)} p^2 \mathbb{P}(|N_G(u) \cap U| \geqslant 1 + 4pd_G(u) | \{u, v\} \subset U) \\ &\leqslant \sum_{u \in V(G)} d(u) p^2/4 \\ &\leqslant p^2 e(G)/2. \end{aligned}$$

Also, using our upper bounds on $|\mathcal{C}_3|$ and $|\mathcal{C}_4|$,

$$\mathbb{E}[Y] = p^4 |\mathcal{C}_4| + p^3 |\mathcal{C}_3| = O(nd^{-1-1/5s}) = O(npd^{-2/5s}),$$

and

$$\mathbb{E}[Z] = p^5 |\mathcal{S}_4| + p^4 |\mathcal{S}_3| = O(nd^{-1}) = O(npd^{-1/5s}).$$

Now, using the assumption that d is sufficiently large with respect to s, δ, we get

$$\mathbb{E}[X - X^* - 5d^{(1/5-2\delta)/s}|U| - 6Y - Z] \geqslant np\Big(d^{(1/5-\delta)/s}/4 - 5d^{(1/5-2\delta)/s} - O(d^{-1/5s})\Big) \geqslant 0.$$

Choosing U such that the above holds, we have that $e(G[U \backslash (U' \backslash U^*)]) \geqslant 5d^{(1/5-2\delta)/s}|U|$ so $H = G[U \backslash (U' \cup U^*)]$ has average degree at least $d' := 5d^{(1/5-2\delta)/s}$ and is C_4-free with $\Delta(H) \leqslant 5d^{1/5s} \leqslant d'^{1/(1-10\delta)}$, as desired. □

We now need the following lemma. Roughly speaking, it states that every graph G with sufficiently large average degree either contains an induced subgraph $G' \subseteq G$ which is almost-regular with still high average degree or we can find a very unbalanced bipartite (not necessarily induced) subgraph which still preserves many edges.

Lemma 3.4. *There exists an absolute constant κ so that the following holds for all $d \geqslant 2$ and $\delta \in (0, 1/5)$ with $d \geqslant (1/\delta)^{\kappa/\delta}$.*

Let G_0 be an n-vertex graph with $d(G_0) = d$ which is d-degenerate. Then either G_0 contains an induced subgraph G^* of average degree $d(G^*) \geqslant 6d^{1-5\delta}$ and maximum degree $\Delta(G^*) \leqslant 6d^{1+3\delta}$, or we can find a partition $V(G_0) = A \cup B$ such that $e(G_0[A, B]) \geqslant nd/4$ and $|B| \leqslant 4|A|/2^{d^\delta}$.

Remark. In the proofs of Lemma 3.4 and Lemma 3.5 below, we include the size assumption $d \geqslant x^{\kappa x}$ (respectively with $x = 1/\delta$ and $x = s/\delta$) so that various inequalities of the form "$C_1 \log d \leqslant d^{1/C_2 x}$" hold. These will be useful later, and also imply that d^δ is at least a large constant.

Proof. Let $V = V(G_0)$, and let $R \subset V$ be the vertices with degree greater than $d2^{d^\delta}$. Note that $e(G_0) \leqslant nd$ and so $|R| < n2^{1-d^\delta}$. If $e(G_0[R, V \setminus R]) \geqslant nd/4$ then we can simply take $B = R, A = V \setminus R$. So we may assume that $e(G_0[R, V \setminus R]) \leqslant nd/4$.

Since $G_0[R]$ is d-degenerate (and we may assume that κ is not too small), $e(G_0[R]) \leqslant d|R| < nd/8$, and so $e(G_0[V \setminus R]) \geqslant nd/8$. Let G be an induced subgraph of $G_0[V \setminus R]$ with maximal average degree: so $d(G) \geqslant d/4$ and $\delta(G) \geqslant d(G)/2 \geqslant d/8$. Note that $\Delta(G) \leqslant d2^{d^\delta}$ as it contains no vertices from R.

We split $V(G)$ dyadically, according to the degree of the vertices. For $-3 \leqslant i \leqslant d^\delta$, take $C_i := \{x \in V(G) : 2^i d \leqslant d_G(x) < 2^{i+1}d\}$; set $d_i := 2^i d$ and $E_i := \sum_{v \in C_i} d_G(v)$. By our bounds on $\delta(G), \Delta(G)$, the sets $C_{-3}, C_{-2}, \ldots, C_{d^\delta}$ partition $V(G)$.

By the pigeonhole principle, we can find j so that $E_j \geqslant 2e(G)/d^{2\delta}$. We now take a random subset $C' \subset C_j$ where each vertex $v \in C_j$ is included independently with probability $1/4$. Let $C'' \subset C'$ be the set of $v \in C'$ such that $|N_G(v) \cap C'| \leqslant d_G(v)/2$. By Markov's inequality, for $v \in C_j$, we have that

$$\mathbb{P}(v \in C'' | v \in C') \geqslant \mathbb{P}(|N_G(v) \cap C'| \leqslant 2\mathbb{E}[|N_G(v) \cap C'|]) \geqslant 1/2.$$

Consequently, we have

$$\mathbb{E}[|C''|] = \sum_{v \in C_j} \mathbb{P}(v \in C')\mathbb{P}(v \in C'' | v \in C') \geqslant |C_j|/8.$$

Fix some choice of C' such that $|C''| \geqslant |C_j|/8$. Note that for all $v \in C''$ and any $S \subset C''$, we have that $|N_G(v) \setminus S| \in [d_j/2, 2d_j]$ (the upper bound by the definition of C_j and the inclusion $C'' \subset C_j$, and the lower bound by the definition of C'' and the inclusion $C'' \subset C'$). Since $G[C'']$ is d-degenerate, we have $e(G[C'']) \leqslant d|C''|$. Thus $R' := \{v \in C'' : |N_G(v) \cap C''| \geqslant 4d\}$ satisfies $|R'| \leqslant |C''|/2$. Let $C''' := C'' \setminus R'$.

It follows that $|C'''| \geqslant |C''|/2$, and thus $e(G[C''', V(G) \setminus C''']) \geqslant |C'''|d_j/2 \geqslant d_j|C_j|/32 \geqslant E_j/64$.

Now consider: $V(G) \setminus C'''$: we partition these dyadically, now according to neighbours in C'''. Specifically, for $-\log d \leqslant i \leqslant d^\delta$, let $C_i^* := \{z \in V(G) \setminus C''' : 2^i d \leqslant |N_G(z) \cap C'| < 2^{i+1}d\}$ and $E_i^* := \sum_{v \in C_i^*} |N_G(v) \cap C'''|$. The sets C_i^* partition $V(G) \setminus (C''' \cup \{v : |N_G(v) \cap C'''| = 0\})$.

By pigeonhole, there exists some k such that $e(G[C''', C_k^*]) \geqslant e(G[C''', V(G) \setminus C'''])/d^{2\delta}$. For $v \in C_k^*$, we have $|N_G(v) \cap C'''| \in [2^k d, 2^{k+1}d]$. Let $R^* := \{v \in C_k^* : |N_G(v) \cap C_k^*| \geqslant 4d\}$. By d-degeneracy, we have that $|R^*| \leqslant |C_k^*|/2$. Thus, taking $C^{**} := C_k^* \setminus R^*$, we have that $e(G[C''', C^{**}]) \geqslant e(G[C''', C_k^*])/4 \geqslant e(G)/128d^{4\delta}$.

For convenience, write $A := C''', B := C_k^{**}$ and consider $\Gamma := G[A, B]$. Since $\Delta(G[A]), \Delta(G[B]) \leqslant 4d$, we have $\Delta(G[S]) \leqslant \Delta(\Gamma[S]) + 4d$ for every $S \subset A \cup B$. Thus we will be done by finding $S \subset A \cup B$ such that $d(\Gamma[S]) \geqslant 6d^{1-5\delta}$ and $\Delta(\Gamma[S]) \leqslant d^{3\delta} d(\Gamma[S]) \leqslant 2d^{1+3\delta}$ (note that $G_0 \supset \Gamma$ being d-degenerate implies $d(\Gamma[S]) \leqslant 2d$).

We now show that Γ is $16d^{2\delta}$-almost-biregular. Since $d(\Gamma) \geqslant d^{1-4\delta}/128$, we are then done by an application of Lemma 3.2. To verify almost-biregularity, we first note

$$\max_{b \in B} \left\{ \frac{d_\Gamma(b)|B|}{e(\Gamma)} \right\} \leqslant \frac{\max_{b \in B}\{d_\Gamma(b)\}}{\min_{b \in B}\{d_\Gamma(b)\}} \leqslant 2$$

(where the last inequality is because $d_\Gamma(b) = |N_G(b) \cap C'''| \in [2^k d, 2^{k+1} d]$ for all $b \in B$). Meanwhile, writing $\Gamma' := G[A, V(G) \setminus A]$ we similarly have that

$$\max_{a \in A} \left\{ \frac{d_{\Gamma'}(a)|A|}{e(\Gamma')} \right\} \leqslant \frac{\max_{a \in A}\{d_{\Gamma'}(a)\}}{\min_{a \in A}\{d_{\Gamma'}(a)\}} \leqslant 4,$$

and so $\max_{a \in A}\left\{\frac{d_\Gamma(a)|A|}{e(\Gamma)}\right\} \leqslant \max_{a \in A}\left\{\frac{d_{\Gamma'}(a)|A|}{e(\Gamma)}\right\} \leqslant 4\frac{e(\Gamma')}{e(\Gamma)} \leqslant 16d^{2\delta}$. □

Lemma 3.5. *There exists an absolute constant κ so that the following holds for all $s \geqslant 2$, $\delta \in (0, 1/2)$ and $d \geqslant (s/\delta)^{\kappa s/\delta}$.*

Let G_0 be a d-degenerate, $K_{s,s}$-free graph with $V(G_0) = A_0 \cup B$. Suppose that $|A_0| \geqslant 2^{d^\delta}|B|$ and $e(G_0[A_0, B]) \geqslant 3d^{1/2}|A_0|$.

Then, for any $r \leqslant d^{\min\{\delta/4, 1/3s\}}$, we can find subsets $A' \subset A_0, B' \subset B$ such that $G_0[A'], G_0[B']$ are independent sets, $|A'| \geqslant 2^{d^{\delta/2}}|B'|$, and $|N(a) \cap B'| = r$ for $a \in A'$.

Remark. In our applications, we will only need the special case where $e(G_0[A_0, B]) \geqslant d|A_0|/10^4$.

Proof. First, we clean G_0 so that our larger part has reasonably bounded degree. Let $A^* := \{a \in A_0 : d_{G_0}(a) \geqslant 10d\}$. We have have that $e(A^*, B) \geqslant 10d|A^*|$ and since G_0 is d-degenerate we must also have $e(A^*, B) \leqslant d(|A^*| + |B|)$, implying $|A^*| \leqslant |B|$ and $e(A^*, B) \leqslant 2d|B|$. Next let $A^{**} := \{a \in A_0 : d_{G_0}(a) \leqslant d^{1/2}\}$. We have that $e(A^{**}, B) \leqslant d^{1/2}|A_0|$.

Let $A_1 := A_0 \setminus (A^* \cup A^{**})$. Then $e(G_0[A_1, B]) \geqslant (3d^{1/2} - 2d|B|/|A_0| - d^{1/2})|A_0| \geqslant d^{1/2}|A_0|$ and so $|A_1| \geqslant e(G_0[A_1, B])/10d \geqslant |A_0|/10d^{1/2}$. Note that $d_{G_0}(a) \in [d^{1/2}, 10d]$ for $a \in A_1$.

As G_0 is d-degenerate, it has chromatic number at most $d+1$. So we can find $A \subset A_1$ with $|A| \geqslant |A_1|/2d \geqslant |A_0|/20d^{3/2} \geqslant |B|2^{2d^{\delta/2}}$, such that $G_0[A]$ is an independent set. Let $G = G_0[A \cup B]$.

Now let D be an orientation of $G[B]$ such that $|N_D^+(b)| \leqslant d$ for each $b \in B$ (this can be done as G is d-degenerate). Let $B_0' \subset B$ be a random subset where each $b \in B$ is kept with probability $p = 1/d^2$. Let B' be the set of $b \in B_0'$ where $|N_D^+(b) \cap B_0'| = 0$.

Consider any $a \in A$. Since $G[N_G(a)]$ is $K_{s,s}$-free and has at least $d^{1/2}$ vertices, we can fix an independent set $I \subset N_G(a)$ of size r (as $r \leqslant d^{1/3s}$) by Kővári-Sós-Turán (Theorem 3.1). Let $\mathcal{E}_a$ be the event that $N_G(a) \cap B' = N_G(a) \cap B_0' = I$. Let $X = (N(a) \setminus I) \cup \bigcup_{b \in I} N_D^+(b)$. Then $|X| \leqslant d(10 + r) \leqslant d^2/2$. Consequently,

$$\mathbb{P}(\mathcal{E}_a) = p^{|I|}(1-p)^{|X|} \geqslant 2^{-2r\log d}(1/2) \geqslant 2^{-d^{\delta/3}}.$$

We set $A' := \{a \in A : \mathcal{E}_a \text{ holds}\}$. We have that $\mathbb{E}[|A'|] \geqslant |A|2^{-d^{\delta/3}} \geqslant 2^{d^{\delta/2}}|B| \geqslant 2^{d^{\delta/2}}|B'|$. Thus, there is some choice of A', B' where this inequality holds. Fixing such a choice, we are done. □

We will also need the following slight variant of a result of Füredi ([11, Theorem 1′]; see also [7]). We say a hypergraph $\mathcal{H}$ is *s-bounded* if $|e| \leqslant s$ for each $e \in E(\mathcal{H})$.

Theorem 3.6. *Let s, r, t be positive integers, and let $T = \sum_{k=0}^{s} \binom{r}{k}$. Let $\mathcal{F} = (V, E)$ be an r-uniform hypergraph. There exist $E^* \subset E$ with $|E^*| \geqslant (tr2^{2r+1})^{-T-1}|E|$ and $c : V \to [r]$ such that:*

- *$c(F) = [r]$ for all $F \in E^*$ (i.e. $\mathcal{F}^* := (V, E^*)$ is r-partite with vertex classes given by the sets $c^{-1}(i)$);*
- *there is an s-bounded hypergraph $\mathcal{H}$ on vertex set $[r]$, such that, for $e \in \binom{[r]}{\leqslant s}$:*
  - *if $e \in E(\mathcal{H})$, then for every $F \in E^*$, there are distinct $F_1, \ldots, F_t \in E^*$ such that $c(F \cap F_i) \supset e$ for each $i \in [t]$;*
  - *if $e \notin E(\mathcal{H})$, then for all distinct $F, F' \in E^*$, we have $c(F \cap F') \not\supset e$.*

  Thus if F_0 is contained in an edge of E^*, then F_0 extends to an edge of $\mathcal{F}^*$ in at least t different ways if $c(F_0)$ is an edge of $\mathcal{H}$, and otherwise it has a unique extension.

Proof. Considering a random coloring $c : V \to [r]$, we have that $\mathbb{P}(c(F) = [r]) = \frac{r!}{r^r} \geqslant \exp(-r)$ for each $F \in E$. Thus, we can choose some c such that $E_0 := \{F \in E : c(F) = [r]\}$ satisfies $|E_0| \geqslant \exp(-r)|E|$.

We run an iterative cleaning process, giving $E_0 \supset E_1 \supset \cdots \supset E_\tau$ (where $\supset$ denotes non-strict containment). For $i = 0, \ldots, \tau$, let S_i be the set of $e \in \binom{[r]}{\leqslant s}$ such that there exist distinct $F, F' \in E_i$ with $c(F \cap F') \supset e$. Note that $S_0 \supset S_1 \supset \cdots \supset S_\tau$, and $|S_0| \leqslant T$. We will ensure that:

(1) for all $e \in S_\tau$ and $F \in E_\tau$, there exist distinct $F_1, \ldots, F_t \in E_\tau$ such that $c(F \cap F_i) \supset e$ for each $i \in [t]$;
(2) for $i < \tau$, we have $|E_{i+1}| \geqslant (2tT^2)^{-1}|E_i|$;
(3) for $i < \tau - 1$, we have $|S_{i+1}| < |S_i|$.

By Items 3 and 2, we see that $\tau \leqslant T + 1$ and thus $|E_\tau| \geqslant (2tT^2)^{-T-1}|E_0|$. Considering Item 1, we see that we can take $E^* = E_\tau$, and that the size condition follows because $|E_0| \geqslant \exp(-r)|E|$ and $2T^2 \leqslant 2^{2r+1}$.

It remains to describe our cleaning process and confirm that Items 2-3 hold.

Suppose we have defined E_i Create a bipartite graph Γ where one part is $A := E_i$, and the other part is $B := \{F \cap c^{-1}(e) : F \in E_i, e \in S_i\}$. For $F \in A, U \in B$, we say $F \sim U$ if $F \supset U$. Note that $d_\Gamma(F) = |S_t| \leqslant T$ for $F \in A$. If $|B| \leqslant |A|/2tT$, then we can iteratively delete $\{U\} \cup N_\Gamma(U)$ for any U with degree $< t$. At the end, we will be left with a graph Γ' with $e(\Gamma') \geqslant e(\Gamma)/2$. We stop the process and take $E_\tau = A(\Gamma')$.

Otherwise, $|B| \geqslant |A|/(2tT)$. In which case, by pigeonhole, we can find $e_i \in S_i$ such such that $|B \cap c^{-1}(e_i)| \geqslant |B|/|S_i| \geqslant |A|/(2tT^2)$. Let $E_{i+1} \subset E_i$ to be a maximal set where $c^{-1}(e_i) \cap F \neq c^{-1}(e_i) \cap F'$ for distinct $F, F' \in E_{i+1}$, we have that $|E_{i+1}| = |B \cap c^{-1}(e_i)| \geqslant |E_i|/(2tT^2)$ and $e_i \notin S_{i+1}$. □

4. An extremal hypergraph problem

We will need some terminology for hypergraphs, some of which is non-standard.

We say a hypergraph $\mathcal{H} = (V, E)$ is *covered* if for every $v \in V$ there is some $e \in E$ where $v \in e$. Recall that we say $\mathcal{H}$ is *ℓ-bounded* if $|e| \leqslant \ell$ for $e \in E$.

For $A, B \subset V$, we say (A, B) is an *induced pair of order* k if $|B| = k$ and for each $b \in B$, there exists $e \in E$ such that $A \cup \{b\} \subset e$, but there is no $e \in E$ with $e \supset A$ and $|e \cap B| > 1$. We define $\alpha(\mathcal{F})$ to be the maximum k such that there exists an induced pair (A, B) of order k.

Let $F(\ell, k)$ denote the minimum n such that every covered ℓ-bounded hypergraph $\mathcal{H}$ on at least n vertices has $\alpha(\mathcal{H}) \geqslant k$. The following lemma tells us this quantity is always finite.

Lemma 4.1. *For $\ell \geqslant 0$, we have that $F(\ell, k) \leqslant \sum_{l=0}^{\ell}(k-1)^l$.*

Proof. We argue by induction on ℓ. For $\ell = 0$, it is vacuously true that $F(0, k) = 1$, as any covered hypergraph $\mathcal{H}$ with $|V(\mathcal{H})| > 0$ has an edge e with $|e| > 0$ (and hence is not 0-bounded). So we may assume $\ell \geqslant 1$.

Given a covered hypergraph $\mathcal{H} = (V, E)$, we consider the graph $G = G_{\mathcal{H}}$ with $V(G) = V(\mathcal{H})$ where u, v are adjacent if there is some $e \in E(\mathcal{H})$ containing both u and v. We note that if $B \subset V$ is an independent set in G, then $(\varnothing, B)$ is an induced pair in $\mathcal{H}$. Indeed, since $\mathcal{H}$ is covered, every vertex of B is contained in some edge of $\mathcal{H}$; and no edge meets B in more than one element, as B is an independent set in G. Thus $\alpha(\mathcal{H}) \geqslant |B|$

Now suppose $\alpha(\mathcal{H}) < k$. Let $I \subset V$ be a maximal independent set in $G = G_{\mathcal{H}}$. We must have $|I| \leqslant k - 1$, and so there must be some $x \in I$ such that $d_G(x) \geqslant \frac{|V|-|I|}{|I|} \geqslant \frac{|V|}{k-1} - 1$.

Let $\mathcal{H}' = (V', E')$, where $V' = N_G(x)$ and $E' = \{e \cap V' : e \in E, x \in e\}$. Then $\alpha(\mathcal{H}') \leqslant \alpha(\mathcal{H})$, as if (A, B) is an induced pair of $\mathcal{H}'$, then $(A \cup \{i\}, B)$ is an induced pair of $\mathcal{H}$. Also, $\mathcal{H}'$ is covered and $\max\{|e'| : e' \in E'\} < \max\{|e| : e \in E\} \leqslant \ell$. So $|V'| < F(\ell - 1, k)$, and the result follows by induction. □

It appears to be an interesting problem to understand how $F(\ell, k)$ grows. We comment on a few different regimes below.

Letting $R(K_a, K_b)$ denote the Ramsey number for a clique of size a or independent set of size b, we observe that $F(\ell, k) \geqslant R(K_{\ell+1}, K_k)$. Indeed, given a graph G with no cliques of size $\ell+1$ or independent sets of size k, we can define $\mathcal{H}_G = (V(G), \{e \subset V(G) : G[e] \text{ is a clique}\})$, which is covered, ℓ-bounded, and has $\alpha(\mathcal{H}_G) < k$. It follows that $F(k, k) \geqslant R(K_k, K_k) \geqslant \exp(\Omega(k))$. Meanwhile, Theorem 4.1 tells us $F(k, k) \leqslant k^k$. So in this regime, our upper bound (roughly) has the right asymptotic shape.

When $k = 2$, Theorem 4.1 tells us that $F(\ell, 2) \leqslant \ell + 1$. It is easy to see that equality holds here (consider a hypergraph with ℓ vertices, which are all contained in one edge). However, for fixed $k \geqslant 3$, Theorem 4.1 only gives an exponential upper bound (i.e., that $F(\ell, k) \leqslant \exp(O_k(\ell))$). Here it seems that a much better upper bound should hold. In fact, we conjecture $F(\ell, k) \leqslant \ell^{O_k(1)}$. But already for $k = 3$, we do not know how to prove $F(\ell, 3) \leqslant \exp(o(\ell))$.

5. Proof of main theorem

We begin in the first subsection by handling a 'model' case in Proposition 5.1; in the following subsection, we prove Theorem 1.1 by using the various cleaning lemmas from Section 3 to reduce to the model case covered by Proposition 5.1

5.1. The model case.

Consider $s, k \geqslant 2$. We want to argue that if G is a $K_{s,s}$-free graph with large average degree, that it contains a C_4-free induced subgraph G' with $d(G') \geqslant k$.

In our proof, all the main difficulties arise when $G = (A, B, E)$ is a bipartite graph that is 'lopsided', meaning $|A|/|B|$ is very large compared to $d(G)$. By applying Lemmas 3.6 and 4.1, we shall prove the following 'model case'.

Proposition 5.1. *The exists some absolute constant C so that the following holds for all $s, k \geqslant 2$ and $r \geqslant k^{Cs}$, and $|A| \geqslant 2^{r^{Cs}}|B|$.*

Let $G = (A, B, E)$ be bipartite graph, with $d_G(a) = r$ for each $a \in A$. Then either G contains a $K_{s,s}$, or an induced subgraph G', which is C_4-free and has $d(G') \geqslant k$.

Proof. We may assume G is $K_{s,s}$-free. Let $\rho := |A|/|B| \geqslant 2^{r^{Cs}}$.

Consider the r-uniform hypergraph $\mathcal{F}$, where $V(\mathcal{F}) = B$ and $e \in E(\mathcal{F})$ if there is $x \in A$ such that $N_G(x) = e$. Since G is $K_{s,s}$-free and $r \geqslant s$, we have that $|E(\mathcal{F})| \geqslant |A|/s$ (because we can't have s different vertices in A corresponding to the same hyperedge of $\mathcal{F}$); very crudely we have

$$|E(\mathcal{F})| \geqslant |A|/s = (\rho/s)|V(\mathcal{F})| \geqslant \rho^{1/2}|V(\mathcal{F})|.$$

Fix an injection $\phi : E(\mathcal{F}) \to A$ so that $f = N_G(\phi(f))$ for each $f \in E(\mathcal{F})$.

We apply Theorem 3.6 with parameters $r^* = r$, $t^* = \max\{k, s\}$ and $s^* = s$ to get an r-partite sub-hypergraph $\mathcal{F}' \subset \mathcal{F}$ satisfying the required properties and with $|E(\mathcal{F}')| \geqslant (t^* r 2^{2r+1})^{-\sum_{i=0}^{s}\binom{r}{i}-1}|E(\mathcal{F})|$; very crudely we have

$$t^* r 2^{2r+1} \leqslant 2^{3r} \quad \text{and} \quad \sum_{i=0}^{s}\binom{r}{s} + 1 \leqslant r^s,$$

and so

$$|E(\mathcal{F}')| \geqslant 2^{-r^{s+3}}|E(\mathcal{F})| \geqslant \rho^{1/4}|V(\mathcal{F})| = \rho^{1/4}|V(\mathcal{F}')|.$$

Note that, as $\rho > 1$, this implies

$$|E(\mathcal{F}')| > |V(\mathcal{F}'|.$$

Let $c : B \to [r]$ be the r-partite coloring associated with $\mathcal{F}'$, and $\mathcal{H}$ be the hypergraph on $[r]$ as in the last bullet of Theorem 3.6.

By Lemma 4.1, we have $|V(\mathcal{H})| = r \geqslant k^{Cs} \geqslant F(s-1, k)$. If $E(\mathcal{H})$ has an edge of size s, then G contains a $K_{s,s}$ which is a contradiction. Thus $\mathcal{H}$ is $(s-1)$-bounded. It is also clear that $\mathcal{H}$ is covered, as $|E(\mathcal{F}')| > |V(\mathcal{F}')|$ (Indeed, consider $i \in [r] = V(\mathcal{H})$. If the singleton $\{i\}$ doe not belong to $E(\mathcal{H})$ then no vertex in the color class $c^{-1}(i)$ is contained in two edges $f, f' \in E(\mathcal{F}')$, by definition of $\mathcal{H}$. Recalling each $f \in E(\mathcal{F}')$ intersects the color class $c^{-1}(i)$, and noting $|c^{-1}(i)| \leqslant |V(\mathcal{F}')| < |E(\mathcal{F}')|$, we get a contradiction due to pigeonhole.)

By definition of $F(s-1, k)$, $\mathcal{H}$ must have an induced pair of order k, say (X, Y).

Fix some $f \in E(\mathcal{F}')$, and take $S := f \cap c^{-1}(X)$ (i.e., its intersection with the parts corresponding to X). Now, we consider the common neighborhood

$$A' := \phi(E(\mathcal{F}')) \cap \bigcap_{v \in S} N_G(v),$$

which is nonempty (as it contains $\phi(f)$). Next, take

$$B' := c^{-1}(Y) \cap \bigcup_{a \in A'} N_G(a).$$

Finally, let G' be the induced subgraph $G[A' \cup B']$.

Claim 5.2. *G' is C_4-free and $d(G') \geqslant k$.*

Proof. Suppose G' contained a C_4. This implies that there are distinct $f_1, f_2 \in E(\mathcal{F}')$ such that $c(f_1 \cap f_2) \supset X \cup \{y_1, y_2\}$ for some distinct $y_1, y_2 \in Y$, contradicting the fact that (X, Y) is an induced pair. It follows that G' must be C_4-free as desired.

Meanwhile, for every $b \in B'$, there are at least $t \geqslant k$ hyperedges $f_1, \ldots, f_t \in \mathcal{F}'$ containing b and S (by the definitions of $\mathcal{F}', \mathcal{H}$, and the fact that (X, Y) is an induced pair) so b has degree at least k within G'. Also, for $a \in A'$, we have that $N_{G'}(a) = N_G(a) \cap c^{-1}(Y)$ implying $d_{G'}(a) = k$ (as $|Y| = k$). Thus $d(G') \geqslant k$ as desired. □

We see that G' has the desired properties, completing the proof. □

Remark. Inspecting the above proof, one can get the same conclusion as long as $r \geqslant F(s-1,k)$, and $\rho := |A|/|B|$ satisfies $\rho > sk^{r^{s+3}}$.

5.2. **The details for general graphs.**

Let G be a graph with average degree $d \geqslant k^{Cs^3}$ (where C is some large absolute constant). We assume G is $K_{s,s}$-free, and wish to deduce that G contains an induced C_4-free subgraph, G', with $d(G') \geqslant k$. We may assume that G is d-degenerate, or else pass to some induced subgraph G^* of G with larger average degree $d^* > d \geqslant k^{Cs^3}$. Thus the assumptions of Lemma 3.4 are satisfied, as the degeneracy of G is at least the average degree of G.

Let $\delta_0 := 1/200s$, and for $i > 0$ set $\delta_i := \delta_0/2^i$. These quantities will be used throughout the proof.

We now apply Lemma 3.4 with parameters $d := d, \delta := \delta_0$. There are two cases.

Case 1: *G contains an induced subgraph G^* with $d(G^*) \geqslant 6d^{1-5\delta_0}$ and $\Delta(G^*) \leqslant 6d^{1+3\delta_0}$.* In this case, we use Lemma 3.3. Indeed, we have $d(G^*) \geqslant d^{1/2}$ and $d(G^*) \geqslant \Delta(G^*)^{1-8\delta_0} = \Delta(G^*)^{1-1/25s}$. Assuming C is large, we are guaranteed that $d(G^*) \geqslant \kappa^{25s}$ with room to spare (where κ is the absolute constant from Lemma 3.3).

We apply Lemma 3.3 with parameters $d := d(G^*), s := s$ and $\delta := 1/25$ (note that $d(G^*) \geqslant \kappa^{25s}$, assuming C is large). We obtain an induced subgraph G' of G, which is C_4-free, and has average degree $d(G') \geqslant d(G^*)^{1/10s} \geqslant d(G)^{1/20s} = k^{(C/20)s^2} \geqslant k$.

Case 2: *There is a partition $V(G) = A_0 \cup B$ such that $e(G[A_0, B]) \geqslant nd/4$ and $|A_0| \geqslant 2^{d^{\delta_1}}|B|$.* Here, we will pass to some induced subgraph where we can apply Proposition 5.1.

Let C_0 be the absolute constant from that proposition. We assume that C is large enough so that $C/1600 \geqslant C_0^2$ (note that $1/1600 = s\delta_1/4 \leqslant s\delta_2$). For any $r \leqslant d^{\min\{\delta_1/4, 1/3s\}} = d^{\delta_1/4}$, we can apply Lemma 3.5 to $G[A_0 \cup B]$ to obtain $A' \subset A_0$ and $B' \subset B$, where:

- $G[A'], G[B']$ are both independent sets;
- $|A'| \geqslant 2^{d^{\delta_2}}|B'| \geqslant 2^{k^{C'^2 s^2}}|B|$ (recalling $d \geqslant k^{Cs^3}$); and
- $|N(a) \cap B'| = r$ for $a \in A'$.

Because C is sufficiently large, we can apply this with $r := k^{C's}$ and obtain $A' \subset A_0$ and $B' \subset B$ such that $|A'| \geqslant 2^{r^{C's}}|B'|$. Writing $G^* := G[A' \cup B']$, we can now apply Proposition 5.1 to find the desired induced subgraph G' inside G^*. This completes the proof. □

Inspecting our proof, one obtains the following technical strengthening of Theorem 1.1.

Theorem 5.3. *There exists an absolute constant C such that the following holds for all $d, s, k \geqslant 2$ and $\delta \in (0, 1/20)$ such that $d \geqslant (k/\delta)^{Cs^3/\delta}$.*

Let G be a graph with average degree $d(G) \geqslant d$ without a $K_{s,s}$. Then G has an induced subgraph G' which is $\{C_3, C_4\}$-free with $d(G') \geqslant k$ and either

- *$d(G') \geqslant \Delta(G')^{1-\delta}$; or*
- *G' is bipartite.*

Furthermore, we can ensure the first outcome if $\Delta(G) \leqslant d2^{d^{\delta/(1000s)}}$.

Proof. Repeating the argument above with $\delta_0 = \frac{\delta}{1000s}$, we will obtain G' satisfying one of the two bullets. In Case 1, it is easy to check that the first outcome holds, by looking at the full statement of Lemma 3.3. And in Case 2, we end up passing to a bipartite graph, causing the second outcome to hold.

Under the assumption that $\Delta(G) \leqslant d2^{d^{\delta/(1000s)}}$, the set R from the proof of Lemma 3.4 will be empty, so we can find an induced subgraph G^* with $d(G^*) \geqslant 6d^{1-5\delta_0}, \Delta(G^*) \leqslant 6d^{1+3\delta_0}$. In this case, we reach Case 1, and thus can ensure the first outcome. □

6. Lower bounds

We require the following well-known extremal bound of Reiman [25] (which slightly improved the bounds of Kővári-Sós-Turán [21]).

Proposition 6.1 ([25, Equation 1.4])**.** *Any n-vertex C_4-free graph G has at most $\frac{1}{2}n^{3/2} + n/4 + 1$ edges.*

Corollary 6.2. *Any C_4-free graph G with average degree $d(G) \geqslant k \geqslant 1$ must have $|V(G)| \geqslant (k-1)^2$.*

Given an integer K and $p \in [0,1]$, let $q(K,p)$ be the probability that $G \sim G(K,p)$ is C_4-free. (We write $G(K,p)$ for the Erdős-Rényi random graph where each edge of the K-vertex complete graph is included independently at random with probability p.)

Proposition 6.3. *We have that*

$$q(K,p) \leqslant (1-p^4)^{\binom{\lfloor K/2 \rfloor}{2}} \leqslant \exp\left(-p^4\binom{\lfloor K/2 \rfloor}{2}\right).$$

Proof. Sample $G \sim G(K,p)$. For $i < j \in [\lfloor K/2 \rfloor]$, $\mathcal{E}_{i,j}$ be the event that $G[\{2i-1, 2i\}, \{2j-1, 2j\}]$ is a C_4. Then $\mathbb{P}(\mathcal{E}_{i,j}) = 1 - p^4 \leqslant \exp(-p^4)$, and these events are independent. The result follows. □

Theorem 6.4. *Let $k, s \geqslant 2$. Choose p, n such that*

$$\max\{n^2p^s, n(1-p^4)^{(k^2-3k-2)/4}, \exp(-\binom{n}{2}p/4)\} \leqslant 1/2.$$

Then there exists an n-vertex graph G with $d(G) \geqslant \frac{n-1}{2}p$, such that G is $K_{s,s}$-free, and every C_4-free induced subgraph $G' \subset G$ has $d(G') < k$.

Proof. Consider $G \sim G(n,p)$ and fix $K := k^2 - 3k$ (which is always even). Let X count the number of K-subsets $S \in \binom{[n]}{K}$ such that $G[S]$ is C_4-free. By Corollary 6.2, if $X = 0$, then G does not contain an induced C_4-free subgraph G' with $d(G') \geqslant k$. We note that (by Proposition 6.3)

$$\mathbb{E}[X] = \binom{n}{K} q(K,p) \leqslant n^K (1-p^4)^{\binom{K/2}{2}} = \left(n(1-p^4)^{(k^2-3k-2)/4}\right)^K.$$

Now let Y count the number of copies of $K_{s,s}$ in G. Then $\mathbb{E}[Y] \leqslant n^{2s}p^{s^2} = (n^2p^s)^s$.

By assumption, $k, s \geqslant 2$ (and so $K \geqslant 2$). Since $\max\{n(1-p^4)^{(k^2-3k-2)/4}, n^2p^s\} \leqslant 1/2$, we get

$$\mathbb{P}(X = Y = 0) \geqslant 1 - \mathbb{E}[X] - \mathbb{E}[Y] \geqslant 1/2.$$

Meanwhile, a Chernoff bound tells us that $\mathbb{P}(e(G) \leqslant \binom{n}{2}p/2) < \exp(-\binom{n}{2}p/4)$.

It follows by our assumptions on n, p that with positive probability, we will have $X = Y = 0$ and $e(G) \geqslant p\binom{n}{2}/2$. Taking such a G gives our result. □

Recall that $f_{\text{Ind}}(s,k)$ denotes the smallest D such that for any $K_{s,s}$-free graph G with $d(G) \geqslant D$, there exists a C_4-free induced subgraph $G' \subset G$ with $d(G) \geqslant k$. We will use Theorem 6.4 to get some lower bounds for this function.

Corollary 6.5 (Theorem 1.3 (a)). *For sufficiently large k, we have that $f_{\text{Ind}}(k,k) \geqslant k^{k/21}$.*

Proof. Assuming k is sufficiently large, we may apply Theorem 6.4 with $n = k^{k/20}$ and $p = k^{-1/5}$.

It is trivial to verify that the first and third terms in the statement of Theorem 6.4 are appropriately bounded, so we only discuss the second condition. Here (for $k \geqslant 10$), the crude bound $(1-p^4)^{(k^2-3k-2)/4} \leqslant \exp(-p^4(k^2-3k-2)/4) \leqslant \exp(-k^{6/5}/8)$ gives us what we need (as $\log n \leqslant k \log k \leqslant k^{6/5}/8 - 1$ for large k). □

Remark. It is not hard to prove that the probabilistic constructions in this section are almost-regular with high probability. Thus, Corollary 6.5 (along with Corollary 6.9, detailed later below) tell us that Lemma 3.3 is essentially tight (up to the constant in our exponent) when we are looking for subgraphs with large average degree. This is rather surprising since in the non-induced setting one gets much better (i.e., polynomial) bounds when G is almost-regular.

Corollary 6.6 (Theorem 1.3 (b)). *Fix $k \geqslant 2$. We have that $f_{\text{Ind}}(s,k) \geqslant s^{(1/4-o(1))(k^2-3k-2)}$.*

Proof. Note that this result is trivial for $k = 2, 3$, as then $k^2 - 3k - 2 < 0$. So now assume $k \geqslant 4$.

We use Theorem 6.4 again, with $n = s^{(1/4-\epsilon)(k^2-3k-2)}, p = 1 - k^2 \log s/s$ where $\epsilon = \epsilon(s)$ tends to zero as $s \to \infty$. Crudely, we have that $n^2 p^s \leqslant \exp(\log s(k^2-3k-2) - k^2 \log s) < 1/2$. Meanwhile, we have that $(1-p^4)^{(k^2-3k-2)/4} \leqslant (4k^2 \log s/s)^{(k^2-3k-2)/4} \leqslant s^{(o(1)-1/4)(k^2-3k-2)} < 1/2n$ (assuming $\epsilon(s)$ does not tend to zero too fast). The third condition holds as $s \to \infty$, because $n \to \infty$ and $p \to 1$. □

Our final lower bound does not require Theorem 6.4. Instead, we need a basic observation. Here we write $\alpha(G)$ for the maximum size of an independent set of a graph G.

Proposition 6.7. *Any n-vertex C_4-free graph G has an independent set of size $\geqslant \frac{n}{3+\sqrt{n}}$.*

Proof. This follows from Proposition 6.1, and the fact that $\alpha(G) \geqslant \frac{|V(G)|}{1+d(G)}$ for any graph G (a well-known corollary of Turán's Theorem). □

Combined with Corollary 6.2, we get the following.

Corollary 6.8. *Any C_4-free graph G with average degree $d(G) \geqslant k \geqslant 1$ must have $\alpha(G) \geqslant \frac{(k-1)^2}{k+2} \geqslant k-4$.*

Corollary 6.9 (Theorem 1.3 (c)). *Fix $s \geqslant 2$. Then $f_{\text{Ind}}(s,k) \geqslant k^{(1/2-o(1))s-1}$.*

Proof. Let $k' = k-4$. We consider $G \sim G(n,p)$, where n, p will be chosen later. Let X count the number of k'-subsets $S \in \binom{[n]}{k'}$ such that $G[S]$ is an independent set. By Corollary 6.8, if $X = 0$, then G will not contain an induced C_4-free subgraph G' with $d(G') \geqslant k$. We note that $\mathbb{E}[X] = \binom{n}{k'}(1-p)^{\binom{k'}{2}} \leqslant (n(1-p)^{k/4})^{k'}$ (assuming $k \geqslant 10$).

Let Y count the number of copies of $K_{s,s}$ in G. We have $\mathbb{E}[Y] \leqslant n^{2s} p^{s^2} \leqslant (n^2 p^s)^s$.

Finally, let $\mathcal{E}$ be the event that $e(G) \leqslant \binom{n}{2}p/2$. A Chernoff bound tells us $\mathbb{P}(\mathcal{E}) < \exp(-\binom{n}{2}p/4)$.

If $\mathbb{E}[X] + \mathbb{E}[Y] + \mathbb{P}(\mathcal{E}) < 1$, a union bound tells us that $f_{\text{Ind}}(s,k) \geqslant \frac{(n-1)p}{2}$. Taking $p = n^{-(2+\varepsilon)/s}$, $n = k^{\frac{1-\varepsilon}{2+\varepsilon}s}$ so that $p = k^{\varepsilon-1}$, one can check that this holds for sufficiently large k. □

Remark. To provide lower bounds for $f_{\mathrm{Ind}}(s,k)$, we found $K_{s,s}$-free graphs G with the stronger property that every set of $(k-1)^2$ vertices contains a C_4. We note that one cannot hope to do better with such an approach (beyond improving the constants in the exponents). Indeed, such G must not contain a clique on $2s$ vertices, nor an independent set on $(k-1)^2$ vertices. Thus, the average degree of G (which is at most $|V(G)|-1$), will be bounded by the off-diagonal Ramsey number $R(K_{2s},K_{k^2})$. A classical bound of Erdős-Szekeres [9] tells us

$$R(K_a,K_b) \leqslant \binom{a+b-2}{b-1} \leqslant (a+b)^{\min\{a,b\}},$$

and in particular we get

$$R(K_{2k},K_{k^2}) \leqslant k^{(4+o(1))k} \quad \text{and} \quad R(K_{2s},K_{k^2}) \leqslant \min\{s^{(1+o(1))k^2}, k^{(4+o(1))s}\}.$$

7. Corollary on subdivisions

In this section we prove Corollary 1.6, which improves the bounds in the theorem of Kühn and Osthus [19] (Theorem 1.5) from triply exponential to singly exponential. We restate the result for convenience.

Corollary 1.6. *There exists a constant C so that for all $k,s\in\mathbb{N}$, every $K_{s,s}$-free graph G with average degree at least k^{Cs^3} contains an induced subdivision of K_k.*

We require the following result, which was proved independently by Bollobás and Thomason [1] and Kómlos and Szemerédi [17].

Theorem 7.1. *There exists an absolute constant $C>0$ such that for all $k\in\mathbb{N}$, if G is a graph with $d(G)\geqslant Ck^2$, then G contains a (not necessarily induced) subdivision of K_k.*

We can now proceed with the proof of the corollary.

Proof of Corollary 1.6. Let G be a $K_{s,s}$-free graph with $d(G)\geqslant k^{Cs^3}$ (where C is a large constant). Applying Theorem 5.3 with $\delta=1/100$, we can find an induced $\{C_3,C_4\}$-free subgraph G' with average degree $C'k^5$ (with $C'=\Omega(C)$), where either G' is bipartite or $d(G')\geqslant \Delta(G')^{1-1/100}$. We handle these in separate cases.

Case 1 *G' is bipartite.* We pass to an induced subgraph $H\subset G'$ with maximum average degree, say d. Then $d\geqslant C'k^2$ and H is d-degenerate with $\delta(H)\geqslant d/2$. Let $H=(A,B)$ where $|A'|\geqslant|B'|$. Let $A'=\{x\in A : d_H(x)\geqslant 4d\}$.

We have $|A'|\leqslant|A|/2$ (by counting edges), and by maximality of average degree, we see that $e(H[A',B])\leqslant(3/4)e(H)$.

Now, let $A^*=A\setminus A'$ and $F=G[A^*,B]$. We note $e(F)\geqslant e(H)/4$ and $|A^*|\geqslant|B|/2$. Furthermore, as $\delta(H)\geqslant d/2$, we have $d_F(a)=d_H(a)\in[d/2,4d]$ for all $a\in A^*$

Let W be a random subset of B, where each vertex is included independently in W with probability $p=1/8d$. Let $U=\{x\in A^* : |N_F(x)\cap W|=2\}$. For $x\in A^*$, we have

$$\mathbb{P}[x\in U]=p^2(1-p)^{d_F(x)-2}\binom{d_F(x)}{2}\geqslant p^2(1-p)^{4d}\binom{d/2}{2}\geqslant p^2d^2/20\geqslant 1/2000.$$

Then $\mathbb{E}[|U|-d|W|/1000]\geqslant|A^*|/2000-|B|/8000>0$ (as $|A^*|\geqslant|B|/2$) and hence there is a choice of W such that $|U|>d|W|/1000$. Fix such a W and define an auxiliary graph J with vertex set W and $E(J)=\{N(z) : z\in U\}$.

Note that $e(J) = |U|$, as F was C_4-free (and so the edges coming from different z are distinct). Thus, we have $d(J) = C''k^2$ for some $C'' \geqslant C'/1000$. Taking C (and thus C'') sufficiently large, it follows that J must contain a subdivision of a K_k by Theoren 7.1. By construction, this corresponds to an induced subdivision of K_k in G.

Case 2 $d(G') \geqslant \Delta(G')^{1-1/100}$. As in the first case, it is enough to find $U, W \subset V(G')$ such that:

- U and W are independent sets;
- $d_{G'}(u) = 2$ for each $u \in U$; and
- $|U| \geqslant C''k^2|W|$ for some appropriately large constant $C'' > 0$.

Theorem 7.1 then gives our induced subdivision of K_k, as desired.

We first pass to a subgraph $H \subset G'$ of maximal average degree. Then $d := d(H) \geqslant \max\{C'k^5, \Delta(H)^{1-1/100}\}$ and also $\delta(H) \geqslant d/2$. Let W_0 be a random subset of $V(H)$ where each vertex $v \in V(H)$ is included in W_0 with probability $p = 1/10d^{8/5}$. Let $W \subset W_0$ be the set $\{w \in W_0 : |N_H(w) \cap W_0| = 0\}$. Finally define $U_0 := \{u \in V(H) : |N_H(u) \cap W| = |N_H(u) \cap W_0| = 2\}$ and let $U \subset U_0 \setminus W_0$ be the set of $u \in U_0$ with $u \notin W_0$ and $|N_H(u) \cap U_0| = 0$.

So for $x \in V(G)$, a union bound tells us

$$\mathbb{P}(x \in U) \geqslant \mathbb{P}(x \notin W_0)\mathbb{P}(x \in U_0 | x \notin W_0)\left(1 - \sum_{y \in N_H(x)} \mathbb{P}(y \in U_0 | x \in U_0)\right),$$

where we have used the fact that if $x \in U_0$ then $x \notin W_0$.

Now $\mathbb{P}(x \notin W_0) = 1 - p \geqslant 1/2$, and

$$\mathbb{P}(x \in U_0 | x \notin W_0) = \sum_{\{y,y'\} \in \binom{N_H(x)}{2}} \mathbb{P}(W_0 \cap N_H(x) = W \cap N_H(x) = \{y, y'\}).$$

Now, for each $\{y, y'\} \in \binom{N_H(x)}{2}$, we have y, y' are not adjacent because $H \subset G'$ is C_3-free (and $x \in N_H(y) \cap N_H(y')$), thus the corresponding summand is at least

$$(1-p)^{d_H(x)-2}p^2(1 - \mathbb{E}[|W_0 \cap N_H(y)| + |W_0 \cap N_H(y')|]) \geqslant p^2(1 - p\Delta(H))(1 - 2p\Delta(H)) \geqslant p^2/2.$$

Consequently, $\mathbb{P}(x \in U_0) \geqslant \binom{d_H(x)}{2}p^2/2 \geqslant d^2p^2/10$ (since $\delta(H) \geqslant d/2$).

Finally, for $y \in N_H(x)$, we have $N_H(y) \cap N_H(x) = \varnothing$ because $H \subset G'$ is C_3-free, whence

$$\begin{aligned}\mathbb{P}(y \in U_0 | x \in U_0 \setminus W_0) &= \mathbb{P}(y \in U_0 | x \notin W_0) \\ &\leqslant p^2\binom{d_H(y) - 1}{2} \\ &\leqslant (p\Delta(H))^2.\end{aligned}$$

Putting these together, we get

$$\mathbb{P}(x \in U) \geqslant \frac{p^2d^2}{10}(1/2)(1 - \Delta(H)(p\Delta(H))^2) \geqslant \frac{p^2d^2}{40}$$

(where in the last step we use $2(1 - 1/100)8/5 > 3$ to deduce $p^2\Delta(H)^3 \leqslant 1/2$).

Consequently, $\mathbb{E}[|U|] \geqslant \frac{pd^2}{40}\mathbb{E}[|W_0|] \geqslant \frac{d^{2/5}}{40}\mathbb{E}[|W|]$. Thus, there is some choice of W_0 such that $|U| \geqslant C''k^2|W|$ (where $C'' = \Omega(C'^{2/5})$). Taking C (and thus C'') sufficiently large, we are done. □

8. Concluding remarks

- It is still very interesting to improve the bounds in the non-induced case. Let $f(k,6)$ denotes the least integer d such that if G is a graph with $d(G) \geqslant d$, then G contains a (not-necessarily induced) C_4-free subgraph $G' \subset G$ with $d(G) \geqslant k$.

  The best known bounds are

$$k^{3-o(1)} \leqslant f(k,6) \leqslant k^{O(k^2)},$$

  which were both established in [23]. It would be very interesting if one could show that a polynomial upper bound held (i.e., that $f(k,6) \leqslant k^{O(1)}$).
- We believe that for every k, there is a polynomial $p_k(s)$ so that every $K_{s,s}$-free graph G with $d(G) \geqslant p_k(s)$ contains an induced C_4-free subgraph with average degree at least k. We could not verify this conjecture, but if true this would show that every degree-bounded hereditary family of graphs is polynomially bounded. We note that recently Briański, Davies, and Walczak [3] showed that there are χ-bounded hereditary families of graphs whose chromatic number can grow arbitrarily fast compared with the clique number.
- An interesting difference between the two problems above is that, in the induced case, things are still difficult when G is almost-regular. Let $f'_{\text{Ind}}(s,k)$ be the smallest integer d_0 such that every $K_{s,s}$-free graph G so that $d(G)/2 \leqslant \delta(G) \leqslant \Delta(G) \leqslant 2d(G)$ contains an induced C_4-free subgraph $G' \subset G$ with $d(G') \geqslant k$.

  Applying Lemma 3.3, we see that $f'_{\text{Ind}}(s,k) \leqslant k^{Cs}$ for some absolute constant C. It would already be nice to get a polynomial bound here (i.e., prove $f'_{\text{Ind}}(s,k) = s^{O_k(1)}$).
- Finally, it would be very nice to better understand the behaviour of the function $F(\ell,k)$ introduced in Section 4. Already for $k=3$ and large ℓ, we only know

$$\ell^{2-o(1)} \leqslant F(\ell,3) \leqslant O(2^\ell),$$

  and improving either of these bounds would be quite interesting. We tentatively expect a polynomial upper bound when $\ell = O(1)$.

References

[1] B. Bollobás and A. Thomason. Proof of a conjecture of Mader, Erdős and Hajnal on topological complete subgraphs. *Europ. J. Comb.*, 19:883–887, 1998. 4, 15

[2] M. Bonamy, N. Bousquet, M. Pilipczuk, P. Rzążewski, S. Thomassé, and B. Walczak. Degeneracy of P_t-free and $C_{\geqslant t}$-free graphs with no large complete bipartite subgraphs. *J. Combin. Theory Ser. B*, 152:353–378, 2022. 3

[3] M. Briański, J. Davies, and B. Walczak. Separating polynomial χ-boundedness from χ-boundedness. *ArXiv:2201.08814v2*. 3, 17

[4] A. Carbonero, P. Hompe, B. Moore, and S. Spirkl. A counterexample to a conjecture about triangle-free induced subgraphs of graphs with large chromatic number. *Journal of Combinatorial Theory, Series B*, 158:63–69, 2023. 3

[5] M. Chudnovsky, A. Scott, and P. Seymour. Induced subgraphs of graphs with large chromatic number. III. Long holes. *Combinatorica*, 37:1057–1072, 2017. 3

[6] D. Dellamonica and V. Rödl. A note on Thomassen's conjecture. *J. Combin. Theory Ser. B*, 101:509–515, 2011. 1

[7] D. Dellamonica Jr, V. Koubek, D. M. Martin, and V. Rödl. On a conjecture of Thomassen concerning subgraphs of large girth. *Journal of Graph Theory*, 67(4):316–331, 2011. 9

[8] P. Erdős. On the combinatorial problems which I would most like to see solved. *Combinatorica*, 1:25–42, 1981. 1

[9] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Mathematica*, 2:463–470, 1935. 15

[10] L. Esperet. *Graph Colorings, Flows and Perfect Matchings*. Habilitation thesis, Université Grenoble Alpes, 2017. 3

[11] Z. Füredi. On finite set-systems whose every intersection is a kernel of a star. *Discrete Math.*, 47:129–132, 1983. 1, 9

[12] A. Girão, F. Illingworth, E. Powierski, M. Savery, A. Scott, Y. Tamitegama, and J. Tan. Induced subgraphs of induced subgraphs of large chromatic number. *ArXiv:2203.03612*. 3

[13] A. Gyárfás. Problems from the world surrounding perfect graphs. *Applicationes Mathematicae*, 19(3-4):413–441, 1987. 2

[14] A. Gyárfás, E. Szemeredi, and Z. Tuza. Induced subtrees in graphs of large chromatic number. *Discrete Mathematics*, 30(3):235–244, 1980. 3

[15] O. Janzer and B. Sudakov. Resolution of the Erdős-Sauer problem on regular subgraphs. *Forum of Mathematics, Pi*, 11:e19, 2023. 1, 4

[16] H. A. Kierstead and S. G. Penrice. Radius two trees specify χ-bounded classes. *J. Graph Theory*, 18:119–129, 1994. 3

[17] J. Komlós and E. Szemerédi. Topological cliques in graphs II. *Combin., Probab. and Comput.*, 5:79–90, 1996. 4, 15

[18] D. Kühn and D. Osthus. Every graph with sufficiently large average degree contains a C_4-free subgraph of large average degree. *Combinatorica*, 24:155–162, 2004. 1, 3

[19] D. Kühn and D. Osthus. Induced subdivisions in $K_{s,s}$-free graphs of large average degree. *Combinatorica*, 24:287–304, 2004. 15

[20] M. Kwan, S. Letzter, B. Sudakov, and T. Tran. Dense induced bipartite subgraphs in triangle-free graphs. *Combinatorica*, 40:283–305, 2020. 2

[21] T. Kóvari, V. Sós, and P. Turán. On a problem of K. Zarankiewicz. *Colloquium Mathematicae*, 3(1):50–57, 1954. 4, 13

[22] R. McCarty. Dense induced subgraphs of dense bipartite graphs. *Siam J. Discrete Math.*, 35(2):661–667, 2021. 2

[23] R. Montgomery, A. Pokrovskiy, and B. Sudakov. C_4-free subgraphs with large average degree. *Israel J. of Math.*, 246:55–71, 2021. 1, 2, 17

[24] L. Pyber, V. Rödl, and E. Szemerédi. Dense graphs without 3-regular subgraphs. *J. Combin. Theory Ser. B*, 63:41–54, 1995. 1

[25] I. Reiman. Über ein problem von K. Zarankiewicz. *Acta Mathematica Academiae Scientiarum Hungarica*, 9:269–273, 1953. 13

[26] A. Scott and P. Seymour. Induced subgraphs of graphs with large chromatic number. x. holes of specific residue. *Combinatorica*, 39(5):1105–1132, 2019. 3

[27] A. Scott and P. Seymour. A survey of χ-boundedness. *J. Graph Theory*, 95:473–504, 2020. 3

[28] A. Scott, P. Seymour, and S. Spirkl. Polynomial bounds for chromatic number. I. Excluding a biclique and an induced tree. *J. Graph Theory*, 102:458–471, 2023. 3

[29] C. Thomassen. Girth in graphs. *J. Combin. Theory Ser. B*, 35:129–141, 1983. 1

(Du) School of Mathematics, Georgia Institute of Technology, Skiles Building, Cherry Street, Atlanta, Georgia, USA.

Email address: xdu90@gatech.edu

(Girão) Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, UK.

Email address: girao@maths.ox.ac.uk

(Hunter) Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, UK.

Email address: zachary.hunter@exeter.ox.ac.uk

(McCarty) Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton, New Jersey, USA.

Email address: rm1850@princeton.edu

(Scott) Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, UK.

Email address: scott@maths.ox.ac.uk