Large cliques or co-cliques in hypergraphs with forbidden order-size pairs

Maria Axenovich* Domagoj Bradač† Lior Gishboliner‡ Dhruv Mubayi§
Lea Weber¶

Abstract

The well-known Erdős-Hajnal conjecture states that for any graph F, there exists $\epsilon > 0$ such that every n-vertex graph G that contains no induced copy of F has a homogeneous set of size at least n^ϵ. We consider a variant of the Erdős-Hajnal problem for hypergraphs where we forbid a family of hypergraphs described by their orders and sizes. For graphs, we observe that if we forbid induced subgraphs on m vertices and f edges for any positive m and $0 \leq f \leq \binom{m}{2}$, then we obtain large homogeneous sets. For triple systems, in the first nontrivial case $m = 4$, for every $S \subseteq \{0, 1, 2, 3, 4\}$, we give bounds on the minimum size of a homogeneous set in a triple system where the number of edges spanned by every four vertices is not in S. In most cases the bounds are essentially tight. We also determine, for all S, whether the growth rate is polynomial or polylogarithmic. Some open problems remain.

1 Introduction

For an integer $r \geq 2$, an r-*graph* or r-uniform hypergraph is a pair $H = (V, E)$, where $V = V(H)$ is the set of vertices and $E = E(H) \subseteq \binom{V}{r}$ is the set of edges. A 2-graph is simply a graph. A *homogeneous set* is a set of vertices that is either a clique or a coclique (independent set). For an r-graph H, let $h(H)$ be the size of a largest homogeneous set. Given r-graphs F, H, say that H is F-*free* if H contains no isomorphic copy of F as an induced subgraph. We say that an r-graph F has the *Erdős-Hajnal-property* or simply EH-*property* if there is a constant $\epsilon = \epsilon_F > 0$ such that every n-vertex F-free r-graph H satisfies $h(H) \geq n^\epsilon$. A conjecture of Erdős and Hajnal [13] states that any 2-graph has the EH-property. The conjecture remains open, see for example a survey by Chudnovsky [8], as well as [1, 5, 17], to name a few central results on the topic. When F is a fixed

*Karlsruhe Institute of Technology, Karlsruhe, Germany, maria.aksenovich@kit.edu. Research supported in part by the DFG grant FKZ AX 93/2-1.
†Department of Mathematics, ETH, Zürich, Switzerland, domagoj.bradac@math.ethz.ch. Research supported by SNSF grant 200021_196965.
‡Department of Mathematics, ETH, Zürich, Switzerland, lior.gishboliner@math.ethz.ch. Research supported by SNSF grant 200021_196965.
§University of Illinois at Chicago, Chicago, USA, mubayi@uic.edu. Research partially supported by NSF grants DMS-1763317, 1952767, 2153576, a Humboldt research award and a Simons fellowship.
¶Karlsruhe Institute of Technology, Karlsruhe, Germany, 0815lea@gmx.net.

graph and G is an F-free n-vertex graph, Erdős and Hajnal proved that $h(G) \geq 2^{c\sqrt{\log n}}$. This was recently improved to $h(G) \geq 2^{c\sqrt{\log n \log\log n}}$ by Bucić, Nguyen, Scott, and Seymour [7].

The Erdős-Hajnal conjecture fails for r-graphs, $r \geq 3$, already when F is a clique of size $r+1$. Indeed, well-known results on off-diagonal hypergraph Ramsey numbers show that there are n-vertex r-graphs that do not have a clique on $r+1$ vertices and do not have cocliques on $f_r(n)$ vertices, where f_r is an iterated logarithmic function (see [25] for the best known results). Moreover, the following result (Claim 1.3 in [19]) tells us exactly which r-graphs, $r \geq 3$, have the EH-property. Here D_2 is the unique 3-graph on 4 vertices with exactly 2 edges.

Theorem 1.1 (Gishboliner and Tomon [19]). *Let $r \geq 3$. If F is an r-graph on at least $r+1$ vertices and $F \neq D_2$, then there is an F-free r-graph H on n vertices such that $h(H) = O(\log n)$. On the other hand, there is a constant $c > 0$ such that if H is an D_2-free n-vertex 3-graph, then $h(H) \geq n^c$.*

It is natural to consider the EH-property for families of r-graphs instead of a single r-graph. In this paper, we consider families determined by a given set of orders and sizes. Several special cases of this have been extensively studied over the years (see, e.g. [12]). For $0 \leq f \leq \binom{m}{r}$, we call an r-graph F on m vertices and f edges an *(m,f)-graph* and we call the pair (m,f) the *order-size pair* for F. Say that H is (m,f)-free if it contains no induced copy of an (m,f)-graph. If $Q = \{(m_1,f_1),\ldots,(m_t,f_t)\}$, say that H is Q-free if H is (m_i,f_i)-free for all $i = 1,\ldots,t$.

Definition 1.2. *Given $r \geq 2$ and $Q = \{(m_1,f_1),\ldots,(m_t,f_t)\}$, let $h(n,Q) = h_r(n,Q)$ be the minimum of $h(H)$, taken over all n-vertex Q-free r-graphs H. Say that Q has the* EH*-property if there exists $\epsilon = \epsilon_Q > 0$ such that $h(n,Q) > n^\epsilon$.*

For example $h_3(n,\{(4,0),(4,2)\}) = k$ means that every n-vertex 3-graph in which any 4 vertices induce 1, 3, or 4 edges has a homogenous set of size k, and there is a 3-graph H as above with $h(H) = k$. We may omit the subscript r in the notation $h_r(n,Q)$ if it is obvious from context. When $Q = \{(m,f)\}$ we use the simpler notation $h(n,m,f)$ instead of $h(n,\{(m,f)\})$. Let us make two simple observations:

$$h_r(n,Q) \leq h_r(n,Q') \qquad \text{if} \qquad Q \subseteq Q', \tag{1}$$

$$h_r(n,Q) = h_r(n,\overline{Q}) \qquad \text{where} \qquad \overline{Q} = \left\{ \left(m, \binom{m}{r} - f\right) : (m,f) \in Q \right\}. \tag{2}$$

Our first result concerns 2-graphs, where we show that forbidding a single order-size pair already guarantees large homogeneous sets.

Proposition 1.3. *For any integers m,f with $m \geq 2$ and $0 \leq f \leq \binom{m}{2}$ there exists $c > 0$ such that $h_2(n,m,f) \geq c\, n^{1/(m-1)}$.*

Proposition 1.3 is proved in Section 2. It seems a challenging problem to give good upper bounds on $h_2(n,m,f)$. For example, determining $h_2(n,m,\binom{m}{2})$ is equivalent to determining off-diagonal Ramsey numbers.

Our second main result concerns the case $r = 3$ and $m = 4$. We shall consider sets Q consisting of pairs $(4,i)$ for $i \in \{0,1,2,3,4\}$. The cases where Q contains both $(4,0)$ and $(4,4)$ are trivial, because Ramsey's theorem guarantees that for sufficiently large n we cannot avoid both $(4,0)$ and $(4,4)$. In all remaining cases, the following theorem determines whether $h_3(n,Q)$ is polynomial or polylogarithmic in n.

Theorem 1.4. *Let $\emptyset \neq S \subseteq \{0,1,2,3,4\}$ and suppose that $\{0,4\} \not\subseteq S$. Set $Q = \{(4,i) : i \in S\}$.*

1. *If $S = \{0\}, \{1\}, \{0,1\}, \{1,3\}$ or $\overline{S} := \{4-i : i \in S\}$ is one of these four sets, then there are constants $c_1, c_2 > 0$ such that $\log^{c_1}(n) \leq h_3(n,Q) \leq \log^{c_2}(n)$.*
2. *In all other cases, there is a constant $c > 0$ such that $h_3(n,Q) \geq n^c$.*

We will prove Theorem 1.4 by considering separately each of the cases (up to complementation, see (2)). Some cases follow from known results, and these are surveyed in Section 1.1. Many cases are new results, and these are presented in Section 1.2.

1.1 Prior work

In this section we review the cases of Theorem 1.4 that follow from prior work. The problem of estimating $h(n,4,0)$ (or, equivalently, of $h(n,4,4)$) is equivalent to estimating the Ramsey number $R_3(4,t)$. Recall that $R_r(s,t)$ is the minimum n such that every n-vertex r-graph contains a clique of size s or an independent set of size t. It is known [9] that $2^{c_1 t \log t} \leq R_3(4,t) \leq 2^{c_2 t^2 \log t}$. This yields positive constants c_1 and c_2 such that

$$c_1 \left(\frac{\log n}{\log\log n}\right)^{1/2} < h_3(n,4,0) < c_2 \frac{\log n}{\log\log n}.$$

Similarly, the case $Q = \{(4,0),(4,1)\}$ is equivalent (due to complementation (2)) to estimating the minimum possible independence number of an n-vertex 3-graph where no 4 vertices span at least 3 edges. This is a well-studied problem in hypergraph Ramsey theory, and an old result of Erdős and Hajnal [12] gives the bound $h_3(n,\{(4,0),(4,1)\}) \geq c_1 \frac{\log n}{\log\log n}$ for some constant $c_1 > 0$. Recently, Fox and He [16] proved a corresponding upper bound, showing that

$$h_3(n,4,1) \leq h_3(n,\{(4,0),(4,1)\}) < c_2 \frac{\log n}{\log\log n} \tag{3}$$

for a suitable constant c_2. It is worth mentioning that the case $Q = \{(4,3),(4,4)\}$ (which is equivalent to $\{(4,0),(4,1)\}$) is the first instance of a (different) conjecture of Erdős and Hajnal [12] about the growth rate of generalized hypergraph Ramsey numbers that correspond to our setting of $h(n,Q)$, where $Q = \{(m,f),(m,f+1),\ldots,(m,\binom{m}{r})\}$. Recent results of Mubayi and Razborov [24] on this problem determine, for each $m > r \geq 4$, the minimum f such that $h_r(n,Q) < c\log^a n$ for some a and $Q = \{(m,f),\ldots,(m,\binom{m}{r})\}$. When $r = 3$, the minimum f was determined by Conlon, Fox, and Sudakov [9] for m being a power of 3 and for growing m, as well as some other values.

For the case $Q = \{(4,2)\}$, we have $h(n,4,2) \geq n^c$ for a suitable constant $c > 0$, by Theorem 1.1.

Finally, we discuss two known cases with $|Q| = 3$. If $Q = \{(4,0),(4,1),(4,2)\}$, then a $\overline{Q}$-free 3-graph is the same as a partial Steiner system (STS), and it is well known [14, 26, 6] that the minimum independence number of an n-vertex partial STS has order of magnitude $\sqrt{n\log n}$. Thus $h_3(n,Q)$ has order of magnitude $\sqrt{n\log n}$.

If $Q=\{(4,1),(4,2),(4,3)\}$ and $n\geq 4$, then it is a simple exercise to show that any Q-free 4-graph on at least four vertices is a clique or coclique and therefore $h_3(n,Q)=n$ for $n\geq 4$.

1.2 New results

In this section we state our new results for the cases not covered in Section 1.1. The results of this section and Section 1.1 immediately imply Theorem 2. Up to complementation, the missing cases correspond to the following sets Q of order-size pairs:

- $\{(4,1)\}$;
- $\{(4,0),(4,2)\},\ \{(4,0),(4,3)\},\ \{(4,1),(4,2)\},\ \{(4,1),(4,3)\}$; and
- $\{(4,0),(4,1),(4,3)\},\ \{(4,0),(4,2),(4,3)\}$.

For $|Q|=1,2$, we summarize our results in the following table. Here, c_1,c_2 always denote suitable positive constants. The table also indicates the section where each result is proved. Note that for $Q=\{(4,1)\}$, the lower bound is proved in Section 3.1 and the upper bound follows from (3).

Q	Lower bound	Upper bound	Appears in
$\{(4,0),(4,2)\}$	$c_1\sqrt{n}$	$c_2\sqrt{n\log n}$	Section 4
$\{(4,1),(4,2)\}$	$c_1 n^{1/3}\log^{1/3} n$	$c_2 n^{1/3}\log^{4/3} n$	Section 5
$\{(4,0),(4,3)\}$	$c_1 n$	$\lceil \frac{n}{3}\rceil+1$	Section 3.2
$\{(4,1),(4,3)\}$	$c_1\log n$	$c_2\log n$	Section 6
$\{(4,1)\}$	$c_1\left(\frac{\log n}{\log\log n}\right)^{1/2}$	$c_2\frac{\log n}{\log\log n}$	Section 3.1 and (3)

Table 1: Bounds for $h_3(n,Q)$

For the two remaining cases with $|Q|=3$, we obtain exact results:

Theorem 1.5. *Let $n\geq 4$. Then $h_3(n,\{(4,0),(4,2),(4,3)\})=n-1$ and*

$$h_3(n,\{(4,0),(4,1),(4,3)\})=\begin{cases}\frac{n}{2} & \textit{if } n\equiv 0 \textit{ (mod 6)}\\ \lceil\frac{n+1}{2}\rceil & \textit{if } n\not\equiv 0 \textit{ (mod 6)}.\end{cases}$$

Theorem 1.5 is proved in Section 7.

Notation: Throughout the paper, for a hypergraph H, let $\omega(H)$ and $\alpha(H)$ denote the size of a largest clique and independent set in H, respectively. Recall that $h(H) = \max\{\omega(H), \alpha(H)\}$. For a 3-graph H and one of its vertices v, we define the *link graph* of v to be the graph $L(v)$ whose vertex set is $V(H) \setminus \{v\}$ and whose edge set is $\{e \subseteq V(H) \setminus \{v\} : e \cup \{v\} \in E(H)\}$. Moreover, for $S \subseteq V(H) \setminus \{v\}$, we use $L_S(v)$ to denote the subgraph of $L(v)$ induced by S. A clique on s vertices is denoted K_s. When denoting edges in 3-graphs, we shall often omit parentheses and commas; for example, instead of writing $\{x, y, z\}$, we shall simply write xyz. A *star* is a hypergraph consisting of a set S and a vertex $v \notin S$ with edge-set $\{vxy : x, y \in S, x \neq y\}$. We will denote this star by (v, S). As usual, we write $f(n) = O(g(n))$ if there is a constant $C > 0$ such that $f(n) \leq Cg(n)$ for all n, and we write $f(n) = \Omega(g(n))$ to mean that $g(n) = O(f(n))$.

2 Graphs: Proof of Proposition 1.3

Proof of Proposition 1.3. We show that $c = 2/\sqrt{5}$ suffices. We shall use induction on m with basis $m = 2$. In this case $f \in \{0, 1\}$. Note that $h(n, 2, 0) = h(n, 2, 1) = n = n^1 = n^{1/(m-1)}$, since forbidden graphs are either a non-edge or an edge. Consider an (m, f)-free graph G on n vertices, $m \geq 3$, and assume that the statement of the proposition holds for smaller values of m. We can also assume that G is not a complete graph or an empty graph. Suppose first that G is an odd cycle or a complement of an odd cycle. Then $\alpha(G)$ or $\omega(G)$ is at least $\frac{n-1}{2}$, so it suffices to check that $\frac{n-1}{2} \geq cn^{1/2}$, as $n^{1/2} \geq n^{1/(m-1)}$. And indeed, by squaring, we get the inequality $(n-1)^2 \geq 4c^2 n = \frac{16}{5}n$, and after rearranging we get $n^2 - \frac{26n}{5} + 1 \geq 0$, which holds for every $n \geq 5$.

So from now on, suppose that G is neither an odd cycle nor the complement of an odd cycle. Let Δ and $\overline{\Delta}$ be the maximum degree of G and of the complement $\overline{G}$ of G, respectively. Using Brooks' theorem, the chromatic number of G and of $\overline{G}$ is at most Δ and $\overline{\Delta}$, respectively. Thus, $\alpha(G) \geq n/\Delta$ and $\omega(G) \geq n/\overline{\Delta}$. Therefore, we can assume that $\Delta \geq n^{(m-2)/(m-1)}$ and $\overline{\Delta} \geq n^{(m-2)/(m-1)}$, otherwise we are done. Thus, there is a vertex with at least $n^{(m-2)/(m-1)}$ edges incident to it and there is a vertex with at least $n^{(m-2)/(m-1)}$ non-edges incident to it.

Assume first that $f \leq m-2$. Then $f \leq \binom{m-1}{2}$. Take $v \in V(G)$ with at least $n^{(m-2)/(m-1)}$ non-edges incident to it, i.e., with a set X of vertices each non-adjacent to v, $|X| \geq n^{(m-2)/(m-1)}$. Since G is (m, f)-free, $G[X]$ is $(m-1, f)$-free. Thus, by induction, $h(G) \geq h(G[X]) \geq c|X|^{1/(m-2)} \geq cn^{1/(m-1)}$.

Now assume that $f \geq m-1$. Consider a vertex v with at least $n^{(m-2)/(m-1)}$ edges incident to it, i.e., with a set X of vertices each adjacent to v, $|X| \geq n^{(m-2)/(m-1)}$. Since G is (m, f)-free, $G[X]$ is $(m-1, f-(m-1))$-free. Thus, by induction, $h(G) \geq h(G[X]) \geq c|X|^{1/(m-2)} \geq cn^{1/(m-1)}$. ☐

3 Two Short Proofs

3.1 $Q = \{(4, 1)\}$

To prove the lower bound on $h(n, 4, 1)$ from Table 1, we shall consider the complementary setting and an arbitrary n-vertex $(4, 3)$-free 3-graph H. We need the following theorem of Fox and He [16].

Theorem 3.1 (Fox and He [16], Thm. 1.4)**.** *For all $t, s \geq 3$, any 3-graph on more than $(2t)^{st}$ vertices contains either a coclique on t vertices or a star (v, S) with $|S| = s$.*

Proposition 3.2. $h(n, 4, 3) \geq c \left(\frac{\log n}{\log\log n}\right)^{1/2}$ *for a constant $c > 0$.*

Proof. We shall apply Theorem 3.1 with the largest possible $t = s$ such that $(2t)^{st} < n$. In this case $t = s \geq c(\log n/ \log\log n)^{1/2}$. If H has a coclique of size t, then $h(H) \geq t$ and we are done. Otherwise H contains a star (v, S) with $|S| = s$. Note that S induces a clique in H, because otherwise v and three vertices of S not inducing an edge give a $(4, 3)$-subgraph. Thus, $h(H) \geq s$. In each case $h(H) \geq c(\log n/ \log\log n)^{1/2}$. □

3.2 $Q = \{(4, 0), (4, 3)\}$

Let us restate our result from Table 1:

Proposition 3.3. $\Omega(n) \leq h_3(n, \{(4, 0), (4, 3)\}) \leq \lceil \frac{n}{3} \rceil + 1$.

Proof. Let H be a $\{(4, 0), (4, 3)\}$-free 3-graph. We may assume that $e(H) = \Omega(n^3)$, else H is not $(4, 0)$-free. (Indeed, if $e(H) = o(n^3)$, then the probability that a random set of 4 vertices contains an edge is $o(1)$, so H contains a $(4, 0)$-subgraph.) Fix $v \in V(H)$ with $e(L(v)) = \Omega(n^2)$. Note that $L(v)$ is induced C_4-free. Indeed, if C is an induced C_4 in $L(v)$, then for each $A \subseteq V(C)$, $|A| = 3$, it holds that $A \notin E(H)$, because else $A \cup \{v\}$ spans exactly 3 edges. This means that $V(C)$ spans 0 edges, a contradiction. By a result of Gyárfás, Hubenko and Solymosi [21], an n-vertex graph with $\Omega(n^2)$ edges and no induced C_4 contains a clique of size $\Omega(n)$. So $L(v)$ contains a clique X of size $\Omega(n)$. For each $A \subseteq X$, $|A| = 3$, we have $A \in E(H)$ because else $A \cup \{v\}$ spans exactly 3 edges. So X is a clique in H, implying $\omega(H) = \Omega(n)$. This proves the lower bound in the proposition.

For the upper bound, let H be a 3-graph on n vertices with vertex set $A \cup B \cup C$, where A, B, and C are pairwise disjoint sets of almost equal sizes. Let $E(H) = \{abc : a \in A,\ b \in B,\ c \in C\} \cup \{abb' : a \in A,\ b, b' \in B\} \cup \{bcc' : b \in B,\ c, c' \in C\} \cup \{caa' : c \in C,\ a, a' \in A\}$. We see that H is $\{(4, 1), (4, 4)\}$-free, $\alpha(H) \leq \lceil n/3 \rceil + 1$, and $\omega(H) = 3$. Using complementation gives the required upper bound. □

4 $Q = \{(4, 0), (4, 2)\}$

It will be convenient to consider $Q = \{(4, 2), (4, 4)\}$ (which is equivalent to $\{(4, 0), (4, 2)\}$ via complementation). Let us restate our result from Table 1:

Theorem 4.1. $\Omega(\sqrt{n}) \leq h_3(n, \{(4, 2), (4, 4)\}) \leq O(\sqrt{n \log n})$.

To lower-bound $h_3(n, \{(4, 2), (4, 4)\})$, we prove the following characterization of $\{(4, 2), (4, 4)\}$-free 3-graphs. A *tight component* is a maximal (with respect to inclusion) set of edges C such that for any distinct $e_1, e_2 \in C$, there is a tight walk from e_1 to e_2, i.e. a sequence of edges $e_1 = f_1, \ldots, f_k = e_2$ with $|f_i \cap f_{i+1}| = 2$. We call a tight component a *star* if it is an edge-set of a star.

Theorem 4.2. *A 3-graph H is $\{(4,2),(4,4)\}$-free if and only if every tight component is a star.*

Proof. Suppose first that every tight component of H is a star. If H contains 4 vertices spanning exactly 2 or 4 edges, then the edges on these vertices are in the same tight component, but a star does not contain 4 vertices spanning exactly 2 or 4 edges, a contradiction. So H is $\{(4,2),(4,4)\}$-free.

We now prove the other direction. Let H be a $\{(4,2),(4,4)\}$-free 3-graph. Observe that for every star (v,S) in H, the set S is independent, because otherwise H would not be $(4,4)$-free.

Claim 4.3. *Let (v,S) be a star in H with $|S|\geq 3$. There is no edge in H of the form uxy with $u\notin\{v\}\cup S$ and $x,y\in S$.*

Proof. Suppose otherwise. The vertices $\{v,u,x,y\}$ must span exactly 3 edges, because $vxy, uxy\in E(H)$ but $\{v,u,x,y\}$ cannot span 2 or 4 edges. Without loss of generality, suppose that $vux\in E(H)$, $vuy\notin E(H)$. Let $z\in S\setminus\{x,y\}$. Suppose first that $vuz\in E(H)$. Then $uyz\in E(H)$ because otherwise $\{v,u,y,z\}$ spans 2 edges. This implies that $uxz\in E(H)$, because else $\{u,x,y,z\}$ spans 2 edges. Now $\{v,u,x,z\}$ spans 4 edges, contradiction. Similarly, suppose that $vuz\notin E(H)$. Then $uyz\notin E(H)$ because else $\{v,u,y,z\}$ spans 2 edges. This implies that $uxz\notin E(H)$, because else $\{u,x,y,z\}$ spans 2 edges. Now, $\{v,u,x,z\}$ spans 2 edges, contradiction. □

Now we complete the proof of the theorem. Let C be a tight component of H, and let us show that C is a star. If $|C|=1$ (i.e. C contains only one edge) then this is immediate, so suppose that C contains at least 2 edges. Let $e,f\in C$ with $|e\cap f|=2$. Write $e=uvx, f=uvy$. Then exactly one of the triples vxy, uxy is an edge, say $vxy\in E(H)$. So C contains the edges of the star $(v,\{u,x,y\})$. Let S be a maximal subset of $V(H)\setminus\{v\}$ such that C contains the edges of the star (v,S), so $|S|\geq 3$. We claim that C contains no other edges. Suppose otherwise. Recall that S induces no edges. So there must be an edge $e\in C$ which contains one vertex w outside $\{v\}\cup S$ and two vertices s,t in $\{v\}\cup S$. By Claim 4.3, it is impossible that $s,t\in S$. So suppose that $s=v, t\in S$. Fix an arbitrary $z\in S\setminus\{t\}$. We have $vzt\in E(H)$. Also, $vwt\in E(H)$ (because $s=v$). By Claim 4.3, $wzt\notin E(H)$, which implies that $vwz\in E(H)$ as otherwise $\{v,w,t,z\}$ spans exactly two edges. As this holds for every $z\in S$, we get that $(v,S\cup\{w\})$ is a star contained in C, contradicting the maximality of S. □

In what follows, for a tight component C that is a star, we denote by $V(C)$ the vertex set of the respective graph and $e(C)=|C|$, the number of edges in C.

Lemma 4.4. *Let C_1,C_2 be distinct tight components of a $\{(4,2),(4,4)\}$-free 3-graph. Then $|V(C_1)\cap V(C_2)|\leq 1$.*

Proof. Suppose by contradiction that there are distinct $x,y\in V(C_1)\cap V(C_2)$. Note that in a star, every pair of vertices is contained in some edge of the star. Let e_i be an edge of C_i containing x,y, $i=1,2$. Then there is a tight walk between every edge of C_1 and every edge of C_2 by using the connection e_1,e_2. It follows that C_1,C_2 are in the same tight component, a contradiction. □

Next, we prove a tight bound for the number of edges in a $\{(4,2),(4,4)\}$-free 3-graph. The extremal case is when H is a star.

Proposition 4.5. *For a $\{(4,2),(4,4)\}$-free n-vertex 3-graph H, it holds that $e(H) \leq \binom{n-1}{2}$.*

Proof. Let $C_1, \dots, C_m$ be the tight connected components of H. Each edge is contained in a unique C_i, and $e(C_i) = \binom{|V(C_i)|-1}{2}$ because C_i is a star. Therefore, $e(H) = \sum_{i=1}^{m} \binom{|V(C_i)|-1}{2}$. Also, $\sum_{i=1}^{m} \binom{|V(C_i)|}{2} \leq \binom{n}{2}$, because each pair of vertices is contained in at most one $V(C_i)$, by Lemma 4.4. Let f be the function $f(x) = x - \frac{1}{2}\sqrt{8x+1} + \frac{1}{2}$, so that $f(\binom{k}{2}) = \binom{k-1}{2}$. Put $x_i = \binom{|V(C_i)|}{2}$, so that $f(x_i) = \binom{|V(C_i)|-1}{2}$. We have $\sum_{i=1}^{m} x_i \leq \binom{n}{2}$. The function f is convex on $[0,\infty)$, so $\sum_{i=1}^{m} f(x_i)$ is maximized when exactly one of the x_i's, say x_1, is non-zero. As $x_1 \leq \binom{n}{2}$, we have $e(H) = \sum_{i=1}^{m} f(x_i) \leq f(\binom{n}{2}) = \binom{n-1}{2}$. $\square$

Proof of Theorem 4.1. The upper bound in the theorem follows from the fact that every linear 3-graph is $\{(4,2),(4,4)\}$-free (this follows e.g. from Theorem 4.2, because every tight component of a linear hypergraph has size 1), and the well-known result that there exist linear 3-graphs with independence number $O(\sqrt{n \log n})$ (which is tight), see [14, 26, 6].

The lower bound in the theorem follows from Proposition 4.5 and the known fact that every n-vertex 3-graph H has an independent set of size $\min\left\{\frac{n}{2}, \frac{cn^{3/2}}{e(H)^{1/2}}\right\}$. (To see this, take a random subset $X \subseteq V(H)$ by keeping each vertex with probability $p = \frac{cn^{1/2}}{e(H)^{1/2}}$, and delete one vertex from each edge inside X.) $\square$

5 $Q = \{(4,1),(4,2)\}$

Here we consider $Q = \{(4,1),(4,2)\}$. By complementation, we may equivalently consider $Q = \{(4,2),(4,3)\}$. Let us restate our result from Table 1.

Theorem 5.1. $\Omega(n^{1/3} \log^{1/3} n) \leq h_3(n, \{(4,2),(4,3)\}) \leq O(n^{1/3} \log^{4/3})$.

For the lower bound in Theorem 5.1, we need the following result of Kostochka, Mubayi, and Verstraëte [23] on independent sets in sparse hypergraphs.

Theorem 5.2 (Kostochka, Mubayi, and Verstraëte [23])**.** *Suppose that H is an n-vertex 3-graph in which every pair of vertices lies in at most d edges, where $0 < d < n/(\log n)^{27}$. Then H has an independent set of size at least $c\sqrt{(n/d)\log(n/d)}$ where c is an absolute constant.*

Proof of the lower bound in Theorem 5.1. Let H be an n-vertex $\{(4,2),(4,3)\}$-free 3-graph, where n is sufficiently large. Let u, v be a pair of vertices in H whose common neighborhood S has maximum size $d > 0$. Given vertices $x, y \in S$, the edges xyu and xyv are both in H, else $\{u,v,x,y\}$ induces a $(4,2)$- or $(4,3)$-graph. Next, any three vertices $x, y, z \in S$ must form an edge of H, otherwise $\{u,x,y,z\}$ induces a $(4,3)$-graph. Therefore S induces a clique in H of size d. If $d > n^{0.4}$, say, then we are done as $h(H) \geq d$. Recalling that n is large enough, we may assume that

$d \le n^{0.4} < n/(\log n)^{27}$. Now Theorem 5.2 yields a coclique in H of size at least $c\sqrt{(n/d)\log n}$ for some positive constant c. Consequently, there is a constant c' such that

$$h(H) \ge \max\{d,\, c\sqrt{(n/d)\log n}\} > c'\,(n\log n)^{1/3}.$$

Replacing c' by a possibly smaller constant c_1 yields the result for all $n > 4$. $\square$

In the rest of this section, we prove the upper bound in Theorem 5.1. We begin with the following two lemmas, giving a structural characterization of $\{(4,2),(4,3)\}$-free 3-graphs and rephrasing the problem of estimating $h_3(n,\{(4,2),(4,3)\})$ in terms of a certain extremal problem for (non-uniform) linear hypergraphs.

Lemma 5.3. *Let H be a $\{(4,2),(4,3)\}$-free 3-graph. Then every two maximal cliques in H intersect in at most one vertex.*

Proof. Let X, Y be maximal cliques and suppose that $|X \cap Y| \ge 2$. Fix $u, v \in X \cap Y$ and $y \in Y \setminus X$. Note that $uvy \in E(H)$. For every $x \in X \setminus \{u, v\}$, we have $uvx \in E(H)$, so we must have $uxy, vxy \in E(H)$, because else $\{u, v, x, y\}$ spans 2 or 3 edges. Next, for every $x_1, x_2 \in X \setminus \{u\}$, we have $ux_1y, ux_2y \in E(H)$, so we must also have $x_1x_2y \in E(H)$. It follows that $X \cup \{y\}$ is a clique, in contradiction to the maximality of X. $\square$

For a (not necessarily uniform) hypergraph $\mathcal{H}$, let $\alpha_2(\mathcal{H})$ be the maximum size of a set $I \subseteq V(\mathcal{H})$ such that $|I \cap e| \le 2$ for every $e \in E(\mathcal{H})$. Denote $g(\mathcal{H}) = \max\left(\max_{e\in E(\mathcal{H})}|e|, \alpha_2(\mathcal{H})\right)$. Denote by $g(n)$ the minimum of $g(\mathcal{H})$ over all linear (not necessarily uniform) hypergraphs with n vertices.

Lemma 5.4. $h_3(n,\{(4,2),(4,3)\}) = g(n)$.

Proof. Let H be an n-vertex Q-free 3-graph with $h(H) = h(n, Q)$, where $Q = \{(4,2),(4,3)\}$. Let $\mathcal{H}$ be the hypergraph on $V(H)$ whose edges are the maximal cliques of H. Then $\mathcal{H}$ is linear by the previous lemma. Also, $\max_{e\in E(\mathcal{H})}|e| = \omega(H)$, and $\alpha_2(\mathcal{H}) = \alpha(H)$, so $h(H) = g(\mathcal{H})$.

In the other direction, let $\mathcal{H}$ be an n-vertex linear hypergraph with $g(\mathcal{H}) = g(n)$. Let H be the 3-graph obtained by making each $e \in E(\mathcal{H})$ a clique. Then $h(H) = g(\mathcal{H})$, and it is easy to check that H is $\{(4,2),(4,3)\}$-free. $\square$

From now on, our goal is to upper-bound $g(n)$. As we will shortly show, the problem can be translated to a problem about C_4-free bipartite graphs. We prove the following.

Theorem 5.5. *For some positive constant C and every large m, there is a C_4-free bipartite graph $G = (X, Y, E)$ with $|X| \ge \frac{1}{2}m^{3/4}\log^2 m$ and $|Y| = (1 + o(1))m$, such that the following holds:*

1. *$d(y) \le 2m^{1/4}\log^2 m$ for every $y \in Y$.*
2. *For every set $X' \subseteq X$ of size at least $Cm^{1/4}\log^2 m$, there is $y \in Y$ with $|N(y) \cap X'| \ge 3$.*

Proof of the upper bound in Theorem 5.1. By Lemma 5.4, it is enough to show that $g(n) = O(n^{1/3}\log^{4/3} n)$. Let $G = (X, Y, E)$ be the graph given by Theorem 5.5. Put $n = |X| = \Omega(m^{3/4}\log^2 m)$. Let $\mathcal{H}$ be the hypergraph whose edges are the sets $N_G(y) \subseteq X$, $y \in Y$. Then $\mathcal{H}$ is linear because G is C_4-free. Also $\max_{e\in E(\mathcal{H})} |e| = O(m^{1/4}\log^2 m) = O(n^{1/3}\log^{4/3} n)$ by Item 1 of Theorem 5.5. Finally, $\alpha_2(\mathcal{H}) = O(m^{1/4}\log^2 m) = O(n^{1/3}\log^{4/3} n)$ by Item 2 of Theorem 5.5. $\square$

5.1 Proof of Theorem 5.5

Let H be the incidence graph of a finite projective plane with $n = (1 + o(1))m$ points and lines; that is, H is a bipartite C_4-free graph with sides X_0, Y of size n, and every pair of vertices in X_0 have exactly one common neighbour in Y. Let X be a random subset of X_0 obtained by including every vertex independently with probability $p = n^{-1/4}\log^2 n$. Let $G = H[X, Y]$. Clearly, with high probability $|X| \geq \frac{3}{4}pn \geq \frac{1}{2}pm \geq \frac{1}{2}m^{3/4}\log^2 m$. Also, we have $d_H(y) = (1+o(1))\sqrt{n}$ for every $y \in Y$, and it is easy to show, using the Chernoff bound, that w.h.p. $d(y) \leq 2\sqrt{n}p = 2n^{1/4}\log^2 n$ for every $y \in Y$. So it remains to show that w.h.p., G satisfies Item 2. To this end, we use the container method. Let $\mathcal{I}$ be the set of all subsets $I \subseteq X_0$ of size $Cn^{1/4}\log^2 n$ such that $|N(y) \cap I| \leq 2$ for every $y \in Y$. We want to show that with high probability X contains no set in $\mathcal{I}$. We will prove the following claim.

Claim 5.6. *There is a positive constant C_0, a set $\mathcal{S} \subseteq \binom{X_0}{C_0 n^{1/4}\log n}$ and a function $f\colon \mathcal{S} \to \binom{X_0}{\leq C_0\sqrt{n}}$ such that for every $I \in \mathcal{I}$, there exists $S = S(I) \in \mathcal{S}$ satisfying $S \subseteq I \subseteq f(S)$.*

Let us first complete the proof given Claim 5.6. Fix an arbitrary $S \in \mathcal{S}$. Note that $|X \cap f(S)|$ is distributed as $\mathrm{Bin}(|f(S)|, p)$. We have $\mathbb{P}[\mathrm{Bin}(N, p) \geq k] \leq \binom{N}{k}p^k \leq (\frac{eNp}{k})^k$. So for $k = Cn^{1/4}\log^2 n \geq \frac{C}{C_0}\cdot p|f(S)|$, we have (assuming $C \gg C_0$),

$$\mathbb{P}\left[|X \cap f(S)| \geq Cn^{1/4}\log^2 n\right] \leq \exp\left(-Cn^{1/4}\log^2 n\right). \tag{4}$$

Taking the union bound over all $S \in \mathcal{S}$, of which there are at most $\binom{n}{C_0 n^{1/4}\log n} \leq \exp(2C_0 n^{1/4}\log^2 n)$, it follows that with high probability, $|X \cap f(S)| < Cn^{1/4}\log^2 n$ holds for every $S \in \mathcal{S}$. Recall that for every $I \in \mathcal{I}$ there is $S \in \mathcal{S}$ such that $I \subseteq f(S(I))$. Hence, for every $I \in \mathcal{I}$, we have $|I \cap X| < Cn^{1/4}\log^2 n \leq |I|$, which implies $I \not\subseteq X$, as required.

Proof of Claim 5.6. We present an algorithm which, given I, produces sets $S(I) \subseteq I$ and $f(S) \supseteq I$. The algorithm maintains sets A^t, S^t. Initially, we set $A^0 = X_0, S^0 = \emptyset$. The algorithm runs for $q = C_0 n^{1/4}\log n$ steps $t = 0, \dots, q-1$ and in step t, obtains an index i^t, to be defined later, and new sets A^{t+1}, S^{t+1}. Recall that for any $I \in \mathcal{I}$, we have $|I| = Cn^{1/4}\log^2 n > q$. Throughout the algorithm we will have $|S^t| = t, S^t \subseteq I \subseteq S^t \cup A^t$ and $S^t \cap A^t = \emptyset$. Now, suppose we are at step t. We define a graph F^t with $V(F^t) = A^t$ and where $aa' \in E(F^t)$ if and only if there exist $s \in S^t$ and $y \in Y$ such that $a, a', s \in N(y)$. Note that F^t only depends on A^t, S^t, but not on I. Let $a^t_1, a^t_2, \dots, a^t_{|A_t|}$ be an ordering of A_t such that for all i, a^t_i is a vertex of maximum degree in $F^t[\{a^t_i, a^t_{i+1}, \dots, a^t_{|A_t|}\}]$, with ties broken according to some fixed ordering of X_0. Let i^t be the minimum index i such that $a^t_i \in I$. We let $S^{t+1} = S^t \cup \{a^t_{i^t}\}$ and $A^{t+1} = A^t \setminus (\{a^t_1, \dots, a^t_{i^t}\} \cup N_{F^t}(a_{i^t}))$. Note that i^t is well-defined since we have $|S^t| < q < |I|$ and $I \subseteq S^t \cup A^t$ (which we will soon prove). After q steps, we let $S(I) = S_q$ and $f(S_q) = S_q \cup A_q$. We denote $\mathcal{S} = \{S(I) \mid I \in \mathcal{I}\}$.

Clearly, we have $S^t \subseteq I, S^t \cap A^t = \emptyset$ for any $t \in \{0, \dots q\}$ and $S^{t+1} \cup A^{t+1} \subseteq S^t \cup A^t$ for any $t \in \{0, \dots, q-1\}$. Let us also verify that $I \subseteq S^t \cup A^t$ throughout, which clearly implies that $I \subseteq f(S(I))$. Indeed, suppose that $I \subseteq S^t \cup A^t$ at some step of the algorithm, let $a_1^t, \dots, a_{|A_t|}^t$ be the ordering of A^t as described in the algorithm, and let $i = i^t$ be the index chosen in the algorithm, i.e. such that $a_i^t \in I$ and $a_1^t, \dots, a_{i-1}^t \notin I$. Consider a neighbour v of a_i^t in F^t. By definition, there exist $s \in S^t \subseteq I$ and $y \in Y$ such that $s, a_i^t, v \in N(y)$. Then, since $I \in \mathcal{I}$ and $s, a_i^t \in I$, it follows that $v \notin I$. Hence, $I \subseteq A^{t+1} \cup S^{t+1}$.

Let us now prove that $f(S)$ is indeed uniquely determined by S. In the following, we will denote by $S^t(I), A^t(I)$ the relevant S^t, A^t when the input of the algorithm is I, and similarly denote by $i^t(I)$ the relevant index i^t. Fix $I, I' \in \mathcal{I}$ such that $S(I) = S(I')$. We show that $S^t(I) = S^t(I')$ and $A^t(I) = A^t(I')$ for all $t \in \{0, \dots, q\}$. This clearly holds for $t = 0$. Suppose that this holds for some t, and let us prove this for $t+1$. Denote $S^t = S^t(I) = S^t(I'), A^t = A^t(I) = A^t(I')$ and $F^t = F^t(I) = F^t(I')$, where the last equality holds since $F^t(J)$ is uniquely determined by $S^t(J)$ and $A^t(J)$. Let $a_1^t, \dots, a_{|A^t|}^t$ be the ordering of $A^t = V(F^t)$ as above. Denote $i = i^t(I)$ and $i' = i^t(I')$. If $i = i'$, then it follows that $S^{t+1}(I) = S^{t+1}(I')$ and from the definition of the algorithm, also $A^{t+1}(I) = A^{t+1}(I')$, as required. So let us assume without loss of generality that $i < i'$. Then, $a_i^t \in S^{t+1}(I) \subseteq S(I)$. On the other hand, by definition of i', we have $a_i^t \notin I'$ which, using that $S(I') \subseteq I'$, implies $a_i^t \notin S(I')$. Hence, $S(I) \neq S(I')$, contradicting our assumption.

Finally, we need to show that $|f(S)| \leq C_0\sqrt{n}$ for every $S \in \mathcal{S}$. We will prove the following claim.

Claim 5.7. *Suppose that $t \geq 2n^{1/4}$ and $|A^t| \geq 10\sqrt{n}$. Then $|A^{t+1}| \leq (1 - n^{-1/4})|A^t|$.*

Let us finish the proof given Claim 5.7. Fix any $I \in \mathcal{I}$ and $S = S(I)$, and suppose for the sake of contradiction that $|f(S)| = |S \cup A^q(I)| \geq 11\sqrt{n}$. As $|S| = q \ll \sqrt{n}$, we must have $|A^q(I)| \geq 10\sqrt{n}$. Then, by Claim 5.7, for any $t \in [2n^{1/4}, q-1]$, we have $|A^{t+1}| \leq (1 - n^{-1/4})|A^t|$, which implies

$$|A^q| \leq n \cdot \left(1 - n^{-1/4}\right)^{q - 2n^{1/4}} < n \cdot e^{-n^{-1/4} \cdot (q/2)} \leq n \cdot e^{-\log n} = 1,$$

a contradiction.

Proof of Claim 5.7. Let $S^t, A^t, F^t, a_1^t, \dots, a_{|A^t|}^t$ and i^t be as given in the algorithm. For $1 \leq j \leq |A^t|$, denote $F_j = F^t[\{a_j^t, \dots, a_{|A^t|}^t\}]$. It is enough to prove that $\Delta(F_j) \geq |V(F_j)|/n^{1/4}$ for every $j \leq |A^t|/2$. Indeed, then if $i^t \leq |A^t|/2$, we obtain $|A^{t+1}| \leq |A^t| - i^t - \Delta(F_{i^t}) \leq |A^t| - i^t - |V(F_{i^t})|/n^{1/4} = |A^t| - i^t - (|A^t| - i^t + 1)/n^{1/4} \leq (1 - n^{-1/4})|A^t|$, and if $i^t \geq |A^t|/2$, then $|A^{t+1}| \leq |A^t|/2$.

Consider a fixed $1 \leq j \leq |A^t|/2$. Denote $A' = \{a_j, \dots, a_{|A^t|}\} = V(F_j)$. We need to show that $\Delta(F_j) \geq |A'|/n^{1/4}$. Fix any $s \in S^t$. Then, for every $y \in N_H(s)$ and distinct $a, a' \in A' \cap N_H(y)$, we have $aa' \in E(F_j)$. Note that the sets $(N_H(y) \setminus \{s\})_{y \in N_H(S^t)}$ partition $X_0 \setminus \{s\}$, since every two vertices in X_0 have exactly one common neighbour in H. The number of pairs $(y, \{a, a'\})$ with $a, a' \in A'$ and $a, a', s \in N_H(y)$ is

$$\sum_{y \in N_H(s)} \binom{|A' \cap N_H(y)|}{2} \geq |N_H(s)| \cdot \binom{|A'|/|N_H(s)|}{2} \geq \frac{|A'|^2}{4\sqrt{n}},$$

where we used Jensen's inequality for the convex function $\binom{x}{2}$, the fact that $N_H(s) = (1+o(1))\sqrt{n}$, and the assumption that $|A'| \geq |A^t|/2 \geq 5\sqrt{n}$. Hence, every $s \in S^t$ contributes at least $\frac{|A'|^2}{4\sqrt{n}}$ edges to F_j. Finally, we prove that for every $aa' \in E(F_j)$, there are unique $s \in S^t, y \in Y$ such that $s, a, a' \in N_H(y)$. Indeed, recall that every pair of vertices in X_0 have a unique common neighbour in Y. Hence, given a, a', the vertex $y \in Y$ is uniquely determined. But then, the vertex $s \in S^t \cap N_H(y)$ is also uniquely determined. Indeed, suppose there are two distinct $s, s' \in S^t \cap N_H(y)$. Without loss of generality, there is an index t_0 such that $\{s\} = S^{t_0} \setminus S^{t_0-1}$ and $s' \in S^{t_0-1}$. Then, by definition, $sa \in E(F^{t_0-1})$, so $a \notin S^{t_0} \cup A^{t_0} \supseteq A^t$, a contradiction.

Therefore, we have $e(F_j) \geq |S^t| \cdot \frac{|A'|^2}{4\sqrt{n}} \geq \frac{|A'|^2}{2n^{1/4}}$, which implies that $\Delta(F_j) \geq |A'|/n^{1/4}$ as required. □

This conclude s the proof of Claim 5.6 and hence the theorem. □

6 $Q = \{(4,1),(4,3)\}$

Here we prove that $h_3(n, \{(4,1),(4,3)\}) = \Theta(\log n)$. We note that $\{(4,1),(4,3)\}$-free 3-graphs are also known as *two-graphs* (not to be confused with 2-graphs, which are just graphs), and have been thoroughly studied in algebraic combinatorics due to their connection to sets of equiangular lines, see e.g. [22, Chapter 11]. Every two-graph H arises from some graph G by taking $x, y, z \in V(G)$ to be an edge of H if and only if $\{x, y, z\}$ induces an odd number of edges in G. This will be used in the proof. We start with the following lemma.

Lemma 6.1. *There is a constant $C > 0$ such that for every n, there is an n-vertex graph in which every set of size $C \log n$ contains a triangle and a coclique of size* 3.

Proof. Take $G \sim G(n, 1/2)$. Fix any $U \subseteq V(G)$, $|U| = k := C \log n$. It is well-known that there is a partial Steiner system on U with $m = (\frac{1}{6} - o(1))k^2 \geq k^2/7$ triples, $T_1, \dots, T_m$. The probability that no T_i is a triangle in G is $(7/8)^m \leq (7/8)^{k^2/7} = (7/8)^{\frac{1}{7}C^2 \log^2 n}$. Taking the union bound over all $\binom{n}{C \log n} \leq e^{C \log^2 n}$ choices for U, and assuming that C is large enough, we get that with high probability, every set of size $C \log n$ contains a triangle. By the same argument, w.h.p. every such set contains a coclique of size 3. □

Theorem 6.2. $h_3(n, \{(4,1),(4,3)\}) = \Theta(\log n)$.

Proof. For the lower bound, let H be an n-vertex $\{(4,1),(4,3)\}$-free 3-graph. Pick a vertex v in H and consider its link graph $L(v)$. Since $R_2(t,t) < 4^{t-1}$ (see Erdős and Szekeres [15]), we see that $L(v)$ has a clique or coclique K of size at least $\frac{1}{2} \log n$. In the first case, K is a clique in H, else we find a $(4,3)$-subgraph in H; and in the second case, K is a coclique in H, else we find a $(4,1)$-subgraph in H.

For the upper bound, let G be the graph from Lemma 6.1. Let H be the 3-graph on vertex set $V(G)$ whose edge set consists of all triples of vertices x, y, z which induce an odd number of edges in G. Lemma 6.1 guarantees that every set of $C \log n$ vertices contains both an edge and a non-edge of H. Hence, $h(H) \leq C \log n$. Let us show that H is Q-free, $Q = \{(4,1),(4,3)\}$. Fix any

$X \subseteq V(G) = V(H)$, $|X| = 4$. For each $A \subseteq X$, $|A| = 3$, we have $A \in E(H)$ if and only if $e_G(A)$ is odd, where $e_G(A)$ is the number of edges spanned by A in G. Note that each edge of $G[X]$ is contained in exactly two sets $A \subseteq X$, $|A| = 3$. Hence, $\sum_{A\subseteq X,|A|=3} e_G(A) = 2e_G(X)$. The right-hand side is even, so there is an even number of A with $e_G(A)$ odd. It follows that every four vertices in H induce an even number of edges. So H is Q-free. □

7 Forbidden sets of size 3: Proof of Theorem 1.5

We will need the following structural characterization of Q-free 3-graphs for $Q = \{(4, 1), (4, 3), (4, 4)\}$.

Theorem 7.1 (Frankl and Füredi [18]). *Let H be an $\{(4, 1), (4, 3), (4, 4)\})$-free 3-graph. Then H is isomorphic to one of the following 3-graphs:*

1. *A blow-up of the 6 vertex 3-graph H' with vertex set $V(H') = [6]$ and edge set $E(H') = \{123, 124, 345, 346, 561, 562, 135, 146, 236, 245\}$. Here for the blow-up we replace every vertex of H' by an independent set, and whenever we have 3 vertices from three distinct of those sets, they induce an edge if and only if the corresponding vertices in H' do.*
2. *The 3-graph whose vertices are the points of a regular n-gon where 3 vertices span an edge if and only if the corresponding points span a triangle whose interior contains the center of the n-gon.*

Proof of Theorem 1.5.

Case $Q = \{(4, 1), (4, 3), (4, 4)\}$.
We are to prove that

$$h(n, \{(4, 0), (4, 1), (4, 3)\}) = h(n, Q) = \begin{cases} \frac{n}{2} & \text{if } n \equiv 0 \pmod 6 \\ \lceil \frac{n+1}{2} \rceil & \text{if } n \not\equiv 0 \pmod 6. \end{cases}$$

First, let us prove that the second 3-graph H in Theorem 7.1 has independence number exactly $\lceil (n + 1)/2) \rceil$. Assume the vertex set is $[n]$ and the vertices are labeled by consecutive integers in clockwise orientation. The lower bound is by taking $\lceil (n + 1)/2) \rceil$ consecutive vertices on the n-gon and noting that no three of them contain the center in their interior. For the upper bound, let us see how many vertices can lie in an independent set containing 1. When n is odd, the triangle formed by $\{1, i, (n-1)/2+i\}$ contains the center and hence is an edge. Therefore we may pair the elements of $[n] \setminus \{1\}$ as $(2, (n + 3)/2), (3, (n + 5)/2), \ldots, ((n + 1)/2, n)$ and note that each pair can have at most one vertex in an independent set containing 1. Hence the maximum size of an independent set containing 1 is at most $(n+1)/2$ and by vertex transitivity of H, the independence number of H is at most $(n+1)/2$. For n even we consider the $n/2-1$ pairs $(2, n/2+1), (3, n/2+2), \ldots, (n/2, n-1)$ and add the vertex n to get an upper bound $n/2 + 1 = \lceil (n + 1)/2) \rceil$.

Next we observe that the 6-vertex 3-graph H' in Theorem 7.1 has independence number exactly 3 (we omit the short case analysis needed for the proof). Hence if we blow-up each vertex of H' into sets of the same size, then we obtain n-vertex 3-graphs with independence number exactly $n/2$ whenever $n \equiv 0 \pmod 6$. This concludes the proof of the upper bound.

For the lower bound, let H be Q-free. Then by Theorem 7.1, H is isomorphic to one of the two graphs described in Theorem 7.1. If H is isomorphic to the second graph, then we have already shown that its independence number is at least $(n+1)/2$, so assume that H is isomorphic to the blow-up of the 6-vertex 10-edge 3-graph H'. There are 10 non-edges in H'. Let $V_1, \ldots, V_6$ be the blown up vertex sets. Since every vertex $i \in [6]$ in H' is contained in exactly 5 non-edges, we obtain

$$5n = 5 \sum_{i \in [6]} |V_i| = \sum_{j_1 j_2 j_3 \notin E(H)} |V_{j_1}| + |V_{j_2}| + |V_{j_3}|.$$

By the pigeonhole principle, there is a non-edge $i_1 i_2 i_3$, such that $|V_{i_1}| + |V_{i_2}| + |V_{i_3}| \geq n/2$. Our bound follows by observing that for any non-edge $i_1 i_2 i_3$ in the original 3-graph H' the set $V_{i_1} \cup V_{i_2} \cup V_{i_3}$ is an independent set. This gives an independent set of size at least $n/2$, and if $n \not\equiv 0 \pmod 6$, then equality cannot hold throughout (a short case analysis, which we omit, is needed to prove this) and we obtain an independent set of size strictly greater than $n/2$ as required.

Case $Q = \{(4,0), (4,2), (4,3)\}$.
We now prove $h(n, \{(4,0), (4,2), (4,3)\}) = n - 1$, for $n \geq 4$. Let H be a 3-graph that is a clique on $n-1$ vertices and a single isolated vertex, then H is Q-free, giving us the upper bound.

For the lower bound, let H be a Q-free 3-graph on n vertices, $n \geq 4$. Assume that H is not a clique. We shall show that H is a clique and a single isolated vertex. Consider a maximal clique S in H. Since $|S| < n$, there is a vertex $v \in V(H) \setminus S$. From the maximality of S, $L_S(v)$ is not a clique. If $L_S(v)$ contains an edge, then we have that for some vertices x, y, y', $xy \in E(L_S(v))$ and $xy' \notin E(L_S(v))$. But then $\{v, x, y, y'\}$ induces a $(4,2)$ or a $(4,3)$-graph, a contradiction. Thus, $L_S(v)$ is an empty graph, i.e., there is no edge in H containing v and two vertices of S. Now assume there exists a second vertex $v' \in V(H) \setminus (S \cup \{v\})$. Then by the same argument as above, v' is also not contained in any edge with two vertices from S. Consider triples $vv'x$, $x \in S$. Since $|S| \geq 3$, by the pigeonhole principle there are two vertices $x, x' \in S$ such that either $vv'x, vv'x' \in E(H)$ or $vv'x, vv'x' \notin E(H)$. Then $\{v, v', x, x'\}$ induces 2 or 0 edges respectively, a contradiction. Thus, $|S| = n - 1$ and v is an isolated vertex. □

8 Concluding Remarks

- In Section 3.1 we showed that $h_3(n,4,1) \geq c_1 (\frac{\log n}{\log \log n})^{1/2}$, and from (3) we have $h_3(n,4,1) \leq c_2 \frac{\log n}{\log \log n}$ (for constants c_1, c_2). It is unclear if either of these bounds represents the correct order of magnitude, but the lower bound certainly seems far off.

  Problem 8.1. *Improve the exponent* $1/2$ *in the lower bound on* $h_3(n,4,1)$.

- In the cases $Q = \{(4,0), (4,2)\}$ and $Q = \{(4,1), (4,2)\}$, there is a polylogarithmic gap between our lower and upper bounds in Table 1, and it would be interesting to close the gap. In particular, it would be interesting to decide whether $h(n, \{(4,2), (4,4)\}) = \Theta(\sqrt{n \log n})$ (this is equivalent to $Q = \{(4,0), (4,2)\}$ by complementation). Recall that in a $\{(4,2), (4,4)\}$-free 3-graph, every tight component is a star (Theorem 4.2). One example of such 3-graphs is linear 3-graphs, and it is well-known that every n-vertex linear 3-graph has an independent set of size $\Omega(\sqrt{n \log n})$, and that this is tight. Another example is to take a projective plane and put a star on each line (so that each star has roughly $\sqrt{n}$ vertices). It would be interesting to estimate the smallest possible independence number of such a hypergraph.

- Fix integers $m > r$. Recall that a set Q of order-size pairs $\{(m, f_1), \ldots, (m, f_t)\}$ has the Erdős-Hajnal (EH) property if there exists $\epsilon = \epsilon_Q$ such that $h_r(n, Q) > n^\epsilon$. As $|Q|$ grows, the collection of Q-free r-graphs is more restrictive, and hence $h_r(n, Q)$ grows (assuming that large Q-free r-graphs are not forbidden to exist by Ramsey's theorem). The case when $h_r(n, Q) = \Omega(n)$ was treated by the first author and Balogh [3] when $r = 2$. A natural question then is to ask what is the smallest t such that every Q of size t has the EH property. Call this minimum value $EH_r(m)$. Our results for $r = 3$ show that for $m = 4$, all Q of size 3 have the EH property, but there are Q of size 2 which do not. Consequently, $EH_3(4) = 3$.

  In order to further study $EH_r(m)$, we need another definition. Given integers $m \geq r \geq 3$, let $g_r(m)$ be the number of edges in an r-graph on m vertices obtained by first taking a partition of the m vertices into almost equal parts, then taking all edges that intersect each part, and then recursing this construction within each part. For example, $g_3(7) = 13$ since we start with a complete 3-partite 3-graph with part sizes $2, 2, 3$ and then add one edge within the part of size 3. It is known (see, e.g. [24]) that as r grows we have

$$g_r(m) = (1 + o(1)) \frac{r!}{r^r - r} \binom{m}{r}.$$

  Note that $\frac{r!}{r^r - r}$ approaches 0 as r grows. The second author and Razborov [24] proved that for all fixed $m > r > 3$, there are n-vertex r-graphs which are Q-free, $Q = \{(m, i) : g_r(m) < i \leq \binom{m}{r}\}$, with $h(G) = O(\log n)$. In other words, there exists Q of size $\binom{m}{r} - g_r(m)$ which does not have the EH property. This proves that $EH_r(m) \geq \binom{m}{r} - g_r(m) + 1$.

  Erdős and Hajnal [12] proved that for all $m > r \geq 3$, the set $Q = \{(m, i) : g_r(m) \leq i \leq \binom{m}{r}\}$ has the EH property. In other words, they proved that every n-vertex r-graph in which every set of m vertices spans less than $g_r(m)$ edges has an independent set of size at least n^ϵ, where ϵ depends only on r and m. This is a particular set Q of size $\binom{m}{r} - g_r(m) + 1$ that has the EH property, and we speculate that every other set Q of this size also has the EH property.

  Problem 8.2. *Prove or disprove that for all $m > r > 2$,*

$$EH_r(m) = \binom{m}{r} - g_r(m) + 1.$$

  We end by noting that $EH_3(4) = 3 = \binom{4}{3} - g_3(4) + 1$.

Acknowledgments: The authors wish to thank the anonymous referee for a careful reading of the paper and useful suggestions which improved the presentation. The second and third authors also thank Benny Sudakov for useful discussions.

References

[1] N. Alon, J. Pach, and J. Solymosi. *Ramsey-type theorems with forbidden subgraphs,* Paul Erdős and his mathematics, Combinatorica 21 (2001), no. 2, 155–170.

[2] N. Alon, and J. H. Spencer. **The probabilistic method**. John Wiley & Sons, 2016.

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

[4] R. C. Baker, G. Harman, and J. Pintz. *The difference between consecutive primes, II.* Proceedings of the London Mathematical Society 83 (2001) (3): 532–562.

[5] N. Bousquet, A. Lagoutte, and S. Thomassé. *The Erdős-Hajnal conjecture for paths and antipaths*, Journal of Combinatorial Theory, Series B 113 (2015) 261–264.

[6] M. de Brandes, K. Phelps, and V. Rödl. *Coloring Steiner triple systems,* SIAM Journal on Algebraic Discrete Methods 3.2 (1982) 241–249.

[7] M. Bucić, T. Nguyen, A. Scott, and P. Seymour. *A loglog step towards Erdős-Hajnal,* arXiv preprint arXiv:2301.10147 (2023).

[8] M. Chudnovsky. *The Erdős-Hajnal conjecture – a survey*, Journal of Graph Theory 75 (2014) 178–190.

[9] D. Conlon, J. Fox, and B. Sudakov. *Hypergraph Ramsey numbers,* J. Amer. Math. Soc. 23 (2010), no. 1, 247–266.

[10] D. de Caen. *Extension of a theorem of Moon and Moser on complete subgraphs,* Ars Combin. 16 (1983), 5–10.

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

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

[13] P. Erdős and A. Hajnal. *Ramsey-type theorems*, Discrete Applied Mathematics 25 (1989) 37–52.

[14] P. Erdős, A. Hajnal, and B. Rothschild. *On chromatic number of graphs and set-systems*, Acta Math. Hungar 16 (1966) 61–99.

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

[16] J. Fox and X. He. *Independent sets in hypergraphs with a forbidden link,* Proceedings of the London Mathematical Society. 123(4), (2021) 384–409.

[17] J. Fox, J. Pach, and A. Suk. *Erdős-Hajnal conjecture for graphs with bounded VC-dimension,* 33rd International Symposium on Computational Geometry, Art. No. 43, 15 pp., LIPIcs. Leibniz Int. Proc. Inform., 77, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern (2017).

[18] P. Frankl and Z. Füredi. *An exact result for 3-graphs*, Discrete Math. 50 no. 2-3, (1984) 323–328.

[19] L. Gishboliner and I. Tomon. *On 3-graphs with no four vertices spanning exactly two edges*, Bulletin of the London Mathematical Society, 54(6) (2022), 2117–2134.

[20] A. Gyárfás. *Reflections on a problem of Erdős and Hajnal,* The Mathematics of Paul Erdős II. Springer, New York, NY, (2013) 135–141.

[21] A. Gyárfás, A. Hubenko, and J. Solymosi. Large cliques in C_4-free graphs. Combinatorica, 22(2) (2002), 269–274.

[22] C. Godsil, and G. F. Royle. **Algebraic graph theory**. Springer Science & Business Media, 2001.

[23] A. Kostochka, D. Mubayi, and J. Verstraëte. *On independent sets in hypergraphs*, Random Structures and Algorithms 44 (2014), no. 2, 224–239.

[24] D. Mubayi and A. Razborov. *Polynomial to exponential transition in Ramsey theory,* Proceedings of the London Mathematical Society 122.1 (2021) 69–92.

[25] D. Mubayi and A. Suk. *New lower bounds for hypergraph Ramsey numbers,* Bull. Lond. Math. Soc. 50 (2018), no. 2, 189–201.

[26] K. Phelps and V. Rödl. *Steiner triple systems with minimum independence number,* Ars Combin 21 (1986) 167–172.