On the maximum density of r-graphs in which every $(r + 1)$-set spans 0 or 2 edges

Vishesh Jain* Haoran Luo* Dhruv Mubayi*

Abstract

In 1984, Frankl and Füredi asked for the maximum density of an n-vertex r-graph in which every $(r+1)$-set of vertices spans 0 or 2 edges. They gave a construction with asymptotic density 2^{1-r}. We significantly improve this bound by constructing such r-graphs with density $\Omega(r^{-3})$, thereby improving the dependence on r from exponential to polynomial. We also obtain lower bounds for the more general problem in which every $(r + 1)$-set spans an even number of edges from $\{0, 2, \ldots, 2k\}$.

1 Introduction

An n-vertex graph in which every set of 3 vertices spans 0 or 2 edges must be complete bipartite, and hence the maximum number of edges in such a graph is $\lfloor n^2/4 \rfloor$. We consider a generalization of this extremal problem to r-uniform hypergraphs (henceforth abbreviated as *r-graphs*) posed by Frankl and Füredi [7] in 1984.

Problem 1 (Problem 1 in [7])**.** What is the maximum number of edges in an n-vertex r-graph in which every set of $r + 1$ vertices spans 0 or 2 edges?

In the same paper, they considered the case $r = 3$ and gave a complete description of such 3-graphs: each such 3-graph is either a blow-up of a fixed 3-graph on 6 vertices with 10 edges or is obtained by taking vertices as points on the unit circle and letting the edges be the triples whose convex hull contains the origin.

For general r, Frankl and Füredi (Construction 3 in [7]) generalized the latter case of $r = 3$: put n points randomly on the unit sphere in $\mathbb{R}^{r-1}$, and let an r-set form an edge if its convex hull contains the origin. This construction gives the lower bound $(1 + o(1))2^{1-r}\binom{n}{r}$ for Problem 1, which was the best known general lower bound before the present work. For the upper bound, a double-counting argument by de Caen [3] gives $\frac{n}{r^2}\binom{n}{r-1} \approx \frac{1}{r}\binom{n}{r}$, see Proposition 14 in [9]. Before the present work, apart from the exact $r = 3$ description, the only improvement over the general Frankl–Füredi construction was for the case $r = 4$. Gunderson and Semeraro [9] report an earlier random-tournament construction of Baber showing that the asymptotically optimal density 1/4 can be attained. They subsequently proved that for every prime power $q \equiv 3 \pmod 4$, there is a 4-graph on $q + 1$ vertices and $\frac{q+1}{16}\binom{q+1}{3}$ edges such that every set of 5 vertices spans 0 or 2 edges, which matches the upper bound mentioned above.

*Department of Mathematics, Statistics and Computer Science, University of Illinois Chicago, United States. Emails: visheshj@uic.edu, haoranl8@uic.edu, mubayi@uic.edu.

In this paper, we give a simple construction which significantly improves the lower bound for Problem 1. We note that we do not try to optimize the constant 32 to the best possible.

Theorem 2. *For every uniformity $r \geqslant 2$ and integer $n \geqslant 4r^2$, there is an n-vertex r-graph $\mathcal{H}$ such that every $(r+1)$-set of vertices in $\mathcal{H}$ spans exactly 0 or 2 edges and*

$$|\mathcal{H}| \geqslant \frac{1}{32r^3}\binom{n}{r}.$$

We remark that Problem 1 is closely related to the classical problem in which every $(r+1)$-set spans at most two edges. Equivalently, this is the Turán problem for H_3^r, the unique r-graph with three edges on $r+1$ vertices, which is also the $(3,2)$-daisy; the broader daisy problem has recently seen a breakthrough, see [5]. For a hypergraph H, let $\pi(H)$ be its Turán density. When $r=3$, H_3^3 is K_4^-, the 3-graph obtained from the clique on 4 vertices by removing one edge. Frankl and Füredi [7] gave the lower bound $\pi(K_4^-) \geqslant 2/7$. Baber and Talbot [1] proved the upper bound $\pi(K_4^-) \leqslant 0.2871$ using flag algebras, and this was later slightly improved to $\pi(K_4^-) \leqslant 0.286889$ by Falgas-Ravry and Vaughan [6]. For general r, the argument of de Caen [3] mentioned above gives $\pi(H_3^r) \leqslant 1/r$, which is also the present general upper bound for Problem 1. On the lower-bound side, the original construction of Frankl and Füredi [7] gives the exponential bound $\pi(H_3^r) \geqslant 2^{1-r}$; Sidorenko [13] recently proved $\pi(H_3^r) \geqslant r^{-2}$ for every r, and in fact $\pi(H_3^r) \geqslant (1.7215 - o(1))r^{-2}$ as $r \to \infty$; Gunderson and Semeraro [10] also obtained the lower bounds $\pi(H_3^6) \geqslant \frac{9}{64}$, $\pi(H_3^7) \geqslant \frac{35}{2^{11}}$, and $\pi(H_3^8) \geqslant \frac{315}{2^{14}}$ via an approach termed tournament switching. More recently, Clemen [2] improved the asymptotic lower bound to $\pi(H_3^r) = \Omega\left(r^{-2}\sqrt{\log r}\right)$ using sparse hypergraph colorings.

We remark that progress on the upper bound for Problem 1 should also be useful for the at-most-two problem. Indeed, the third author [12] used the exact Frankl–Füredi result for Problem 1 when $r=3$ to improve de Caen's upper bound for $\pi(K_4^-)$. The idea is that, in an H_3^r-free r-graph, every $(r+1)$-set spans 0, 1, or 2 edges. If a dense H_3^r-free r-graph exceeds the corresponding 0/2 threshold on many small vertex sets, then many of those small sets must contain an $(r+1)$-set spanning exactly one edge. These one-edge local configurations can then be inserted into the extension-counting inequality to improve the upper bound. Therefore, an improved upper bound for Problem 1 would, by the same supersaturation mechanism, give an improved upper bound for $\pi(H_3^r)$.

Our methods in Theorem 2 apply to the following rather natural generalization of Problem 1.

Definition 3. For fixed $k \geqslant 2$, let $M_{\leqslant 2k}(n,r)$ be the maximum number of edges in an n-vertex r-graph in which the number of edges contained in every $(r+1)$-set lies in $\{0, 2, \ldots, 2k\}$.

For every fixed r and k, the limit

$$m_{\leqslant 2k}(r) \coloneqq \lim_{n\to\infty} \frac{M_{\leqslant 2k}(n,r)}{\binom{n}{r}}$$

exists by the standard averaging argument of Katona, Nemetz, and Simonovits [11]. Indeed, if $n \geqslant m$, then averaging an extremal n-vertex example over all m-vertex subsets gives

$$\frac{M_{\leqslant 2k}(m,r)}{\binom{m}{r}} \geqslant \frac{M_{\leqslant 2k}(n,r)}{\binom{n}{r}}.$$

Thus the sequence $M_{\leqslant 2k}(n,r)/\binom{n}{r}$ is nonincreasing. We give the following lower bounds for $m_{\leqslant 2k}(r)$.

Theorem 4. *For every fixed integer $k \geqslant 2$, there is a constant $c_k > 0$ such that, for all sufficiently large r,*

$$m_{\leqslant 2k}(r) \geqslant c_k r^{-1-4/k}(\log r)^{1/k}.$$

For example, the special case $k = 2$ gives the following logarithmic improvement over the order r^{-3} bound in Theorem 2.

Corollary 5. *There is an absolute constant $c > 0$ such that, for all sufficiently large r,*

$$m_{\leqslant 4}(r) \geqslant cr^{-3}\sqrt{\log r}.$$

The proof of Theorem 4 is inspired by Clemen's recent improvement [2] for the Turán density of H_3^r. In both arguments, one first constructs a dense family of cores, then encodes the remaining bad local configurations as edges of an auxiliary sparse hypergraph, and finally passes to a large independent subfamily. In our setting, the auxiliary hypergraph is a linear hypergraph, so a well-known theorem of Duke, Lefmann, and Rödl [4] is sufficient.

The remaining part of this paper is organized as follows. In Section 2, we give the proof of Theorem 2. In Section 3, we prove Theorem 4 and Corollary 5. In Section 4, we give some concluding remarks on related problems.

2 Proof of Theorem 2

The main difficulty in the construction for Theorem 2 is that the property that every $(r+1)$-set spans exactly 0 or 2 edges imposes strong restrictions on the r-graph. Our main observation is the following lemma which enables us to construct a large collection of r-graphs that satisfy this property.

For an $(r-1)$-graph $\mathcal{H}$, its *upper shadow* $\partial^+(\mathcal{H})$ is the r-graph on the same vertex set as $\mathcal{H}$ whose edges are the r-sets containing at least one edge in $\mathcal{H}$.

Lemma 6. *For every $(r-1)$-graph $\mathcal{H}$ with the property that every $(r+1)$-set of vertices spans at most one edge, $\partial^+(\mathcal{H})$ has the property that every $(r+1)$-set of vertices spans exactly 0 or 2 edges.*

Proof. Let S be an arbitrary $(r+1)$-set of vertices. By assumption, S spans 0 or 1 edge in $\mathcal{H}$. If S spans no edge in $\mathcal{H}$, then S spans no edge in $\partial^+(\mathcal{H})$. If S spans one edge in $\mathcal{H}$, then exactly two r-subsets of S contain this edge, and hence S spans exactly two edges in $\partial^+(\mathcal{H})$. □

We use the following standard terminology. If $\mathcal{H}_0$ is an r-graph with vertex set $[m]$, then a *blow-up* of $\mathcal{H}_0$ is obtained by replacing each vertex $i \in [m]$ by a vertex class V_i, and replacing each edge $\{i_1, \ldots, i_r\} \in \mathcal{H}_0$ by all transversal r-sets with one vertex in each of $V_{i_1}, \ldots, V_{i_r}$. The blow-up is *balanced* if the classes $V_1, \ldots, V_m$ have sizes differing by at most one.

Lemma 7. *Let $\mathcal{H}_0$ be an r-graph with the property that every $(r+1)$-set of vertices spans exactly 0 or 2 edges. Then every blow-up of $\mathcal{H}_0$ also has this property.*

Proof. Let $\mathcal{H}$ be a blow-up of $\mathcal{H}_0$, and let π be the natural projection from $V(\mathcal{H})$ to $V(\mathcal{H}_0)$. For every $(r+1)$-set $S \subseteq V(\mathcal{H})$, consider $|\pi(S)|$. If $|\pi(S)| \leqslant r-1$, then S spans no edge in $\mathcal{H}$. If $|\pi(S)| = r+1$, then all vertices of S lie in distinct classes, and hence

$$|\mathcal{H}[S]| = |\mathcal{H}_0[\pi(S)]| \in \{0, 2\}.$$

Finally, suppose that $|\pi(S)| = r$. Then exactly one class contributes two vertices of S, and the other $r-1$ classes contribute one vertex each. If $\pi(S)$ is not an edge of $\mathcal{H}_0$, then S spans no edge in $\mathcal{H}$. If $\pi(S)$ is an edge of $\mathcal{H}_0$, then the only edges of $\mathcal{H}$ inside S are obtained by choosing one of the two vertices in the repeated class together with the unique vertex from each of the other $r-1$ classes, so S spans exactly two edges. □

Lemma 8. *Let N, m, r be integers such that $r \geqslant 2$, $m \geqslant 2r^2$, and $m \mid N$. Let α be a positive real number. Suppose $\mathcal{H}_0$ is an r-graph with m vertices and $\alpha\binom{m}{r}$ edges. Then there is an N-vertex blow-up $\mathcal{H}_N$ of $\mathcal{H}_0$ such that*

$$|\mathcal{H}_N| \geqslant \frac{\alpha}{2}\binom{N}{r}.$$

Proof. Replace each vertex of $\mathcal{H}_0$ by a set of size N/m, where distinct vertices are replaced by disjoint sets. Then

$$|\mathcal{H}_N| = \alpha\binom{m}{r}\left(\frac{N}{m}\right)^r.$$

Since $m \geqslant 2r^2$, we have

$$\binom{m}{r} = \frac{m^r}{r!}\prod_{i=0}^{r-1}\left(1-\frac{i}{m}\right) \geqslant \frac{m^r}{r!}\left(1-\sum_{i=0}^{r-1}\frac{i}{m}\right) \geqslant \frac{3}{4}\frac{m^r}{r!}.$$

Therefore

$$|\mathcal{H}_N| \geqslant \frac{3\alpha}{4}\frac{N^r}{r!} \geqslant \frac{\alpha}{2}\binom{N}{r},$$

as claimed. □

Lemma 9. *Let $N \geqslant n \geqslant r$, and let $\beta > 0$. Suppose that $\mathcal{H}_N$ is an N-vertex r-graph such that every $(r+1)$-set spans exactly 0 or 2 edges and*

$$|\mathcal{H}_N| \geqslant \beta\binom{N}{r}.$$

Then there is an n-vertex r-graph $\mathcal{H}$ such that every $(r+1)$-set spans exactly 0 or 2 edges and

$$|\mathcal{H}| \geqslant \beta\binom{n}{r}.$$

Proof. Choose an n-vertex subset $U \subseteq V(\mathcal{H}_N)$ uniformly at random. Each edge of $\mathcal{H}_N$ is contained in U with probability $\binom{n}{r}/\binom{N}{r}$. Hence

$$\mathbb{E}|\mathcal{H}_N[U]| = |\mathcal{H}_N|\frac{\binom{n}{r}}{\binom{N}{r}} \geqslant \beta\binom{n}{r}.$$

Therefore some choice of U satisfies

$$|\mathcal{H}_N[U]| \geqslant \beta\binom{n}{r}.$$

Since the property that every $(r+1)$-set spans exactly 0 or 2 edges is inherited by induced subgraphs, $\mathcal{H} = \mathcal{H}_N[U]$ has the desired local property. □

We also need the following construction, for which we give two proofs.

Proposition 10. *For every uniformity $r \geqslant 2$ and integer $n \geqslant r$, there is an n-vertex $(r-1)$-graph $\mathcal{H}$ with the property that every $(r+1)$-set of vertices spans at most one edge and*

$$|\mathcal{H}| \geqslant \frac{1}{4n^2}\binom{n}{r-1}.$$

We remark that, in the range $n \geqslant 2r$, Proposition 10 is best possible up to the absolute constant. Indeed, suppose that $r \geqslant 3$, $n \geqslant 2r$, and that an $(r-1)$-graph $\mathcal{F}$ on n vertices has the property that every $(r+1)$-set spans at most one edge. If two distinct edges of $\mathcal{F}$ contain the same $(r-3)$-set, then their union would have size at most $r+1$, and hence some $(r+1)$-set would contain both of them, a contradiction. Therefore every $(r-3)$-set is contained in at most one edge of $\mathcal{F}$, and so

$$|\mathcal{F}|\binom{r-1}{2} \leqslant \binom{n}{r-3}.$$

Equivalently,

$$|\mathcal{F}| \leqslant \frac{2}{(n-r+2)(n-r+3)}\binom{n}{r-1}.$$

In particular, if $n \geqslant 2r$, then $|\mathcal{F}| \leqslant \frac{8}{n^2}\binom{n}{r-1}$.

First proof of Proposition 10. We use the following theorem of Graham and Sloane [8].

Let $A(n, 2\delta, w)$ be the maximum number of codewords in any binary code of length n, constant weight w, and minimum Hamming distance at least 2δ.

Theorem 11 (Theorem 4 in [8]). *Let q be a prime power such that $q \geqslant n$. Then,*

$$A(n, 2\delta, w) \geqslant \frac{1}{q^{\delta-1}}\binom{n}{w}.$$

Now let $\delta = 3$, $w = r-1$, and let q be a prime number in $[n, 2n]$. By Theorem 11, we have

$$A(n, 6, r-1) \geqslant \frac{1}{q^{3-1}}\binom{n}{r-1} \geqslant \frac{1}{4n^2}\binom{n}{r-1}.$$

Viewing the codewords as the indicator vectors of the $(r-1)$-edges, we get an $(r-1)$-graph $\mathcal{H}$. Suppose for a contradiction that there is an $(r+1)$-set S spanning at least two edges, say e_1, e_2. Since $|e_1| = |e_2| = r-1$ and $e_1, e_2 \subseteq S$, we have $|e_1 \cap e_2| \geqslant r-3$. Hence, the corresponding codewords for e_1, e_2 have Hamming distance at most 4, a contradiction. $\square$

Second proof of Proposition 10. For $r = 2$, taking a single vertex as the only 1-edge gives the desired construction. Hence, we may assume $r \geqslant 3$.

Let q be a prime number in $[n, 2n]$, and identify the vertex set with an arbitrary n-element subset $U \subseteq \mathbb{F}_q$. Define

$$\lambda : \mathbb{F}_q \to \mathbb{F}_q^2, \qquad \lambda(t) = (t, t^2).$$

Note that if

$$\lambda(a) + \lambda(b) = \lambda(c) + \lambda(d),$$

then $a+b=c+d$ and $a^2+b^2=c^2+d^2$. Since q is odd, this implies $ab=cd$, and then $\{a,b\}$ and $\{c,d\}$ have the same sum and product, so they are equal.

For each $z \in \mathbb{F}_q^2$, define

$$\mathcal{G}_z = \left\{ A \in \binom{U}{r-1} : \sum_{a \in A} \lambda(a) = z \right\}.$$

The families $\mathcal{G}_z$ partition $\binom{U}{r-1}$ as z ranges over $\mathbb{F}_q^2$. Hence, by averaging, there exists $z \in \mathbb{F}_q^2$ such that

$$|\mathcal{G}_z| \geqslant \frac{1}{q^2}\binom{n}{r-1} \geqslant \frac{1}{4n^2}\binom{n}{r-1}.$$

We claim that this $\mathcal{G}_z$ has the required local property.

Let $S \in \binom{U}{r+1}$. Suppose that $A, B \in \mathcal{G}_z$ and $A, B \subseteq S$. Write

$$S \setminus A = \{u, v\}, \qquad S \setminus B = \{u', v'\}.$$

Since $\sum_{a \in A} \lambda(a) = \sum_{b \in B} \lambda(b) = z$, we have

$$\lambda(u) + \lambda(v) = \sum_{s \in S} \lambda(s) - z = \lambda(u') + \lambda(v').$$

By the Sidon property of λ, we get $\{u,v\} = \{u',v'\}$, and hence $A = B$. Therefore, every $(r+1)$-set contains at most one member of $\mathcal{G}_z$, as claimed. □

Proof of Theorem 2. By Proposition 10, there exists an $(r-1)$-graph $\mathcal{H}^{(r-1)}$ with $2r^2$ vertices and at least $\frac{1}{16r^4}\binom{2r^2}{r-1}$ edges such that every $(r+1)$-set of vertices spans at most one edge. Let $\mathcal{H}_0 = \partial^+(\mathcal{H}^{(r-1)})$. By Lemma 6, every $(r+1)$-set of vertices in $\mathcal{H}_0$ spans exactly 0 or 2 edges.

Claim 12. $|\mathcal{H}_0| \geqslant \frac{1}{16r^3}\binom{2r^2}{r}$.

Proof. We first note that every r-set contains at most one edge of $\mathcal{H}^{(r-1)}$. Indeed, if an r-set S contained two distinct edges $A, B \in \mathcal{H}^{(r-1)}$, then, since $\mathcal{H}^{(r-1)}$ has $2r^2 \geqslant r+1$ vertices, we could choose a vertex outside S and extend S to an $(r+1)$-set containing both A and B, which would contradict the defining property of $\mathcal{H}^{(r-1)}$.

Therefore, each edge of $\mathcal{H}_0 = \partial^+(\mathcal{H}^{(r-1)})$ contains a unique edge of $\mathcal{H}^{(r-1)}$. Since each $(r-1)$-edge of $\mathcal{H}^{(r-1)}$ is contained in exactly $2r^2-(r-1)$ many r-sets, we have

$$\begin{aligned} |\mathcal{H}_0| &= (2r^2-(r-1))|\mathcal{H}^{(r-1)}| \\ &\geqslant \frac{2r^2-r+1}{16r^4}\binom{2r^2}{r-1} = \frac{2r^2-r+1}{16r^4} \cdot \frac{r}{2r^2-r+1}\binom{2r^2}{r} \\ &= \frac{1}{16r^3}\binom{2r^2}{r}. \end{aligned}$$

□

Choose an integer $N \geqslant n$ divisible by $2r^2$. Applying Lemma 8 with $m = 2r^2$ and $\alpha = 1/(16r^3)$, we obtain an N-vertex blow-up $\mathcal{H}_N$ of $\mathcal{H}_0$ such that

$$|\mathcal{H}_N| \geqslant \frac{1}{32r^3}\binom{N}{r}.$$

By Lemma 7, every $(r+1)$-set of vertices in $\mathcal{H}_N$ spans exactly 0 or 2 edges. Applying Lemma 9 with $\beta = 1/(32r^3)$, we obtain an n-vertex r-graph $\mathcal{H}$ such that every $(r+1)$-set spans exactly 0 or 2 edges and

$$|\mathcal{H}| \geqslant \frac{1}{32r^3}\binom{n}{r}. \qquad \square$$

3 Proofs for relaxed even local counts

We prove Theorem 4; Corollary 5 then follows by taking $k = 2$. Throughout this section $k \geqslant 2$ is fixed, and $c_k, C_k > 0$ denote constants depending only on k, whose values may change from line to line. We assume that r is sufficiently large in terms of k; in particular, $r+1 \geqslant 2k+4$. Let q be an odd prime satisfying $r^2 \leqslant q \leqslant 2r^2$, which exists by Bertrand's postulate. For every $c \in \mathbb{F}_q$, define

$$\mathcal{G}_c = \left\{ A \in \binom{\mathbb{F}_q}{r-1} : \sum_{a \in A} a = c \right\}.$$

By averaging over $c \in \mathbb{F}_q$, we may fix c such that

$$|\mathcal{G}_c| \geqslant \frac{1}{q}\binom{q}{r-1}. \tag{1}$$

Lemma 13. *The family $\mathcal{G}_c$ has the following two properties.*

- *If $S \in \binom{\mathbb{F}_q}{r+1}$, then the members of $\mathcal{G}_c$ contained in S are in bijection with the pairs $P \in \binom{S}{2}$ satisfying*
$$\sum_{u \in P} u = \sum_{s \in S} s - c,$$
via $A = S \setminus P$. In particular, the corresponding pairs are pairwise disjoint.

- *Every r-set contains at most one member of $\mathcal{G}_c$.*

Proof. The first assertion follows directly from

$$\sum_{a \in S \setminus P} a = c \qquad \Longleftrightarrow \qquad \sum_{u \in P} u = \sum_{s \in S} s - c.$$

Since q is odd, the two-element subsets of $\mathbb{F}_q$ with a fixed sum form a matching; hence the corresponding pairs in S are pairwise disjoint.

For the second assertion, suppose that $X \in \binom{\mathbb{F}_q}{r}$ and $X \setminus \{x\}, X \setminus \{y\} \in \mathcal{G}_c$. Then

$$\sum_{z \in X} z - x = c = \sum_{z \in X} z - y,$$

so $x = y$. $\square$

For $s \geqslant 1$, call a collection of s distinct members of $\mathcal{G}_c$ *bad* if all its members are contained in a common $(r+1)$-set. Let B_s denote the number of bad s-collections.

Lemma 14. *For each $s \in \{k+1, k+2\}$,*

$$B_s \leqslant C_k r^4 |\mathcal{G}_c|.$$

Proof. Fix $s \in \{k+1, k+2\}$, and suppose that $A_1, \dots, A_s \in \mathcal{G}_c$ are distinct and contained in some $S \in \binom{\mathbb{F}_q}{r+1}$. Write $P_i = S \setminus A_i$. By Lemma 13, the pairs P_i are pairwise disjoint and have a common sum, say

$$\sum_{x \in P_i} x = L \qquad \text{for all } i \in [s].$$

Let $R = S \setminus (P_1 \cup \cdots \cup P_s)$. Then $|R| = r + 1 - 2s$. Since

$$\sum_{x \in S} x = \sum_{x \in R} x + sL \qquad \text{and} \qquad c = \sum_{x \in A_i} x = \sum_{x \in S} x - L,$$

we have

$$(s-1)L = c - \sum_{x \in R} x.$$

For sufficiently large r, $s - 1 \neq 0$ in $\mathbb{F}_q$, so R determines L. Once R and L are fixed, the pairs $P_1, \dots, P_s$ form an s-element submatching of the matching of pairs with sum L. This matching has at most $q/2$ edges, and hence the number of possible unordered collections $\{P_1, \dots, P_s\}$ is at most

$$\binom{\lfloor q/2 \rfloor}{s} \leqslant q^s.$$

Therefore

$$B_s \leqslant \binom{q}{r+1-2s} q^s.$$

Using (1) and $q - r \geqslant q/2$, we get

$$\frac{B_s}{|\mathcal{G}_c|} \leqslant q \frac{\binom{q}{r+1-2s}}{\binom{q}{r-1}} q^s \leqslant C_k q^{s+1} \left(\frac{r}{q}\right)^{2s-2} = C_k r^{2s-2} q^{3-s} \leqslant C_k r^4,$$

where the last inequality uses $s \geqslant k+1 \geqslant 3$ and $q \geqslant r^2$. $\square$

Lemma 15. *There is a subfamily $\mathcal{G}_1 \subseteq \mathcal{G}_c$ such that*

$$|\mathcal{G}_1| \geqslant c_k r^{-4/(k+1)} |\mathcal{G}_c|, \tag{2}$$

no $(r+1)$-set contains more than $k+1$ members of $\mathcal{G}_1$, and the number of bad $(k+1)$-collections in $\mathcal{G}_1$ is at most $C_k |\mathcal{G}_c|$.

Proof. Choose a sufficiently small constant $\alpha = \alpha(k) > 0$, and set

$$\rho = \alpha r^{-4/(k+1)}.$$

Let $\mathcal{G}' \subseteq \mathcal{G}_c$ be obtained by retaining each member independently with probability ρ. Let Y_s be the number of bad s-collections contained in $\mathcal{G}'$. By Lemma 14,

$$\mathbb{E}|\mathcal{G}'| = \rho |\mathcal{G}_c|, \qquad \mathbb{E}Y_{k+1} \leqslant C_k \alpha^{k+1} |\mathcal{G}_c|, \qquad \mathbb{E}Y_{k+2} \leqslant C_k \alpha^{k+1} \rho |\mathcal{G}_c|.$$

Since $\rho|\mathcal{G}_c| \to \infty$ as $r \to \infty$, the Chernoff bound for $|\mathcal{G}'|$, Markov's inequality for Y_{k+1}, Y_{k+2}, and the choice of α imply that, with positive probability,

$$|\mathcal{G}'| \geqslant \frac{1}{2}\rho|\mathcal{G}_c|, \qquad Y_{k+2} \leqslant \frac{1}{4}\rho|\mathcal{G}_c|, \qquad Y_{k+1} \leqslant C_k|\mathcal{G}_c|.$$

Fix such a realization.

For each bad $(k+2)$-collection in $\mathcal{G}'$, choose one of its members and delete all members chosen in this way. Let the resulting family be $\mathcal{G}_1$. Then

$$|\mathcal{G}_1| \geqslant |\mathcal{G}'| - Y_{k+2} \geqslant \frac{1}{4}\rho|\mathcal{G}_c|,$$

which gives (2). No bad $(k+2)$-collection remains, so no $(r+1)$-set contains more than $k+1$ members of $\mathcal{G}_1$. Finally, deleting members cannot create bad $(k+1)$-collections, so the number of such collections in $\mathcal{G}_1$ is at most $Y_{k+1} \leqslant C_k|\mathcal{G}_c|$. □

Lemma 16. *There is a subfamily $\mathcal{G}_2 \subseteq \mathcal{G}_1$ such that every $(r+1)$-set contains at most k members of $\mathcal{G}_2$, and*

$$|\mathcal{G}_2| \geqslant c_k|\mathcal{G}_c|r^{-4/k}(\log r)^{1/k}. \tag{3}$$

Proof. Define a $(k+1)$-uniform hypergraph $\mathcal{T}$ on vertex set $\mathcal{G}_1$ by declaring a $(k+1)$-set to be an edge if its members are contained in a common $(r+1)$-set.

We first show that $\mathcal{T}$ is linear. Suppose that two edges share two vertices $A, B \in \mathcal{G}_1$. Then A and B are contained in a common $(r+1)$-set, so $|A \cap B| \geqslant r-3$. The case $|A \cap B| = r-2$ is impossible, since then the r-set $A \cup B$ would contain two distinct members of $\mathcal{G}_c$, contradicting Lemma 13. Thus $|A \cap B| = r-3$, and $A \cup B$ is the unique $(r+1)$-set containing both A and B. By Lemma 15, this $(r+1)$-set contains at most $k+1$ members of $\mathcal{G}_1$. Hence there is at most one edge of $\mathcal{T}$ containing A and B, proving linearity.

The number of edges of $\mathcal{T}$ is at most the number of bad $(k+1)$-collections in $\mathcal{G}_1$, hence at most $C_k|\mathcal{G}_c|$. Together with (2), this shows that the average degree of $\mathcal{T}$ is at most

$$D = C_k r^{4/(k+1)}.$$

By the linear-hypergraph form of the theorem of Duke, Lefmann, and Rödl [4], every linear $(k+1)$-uniform hypergraph on N vertices with average degree at most D has an independent set of size at least

$$c_k N \left(\frac{\log D}{D}\right)^{1/k}.$$

Applying this to $\mathcal{T}$, and using $D = C_k r^{4/(k+1)}$, gives an independent set $\mathcal{G}_2 \subseteq \mathcal{G}_1$ satisfying

$$|\mathcal{G}_2| \geqslant c_k|\mathcal{G}_1| \left(r^{-4/(k+1)} \log r\right)^{1/k} \geqslant c_k|\mathcal{G}_c|r^{-4/k}(\log r)^{1/k}.$$

Since $\mathcal{G}_2$ is independent in $\mathcal{T}$, no $(r+1)$-set contains $k+1$ members of $\mathcal{G}_2$. □

Lemma 17. *Let*

$$\mathcal{H}_0 = \partial^+\mathcal{G}_2 = \left\{ X \in \binom{\mathbb{F}_q}{r} : \exists A \in \mathcal{G}_2 \text{ such that } A \subseteq X \right\}.$$

Then every $(r+1)$-set spans one of $0, 2, \ldots, 2k$ edges in $\mathcal{H}_0$, and

$$|\mathcal{H}_0| \geqslant c_k r^{-4/k} (\log r)^{1/k} \frac{r}{q} \binom{q}{r}. \tag{4}$$

Proof. By Lemma 13, every r-set contains at most one member of $\mathcal{G}_c$. Hence every edge of $\mathcal{H}_0$ contains a unique member of $\mathcal{G}_2$, and therefore

$$|\mathcal{H}_0| = (q - r + 1)|\mathcal{G}_2|.$$

Combining this with (1) and (3) gives

$$|\mathcal{H}_0| \geqslant (q - r + 1) c_k r^{-4/k} (\log r)^{1/k} \frac{1}{q} \binom{q}{r-1} = c_k r^{-4/k} (\log r)^{1/k} \frac{r}{q} \binom{q}{r},$$

which proves (4).

Now fix $S \in \binom{\mathbb{F}_q}{r+1}$, and suppose that S contains exactly m members of $\mathcal{G}_2$. By Lemma 16, $m \leqslant k$. Writing these members as $A_i = S \setminus P_i$, the pairs P_i are pairwise disjoint by Lemma 13. Each A_i contributes exactly the two r-subsets of S obtained by adding one element of P_i, and these contributions are disjoint because every r-set contains at most one member of $\mathcal{G}_c$. Thus

$$|\mathcal{H}_0 \cap \binom{S}{r}| = 2m \in \{0, 2, \ldots, 2k\}. \qquad \square$$

Proof of Theorem 4. Let $\mathcal{H}_0$ be the q-vertex r-graph from Lemma 17. For each N, take a balanced blow-up of the q vertices into N vertices, and include all transversal r-sets whose labels form an edge of $\mathcal{H}_0$. Call the resulting r-graph $\mathcal{H}_N$.

Let π be the projection to the q labels, and let S be an $(r+1)$-set in the blow-up. If $|\pi(S)| \leqslant r-1$, then S spans no edge. If $|\pi(S)| = r$, then exactly one label is repeated, so S spans either no edge or exactly two edges. If $|\pi(S)| = r+1$, then S is transversal and the number of edges spanned by S equals

$$|\mathcal{H}_0 \cap \binom{\pi(S)}{r}| \in \{0, 2, \ldots, 2k\}.$$

Thus $\mathcal{H}_N$ satisfies the required local condition. Hence

$$M_{\leqslant 2k}(N, r) \geqslant |\mathcal{H}_N|.$$

It remains to estimate the limiting density as $N \to \infty$. For the balanced blow-up,

$$\lim_{N \to \infty} \frac{|\mathcal{H}_N|}{\binom{N}{r}} = |\mathcal{H}_0| \frac{r!}{q^r}.$$

By (4),

$$|\mathcal{H}_0| \frac{r!}{q^r} \geqslant c_k r^{-4/k} (\log r)^{1/k} \frac{r}{q} \prod_{i=0}^{r-1} \left(1 - \frac{i}{q}\right).$$

Since $q \geqslant r^2$,

$$\prod_{i=0}^{r-1} \left(1 - \frac{i}{q}\right) \geqslant 1 - \sum_{i=0}^{r-1} \frac{i}{q} \geqslant \frac{1}{2},$$

and since $q \leqslant 2r^2$, we also have $r/q \geqslant 1/(2r)$. Therefore

$$\lim_{N \to \infty} \frac{M_{\leqslant 2k}(N, r)}{\binom{N}{r}} \geqslant \lim_{N \to \infty} \frac{|\mathcal{H}_N|}{\binom{N}{r}} \geqslant c_k r^{-1-4/k} (\log r)^{1/k}. \qquad \square$$

4 Concluding remarks

It is natural to consider the following more general local problem. For a set $\Lambda \subseteq \{0, 1, \ldots, r+1\}$, let $M_\Lambda(n, r)$ be the maximum size of an r-graph $\mathcal{H}$ on n vertices such that $|\mathcal{H} \cap \binom{S}{r}| \in \Lambda$ for every $(r+1)$-set S. For every fixed r and Λ, the limit

$$m_\Lambda(r) := \lim_{n\to\infty} \frac{M_\Lambda(n, r)}{\binom{n}{r}}$$

exists by the same averaging argument as before. In this notation, Problem 1 asks for lower and upper bounds on $m_{\{0,2\}}(r)$.

The results in Theorem 4 and Corollary 5 suggest that the structure of the allowed set Λ is important. A basic open direction is to determine which choices of Λ allow density of order r^{-2}, or even larger, and which choices are constrained to the r^{-3} scale by upper-shadow-type barriers.

Acknowledgments and declaration on the use of generative AI

A slightly weaker lower bound for Theorem 2 of the form $\Omega(r^{-5})$ was obtained by the authors without any use of AI tools, where instead of Proposition 10, a simple probabilistic deletion method was used. Subsequently, ChatGPT 5.4 Pro was used to find the better constructions in the proof of Proposition 10 to obtain the $\Omega(r^{-3})$ bound; additionally, upon being given a sketch of the argument, it produced a proof of the bound in Theorem 4. All proofs and calculations were checked by the authors.

The authors are grateful to Huy Tuan Pham for helpful discussions.

V.J. is partially supported by NSF grant DMS-2237646. D.M. is partially supported by NSF grant DMS-2153576.

References

[1] R. Baber and J. Talbot. Hypergraphs do jump. *Combin. Probab. Comput.*, 20(2):161–171, 2011.

[2] F. C. Clemen. Applications of sparse hypergraph colorings. *Discrete Math.*, 349(2):Paper No. 114822, 2026.

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

[4] R. A. Duke, H. Lefmann, and V. Rödl. On uncrowded hypergraphs. *Random Structures Algorithms*, 6(2-3):209–212, 1995.

[5] D. Ellis, M.-R. Ivan, and I. Leader. Turán densities for daisies and hypercubes. *Bull. Lond. Math. Soc.*, 56(12):3838–3853, 2024.

[6] V. Falgas-Ravry and E. R. Vaughan. Applications of the semi-definite method to the Turán density problem for 3-graphs. *Combin. Probab. Comput.*, 22(1):21–54, 2013.

[7] P. Frankl and Z. Füredi. An exact result for 3-graphs. *Discrete Math.*, 50(2-3):323–328, 1984.

[8] R. L. Graham and N. J. A. Sloane. Lower bounds for constant weight codes. *IEEE Trans. Inform. Theory*, 26(1):37–43, 1980.

[9] K. Gunderson and J. Semeraro. Tournaments, 4-uniform hypergraphs, and an exact extremal result. *J. Combin. Theory Ser. B*, 126:114–136, 2017.

[10] K. Gunderson and J. Semeraro. Turán numbers and switching. *Discrete Math.*, 348(2):Paper No. 114275, 2025.

[11] G. Katona, T. Nemetz, and M. Simonovits. On a problem of Turán in the theory of graphs. *Mat. Lapok*, 15:228–238, 1964. In Hungarian.

[12] D. Mubayi. On hypergraphs with every four points spanning at most two triples. *Electron. J. Combin.*, 10(1):Note 10, 4 pp., 2003.

[13] A. Sidorenko. Turán numbers of r-graphs on $r + 1$ vertices. *J. Combin. Theory Ser. B*, 169:150–160, 2024.