Independent sets in hypergraphs omitting an intersection

Tom Bohman * Xizhi Liu † Dhruv Mubayi ‡

October 19, 2021

Abstract

A k-uniform hypergraph with n vertices is an (n,k,ℓ)-omitting system if it has no two edges with intersection size ℓ. If in addition it has no two edges with intersection size greater than ℓ, then it is an (n,k,ℓ)-system. Rödl and Šiňajová proved a sharp lower bound for the independence number of (n,k,ℓ)-systems. We consider the same question for (n,k,ℓ)-omitting systems. Our proofs use adaptations of the random greedy independent set algorithm, and pseudorandom graphs. We also prove related results where we forbid more than two edges with a prescribed common intersection size leading to some applications in Ramsey theory. For example, we obtain good bounds for the Ramsey number $r_k(F,t)$, where F is the k-uniform Fan. The behavior is quite different than the case $k=2$ which is the classical Ramsey number $r(3,t)$.

1 Introduction

For a finite set V and $k \geq 2$ denote by $\binom{V}{k}$ the collection of all k-subsets of V. A k-uniform hypergraph (k-graph) $\mathcal{H}$ is a family of k-subsets of finite set which is called the vertex set of $\mathcal{H}$ and is denoted by $V(\mathcal{H})$. A set $I \subset V(\mathcal{H})$ is *independent* in $\mathcal{H}$ if it contains no edge of $\mathcal{H}$. The *independence number* of $\mathcal{H}$, denoted by $\alpha(\mathcal{H})$, is the maximum size of an independent set in $\mathcal{H}$. For every $v \in V(\mathcal{H})$ the *degree* $d_{\mathcal{H}}(v)$ of v in $\mathcal{H}$ is the number of edges in $\mathcal{H}$ that contain v. Denote by $d(\mathcal{H})$ and $\Delta(\mathcal{H})$ the average degree and the maximum degree of $\mathcal{H}$, respectively.

An old result of Turán [25] implies that $\alpha(G) \geq n/(d+1)$ for every graph G on n vertices with average degree d. Later, Spencer [23] extended Turán's result and proved that for all $k \geq 3$ every n-vertex k-graph $\mathcal{H}$ with average degree d satisfies

$$\alpha(\mathcal{H}) \geq c_k \frac{n}{d^{1/(k-1)}} \tag{1}$$

for some constant $c_k > 0$.

The bound for $\alpha(\mathcal{H})$ can be improved if we forbid some family $\mathcal{F}$ of hypergraphs in $\mathcal{H}$. For $\ell \geq 2$ a (Berge) *cycle* of length ℓ in $\mathcal{H}$ is a collection of ℓ edges $E_1, \ldots, E_\ell \in \mathcal{H}$ such that

*Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA, 15213 USA. email: tbohman@math.cmu.edu.

†Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL, 60607 USA. email: xliu246@uic.edu. Research partially supported by NSF awards DMS-1763317 and DMS-1952767.

‡Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL, 60607 USA. email: mubayi@uic.edu. Research partially supported by NSF awards DMS-1763317 and DMS-1952767.

there exists ℓ distinct vertices $v_1, \ldots, v_\ell$ with $v_i \in E_i \cap E_{i+1}$ for $i \in [\ell - 1]$ and $v_\ell \in E_\ell \cap E_1$. A seminal result of Ajtai, Komlós, Pintz, Spencer, and Szemerédi [2] states that for every n-vertex k-graph $\mathcal{H}$ with average degree d that contains no cycles of length $2, 3$, and 4, there exists a constant $c'_k > 0$ such that

$$\alpha(\mathcal{H}) \geq c'_k \frac{n}{d^{1/(k-1)}} (\log d)^{1/(k-1)}. \tag{2}$$

Moreover, this is tight apart from c'_k.

Spencer [20] conjectured and Duke, Lefmann, and Rödl [9] proved that the same conclusion holds even if $\mathcal{H}$ just contains no cycles of length 2. Their result was further extended by Rödl and Šiňajová [21] to the larger family of (n, k, ℓ)-systems defined in the following section.

1.1 (n, k, ℓ)-systems and (n, k, ℓ)-omitting systems

Let $k > \ell \geq 1$. An n-vertex k-graph $\mathcal{H}$ is an (n, k, ℓ)-*system* if the intersection of every pair of edges in $\mathcal{H}$ has size less than ℓ, and $\mathcal{H}$ is an (n, k, ℓ)-*omitting system* if it has no two edges whose intersection has size exactly ℓ. It is clear from the definition that an (n, k, ℓ)-system is an (n, k, ℓ)-omitting system, but not vice versa, since an (n, k, ℓ)-omitting system may have pairwise intersection sizes greater than ℓ.

Define

$$f(n, k, \ell) = \min \{\alpha(\mathcal{H}) : \mathcal{H} \text{ is an } (n, k, \ell)\text{-system}\}, \quad \text{and}$$
$$g(n, k, \ell) = \min \{\alpha(\mathcal{H}) : \mathcal{H} \text{ is an } (n, k, \ell)\text{-omitting system}\}.$$

We will use the standard asymptotic notations O, Ω, Θ, o to simplify the formulas used in the present paper. Recall that given two positive functions $f(n)$ and $g(n)$ we write $f(n) = O(g(n))$, or equivalently, $g(n) = \Omega(f(n))$ if there exists a constant $C > 0$ such that $f(n) \leq Cg(n)$ for all sufficiently large n, we write $f(n) = o(g(n))$ if $\lim_{n\to\infty} f(n)/g(n) = 0$, and we write $f(n) = \Theta(g(n))$ if both $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$ hold.

The study of $f(n, k, \ell)$ has a long history (e.g. [21, 16, 10, 24]) and, in particular, Rödl and Šiňajová [21] proved that

$$f(n, k, \ell) = \Theta\left(n^{\frac{k-\ell}{k-1}} (\log n)^{\frac{1}{k-1}}\right) \text{ for all fixed } k > \ell \geq 2. \tag{3}$$

It follows that

$$g(n, k, \ell) \leq f(n, k, \ell) = O\left(n^{\frac{k-\ell}{k-1}} (\log n)^{\frac{1}{k-1}}\right). \tag{4}$$

One important difference between (n, k, ℓ)-systems and (n, k, ℓ)-omitting systems is their maximum sizes. By definition, every set of ℓ vertices in an (n, k, ℓ)-system is contained in at most one edge, thus every (n, k, ℓ)-system has size at most $\binom{n}{\ell}/\binom{k}{\ell} = O(n^\ell)$. However, this is not true for (n, k, ℓ)-omitting systems. Indeed, the following result of Frankl and Füredi [11] shows that the maximum size of an (n, k, ℓ)-omitting system can be much larger than that of an (n, k, ℓ)-system when $k > 2\ell + 1$.

Let $k > \ell \geq 1$ and $\lambda \geq 1$ be integers. The k-graph $S_\lambda^k(\ell)$ consists of λ edges $E_1, \ldots, E_\lambda$ such that $E_i \cap E_j = S$ for $1 \leq i < j \leq \lambda$ and some fixed set S (called the center) of size ℓ. When $\ell = 1$ we just write S_λ^k, and we will omit the superscript k in $S_\lambda^k(\ell)$ if it is obvious. It is easy to see that an n-vertex k-graph is an (n, k, ℓ)-omitting system iff it is $S_2(\ell)$-free, and is an (n, k, ℓ)-system iff it is $\{S_2(\ell), \ldots, S_2(k-1)\}$-free.

Theorem 1.1 (Frankl–Füredi [11]). *Let $k > \ell \geq 1$ and $\lambda > 1$ be fixed integers and $\mathcal{H}$ be an $S_\lambda(\ell)$-free k-graph on n vertices. Then $|\mathcal{H}| = O\left(n^{\max\{\ell, k-\ell-1\}}\right)$. Moreover, the bound is tight up to a constant multiplicative factor.*

Theorem 1.1 together with (1) imply that for fixed k, ℓ,

$$g(n,k,\ell) = \begin{cases} \Omega\left(n^{\frac{k-\ell}{k-1}}\right) & k \leq 2\ell+1, \\ \Omega\left(n^{\frac{\ell+1}{k-1}}\right) & k > 2\ell+1. \end{cases} \tag{5}$$

Notice that for $k \leq 2\ell + 1$ the bounds given by (4) and (5) match except for a factor of $(\log n)^{1/(k-1)}$, but for $k > 2\ell + 1$, these two bounds have a gap in the exponent of n.

Our main goal in this paper is to extend the results of Rödl and Šiňajová to the larger class of (n, k, ℓ)-omitting systems and improve the bounds given by (4) and (5). In other words, the question we focus on is the following:

What is the value of $g(n, k, \ell)$?

Our results for (n, k, ℓ)-omitting systems are divided into two parts. For $k \leq 2\ell + 1$, we believe that the behavior is similar to that of (n, k, ℓ)-systems and prove a nontrivial lower bound for the first open case $\ell = k - 2$. For $k > 2\ell + 1$ we give new lower and upper bounds which show that the minimum independence number of (n, k, ℓ)-omitting systems has a very different behavior than for (n, k, ℓ)-systems.

1.2 $k \leq 2\ell + 1$

As mentioned above, for this range of ℓ and k, the issue at hand is only the polylogarithmic factor in $g(n, k, \ell)$. It follows from the definition that an $(n, k, k-1)$-omitting system is also an $(n, k, k-1)$-system, thus Rödl and Šiňajová's result (3) implies that

$$g(n,k,k-1) = f(n,k,k-1) = \Theta\left(n^{\frac{1}{k-1}} (\log n)^{\frac{1}{k-1}}\right).$$

So, the first open case in the range of $k \leq 2\ell + 1$ is $\ell = k - 2$, and for this case we prove the following nontrivial lower bound for $g(n, k, k-2)$, which improves (5).

Theorem 1.2. *Suppose that $k \geq 4$. Then every $(n, k, k-2)$-omitting system has an independent set of size $\Omega\left(n^{2/(k-1)} (\log\log n)^{1/(k-1)}\right)$. In other words,*

$$g(n,k,k-2) = \Omega\left(n^{\frac{2}{k-1}} (\log\log n)^{\frac{1}{k-1}}\right).$$

Unfortunately, our method for proving Theorem 1.2 cannot be extended to the entire range of $k \leq 2\ell + 1$, but we make the following conjecture.

Conjecture 1.3. *For all fixed integers $k > \ell \geq 2$ that satisfy $k \leq 2\ell + 1$ there exists a function $\omega(n) \to \infty$ as $n \to \infty$ such that $g(n, k, \ell) = \Omega\left(n^{\frac{k-\ell}{k-1}} \omega(n)\right)$.*

Theorem 1.2 shows that Conjecture 1.3 is true for $\ell = k - 2$. The smallest open case is $k = 5$ and $\ell = 2$.

1.3 $k > 2\ell + 1$

Recall that in the range of $k > 2\ell + 1$ the bounds given by (4) and (5) leave a gap in the exponent of n. The following result shows that for a wide range of k and ℓ neither of them gives the correct order of magnitude.

Theorem 1.4. *Let $\ell \geq 2$ and $k > 2\ell + 1$ be fixed. Then*

$$\Omega\left(\max\left\{n^{\frac{\ell+1}{3\ell-1}}, n^{\frac{\ell+1}{k-1}}\right\}\right) = g(n,k,\ell) = O\left(n^{\frac{\ell+1}{2\ell}} (\log n)^{\frac{1}{\ell}}\right).$$

Remark.

(a) The lower bound $n^{\frac{\ell+1}{3\ell-1}}$ can be improved to $n^{\frac{3-\sqrt{5}}{2}+o_\ell(1)} \sim n^{0.38196+o_\ell(1)}$. See the remark in the end of Section 3 for details.

(b) It is clear that Theorem 1.4 improves the bound given by (5) for $k > 3\ell$, and it also improves the bound given by (4) for $k > 2\ell + 1$ as $\frac{k-\ell}{k-1} - \frac{\ell+1}{2\ell} = \frac{(\ell-1)(k-2\ell-1)}{2\ell(k-1)} > 0$ for $k > 2\ell + 1$.

It would be interesting to determine $g(n,k,\ell)$ for $k > 2\ell + 1$. Here, we are not able to offer a conjecture for the exponent of n.

Problem 1.5. *Determine the order of magnitude of $g(n,k,\ell)$ for $k > 2\ell + 1$.*

For the first open case $(k,\ell) = (6,2)$ Theorem 1.4 gives $\Omega\left(n^{3/5}\right) = g(n,6,2) = O\left(n^{3/4+o(1)}\right)$. Similar to Remark (a) above the lower bound for $g(n,6,2)$ can be improved to $\Omega\left(n^{2/3}\right)$. See the remark in the end of Section 3 for details.

1.4 (n,k,ℓ,λ)-systems and (n,k,ℓ,λ)-omitting systems

We consider the following generalization of (n,k,ℓ)-omitting systems and (n,k,ℓ)-systems in this section.

An n-vertex k-graph $\mathcal{H}$ is an (n,k,ℓ,λ)-*system* if every set of ℓ vertices is contained in at most λ edges, and $\mathcal{H}$ is an (n,k,ℓ,λ)-*omitting system* if it does not contain $S_{\lambda+1}(\ell)$ as a subgraph.

Define

$$f(n,k,\ell,\lambda) = \min\{\alpha(\mathcal{H}) : \mathcal{H} \text{ is an } (n,k,\ell,\lambda)\text{-system}\}, \quad \text{and}$$
$$g(n,k,\ell,\lambda) = \min\{\alpha(\mathcal{H}) : \mathcal{H} \text{ is an } (n,k,\ell,\lambda)\text{-omitting system}\}.$$

When λ is a fixed constant, the value of $f(n,k,\ell,\lambda)$ is essentially the same as $f(n,k,\ell)$ (e.g. see [21]), i.e. $f(n,k,\ell,\lambda) = \Theta(f(n,k,\ell))$. Similarly, the same conclusions as in Theorems 1.2 and 1.4 also hold for $g(n,k,\ell,\lambda)$, since Theorem 1.1 holds for all $S_\lambda(\ell)$-free hypergraphs and using it one can easily extend the proof for the case $\lambda = 1$ to the case $\lambda > 1$. For the sake of simplicity, we will prove Theorem 1.2 only for the case $\lambda = 1$.

When λ is not a constant, even the value of $f(n,k,\ell,\lambda)$ is not known in general. Here is a summary of the known results.

- $\ell = 1$: An $(n,k,1,\lambda)$-system is just a k-graph with maximum degree λ and here complete k-graphs and (1) yield

$$f(n,k,1,\lambda) = \Theta\left(\frac{n}{\lambda^{1/(k-1)}}\right).$$

On the other hand a result of Loh [17] implies

$$g(n,k,1,\lambda) = \frac{n}{\lambda+1} \qquad \text{whenever} \qquad (\lambda+1)(k-1) \mid n.$$

If the divisibility condition fails then we have a small error term above.

- $\ell = k-1$: Kostochka, Mubayi, and Verstraëte [16] proved that

$$f(n,k,k-1,\lambda) = \Theta\left(\left(\frac{n}{\lambda}\right)^{\frac{1}{k-1}}\left(\log\frac{n}{\lambda}\right)^{\frac{1}{k-1}}\right) \quad \text{for} \quad 1 \le \lambda \le \frac{n}{(\log n)^{3(k-1)^2}}.$$

- $2 \le \ell \le k-2$: Tian and Liu [24] proved that

$$f(n,k,\ell,\lambda) = \Omega\left(\left(\frac{n}{\lambda}\log\frac{n}{\lambda}\right)^{1/\ell}\right) \quad \text{for} \quad k \ge 5,\ \frac{2k+4}{5} < \ell \le k-2,\ \lambda = o\left(n^{\frac{5\ell-2k-4}{3k-9}}\right).$$

They also gave a construction which implies that

$$f(n,k,\ell,\lambda) = O\left(\left(\frac{n^{k-\ell}}{\lambda}\right)^{\frac{1}{k-1}}\left(\log\frac{n}{\lambda}\right)^{\frac{1}{k-1}}\right) \quad \text{for} \quad 2 \le \ell \le k-1,\ \log n \ll \lambda \ll n.$$

Since for every $\lambda > 0$ an (n,k,ℓ,λ)-system has size $O\left(\lambda n^{\ell}\right)$, it follows from (1) that

$$f(n,k,\ell,\lambda) = \Omega\left(\left(\frac{n^{k-\ell}}{\lambda}\right)^{\frac{1}{k-1}}\right),$$

which, by Tian and Liu's upper bound, is tight up to a factor of $(\log n)^{1/(k-1)}$ when $\log n \ll \lambda \ll n$.

Using a result of Duke, Lefmann, and Rödl [9] we are able to improve the lower bound for $f(n,k,\ell,\lambda)$ to match the upper bound obtained by Tian and Liu for a wide range of λ.

Theorem 1.6. *Let $k > \ell \ge 2$ be fixed. If there exists a constant $\delta > 0$ such that $0 < \lambda < n^{\frac{\ell-1}{k-2}-\delta}$, then*

$$f(n,k,\ell,\lambda) = \Omega\left(\left(\frac{n^{k-\ell}}{\lambda}\right)^{\frac{1}{k-1}}(\log n)^{\frac{1}{k-1}}\right).$$

Remark. It remains open to determine $f(n,k,\ell,\lambda)$ for $\Omega\left(n^{\frac{\ell-1}{k-2}-o(1)}\right) = \lambda = O\left(n^{k-\ell}\right)$.

Since Theorem 1.1 does not hold when λ is not a constant, our method of proving Theorems 1.2 and 1.4 cannot be extended to this case.

1.5 Applications in Ramsey theory

For a k-graph $\mathcal{F}$ the *Ramsey number* $r_k(\mathcal{F}, t)$ is the smallest integer n such that every $\mathcal{F}$-free k-graph on n vertices has an independent set of size at least t. Determining the minimum independence number of an $\mathcal{F}$-free k-graph on n vertices is essentially the same as determining the value of $r_k(\mathcal{F}, t)$. So, our results above can be applied to determine the Ramsey number of some hypergraphs.

First, Theorem 1.2 and (4) imply the following corollary.

Corollary 1.7. *Let $k \geq 4$ and $\lambda \geq 2$ be fixed integers. Then*

$$\Omega\left(\frac{t^{(k-1)/2}}{(\log t)^{1/2}}\right) = r_k(S_\lambda(k-2), t) = O\left(\frac{t^{(k-1)/2}}{(\log\log t)^{1/2}}\right).$$

Similarly, Theorem 1.4 gives the following corollary.

Corollary 1.8. *Let $\ell \geq 2$, $k > 2\ell + 1$, and $\lambda \geq 2$ be fixed integers. Then*

$$\Omega\left(\frac{t^{2\ell/(\ell+1)}}{(\log t)^{2/(\ell+1)}}\right) = r_k(S_\lambda(\ell), t) = O\left(\min\left\{t^{\frac{3\ell-1}{\ell+1}}, t^{\frac{k-1}{\ell+1}}\right\}\right).$$

Remark. According to Remark (a) after Theorem 1.4, the upper bound $t^{\frac{3\ell-1}{\ell+1}}$ above can be improved to $t^{\frac{3+\sqrt{5}}{2}+o_\ell(1)} \sim t^{2.61803+o_\ell(1)}$.

The following result about $r_k(S^k_\lambda, t)$ follows from a more general result of Loh [17].

Theorem 1.9 (Loh [17])**.** *Let $t \geq k \geq 2$, $t - 1 = q(k-1) + r$ for some $q, r \in \mathbb{N}$ with $0 \leq r \leq k-2$. Then for every $\lambda \geq 2$*

$$\lambda q(k-1) + r + 1 \leq r_k(S^k_\lambda, t) \leq \lambda q(k-1) + \lambda r + 1.$$

In particular, $r_k(S^k_\lambda, t) = \lambda(t-1) + 1$ whenever $(k-1) \mid (t-1)$.

The k-*Fan*, denoted by F^k, is the k-graph consisting of $k+1$ edges $E_1, \ldots, E_k, E$ such that $E_i \cap E_j = v$ for all $1 \leq i < j \leq k$, where $v \notin E$, and $|E_i \cap E| = 1$ for $1 \leq i \leq k$. In other words, F^k is obtained from S^k_k by adding an edge omitting v that intersects each edge of S^k_k. It is easy to see that F^2 is just the triangle K_3. The k-graph F^k was first introduced by Mubayi and Pikhurko [19] in order to extend Mantel's theorem to hypergraphs. Unlike the case $k = 2$, where it is well known that $r_2(K_3, t) = \Theta\left(t^2/\log t\right)$ (e.g. see [3, 14]), the following result shows that $r_k(F^k, t) = \Theta(t^2)$ for all $k \geq 3$.

Theorem 1.10. *Suppose that $t \geq k \geq 3$. Then*

$$\left\lfloor\frac{t}{2}\right\rfloor\left\lfloor\frac{t-1}{2(k-2)}\right\rfloor < r_k(F^k, t) \leq t(t-1) + 1.$$

As $t \to \infty$, it remains open to determine $\lim r_k(F^k, t)/t^2$.

In Section 2, we prove Theorem 1.2. In Section 3, we prove Theorem 1.4. In Section 4, we prove Theorem 1.6. In Section 5, we prove Theorem 1.10. Throughout the paper we will omit floors and ceilings when they do not affect the proofs.

2 Proof of Theorem 1.2

In this section we prove Theorem 1.2. Let us show some preliminary results first.

2.1 Preliminaries

For a k-graph $\mathcal{H}$ and $i \in [k-1]$ the *i-th shadow* of $\mathcal{H}$ is

$$\partial_i \mathcal{H} = \left\{ A \in \binom{V(\mathcal{H})}{k-i} : \exists E \in \mathcal{H} \text{ such that } A \subset E \right\}.$$

The *shadow* of $\mathcal{H}$ is $\partial \mathcal{H} = \partial_1 \mathcal{H}$. For a set $S \subset V(\mathcal{H})$ the *neighborhood* of S in $\mathcal{H}$ is

$$N_{\mathcal{H}}(S) = \{ v \in V(\mathcal{H}) \setminus S : \exists E \in \mathcal{H} \text{ such that } S \cup \{v\} \subset E \},$$

the *link* of S in $\mathcal{H}$ is

$$L_{\mathcal{H}}(S) = \{ E \setminus S : E \in \mathcal{H} \text{ and } S \subset E \},$$

and $d_{\mathcal{H}}(S) = |L_{\mathcal{H}}(S)|$ is the *degree* of S in $\mathcal{H}$. For $i \in [k-1]$ the *maximum i-degree* of $\mathcal{H}$ is

$$\Delta_i(\mathcal{H}) = \max \left\{ d_{\mathcal{H}}(A) : A \in \binom{V(\mathcal{H})}{i} \right\},$$

and note that $\Delta(\mathcal{H}) = \Delta_1(\mathcal{H})$.

For a pair of distinct vertices $u, v \in V(\mathcal{H})$ the $(k-1)$-*codegree* of u and v is the number of $(k-1)$-sets $S \subset V(\mathcal{H})$ such that $S \cup \{u\} \in \mathcal{H}$ and $S \cup \{v\} \in \mathcal{H}$. Denoted by $\Gamma(\mathcal{H})$ the maximum $(k-1)$-codegree of $\mathcal{H}$.

The random greedy independent set algorithm. We begin with $\mathcal{H}(0) = \mathcal{H}, V(0) = V(\mathcal{H})$ and $I(0) = \emptyset$. Given independent set $I(i)$ and hypergraph $\mathcal{H}(i)$ on vertex set $V(i)$, a vertex $v \in V(i)$ is chosen uniformly at random and added to $I(i)$ to form $I(i+1)$. The vertex set $V(i+1)$ is set equal to $V(i)$ less v and all vertices u such that $\{u, v\}$ is an edge in $\mathcal{H}(i)$. The hypergraph $\mathcal{H}(i+1)$ is formed form $\mathcal{H}_i$ by

1. removing v from all edges of size at least three in $\mathcal{H}(i)$ that contain v, and
2. removing every edge that contains a vertex u such that the pair $\{u, v\}$ is an edge of $\mathcal{H}(i)$.

The process terminates when $V(i) = \emptyset$. At this point $I(i)$ is a maximal independent set in $\mathcal{H}$. Let $i_{\max}$ denote the step where the algorithm terminates.

In [5], Bennett and Bohman analyzed the random greedy independent set algorithm using the differential equation method, and they proved that if a k-graph satisfies certain degree and codegree conditions, then the random greedy independent set algorithm produces a large independent set with high probability.

Theorem 2.1 (Bennett–Bohman [5])**.** *Let k and $\epsilon > 0$ be fixed. Let $\mathcal{H}$ be a D-regular k-graph on n vertices such that $D > n^{\epsilon}$. If*

$$\Delta_i(\mathcal{H}) < D^{\frac{k-i}{k-1} - \epsilon} \quad \text{for} \quad 2 \le i \le k-1, \quad \text{and} \quad \Gamma(\mathcal{H}) < D^{1-\epsilon},$$

then the random greedy independent set algorithm produces an independent set I in $\mathcal{H}$ of size $\Omega\left((\log n)^{1/(k-1)} \cdot n / D^{1/(k-1)} \right)$ with probability $1 - o(1)$.

The lower bound on independence number in Theorem 2.1 can easily be proved by applying a theorem of Duke-Lefmann-Rödl [9] (see Theorem 2.3), so the main novelty of Theorem 2.1 is the fact that the random greedy independent set algorithm produces an independent set of this size with high probability.

Let $S \subset V(\mathcal{H})$ be a set of bounded size s such that S contains no edge in $\mathcal{H}$. A nice property of the random greedy independent set algorithm is that S is contained in the set $I(i)$ with probability $(1+o(1))\,(i/n)^s$, which is almost the probability that S is contained in a random i-subset of $V(\mathcal{H})$.

Using this property we can easily control the size of the induced subgraph of $\mathcal{G}$ on $I(i)$, where $\mathcal{G}$ is a hypergraph that has the same vertex set with $\mathcal{H}$.

Proposition 2.2 (Bennett–Bohman [5]). *Let $\mathcal{H}$ be a hypergraph that satisfies the conditions in Theorem 2.1 and $\mathcal{G}$ be a k'-graph on $V(\mathcal{H})$ (i.e. $\mathcal{G}$ and $\mathcal{H}$ are on the same vertex set). If $i \le i_{\max}$ is fixed, then the expected number of edges of $\mathcal{G}$ contained in $I(i)$ is at most $(1+o(1))\,(i/n)^{k'} \cdot |\mathcal{G}|$.*

For $2 \le j \le k-1$ and two edges E, E' in a k-graph $\mathcal{H}$ we say $\{E, E'\}$ is a $(2,j)$-*cycle* if $|E \cap E'| = j$. Denote by $C_{\mathcal{H}}(2,j)$ the number of $(2,j)$-cycles in $\mathcal{H}$. Duke, Lefmann, and Rödl [9] proved the following result for hypergraphs with few $(2,j)$-cycles.

Theorem 2.3 (Duke–Lefmann–Rödl [9]). *Let $\mathcal{H}$ be a k-graph on n vertices satisfying $\Delta(\mathcal{H}) \le t^{k-1}$, where $t \gg k$. If $C_{\mathcal{H}}(2,j) \le nt^{2k-j-1-\epsilon}$ for $2 \le j \le k-1$ and some constant $\epsilon > 0$, then $\alpha(\mathcal{H}) \ge c(k,\epsilon)\,(\log t)^{1/(k-1)} \cdot n/t$.*

Recall that a hypergraph is *linear* if every pair of edges has at most one vertex in common. It is easy to see that $\mathcal{H}$ is linear iff $C_{\mathcal{H}}(2,j) = 0$ for $2 \le j \le k-1$. The following easy corollary of Theorem 2.3 will be handy for proofs in the next section.

Corollary 2.4 (see e.g. [13]). *Suppose that $\mathcal{H}$ is a linear k-graph with n vertices and average degree d. Then $\alpha(\mathcal{H}) = \Omega\left((\log d)^{1/(k-1)} \cdot n/d^{1/(k-1)}\right)$.*

For a (not necessarily uniform) hypergraph $\mathcal{H}$ on n vertices (assuming that $V(\mathcal{H}) = [n]$) and a family $\mathcal{F} = \{\mathcal{G}_1, \ldots, \mathcal{G}_n\}$ of m-vertex k-graphs with $V(\mathcal{G}_1) = \cdots = V(\mathcal{G}_n) = V_{\mathcal{F}}$ the *Cartesian product* of $\mathcal{H}$ and $\mathcal{F}$, denoted by $\mathcal{H}\Box\mathcal{F}$, is a hypergraph on $V(\mathcal{H}) \times V_{\mathcal{F}}$ and

$$\mathcal{H}\Box\mathcal{F} = \{(E,v) : E \in \mathcal{H} \text{ and } v \in V_{\mathcal{F}}\} \cup \{(i,F) : i \in [n] \text{ and } F \in \mathcal{G}_i\}.$$

Since the hypergraphs we considered here are not necessarily regular, Theorem 2.1 cannot be applied directly to our situations. To overcome this issue we use an adaption of a trick used by Shearer in [22], that is, for every nonregular hypergraph $\mathcal{H}$ we take the Cartesian product of $\mathcal{H}$ and a family of linear hypergraphs to get a new hypergraph $\widehat{\mathcal{H}}$ that is regular. Then we apply Theorem 2.1 to $\widehat{\mathcal{H}}$ to get a large independent set, and by the Pigeonhole principle, this ensures that $\mathcal{H}$ has a large independent set.

First, we need the following theorem to show the existence of sparse regular linear hypergraphs.

Given two k-graphs $\mathcal{H}_1$ and $\mathcal{H}_2$ with the same number of vertices a *packing* of $\mathcal{H}_1$ and $\mathcal{H}_2$ is a bijection $\phi : V(\mathcal{H}_1) \to V(\mathcal{H}_2)$ such that $\phi(E) \notin \mathcal{H}_2$ for all $E \in \mathcal{H}_1$.

Theorem 2.5 (Lu–Székely [18]). *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two k-graphs on n vertices. If*

$$\Delta(\mathcal{H}_1)|\mathcal{H}_2| + \Delta(\mathcal{H}_2)|\mathcal{H}_1| < \frac{1}{ek}\binom{n}{k},$$

then there is a packing of $\mathcal{H}_1$ and $\mathcal{H}_2$.

Theorem 2.5 enables us to construct sparse regular linear hypergraphs inductively.

Lemma 2.6. *For every positive integer n that satisfies $k \mid n$ and every positive integer d that satisfies*

$$d \le \frac{(n-k+2)(n-k+1)}{ek^2(k-1)^2 n} + 1,$$

there exists a d-regular linear k-graph with n vertices.

Proof of Lemma 2.6. We proceed by induction on d and note that the case $d = 1$ is trivial since a perfect matching on n vertices is a 1-regular linear k-graph. Now suppose that $d \ge 2$. By the induction hypothesis, there exists a $(d-1)$-regular linear k-graph on n vertices, and let $\mathcal{H}_{d-1}$ be such a k-graph. Let $\mathcal{H}_1$ be a perfect matching on n vertices. Define the extended k-graph $\widehat{\mathcal{H}}_1$ of $\mathcal{H}_1$ as

$$\widehat{\mathcal{H}}_1 = \left\{ \{u,v\} \cup A : \{u,v\} \in \partial_{k-2}\mathcal{H}_1 \text{ and } A \in \binom{V(\mathcal{H}_1) \setminus \{u,v\}}{k-2} \right\}.$$

It is clear from the definition that $\mathcal{H}_1 \subset \widehat{\mathcal{H}}_1$, $|\widehat{\mathcal{H}}_1| < \frac{n}{k}\binom{k}{2}\binom{n}{k-2}$, and $\widehat{\mathcal{H}}_1$ is regular. So,

$$\Delta(\widehat{\mathcal{H}}_1) = \frac{k|\widehat{\mathcal{H}}_1|}{n} < \frac{k}{n}\frac{n}{k}\binom{k}{2}\binom{n}{k-2} = \binom{k}{2}\binom{n}{k-2}.$$

By assumption

$$\begin{aligned}\Delta(\mathcal{H}_{d-1})|\widehat{\mathcal{H}}_1| + \Delta(\widehat{\mathcal{H}}_1)|\mathcal{H}_{d-1}| &< (d-1)\frac{n}{k}\binom{k}{2}\binom{n}{k-2} + \frac{(d-1)n}{k}\binom{k}{2}\binom{n}{k-2} \\ &= 2(d-1)\frac{n}{k}\binom{k}{2}\binom{n}{k-2} \le \frac{1}{ek}\binom{n}{k}.\end{aligned}$$

Therefore, by Theorem 2.5, there exist a bijection $\phi : V(\mathcal{H}_{d-1}) \to V(\mathcal{H}_1)$ such that $|\phi(E) \cap E'| \le k-1$ for all $E \in \mathcal{H}_{d-1}$ and all $E' \in \widehat{\mathcal{H}}_1$, and this implies that $|\phi(E) \cap E''| \le 1$ for all $E \in \mathcal{H}_{d-1}$ and all $E'' \in \mathcal{H}_1$. Therefore, $\mathcal{H}_1 \cup \phi(\mathcal{H}_{d-1})$ is a d-regular linear k-graph on n vertices. ∎

2.2 Proofs

First we use Theorem 2.1 and Proposition 2.2 to prove a result about the common independent set of two hypergraphs on the same vertex set.

Theorem 2.7. *Let $k_1, k_2 \ge 2$ be integers, $\epsilon > 0$, $n, D \in \mathbb{N}$, and $d > 0$. Suppose that*

(a) $\mathcal{H}$ is an n-vertex k_1-graph, $\mathcal{G}$ is an n-vertex k_2-graph, and $V(\mathcal{H}) = V(\mathcal{G}) = V$,

(b) $D > n^{\epsilon}$ and $d(\log n/D)^{\frac{k_2-1}{k_1-1}} \gg 1$,

(c) $\mathcal{H}$ *satisfies that* $\Delta(\mathcal{H}) \le D$,

$$\Delta_i(\mathcal{H}) < D^{\frac{k_1-i}{k_1-1}-\epsilon} \quad \text{for} \quad 2 \le i \le k_1 - 1, \quad \text{and} \quad \Gamma(\mathcal{H}) < D^{1-\epsilon},$$

(d) $\mathcal{G}$ *satisfies that* $d(\mathcal{G}) \le d$ *and*

$$C_{\mathcal{G}}(2,i) \ll n\,(D/\log n)^{\frac{2k_2-i-1}{k_1-1}} \quad \text{for} \quad 2 \le i \le k_2 - 1.$$

Then, $\alpha(\mathcal{H} \cup \mathcal{G}) = \Omega\left(\omega \cdot n/d^{1/(k_2-1)}\right)$, *where*

$$\omega = \omega(n, D, d, k_1, k_2) = \left(\log\left((\log n/D)^{\frac{k_2-1}{k_1-1}}\, d\right)\right)^{1/(k_2-1)}.$$

Remarks.

- Although Theorem 2.7 imposes no condition on k_1 and k_2, we will only apply the result in the case $k_2 = k_1 + 1$.
- Spencer's bound (1) implies that $\alpha(\mathcal{G}) = \Omega\left(n/d^{1/(k_2-1)}\right)$. Theorem 2.7 improves it in two ways: first it improves the bound by a factor of ω, second it is a lower bound for the independence number of $\mathcal{G} \cup \mathcal{H}$. Ajtai, Komlós, Pintz, Spencer, and Szemerédi's result (2) implies that the upper bound for ω is $(\log n)^{1/(k_2-1)}$. However, we are not able to show that $\omega = \Omega\left((\log n)^{1/(k_2-1)}\right)$ in general, and it would be interesting to determine the optimal value of ω.
- If $\mathcal{H}$ and $\mathcal{G}$ satisfy conditions (a) and (c) in Theorem 2.7 and also satisfy

  (b') $D > n^{\epsilon}$ and $d\,(\log n/D)^{\frac{k_2-1}{k_1-1}} \ll 1$,

  then $\alpha(\mathcal{H} \cup \mathcal{G}) = \Omega\left((\log n)^{1/(k_1-1)} \cdot n/D^{1/(k_1-1)}\right)$. Moreover, if $\mathcal{G} = \emptyset$, then $\alpha(\mathcal{H}) = \Omega\left((\log n)^{1/(k_1-1)} \cdot n/D^{1/(k_1-1)}\right)$ which is the bound in Theorem 2.1. The proof is similar to the proof of Theorem 2.7.

Proof of Theorem 2.7. For $2 \le i \le k_2 - 1$ define

$$\mathcal{G}^i = \left\{ S \in \binom{V}{2k_2 - i} : \mathcal{G}[S] \text{ contains a } (2,i)\text{-cycle} \right\}.$$

Fix $m \in \mathbb{N}$ such that $D \ll m = O(n^{k_1})$, and $k_1 \mid m$. Notice that D has a trivial upper bound n^{k_1-1}, so such an integer m exists. For every $v \in V$ let $D_v = D - d_{\mathcal{H}}(v)$. Since $m \gg D$ and $k_1 \mid m$, by Lemma 2.6, there exists a D_v-regular linear k_1-graph $\mathcal{F}(v)$ on $[m]$ for every $v \in V$. Let

$$\mathcal{H}' = \mathcal{H} \cup \mathcal{G} \cup \left(\bigcup_{2 \le i \le k_2 - 1} \mathcal{G}^i \right), \quad \mathcal{F} = \{\mathcal{F}(v) : v \in V\}, \quad \text{and} \quad \widehat{\mathcal{H}'} = \mathcal{H}' \square \mathcal{F}.$$

Note that $\widehat{\mathcal{H}'}$ is consisting of

1. the k_1-graph $\widehat{\mathcal{H}} = \mathcal{H} \square \mathcal{F}$,
2. the k_2-graph $\widehat{\mathcal{G}}$ that is the union of m pairwise vertex-disjoint copies of $\mathcal{G}$, and

3. the $(2k_2-i)$-graph $\widehat{\mathcal{G}}^i$ that is the union of m pairwise vertex-disjoint copies of $\mathcal{G}^i$ for $2 \le i \le k_2-1$.

For every $v \in V(\widehat{\mathcal{H}})$ we have $d_{\widehat{\mathcal{H}}}(v) = d_{\mathcal{H}}(v) + D_v = D$. Moreover,

$$\Delta_i(\widehat{\mathcal{H}}) = \Delta_i(\mathcal{H}) < D^{\frac{k_1-i}{k_1-1}-\epsilon} \quad \text{for} \quad 2 \le i \le k_1-1, \quad \text{and} \quad \Gamma(\widehat{\mathcal{H}}) = \Gamma(\mathcal{H}) < D^{1-\epsilon},$$

Applying the random greedy independent set algorithm and Theorem 2.1 to $\widehat{\mathcal{H}}$, we obtain an independent set $\hat{I}$ of size at least $c\,(\log nm)^{1/(k_1-1)} \cdot nm/D^{1/(k_1-1)}$ for some constant $c > 0$ with probability $1-o(1)$. Let $p = c\,((\log nm)/D)^{1/(k_1-1)}$ and we may assume that $|\hat{I}| = pnm$ since otherwise we can take the set of the first pnm vertices generated by the random greedy independent set algorithm instead.

Applying Proposition 2.2 to $\widehat{\mathcal{G}}, \widehat{\mathcal{G}}^2, \ldots, \widehat{\mathcal{G}}^{k_2-1}$ and by assumption (d) we obtain

$$\mathbb{E}\left[\left|\widehat{\mathcal{G}}[\hat{I}]\right|\right] \le (1+o(1))p^{k_2}|\widehat{\mathcal{G}}| < 2dnmp^{k_2},$$

and for $2 \le i \le k_2-1$

$$\mathbb{E}\left[\left|\widehat{\mathcal{G}}^i[\hat{I}]\right|\right] = (1+o(1))p^{2k_2-i} \cdot m \cdot C_{\mathcal{G}}(2,i) = o(pnm).$$

So, by Markov's inequality and the union bound, with probability at least 1/2 both

$$\left|\widehat{\mathcal{G}}[\hat{I}]\right| \le 10dnmp^{k_2} \quad \text{and} \quad \left|\widehat{\mathcal{G}}^i[\hat{I}]\right| = o(pnm) \quad \forall\, 2 \le i \le k_2-1$$

hold.

Fix a set $\hat{I}$ such that $|\hat{I}| = pnm$ and the events above hold. Then by removing $o(pnm)$ vertices we obtain a subset $\hat{I}' \subset \hat{I}$ such that

$$\left|\widehat{\mathcal{G}}^i[\hat{I}]\right| = 0 \quad \text{for} \quad 2 \le i \le k_2-1.$$

In other words, the k_2-graph $\widehat{\mathcal{G}}[\hat{I}']$ is linear. Since

$$d\left(\widehat{\mathcal{G}}[\hat{I}']\right) \le \frac{k_2 \cdot 10dnmp^{k_2}}{(1-o(1))pnm} \le 20k_2dp^{k_2-1},$$

by Corollary 2.4, it has an independent set I' of size at least

$$\Omega\left(\frac{pnm}{(20k_2dp^{k_2-1})^{1/(k_2-1)}}\left(\log 20dp^{k_2-1}\right)^{\frac{1}{k_2-1}}\right) = \Omega\left(m\frac{n}{d^{1/(k_2-1)}}\left(\log p^{k_2-1}d\right)^{\frac{1}{k_2-1}}\right)$$
$$= \Omega\left(m\frac{n}{d^{1/(k_2-1)}}\omega\right).$$

Here we used assumption (b) to ensure that $20k_2dp^{k_2-1} \ge 1$.

By the Pigeonhole principle, there exists $j \in [m]$ such that $I = I' \cap (V \times \{j\})$ has size at least $|\hat{I}|/m = \Omega\left(\omega \cdot n/d^{1/(k_2-1)}\right)$, and it is clear that I is an independent set in both $\mathcal{H}$ and $\mathcal{G}$. ∎

Next we use Theorem 2.7 to prove Theorem 1.2. The idea is to first decompose an $(n,k,k-2)$-omitting system $\mathcal{H}$ into two parts: $\mathcal{H}_{k-1} \subset \partial\mathcal{H}$ and $\mathcal{H}_k \subset \mathcal{H}$, and then apply Theorem 2.7 to $\mathcal{H}_{k-1}$ and $\mathcal{H}_k$ to find a large set $I \subset V$ that is independent in both of them. It will be easy to see that the set I is independent in $\mathcal{H}$.

Proof of Theorem 1.2. Let $\mathcal{H}$ be an $(n,k,k-2)$-omitting system and let $V = V(\mathcal{H})$. By Theorem 1.1, there exists a constant C_1 such that $|\mathcal{H}| \le C_1 n^{k-2}$. Let $\beta = \beta(k) > 0$ be a constant such that $\frac{k}{2(k-1)} < \beta < 1$, for example, take $\beta = 4/5$. Define

$$\mathcal{H}_{k-1} = \left\{ A \in \partial\mathcal{H} : d_{\mathcal{H}}(A) \ge \frac{n^{\frac{k-3}{k-1}}}{(\log n)^\beta} \right\} \quad \text{and} \quad \mathcal{H}_k = \left\{ E \in \mathcal{H} : \binom{E}{k-1} \cap \mathcal{H}_{k-1} = \emptyset \right\}.$$

Let $k_1 = k-1$, $k_2 = k$, $D = n^{k-4+2/(k-1)}(\log n)^\beta$, $d = C_1 n^{k-3}$, and ϵ be a constant such that $0 < \epsilon < 1/(k-1)$. Then $D > n^\epsilon$ and

$$d\left(\frac{\log n}{D}\right)^{\frac{k_2-1}{k_1-1}} = C_1 n^{k-3}\left(\frac{\log n}{n^{k-4+\frac{2}{k-1}}(\log n)^\beta}\right)^{\frac{k-1}{k-2}} = C_1 (\log n)^{(1-\beta)\frac{k-1}{k-2}} \gg 1.$$

Therefore, condition (b) in Theorem 2.7 is satisfied. Next we show that $\mathcal{H}_{k-1}$ and $\mathcal{H}_k$ satisfy (c) and (d) in Theorem 2.7 with our choice of k_1, k_2, D, d, ϵ.

Claim 2.8. *The $(k-1)$-graph $\mathcal{H}_{k-1}$ is an $(n,k-1,k-2)$-system with $\Delta_{k-3}(\mathcal{H}_{k-1}) \le n^{2/(k-1)}(\log n)^\beta$.*

Proof of Claim 2.8. First we prove that $\mathcal{H}_{k-1}$ is an $(n,k-1,k-2)$-system. Indeed, suppose to the contrary that there exist $e_1, e_2 \in \mathcal{H}_{k-1}$ such that $S = e_1 \cap e_2$ has size $k-2$. By the definition of $\mathcal{H}_{k-1}$, $|N_{\mathcal{H}}(e_i)| \ge n^{\frac{k-3}{k-1}}/(\log n)^\beta > 2k$ for $i = 1,2$. So there exist $v_1, v_2 \in V \setminus (e_1 \cup e_2)$ such that $E_i = e_i \cup \{v_i\} \in \mathcal{H}$ for $i = 1,2$. However, $E_1 \cap E_2 = S$ has size $k-2$, contradicting the assumption that $\mathcal{H}$ is an $(n,k,k-2)$-omitting system. Therefore, $\mathcal{H}_{k-1}$ is an $(n,k-1,k-2)$-system.

Now suppose to the contrary that there exists a set $A \subset V$ of size $k-3$ with $d_{\mathcal{H}_{k-1}}(A) = m > n^{2/(k-1)}(\log n)^\beta$. Since $\mathcal{H}_{k-1}$ is an $(n,k-1,k-2)$-system, $L_{\mathcal{H}_{k-1}}(A)$ is a matching consisting of m edges. Suppose that $L_{\mathcal{H}_{k-1}}(A) = \{e_1, \dots, e_m\}$, and let $B_i = A \cup e_i$ for $1 \le i \le m$. Since $B_i \in \mathcal{H}_{k-1}$, by definition, there exists a set $N_i \subset V$ of size at least $n^{\frac{k-3}{k-1}}/(\log n)^\beta$ such that $B_i \cup \{u\} \in \mathcal{H}$ for all $u \in N_i$.

Suppose that there exists $v \in N_i \cap N_j$ for some distinct $i,j \in [m]$. Then the two sets $A \cup e_i \cup \{v\}$ and $A \cup e_j \cup \{v\}$ are edges in $\mathcal{H}$ and have an intersection of size $k-2$, a contradiction. Therefore, $N_i \cap N_j = \emptyset$ for all distinct $i,j \in [m]$. It follows that

$$n = |V| \ge \sum_{i\in[m]} |N_i| \ge m n^{\frac{k-3}{k-1}}/(\log n)^\beta > n^{\frac{2}{k}}(\log n)^\beta n^{\frac{k-3}{k-1}}/(\log n)^\beta > n,$$

a contradiction. Therefore, $\Delta_{k-3}(\mathcal{H}_{k-1}) \le n^{2/(k-1)}(\log n)^\beta$. ∎

Since $\Delta_{k-3}(\mathcal{H}_{k-1}) \le n^{2/(k-1)}(\log n)^\beta$, for every set $S \subset V$ of size i with $i \in [k-4]$ the link $L_{\mathcal{H}_{k-1}}(S)$ is an $\big(n, k-1-i, k-3-i, n^{2/(k-1)}(\log n)^\beta\big)$-system. Therefore, for $i \in [k-4]$

$$\Delta_i(\mathcal{H}_{k-1}) \le n^{\frac{2}{k-1}}(\log n)^\beta \binom{n}{k-3-i} \Big/ \binom{k-1-i}{k-3-i} < n^{k-3-i+\frac{2}{k-1}}(\log n)^\beta.$$

Since

$$\left(k-4+\frac{2}{k-1}\right)\frac{k-1-i}{k-2} - \left(k-3-i+\frac{2}{k-1}\right) = \frac{2(i-1)}{k-1} > \epsilon,$$

we obtain

$$\Delta_i(\mathcal{H}_{k-1}) < n^{k-3-i+\frac{2}{k-1}}(\log n)^\beta < D^{\frac{k-1-i}{k-1-1}-\epsilon} \quad \text{for} \quad 2 \le i \le k-3.$$

On the other hand, since $\mathcal{H}$ is an $(n, k-1, k-2)$-system, $\Delta_{k-2}(\mathcal{H}_{k-1}) \le 1 < D^{\frac{k-1-(k-2)}{k-1-1}-\epsilon}$ and $\Gamma(\mathcal{H}_{k-1}) = 0 < D^{1-\epsilon}$. Therefore, $\mathcal{H}_{k-1}$ satisfies condition (c) in Theorem 2.7.

Claim 2.9. *The k-graph $\mathcal{H}_k$ satisfies $d(\mathcal{H}_k) \le C_1 k n^{k-2}$,*

$$C_{\mathcal{H}_k}(2, i) = O\left(n^{2k-4-i}\right) \quad \text{for} \quad 2 \le i \le k-3,$$

$C_{\mathcal{H}_k}(2, k-2) = 0$, *and* $C_{\mathcal{H}_k}(2, k-1) = O\left(n^{k-2+\frac{k-3}{k-1}}/(\log n)^\beta\right)$.

Proof of Claim 2.9. First, it is clear that $C_{\mathcal{H}_k}(2, k-2) = 0$ since there is no pair of edges in $\mathcal{H}_k$ with an intersection of size $k-2$.

Let $2 \le i \le k-3$ and $S \subset V$ be a set of size i. Since $\mathcal{H}_k$ is an $(n, k, k-2)$-omitting system, the link $L_{\mathcal{H}_k}(S)$ is an $(n, k-i, k-2-i)$-omitting system. So, by Theorem 1.1, $|L_{\mathcal{H}_k}(S)| = O\left(n^{k-2-i}\right)$, which implies that

$$C_{\mathcal{H}_k}(2, i) \le |\mathcal{H}_k| \cdot \binom{k}{i} \cdot O\left(n^{k-2-i}\right) = O\left(n^{2k-4-i}\right) \quad \text{for} \quad 2 \le i \le k-3.$$

Now let $S \subset V$ be a set of size $k-1$. By the definition of $\mathcal{H}_k$, $d_{\mathcal{H}_k}(S) \le n^{2/(k-1)}/(\log n)^\beta$. Therefore,

$$C_{\mathcal{H}_k}(2, k-1) \le |\mathcal{H}_k| \cdot \binom{k}{k-1} \cdot n^{\frac{k-3}{k-1}}/(\log n)^\beta = O\left(n^{k-2+\frac{k-3}{k-1}}/(\log n)^\beta\right).$$

∎

Since

$$1 + \left(k - 4 + \frac{2}{k-1}\right)\frac{2k-1-i}{k-2} - (2k-4-i) = \frac{2(i-1)}{k-1} > \epsilon,$$

by Claim 2.9,

$$C_{\mathcal{H}_k}(2, i) = O\left(n^{2k-4-i}\right) = o\left(n\,(D/\log n)^{\frac{2k-i-1}{k-1-1}}\right) \quad \text{for} \quad 2 \le i \le k-3.$$

Moreover, $C_{\mathcal{H}_k}(2, k-2) = 0 \ll n\,(D/\log n)^{\frac{2k-(k-2)-1}{k-1-1}}$, and

$$C_{\mathcal{H}_k}(2, k-1) = O\left(\frac{n^{k-2+\frac{k-3}{k-1}}}{(\log n)^\beta}\right) \ll \frac{n^{k-2+\frac{k-3}{k-1}}}{(\log n)^{(1-\beta)\frac{k}{k-2}}} = n\left(\frac{D}{\log n}\right)^{\frac{2k-(k-1)-1}{k-1-1}},$$

where the inequality follows from the assumption that $\beta > \frac{k}{2(k-1)}$. Therefore, $\mathcal{H}_k$ satisfies condition (d) in Theorem 2.7.

So, by Theorem 2.7, there exists a set $I \subset V$ of size $\Omega\left(\omega \cdot n/n^{\frac{k-3}{k-1}}\right) = \Omega\left(n^{2/(k-1)}\omega\right)$ such that I is independent in both $\mathcal{H}_{k-1}$ and $\mathcal{H}_k$. Here

$$\begin{aligned}\omega = \left(\log\left(((\log n)/D)^{\frac{k_2-1}{k_1-1}}\,d\right)\right)^{1/(k_2-1)} &= \left(\log\,(\log n)^{(1-\beta)\frac{k-1}{k-2}}\right)^{1/(k-1)} \\ &= \Omega\left((\log\log n)^{1/(k-1)}\right).\end{aligned}$$

∎

3 Proof of Theorem 1.4

3.1 Lower bound

We prove the lower bound in Theorem 1.4 in this section. The proof idea is similar to that used in the proof of Theorem 1.2, that is, we decompose an (n,k,ℓ)-omitting system into many different hypergraphs so that each hypergraph contains the information of a certain subset of edges in the original hypergraph. Then we use a probabilistic argument to show that there exists a large common independent set of these hypergraphs.

Recall that an n-vertex k-graph $\mathcal{H}$ is an (n,k,ℓ,λ)-omitting system iff it is $S_{\lambda+1}(\ell)$-free. While Theorem 1.4 as stated provides a lower bound on the independence number of (n,k,ℓ)-omitting systems, the result holds in the more general setting of (n,k,ℓ,λ)-omitting systems. We present the proof in this more general setting.

Let $k \geq k_0 > \ell \geq 1$, $\lambda \geq 2$, and $\mathcal{H}$ be an $S_\lambda(\ell)$-free k-graph. We say $\mathcal{H}$ is (k_0,λ)-*indecomposable* if

- $k = k_0$, or
- $k > k_0$ and $\mathcal{H}$ is $\{S_{\lambda_1}(k-1),\ldots,S_{\lambda_{k-k_0}}(k_0)\}$-free, where $\lambda_i = (k\lambda)^{2^{i-1}}$ for $i \in [k-k_0]$.

Otherwise, we say $\mathcal{H}$ is (k_0,λ)-*decomposable*.

Call a family $\mathcal{F}$ of hypergraphs (k_0,λ)-indecomposable if every member in it is (k_0,λ)-indecomposable. Otherwise, we say $\mathcal{F}$ is (k_0,λ)-decomposable.

The decomposition algorithm.
Input: An $S_\lambda(\ell)$-free k-graph $\mathcal{H}$ and a threshold k_0 with $k \geq k_0 > \ell$.
Output: A family $\mathcal{F}$ of $S_\lambda(\ell)$-free (k_0,λ)-indecomposable hypergraphs.
Operation: We start with the family $\mathcal{F} = \{\mathcal{H}\}$. If $\mathcal{F}$ is (k_0,λ)-indecomposable, then we terminate this algorithm. Otherwise, let $\mathcal{G} \in \mathcal{F}$ be a (k_0,λ)-decomposable hypergraph and let k' denote the size of each edge in $\mathcal{G}$. Let $i_0 \in \{1,\ldots,k'-k_0\}$ be the smallest integer such that $\mathcal{G}$ contains a copy of $S_{\lambda_{i_0}}(k'-i_0)$, where $\lambda_{i_0} = (k\lambda)^{2^{i_0-1}}$. Define

$$\mathcal{G}_{k'-i_0} = \left\{A \in \binom{V(\mathcal{H})}{k'-i_0} : d_{\mathcal{G}}(A) \geq \lambda_{i_0}\right\} \quad \text{and} \quad \mathcal{G}_{k'} = \left\{B \in \mathcal{G} : \binom{B}{k'-i_0} \cap \mathcal{G}_{k'-i_0} = \emptyset\right\}.$$

Update $\mathcal{F}$ by removing $\mathcal{G}$ and adding $\mathcal{G}_{k'-i_0}$ and $\mathcal{G}_{k'}$. Repeat this operation until $\mathcal{F}$ is (k_0,λ)-indecomposable.

We need the following lemmas to show that the algorithm defined above always terminates. Write $\nu(\mathcal{H})$ for the size of a maximum matching in $\mathcal{H}$.

Lemma 3.1. *Let $\mathcal{H}$ be an $\{S_{\lambda_1}(k-1),\ldots,S_{\lambda_{k-1}}(1)\}$-free k-graph with m edges. Then*

$$\nu(\mathcal{H}) \geq \frac{m}{\prod_{i=1}^{k-1}(i+1)\lambda_i}.$$

Proof of Lemma 3.1. For $j \in [k-1]$ let $\Lambda_j = \prod_{i=1}^{j}(i+1)\lambda_i$. We prove this lemma by induction on k. Suppose that $k=2$. Since $\mathcal{H}$ is $S_{\lambda_1}(1)$-free, $d_{\mathcal{H}}(v) \leq \lambda_1 - 1$ for all $v \in V(\mathcal{H})$. Therefore, by greedily choosing an edge e and removing all edges that have nonempty intersection with e, we obtain at least $m/(2\lambda_1)$ pairwise disjoint edges in $\mathcal{H}$.

Now suppose that $k \geq 3$. We claim that $d_{\mathcal{H}}(v) \leq (\lambda_{k-1}-1)\Lambda_{k-2}$ for all $v \in V(\mathcal{H})$. Indeed, suppose to the contrary that there exists $v_0 \in V(\mathcal{H})$ with $d_{\mathcal{H}}(v_0) \geq (\lambda_{k-1}-1)\Lambda_{k-2}+1$. Since $\mathcal{H}$ is $\left\{S_{\lambda_1}(k-1), \ldots, S_{\lambda_{k-2}}(2)\right\}$-free, the link $L_{\mathcal{H}}(v_0)$ is $\left\{S_{\lambda_1}(k-2), \ldots, S_{\lambda_{k-2}}(1)\right\}$-free. By the induction hypothesis,

$$\nu(L_{\mathcal{H}}(v_0)) \geq \frac{(\lambda_{k-1}-1)\Lambda_{k-2}+1}{\Lambda_{k-2}} > \lambda_{k-1}-1,$$

but this contradicts the assumption that $\mathcal{H}$ is $S_{\lambda_{k-1}}(1)$-free. Therefore, $d_{\mathcal{H}}(v) \leq (\lambda_{k-1} - 1)\Lambda_{k-2}$ for all $v \in V(\mathcal{H})$. Then, similar to the case of $k=2$, by greedily choosing an edge e and removing all edges that have nonempty intersection with e, we obtain

$$\nu(\mathcal{H}) \geq \frac{m}{k(\lambda_{k-1}-1)\Lambda_{k-2}+1} > \frac{m}{\Lambda_{k-1}}$$

completing the proof. ∎

Let $\mathcal{H}$ be an $S_\lambda(\ell)$-free k-graph. Define

$$\mathcal{H}_{k-1} = \left\{A \in \binom{V(\mathcal{H})}{k-1} : d_{\mathcal{H}}(A) \geq k\lambda\right\}.$$

If $\mathcal{H}$ is $\left\{S_{\lambda'_1}(k-1), \ldots, S_{\lambda'_{k-k'-1}}(k'+1), S_\lambda(\ell)\right\}$-free for some $\ell < k' \leq k-2$, then also define

$$\mathcal{H}_{k'} = \left\{A \in \binom{V}{k'} : d_{\mathcal{H}}(A) \geq k\lambda \prod_{i=1}^{k-k'-1}(i+1)\lambda'_i\right\}.$$

Lemma 3.2. *The hypergraphs $\mathcal{H}_{k'}$ and $\mathcal{H}_{k-1}$ defined above are $S_\lambda(\ell)$-free.*

Proof of Lemma 3.2. We may only prove that $\mathcal{H}_{k'}$ is $S_\lambda(\ell)$-free, since the proof for $\mathcal{H}_{k-1}$ is basically the same. Suppose to the contrary that there exists $\{A_1, \ldots, A_\lambda\} \subset \mathcal{H}_{k'}$ forming a copy of $S_\lambda(\ell)$. Since $\mathcal{H}$ is $\{S_{\lambda'_1}(k-1), \ldots, S_{\lambda'_{k-k'-1}}(k'+1)\}$-free, the link $L_{\mathcal{H}}(A_i)$ is $\{S_{\lambda'_1}(k-k'-1), \ldots, S_{\lambda'_{k-k'-1}}(1)\}$-free for $i \in [\lambda]$. Let $\Lambda' = \prod_{i=1}^{k-k'-1}(i+1)\lambda'_i$. It follows from the definition of $\mathcal{H}_{k'}$ that $|L_{\mathcal{H}}(A_i)| \geq k\lambda\Lambda'$ for $i \in [\lambda]$. So, by Lemma 3.1, there are at least $k\lambda\Lambda'/\Lambda' \geq k\lambda$ pairwise disjoint edges in $L_{\mathcal{H}}(A_i)$ for $i \in [\lambda]$. Therefore, there exist λ pairwise disjoint $(k-k')$-sets $B_1, \ldots, B_\lambda$ such that $B_i \subset V \setminus \left(\bigcup_{i=1}^{\lambda} A_i\right)$ and $E_i = A_i \cup B_i \in \mathcal{H}$ for $i \in [\lambda]$. It is clear that $\{E_1, \ldots, E_\lambda\}$ is a copy of $S_\lambda(\ell)$ in $\mathcal{H}$, a contradiction. ∎

Recall that in the decomposition algorithm we defined

$$\mathcal{G}_{k'-i_0} = \left\{A \in \binom{V(\mathcal{H})}{k'-i_0} : d_{\mathcal{G}}(A) \geq \lambda_{i_0}\right\}, \quad \text{and} \quad \mathcal{G}_{k'} = \left\{B \in \mathcal{G} : \binom{B}{k'-i_0} \cap \mathcal{G}_{k'-i_0} = \emptyset\right\},$$

where $i_0 \in \{1, \ldots, k'-k_0\}$ is the smallest integer such that $\mathcal{G}$ contains a copy of $S_{\lambda_{i_0}}(k'-i_0)$ and $\lambda_{i_0} = (k\lambda)^{2^{i_0-1}}$. It is clear from the definition that $\mathcal{G}_{k'}$ is $S_{\lambda_{i_0}}(k'-i_0)$-free. On the other hand, Lemma 3.2 implies that both $\mathcal{G}_{k'}$ and $\mathcal{G}_{k'-i_0}$ are $S_\lambda(\ell)$-free. Therefore, the new hypergraphs $\mathcal{G}_{k'-i_0}$ and $\mathcal{G}_{k'}$ we added into $\mathcal{F}$ either have a smaller edge size (the case $\mathcal{G}_{k'-i_0}$) or forbid one more hypergraph (the case $\mathcal{G}_{k'}$). So the algorithm must terminate after finite many steps, and it is easy to see that the outputted family $\mathcal{F}$ has size at most 2^{k-k_0}. Indeed, the latter statement can be proved by associating a binary tree $T_{\mathcal{H}}$ to the algorithm: the vertex set of $T_{\mathcal{H}}$ is the collection of all hypergraphs (including $\mathcal{H}$) generated in each operation of the algorithm, the root of $T_{\mathcal{H}}$ is $\mathcal{H}$, and the children of a vertex $\mathcal{G}$ are $\mathcal{G}_{k'-i_0}$ and $\mathcal{G}_{k'}$ (if they are defined). It is easy to see that the height of $T_{\mathcal{H}}$ is at most $k-k_0$ and the outputted family $\mathcal{F}$ is the collection of hypergraphs that are leaf vertices of $T_{\mathcal{H}}$. Therefore, $|\mathcal{F}| \le 2^{k-k_0}$.

The following lemma shows that in order to find a large independent set in $\mathcal{H}$ it suffices to find a large common independent set of all hypergraphs in $\mathcal{F}$.

Lemma 3.3. *Let $\mathcal{H}$ be an $S_\lambda(\ell)$-free k-graph and $\mathcal{F}$ be the outputted family after applying the decomposition algorithm to $\mathcal{H}$. Then*

$$\alpha(\mathcal{H}) \ge \alpha\left(\bigcup_{\mathcal{G}\in\mathcal{F}} \mathcal{G}\right).$$

Proof of Lemma 3.3. Suppose that $\mathcal{F} = \{\mathcal{H}_1,\dots,\mathcal{H}_m\}$ and $I \subset V(\mathcal{H})$ is independent in $\mathcal{H}_i$ for $i \in [m]$. It is clear from the definition that for every $E \in \mathcal{H}$ there is a subset $E' \subset E$ such that $E' \in \mathcal{H}_i$ for some $i \in [m]$. Since I is independent $\mathcal{H}_i$, $E' \not\subset I$ and it follows that $E \not\subset I$. Therefore, I is independent in $\mathcal{H}$. ∎

We also need the following lemma which gives an upper bound for the size of an indecomposable hypergraph.

Theorem 3.4 (Deza–Erdős–Frankl [8])**.** *Let $r \ge 1$, $t \ge 2$ be integers and $L = \{\ell_1,\dots,\ell_r\}$ be a set of integers with $0 \le \ell_1 < \cdots < \ell_r < k$. If an n-vertex k-graph $\mathcal{H}$ is $S_t(\ell)$-free for every $\ell \in [k]\setminus L$, then $|\mathcal{H}| = O(n^{r-1})$ unless $(\ell_2-\ell_1) \mid \cdots \mid (\ell_r-\ell_{r-1}) \mid (k-\ell_r)$.*

Lemma 3.5. *Let $k \ge k_0 > \ell \ge 1$, $\lambda \ge 2$ be integers, $k > 2\ell+1$, $k_0 \ge \ell+3$, and $\mathcal{H}$ be a $S_\lambda(\ell)$-free (k_0,λ)-indecomposable k-graph with n vertices. Then there exists a constant $C_{k,\ell,\lambda}$ such that $|\mathcal{H}| \le C_{k,\ell,\lambda} n^{\min\{k_0-2,k-\ell-1\}}$.*

Proof of Lemma 3.5. Since $\mathcal{H}$ is $S_\lambda(\ell)$-free and $k > 2\ell+1$, by the results in [11], $|\mathcal{H}| = O\left(n^{k-\ell-1}\right)$. On the other hand, since $\mathcal{H}$ is $\{S_{\lambda_1}(k-1),\dots,S_{\lambda_{k-k_0}}(k_0),S_\lambda(\ell)\}$-free, applying Theorem 3.4 to $\mathcal{H}$ with $t = \max\{\lambda_1,\dots,\lambda_{k-k_0},\lambda\}$ and $L = \{0,1,\dots,\ell-1,\ell+1,\dots,k_0-1\}$ we obtain $|\mathcal{H}| = O\left(n^{k_0-2}\right)$. ∎

Now we are ready to prove the lower bound in Theorem 1.4.

Proof of the lower bound in Theorem 1.4. We may assume that $k > 3\ell$ since otherwise by (5) we are done. Let $\mathcal{H}$ be an $S_\lambda(\ell)$-free k-graph on n vertices and $V = V(\mathcal{H})$. Apply the decomposition algorithm to $\mathcal{H}$ with the threshold $k_0 = 2\ell+1$, and let $\mathcal{F}$ denote the outputted family. Suppose that $\mathcal{F} = \{\mathcal{H}_1,\dots,\mathcal{H}_m\}$ for some integer m. For $i \in [m]$ let k_i denote the size of each edge in $\mathcal{H}_i$ and note from the definition of the algorithm that $2\ell+1 \le k_i \le k$. Let $C = \max\{C_{k_i,\ell,\lambda} : 2\ell+1 \le k_i \le k\}$, where $C_{k_i,\ell,\lambda}$ is the constant given by Lemma 3.5. Choose a set $I \subset V$ such that every vertex is included in I independently with probability $p = \delta n^{-\frac{2\ell-2}{3\ell-1}}$, where $\delta > 0$ is a small constant that satisfies

$Cm\delta^{3\ell-2} \le 1/4$. Then by Lemma 3.5,

$$\begin{aligned}
\mathbb{E}\left[|I| - \sum_{i=1}^{m} |\mathcal{H}_i[I]|\right] &= \mathbb{E}[|I|] - \sum_{i=1}^{m} \mathbb{E}[|\mathcal{H}_i[I]|] \\
&\ge pn - \sum_{i=1}^{m} Cp^{k_i} n^{\min\{2\ell-1, k_i-\ell-1\}} \\
&= pn - C\left(\sum_{i\in[m]:k_i\ge 3\ell} p^{k_i} n^{2\ell-1} + \sum_{i\in[m]:k_i\le 3\ell-1} p^{k_i} n^{k_i-\ell-1}\right) \\
&\ge \delta n^{\frac{\ell+1}{3\ell-1}} - Cm\delta^{3\ell} n^{\frac{\ell+1}{3\ell-1}} - Cm\delta^{3\ell-1} n^{\frac{\ell+1}{3\ell-1}} \ge \delta n^{\frac{\ell+1}{3\ell-1}}/2.
\end{aligned}$$

Therefore, there exists a set I of size $\Omega\left(n^{\frac{\ell+1}{3\ell-1}}\right)$ such that $\mathcal{H}_i[I] = \emptyset$ for $i \in [m]$, and it follow from Lemma 3.3 that $\alpha(\mathcal{H}) \ge |I| = \Omega\left(n^{\frac{\ell+1}{3\ell-1}}\right)$. ∎

Remark. The lower bound $n^{\frac{\ell+1}{3\ell-1}}$ can be improved by optimizing the choice of k_0. Indeed, suppose that ℓ is sufficiently large. Let

$$k_0 = \left(\frac{\sqrt{5}+1}{2} + o_\ell(1)\right)\ell, \quad s = \left(\frac{\sqrt{5}+3}{2} + o_\ell(1)\right)\ell, \quad \text{and} \quad p = \delta n^{-\left(\frac{\sqrt{5}-1}{2} + o_\ell(1)\right)},$$

where $\delta > 0$ is a sufficiently small constant. Repeating the argument above we obtain

$$\begin{aligned}
\mathbb{E}\left[|I| - \sum_{i=1}^{m} |\mathcal{H}_i[I]|\right] &= \mathbb{E}[|I|] - \sum_{i=1}^{m} \mathbb{E}[|\mathcal{H}_i[I]|] \\
&= pn - \left(\sum_{k_i\ge s} \mathbb{E}[|\mathcal{H}_i[I]|] + \sum_{2\ell+1\le k_i\le s} \mathbb{E}[|\mathcal{H}_i[I]|] + \sum_{k_0<k_i\le 2\ell} \mathbb{E}[|\mathcal{H}_i[I]|]\right).
\end{aligned}$$

By Lemma 3.5, we have

$$\sum_{k_i\ge s} \mathbb{E}[|\mathcal{H}_i[I]|] \le C\sum_{k_i\ge s} p^{k_i} n^{k_0-2}.$$

By Theorem 1.1, we have

$$\sum_{2\ell+1\le k_i\le s} \mathbb{E}[|\mathcal{H}_i[I]|] \le C \sum_{2\ell+1\le k_i<s} p^{k_i} n^{k_i-\ell-1} \quad \text{and} \quad \sum_{k_0<k_i\le 2\ell} \mathbb{E}[|\mathcal{H}_i[I]|] \le C \sum_{k_0\le k_i\le 2\ell} p^{k_i} n^{\ell}.$$

Therefore,

$$\begin{aligned}
\mathbb{E}\left[|I| - \sum_{i=1}^{m} |\mathcal{H}_i[I]|\right] &\ge pn - C\left(\sum_{k_i\ge s} p^{k_i} n^{k_0-2} + \sum_{2\ell+1\le k_i<s} p^{k_i} n^{k_i-\ell-1} + \sum_{k_0\le k_i\le 2\ell} p^{k_i} n^{\ell}\right) \\
&\ge pn - Cm\left(p^s n^{k_0-2} + p^{s-1} n^{s-\ell-2} + p^{k_0} n^{\ell}\right) \\
&\ge \delta n^{\left(\frac{3-\sqrt{5}}{2} + o_\ell(1)\right)}/2,
\end{aligned}$$

which implies that $\mathcal{H}$ contains an independent set I of size $\Omega\left(n^{\left(\frac{3-\sqrt{5}}{2}+o_\ell(1)\right)}\right)$.

Similarly, the lower bound for $g(n,6,2)$ can be improved from $\Omega\left(n^{3/5}\right)$ to $\Omega\left(n^{2/3}\right)$ by letting $k_0 = 4$. Indeed, it is easy to see that when applying the decomposition algorithm to an n-vertex $S_\lambda(2)$-free 6-graph $\mathcal{H}$ with the threshold $k_0 = 4$, the outputted family $\mathcal{F}$ consists of three hypergraphs: an $S_\lambda(2)$-free $(4, \lambda)$-indecomposable 6-graph $\mathcal{H}_1$, an $S_\lambda(2)$-free $(4, \lambda)$-indecomposable 5-graph $\mathcal{H}_2$, and an $S_\lambda(2)$-free 4-graph $\mathcal{H}_3$. By Theorem 1.1 (the stronger version in [11]), $|\mathcal{H}_2| = O\left(n^2\right)$ and $|\mathcal{H}_3| = O\left(n^2\right)$. By Theorem 3.4, $\mathcal{H}_1 = O\left(n^2\right)$. So, it follows from a similar probabilistic argument as above that $\alpha(\mathcal{H}) \geq \alpha(\mathcal{H}_1 \cup \mathcal{H}_2 \cup \mathcal{H}_3) = \Omega\left(n^{2/3}\right)$.

3.2 Pseudorandom bipartite graphs

Our construction for the upper bound in Theorem 1.4 is related to some pseudorandom bipartite graphs, so it will be convenient to introduce some definitions and results related to pseudorandom bipartite graphs.

For a graph G on n vertices (assuming that $V(G) = [n]$) the *adjacency matrix* A_G of G is an $n \times n$ matrix whose (i, j)-th entry is

$$A_G(i, j) = \begin{cases} 1, & \text{if } \{i, j\} \in E(G), \\ 0, & \text{otherwise.} \end{cases}$$

Denote by $G(V_1, V_2)$ a bipartite graph with two parts V_1 and V_2, and that say $G(V_1, V_2)$ is (d_1, d_2)*-regular* if $d_G(v) = d_i$ for all $v \in V_i$ and $i = 1, 2$.

For a bipartite $G = G(V_1, V_2)$ denote by $\lambda(G)$ the second largest eigenvalue of A_G. Suppose that G is (d_1, d_2)-regular. Then we say G is *pseudorandom* if $\lambda(G) = O\left(\max\{\sqrt{d_1}, \sqrt{d_2}\}\right)$.

The *Zarankiewicz number* $z(m, n, s, t)$ is the maximum number of edges in a bipartite graph $G(V_1, V_2)$ with $|V_1| = m$, $|V_2| = n$ such that G contains no complete bipartite graph with s vertices in V_1 and t vertices in V_2.

Our construction of (n, k, ℓ)-systems is related to the lower bound (construction) for $z(m, n, s, t)$. More specifically, it is related to a construction defined by Alon, Mellinger, Mubayi and Verstraëte in [4], which was used to show that $z(n^{\ell/2}, n, 2, \ell) = \Omega\left(n^{(\ell+1)/2}\right)$.

Let q be a prime power and $\mathbb{F} = GF(q)$ be the finite field of size q. Denote by $\mathbb{F}[X]$ the collection of all polynomials over $\mathbb{F}$. The graph $G(q^\ell, q^2, 2, \ell)$ is a bipartite graph with two parts V_1 and V_2, where

$$V_1 = \{P(x) : P(x) \in \mathbb{F}[X], \deg(P(x)) \leq \ell - 1\}, \quad \text{and} \quad V_2 = \mathbb{F} \times \mathbb{F},$$

and for every $P(x) \in V_1$ and every $(x, y) \in V_2$, the pair $\{P(x), (x, y)\}$ is an edge in $G(q^\ell, q^2, 2, \ell)$ iff $y = P(x)$.

It is clear that $G(q^\ell, q^2, 2, \ell)$ does not contain a complete bipartite graph with two vertices in V_1 and ℓ vertices in V_2 since two distinct polynomials of degree at most $\ell - 1$ over $\mathbb{F}$ can have the same value in at most $\ell - 1$ points. It is also easy to see that $G(q^\ell, q^2, 2, \ell)$ is $(q, q^{\ell-1})$-regular.

The proof of the following result concerning the eigenvalues of $G(q^\ell, q^2, 2, \ell)$ can be found in [10].

Lemma 3.6 ([10])**.** *The eigenvalues of the adjacency matrix of $G(q^\ell, q^2, 2, \ell)$ are*

$$q^{\ell/2}, \underbrace{q^{(\ell-1)/2}, \ldots, q^{(\ell-1)/2}}_{q^2-q \text{ times}}, 0, \ldots, 0, \underbrace{-q^{(\ell-1)/2}, \ldots, -q^{(\ell-1)/2}}_{q^2-q \text{ times}}, -q^{\ell/2}.$$

In particular, $G(q^\ell, q^2, 2, \ell)$ is pseudorandom.

3.3 Upper bound

In this section we prove the existence of (n,k,ℓ)-systems with independence number $O\left(n^{\frac{\ell+1}{2\ell}}(\log n)^{\frac{1}{\ell}}\right)$. Our construction is obtained from a random subgraph of the bipartite graph $G(q^\ell, q^2, 2, \ell)$ defined in the last section, and the method we used here is similar to that used in [15, 10].

First let us summarize the constructions used in [15] and [10] into a more general form.

Since we cannot ensure the random subgraph chosen from $G(q^\ell, q^2, 2, \ell)$ is exactly (d_1, d_2)-regular for some $d_1, d_2 \in \mathbb{N}$, it will be useful to consider the following more general setting.

Let $C, d_1, d_2 \geq 1$ be real numbers. A hypergraph $\mathcal{H}$ is

(a) (C, d_1)-*uniform* if $d_1/C \leq |E| \leq Cd_1$ for all $E \in \mathcal{H}$, and

(b) (C, d_2)-*regular* if $d_2/C \leq d_{\mathcal{H}}(v) \leq Cd_2$ for all $v \in V(\mathcal{H})$.

The *edge density* of a k-graph $\mathcal{H}$ with n vertices is $\rho(\mathcal{H}) = |\mathcal{H}|/\binom{n}{k}$. The *bipartite incidence graph* $G_{\mathcal{H}}$ of $\mathcal{H}$ is a bipartite graph with two parts $V_1 = E(\mathcal{H})$ and $V_2 = V(\mathcal{H})$, and for every $E \in E(\mathcal{H})$ and $v \in V(\mathcal{H})$ the pair $\{E, v\}$ is an edge in $G_{\mathcal{H}}$ iff $v \in E$. Denote by $A_{\mathcal{H}}$ the adjacency matrix of $G_{\mathcal{H}}$.

Let $n = |V(\mathcal{H})|$, $m = |\mathcal{H}|$ and labelling the edges in $\mathcal{H}$ with $E_1, \ldots, E_m$ We say a family $\mathcal{F}$ of hypergraphs *fits* $\mathcal{H}$ if $\mathcal{F} = \{\mathcal{G}_i : 1 \leq i \leq m\}$ and $\mathcal{G}_i$ is a hypergraph with $|V(\mathcal{G}_i)| = |E_i|$ for $i \in [m]$.

Given a hypergraph $\mathcal{H}$ and a family $\mathcal{F}$ that fits $\mathcal{H}$ we let $\mathcal{H}(\mathcal{F})$ be the random hypergraph obtained from $\mathcal{H}$ by taking independently for every $i \in [m]$ a bijection $\psi_i : E_i \to V(\mathcal{G}_i)$ and letting a set $S \subset E_i$ be an edge in $\mathcal{H}(\mathcal{F})$ if $\psi_i(S) \in \mathcal{G}_i$.

Let $\tau \geq 1$ be an integer and denote by $B_\tau(\mathcal{G})$ the collection of τ-subsets of $V(\mathcal{G})$ that are not independent in $\mathcal{G}$. Let $b_\tau(\mathcal{G}) = |B_\tau(\mathcal{G})|$ and $p_\tau(\mathcal{G}) = b_\tau(\mathcal{G})/\binom{v(\mathcal{G})}{\tau}$. In other words, $p_\tau(\mathcal{G})$ is the probability that a random τ-subset of $V(\mathcal{G})$ is not independent in $\mathcal{G}$. For a family $\mathcal{F}$ of hypergraphs define

$$p_\tau(\mathcal{F}) = \min\{p_\tau(\mathcal{G}) : \mathcal{G} \in \mathcal{F}\}.$$

We extend the definition of $C_{\mathcal{G}}(2, j)$ in Section 2 by letting $C_{\mathcal{G}}(2, j)$ denote the number of pairs of edges $\{E, E'\}$ in a k-graph $\mathcal{G}$ with $|E \cap E'| = j$ for all $0 \leq j \leq k-1$.

The following lemma gives an upper bound for the independence number of $\mathcal{H}(\mathcal{F})$.

Lemma 3.7. *Let $C, d_1, d_2 \geq 1$ be real numbers and $k \geq 2$ be an integer. Suppose that $\mathcal{H}$ is a hypergraph with n vertices, m edges, and is (C, d_1)-uniform, (C, d_2)-regular. Let $\mathcal{F} = \{\mathcal{G}_i : i \in [m]\}$ be a family of k-graphs that fits $\mathcal{H}$. Suppose there exists $\lambda \geq 0$ such that the bipartite graph $G_{\mathcal{H}}$ satisfies*

$$\left| e_{G_{\mathcal{H}}}(X, Y) - \frac{d_1}{n}|X||Y| \right| \leq \lambda\sqrt{|X||Y|} \tag{6}$$

for all $X \subset V(\mathcal{H})$ and $Y \subset E(\mathcal{H})$. Then, w.h.p. $\alpha(H(\mathcal{F})) \leq 2\tau n/d_1$, if τ satisfies

$$\frac{p_\tau(\mathcal{F})}{\tau} \geq \frac{8C^2 \log n}{d_2} \quad \text{and} \quad \tau \geq \frac{8C^2\lambda^2}{d_2}. \tag{7}$$

Proof of Lemma 3.7. Let τ be a real number that satisfies (7), $V = V(\mathcal{H})$, and $I \subset V$ be a set of size $2\tau n/d_1$.

Let $m = |\mathcal{H}|$ and label the edges in $\mathcal{H}$ by $\{E_1, \ldots, E_m\}$. Let $m_i = |E_i|$ for $i \in [m]$. Since $\mathcal{H}$ is (C, d_1)-uniform and (C, d_2)-regular, we obtain $d_1|\mathcal{H}|/C \leq \sum_{v \in V} d_{\mathcal{H}}(v) \leq Cd_1|\mathcal{H}|$, and consequently,

$$md_1/C^2 \leq nd_2 \leq C^2 md_1. \tag{8}$$

Define

$$\mathcal{E}_1 = \{E \in \mathcal{H} : |E \cap I| < \tau\}, \quad \text{and} \quad \mathcal{E}_2 = \{E \in \mathcal{H} : |E \cap I| > 3\tau\}.$$

Claim 3.8. $|\mathcal{E}_i| \leq 2C^2\lambda^2 m/d_2\tau \leq m/4$ *for* $i = 1, 2$.

Proof of Claim 3.8. It follows from (6) that

$$\sum_{E \in \mathcal{E}_1} |E \cap I| = e_{G_{\mathcal{H}}}(I, \mathcal{E}_1) \geq d_1|I||\mathcal{E}_1|/n - \lambda \left(|I||\mathcal{E}_1|\right)^{1/2},$$

and by definition, $\sum_{E \in \mathcal{E}_1} |E \cap I| < \tau|\mathcal{E}_1|$. Therefore,

$$\tau|\mathcal{E}_1| > d_1|I||\mathcal{E}_1|/n - \lambda \left(|I||\mathcal{E}_1|\right)^{1/2}.$$

Since $|I| = 2\tau n/d_1$, we obtain

$$|\mathcal{E}_1| < \left(\frac{\lambda|I|^{1/2}}{d_1|I|/n - \tau}\right)^2 = \left(\frac{\lambda|I|^{1/2}}{d_1|I|/2n}\right)^2 = \frac{2\lambda^2 n}{\tau d_1},$$

which together with (8) implies $|\mathcal{E}_1| < 2C^2\lambda^2 m/d_2\tau$. Notice that (7) implies that $C^2\lambda^2/d_2\tau \leq 1/8$, so $|\mathcal{E}_1| < m/4$.

Now consider $\mathcal{E}_2$. Similarly, By (6),

$$\sum_{E \in \mathcal{E}_2} |E \cap I| = e_{G_{\mathcal{H}}}(I, \mathcal{E}_2) \leq d_1|I||\mathcal{E}_1|/n + \lambda \left(|I||\mathcal{E}_1|\right)^{1/2},$$

and by definition, $\sum_{E \in \mathcal{E}_2} |E \cap I| > 3\tau|\mathcal{E}_2|$. Therefore,

$$3\tau|\mathcal{E}_2| < d_1|I||\mathcal{E}_2|/n + \lambda \left(|I||\mathcal{E}_2|\right)^{1/2},$$

Since $|I| = 2\tau n/d_1$, we obtain

$$|\mathcal{E}_2| < \left(\frac{\lambda|I|^{1/2}}{3\tau - d_1|I|/n}\right)^2 = \left(\frac{\lambda|I|^{1/2}}{d_1|I|/2n}\right)^2 = \frac{2\lambda^2 n}{\tau d_1} \leq \frac{2C^2\lambda^2 m}{\tau d_2} \leq \frac{m}{4}.$$

∎

For $i \in [m]$ let $I_i = I \cap E_i$. By Claim 3.8 the number of set I_i that satisfies $\tau \leq |I_i| \leq 3\tau$ (in fact, $|I_i| \geq \tau$ is sufficient for the proof) is at least $m - 2m/4 = m/2$. By the definition of $p_\tau(\mathcal{F})$, for every I_i that satisfies $\tau \leq |I_i| \leq 3\tau$ we have

$$P\left(\psi_i(I_i) \text{ is independent in } \mathcal{G}_i\right) \leq 1 - p_\tau(\mathcal{F}).$$

Since, by definition, the bijections $\{\psi_i : i \in [m]\}$ are mutually independent, the events

$$\{\psi_i(I_i) \text{ is independent in } \mathcal{G}_i : i \in [m]\}$$

are mutually independent. Therefore,

$$\begin{aligned} P\left(I \text{ is independent in } \mathcal{H}(\mathcal{F})\right) &\leq P\left(\bigwedge_{i\in[m]} \psi_i(I_i) \text{ is independent in } \mathcal{G}_i\right) \\ &= \prod_{i\in[m]} P\left(\psi_i(I_i) \text{ is independent in } \mathcal{G}_i\right) \leq \left(1 - p_\tau(\mathcal{F})\right)^{m/2}. \end{aligned}$$

So the expected number of independent $2n\tau/d_1$-sets in $\mathcal{H}(\mathcal{G})$ is at most

$$\begin{aligned} (1 - p_\tau(\mathcal{F}))^{m/2}\binom{n}{2n\tau/d_1} &< \exp\left(-p_\tau(\mathcal{F})\frac{m}{2} + \frac{2\tau n}{d_1}\log\left(\frac{en}{2n\tau/d_1}\right)\right) \\ &< \exp\left(-p_\tau(\mathcal{F})\frac{m}{2} + \frac{2C^2\tau m}{d_2}\log n\right) \\ &< \exp\left(-p_\tau(\mathcal{F})\frac{m}{4}\right) \to 0 \quad \text{as} \quad m \to \infty. \end{aligned}$$

Therefore, $\alpha(\mathcal{H}(\mathcal{F})) \leq 2\tau n/d_1$ holds with high probability. ∎

The following corollary may be a simpler form to use Lemma 3.7.

Corollary 3.9. *Let $C, d_1, d_2 \geq 1$ be real numbers and $k \geq 2$ be an integer. Suppose that $\mathcal{H}$ is a hypergraph with n vertices, m edges, and is (C, d_1)-uniform, (C, d_2)-regular. Let $\mathcal{F} = \{\mathcal{G}_i : i \in [m]\}$ be a family of k-graphs that fits $\mathcal{H}$. Suppose there exists $\lambda \geq 0$ such that the bipartite graph $G_\mathcal{H}$ satisfies (6) for all $X \subset V(\mathcal{H})$ and $Y \subset E(\mathcal{H})$. Suppose further that*

- *there exists $\rho > 0$ such that $\rho(\mathcal{G}_i) \geq \rho$ for $i \in [m]$, and*
- *$\lambda < (d_2\tau/8C^2)^{1/2}$, where $\tau = 2\left(16k!C^2\log n/\rho d_2\right)^{1/(k-1)} \gg 1$, and*
- *$C_{\mathcal{G}_i}(2, j) \leq |\mathcal{G}_i|\left(v(\mathcal{G}_i)/3\tau\right)^{k-j}$ for $0 \leq j \leq k-1$ and $i \in [m]$.*

Then, w.h.p. $\alpha\left(\mathcal{H}(\mathcal{F})\right) \leq 2\tau n/d_1$.

Proof of Corollary 3.9. It suffices to show that $\tau = 2\left(16k!C^2\log n/\rho d_2\right)^{1/(k-1)}$ satisfies (7). First let us calculate $p_\tau(\mathcal{F})$. Fix $i \in [m]$ and for every edge set $\mathcal{E} \subset \mathcal{G}_i$ let $T_\mathcal{E}$ denote the collection of τ-sets in $V(\mathcal{G}_i)$ containing the vertex set $\bigcup_{E\in\mathcal{E}} E$. By the definition of $B_\tau(\mathcal{G}_i)$, we have

$$B_\tau(\mathcal{G}_i) = \bigcup_{E\in\mathcal{G}_i} T_{\{E\}}.$$

It follows from the Bonferroni inequalities [6] that

$$\begin{aligned} b_\tau(\mathcal{G}_i) = \left|\bigcup_{E\in\mathcal{G}_i} T_{\{E\}}\right| &\geq \sum_{E\in\mathcal{G}_i} |T_{\{E\}}| - \sum_{\{E,E'\}\in\binom{\mathcal{G}_i}{2}} |T_{\{E,E'\}}| \\ &= |\mathcal{G}_i|\binom{v(\mathcal{G}_i) - k}{\tau - k} - \sum_{j=0}^{k-1} C_{\mathcal{G}_i}(2, j) \cdot \binom{v(\mathcal{G}_i) - 2k + j}{\tau - 2k + j} \end{aligned}$$

Since $C_{\mathcal{G}_i}(2,j) \le |\mathcal{G}_i|\,(v(\mathcal{G}_i)/\tau)^{k-j}$ for $0 \le j \le k-1$, we obtain

$$\begin{aligned}
\sum_{j=0}^{k-1} C_{\mathcal{G}_i}(2,j) \cdot \binom{v(\mathcal{G}_i)-2k+j}{\tau-2k+j} &\le \sum_{j=0}^{k-1} |\mathcal{G}_i| \left(\frac{v(\mathcal{G}_i)}{3\tau}\right)^{k-j} \binom{v(\mathcal{G}_i)-2k+j}{\tau-2k+j} \\
&= \sum_{j=0}^{k-1} |\mathcal{G}_i| \left(\frac{v(\mathcal{G}_i)}{3\tau}\right)^{k-j} \frac{(\tau-k)_{k-j}}{(v(\mathcal{G}_i)-k)_{k-j}} \binom{v(\mathcal{G}_i)-k}{\tau-k} \\
&\le \sum_{j=0}^{k-1} |\mathcal{G}_i| \left(\frac{v(\mathcal{G}_i)}{3\tau}\right)^{k-j} \left(\frac{\tau}{v(\mathcal{G}_i)}\right)^{k-j} \binom{v(\mathcal{G}_i)-k}{\tau-k} \\
&\le \frac{1}{2}|\mathcal{G}_i| \binom{v(\mathcal{G}_i)-k}{\tau-k}.
\end{aligned}$$

Therefore, $b_\tau(\mathcal{G}_i) \ge \frac{1}{2}|\mathcal{G}_i|\binom{v(\mathcal{G}_i)-k}{\tau-k}$. Consequently,

$$p_\tau(\mathcal{G}_i) = \frac{b_\tau(\mathcal{G}_i)}{\binom{v(\mathcal{G}_i)}{\tau}} \ge \frac{1}{2}\frac{|\mathcal{G}_i|\binom{v(\mathcal{G}_i)-k}{\tau-k}}{\binom{v(\mathcal{G}_i)}{\tau}} = \frac{1}{2}\frac{|\mathcal{G}_i|}{\binom{v(\mathcal{G}_i)}{k}}\frac{(\tau)_k}{k!} = \frac{\rho(\mathcal{G}_i)}{2}\frac{(\tau)_k}{k!} \ge \frac{\rho}{2k!}(\tau)_k.$$

So we obtain

$$\frac{p_\tau(\mathcal{F})}{\tau} \ge \frac{\rho}{2k!}(\tau-1)_{k-1} \ge \frac{\rho}{2k!}\left(\frac{\tau}{2}\right)^{k-1} \ge \frac{8C^2\log n}{d_2}.$$

On the other hand, our assumption on λ clearly implies $\tau \ge 8C^2\lambda/d_2$. Therefore, by Lemma 3.7, *w.h.p.* $\alpha(\mathcal{H}(\mathcal{F})) \le 2\tau n/d_1$. ∎

We will also need the following result in our proof.

Lemma 3.10 ([15])**.** *Let $\mathcal{H}$ be a d_1-uniform d_2-regular hypergraph on n vertices. Then for every $V' \subset V(\mathcal{H})$ and $\mathcal{E} \subset E(\mathcal{H})$,*

$$\left|\sum_{E\in\mathcal{E}} |E\cap V'| - \frac{d_1}{n}|V'||\mathcal{E}|\right| \le \lambda(G_{\mathcal{H}})\sqrt{|V'||\mathcal{E}|}.$$

We also need the following Chernoff's inequality (e.g. see Theorem 22.6 in [12]).

Theorem 3.11 (Chernoff's inequality)**.** *Suppose that $S_n = X_1 + \cdots + X_n$ where $0 \le X_i \le 1$ for $i \in [n]$ are independent random variables. Let $\mu = \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n]$. Then for every $0 \le t \le \mu$,*

$$P(|S_n - \mu| \ge t) \le e^{-\frac{t^2}{3\mu}}.$$

Now we are ready to prove the upper bound in Theorem 1.4.

Proof of the upper bound in Theorem 1.4. Let q be a prime power and $G = G(q^\ell, q^2, 2, \ell)$ be the bipartite graph on $V_1 \cup V_2$ with $|V_1| = q^\ell$ and $|V_2| = q^2$. Let $\mathcal{G}$ denote the hypergraph on q^2 vertices whose bipartite incident graph is G. Note that $\mathcal{G}$ is a $q^{\ell-1}$-regular q-graph, and by Lemmas 3.6 and 3.10,

$$\left|\sum_{E\in\mathcal{E}} |E\cap V'| - \frac{1}{q}|V'||\mathcal{E}|\right| \le q^{(\ell-1)/2}\sqrt{|V'||\mathcal{E}|} \tag{9}$$

holds for all $V' \subset V(\mathcal{G})$ and $\mathcal{E} \subset \mathcal{G}$.

Let $U \subset V(\mathcal{G})$ be a random set such that every vertex in $V(\mathcal{G})$ is included in U independently with probability $p = q^{-\frac{2}{\ell+1}}$. Then $\mathbb{E}[|U|] = pq^2 = q^{\frac{2\ell}{\ell+1}}$, and by the Chernoff inequality,

$$P\left(\left||U| - pq^2\right| > pq^2/2\right) < e^{-\frac{\left(pq^2/2\right)^2}{3pq^2}} = e^{-pq^2/12} \to 0 \quad \text{as} \quad q \to \infty.$$

For every $E \in \mathcal{G}$ we have $\mathbb{E}[|E \cap U|] = pq = q^{\frac{\ell-1}{\ell+1}}$, and by the Chernoff inequality,

$$P\left(||E \cap U| - pq| > pq/2\right) < e^{-\frac{(pq/2)^2}{3pd_1}} = e^{-pq/12}.$$

Let B denote the collection of edges $E \in \mathcal{G}$ such that $||E \cap U| - pq| > pq/2$. Then

$$\mathbb{E}[|B|] \leq q^\ell e^{-pq/12} = q^\ell e^{-q^{\frac{\ell-1}{\ell+1}}/12} \to 0 \quad \text{as} \quad q \to \infty.$$

Therefore, *w.h.p.* the set U satisfies that $q^{\frac{2\ell}{\ell+1}}/2 \leq |U| \leq 3q^{\frac{2\ell}{\ell+1}}/2$ and $q^{\frac{\ell-1}{\ell+1}}/2 \leq |E \cap U| \leq 3q^{\frac{\ell-1}{\ell+1}}/2$ for all $E \in \mathcal{G}$.

Fix such a set U that satisfies the conclusion above, and let $c \in [1/2, 3/2]$ be the real number such that $|U| = cq^{\frac{2\ell}{\ell+1}}$. Let $n = |U| = cq^{\frac{2\ell}{\ell+1}}$, $m = |\mathcal{G}| = q^\ell$, $d_1 = q^{\frac{\ell-1}{\ell+1}}$, and $d_2 = q^{\ell-1}$. Let $\mathcal{H}$ be the hypergraph on U with

$$\mathcal{H} = \{E \cap U : E \in \mathcal{G}\}.$$

Since $d_1/2 \leq |E \cap U| \leq 3d_1/2$ for all $E \in \mathcal{G}$, the hypergraph $\mathcal{H}$ is a $(2, d_1)$-uniform. Moreover, for every pair of edges $E, E' \in \mathcal{G}$, since $|E \cap E'| < \ell < d_1/2$, we have $E \cap U \neq E' \cap U$. So, $d_{\mathcal{H}}(u) = d_{\mathcal{G}}(u) = d_2$ for all $u \in U$. In addition, (9) also holds for all $V' \subset U$ and $\mathcal{E} \subset \mathcal{G}$.

Label the edges in $\mathcal{H}$ with $\{E_1, \ldots, E_m\}$ and let $m_i = |E_i|$ for $i \in [m]$. Let $\mathcal{F} = \{\mathcal{S}_i : i \in [m]\}$, where $\mathcal{S}_i$ is the k-graph on $[m_i]$ whose edge set is the collection of all k-subsets of $[m_i]$ that contain $[\ell+1]$. Our construction of the (n, k, ℓ)-omitting system is simply $\mathcal{H}(k, \ell) = \mathcal{H}(\mathcal{F})$, and indeed, one can easily check that $|e' \cap e'| \neq \ell$ for all distinct edges $e, e' \in \mathcal{H}(k, \ell)$.

Let $\tau = 100\,(\log n)^{1/\ell}$.

Claim 3.12. $p_\tau(\mathcal{F}) \geq \left(\frac{\tau}{3d_1/2}\right)^{\ell+1}/2$.

Proof of Claim 3.12. Fix $i \in [m]$ and let I be a random τ-subset of $[m_i]$. It is easy to see that I is not independent in $\mathcal{S}_i$ iff $[\ell+1] \subset I$. Since

$$P\left([\ell+1] \subset I\right) = \frac{\binom{m_i-\ell-1}{\tau-\ell-1}}{\binom{m_i}{\tau}} = \frac{\tau \cdots (\tau-\ell)}{m_i \cdots (m_i-\ell)} \geq (1-o(1))\left(\frac{\tau}{m_i}\right)^{\ell+1} > \frac{1}{2}\left(\frac{\tau}{3d_i/2}\right)^{\ell+1},$$

we obtain

$$p_\tau(\mathcal{F}) > \frac{1}{2}\left(\frac{\tau}{3d_i/2}\right)^{\ell+1}.$$

∎

Observe that τ satisfies

$$\frac{p_\tau(\mathcal{F})}{\tau} > \frac{\left(\frac{\tau}{3d_1/2}\right)^{\ell+1}/2}{\tau} = \frac{100^\ell \log n}{2(3/2)^{\ell+1}d_1^{\ell+1}} = \frac{100^\ell}{2(3/2)^{\ell+1}}\frac{\log n}{d_2} > \frac{32\log n}{d_2}$$

(here we used the fact that $d_2 = d_1^{\ell+1}$) and

$$\tau = 100(\log n)^{1/\ell} > \frac{32\left(q^{(\ell-1)/2}\right)^2}{q^{\ell-1}}.$$

We may therefore apply Lemma 3.7 with $C = 2$ to obtain

$$\alpha\left(\mathcal{H}(k,\ell)\right) \le 2\tau n/d_1 \le 400 n^{\frac{\ell+1}{2\ell}}(\log n)^{1/\ell}.$$

∎

4 Independent sets in (n,k,ℓ,λ)-systems

In this section we prove Theorem 1.6. Our proof is a direct application of Theorem 2.3.

Proof of Theorem 1.6. Fix $\delta > 0$, and let $\epsilon > 0$ be sufficiently small such that $\frac{\ell-1}{k-2} - \delta < \frac{(\ell-1)(1-\epsilon)}{k-2+\epsilon}$ holds. Let $t = \lambda^{\frac{1}{k-1}} n^{\frac{\ell-1}{k-1}}$ and $\mathcal{H}$ be a (n,k,ℓ,λ)-system, where $0 < \lambda < n^{\frac{\ell-1}{k-2}-\delta}$. Let $j \in [\ell-1]$ and $S \subset V(\mathcal{H})$ be a set of size j. Since $\mathcal{H}$ is an (n,k,ℓ,λ)-system, $L_{\mathcal{H}}(S)$ is an $(n,k-j,\ell-j,\lambda)$-system. Therefore,

$$\Delta(\mathcal{H}) \le \lambda\binom{n}{\ell-1}\Big/\binom{k-1}{\ell-1} < t^{k-1}, \quad \text{and}$$
$$|\Delta_j(\mathcal{H})| \le \lambda\binom{n}{\ell-j}\Big/\binom{k-j}{\ell-j} = O(\lambda n^{\ell-j}) \quad \text{for} \quad 2 \le j \le \ell-1.$$

It follows that

$$C_{\mathcal{H}}(2,j) = O\left(\lambda n^{\ell-j}|\mathcal{H}|\right) = O\left(\lambda^2 n^{2\ell-j}\right) \le n t^{2k-j-1-\epsilon} \quad \text{for} \quad 2 \le j \le \ell-1.$$

On the other hand, for $\ell \le j' \le k-1$ and a set $S \subset V(\mathcal{H})$ of size j' the link $L_{\mathcal{H}}(S)$ has size at most λ. Therefore,

$$C_{\mathcal{H}}(2,j') = O\left(\lambda|\mathcal{H}|\right) = O\left(\lambda^2 n^{\ell}\right) \le n t^{2k-j'-1-\epsilon} \quad \text{for} \quad \ell \le j' \le k-1.$$

Therefore, by Theorem 2.3, $\alpha(\mathcal{H}) = \Omega\left((\log t)^{1/(k-1)} n/t\right) = \Omega\left(\lambda^{-\frac{1}{k-1}} n^{\frac{k-\ell}{k-1}} (\log n)^{\frac{1}{k-1}}\right)$. ∎

5 The Ramsey number of the k-Fan

In this section we prove Theorem 1.10. The lower bound (construction) is given by the so called *L-constructions.* These were introduced in [7], where they were used to answer an old Ramsey-type question of Ajtai–Erdős–Komlós–Szemerédi [1].

Let $m,n \ge 2$ and let $\mathcal{L}_{m,n}$ be the k-graph with vertex set $[m]\times[n]$ and edge set

$$\left\{\{(x_1,y_1),(x_1,y_2),\ldots,(x_{k-1},y_2)\} : x_1 < \cdots < x_{k-1}, y_1 > y_2\right\}.$$

Proposition 5.1. *For every $m, n \geq 2$ the hypergraph $\mathcal{L}_{m,n}$ is F^k-free.*

Proof of Proposition 5.1. Suppose that $\mathcal{L}_{m,n}$ contains a copy of $F^k = \{E_1, \ldots, E_k, E\}$. Let $v = \bigcap_{i=1}^{k} E_i$ and assume that $v = (x_0, y_0)$, $E = \{(x_1, y_1), (x_1, y_2), \ldots, (x_{k-1}, y_2)\}$, where $x_1 < \cdots < x_{k-1}$ and $y_1 > y_2$.

By the definition of F^k, for every vertex $u \in E$, there exists an edge E_i that contains both u and v. It is easy to see that if $x_1' < x_2'$ and $y_1' < y_2'$, then there is no edge in $\mathcal{L}_{m,n}$ containing both (x_1', y_1') and (x_2', y_2'). Therefore, we must have (see Figure 1)

(1) $x_0 \leq x_1$ and $y_0 \geq y_1$, or

(2) $x_0 \geq x_{k-1}$ and $y_0 \leq y_2$, or

(3) $x_0 = x_1$ and $y_2 < y_0 < y_1$, or

(4) $y_0 = y_2$ and $x_1 < x_0 < x_{k-1}$.

If $x_0 \leq x_1$ and $y_0 \geq y_1$, then by the definition of $\mathcal{L}_{m,n}$, there is a $(k-1)$-set $J \subset [k]$ such that $\bigcap_{j \in J} E_j = (x_0, y_2)$, a contradiction. If $x_0 \geq x_{k-1}$ and $y_0 \leq y_2$, then by the definition of $\mathcal{L}_{m,n}$, there exist $\{i, j\} \subset [k]$ such that $E_i \cap E_j = (x_1, y_0)$, a contradiction. Similarly, if Case (3) or Case (4) happens, then there exist $\{i, j\} \subset [k]$ such that $E_i \cap E_j = (x_1, y_2)$, a contradiction. ∎

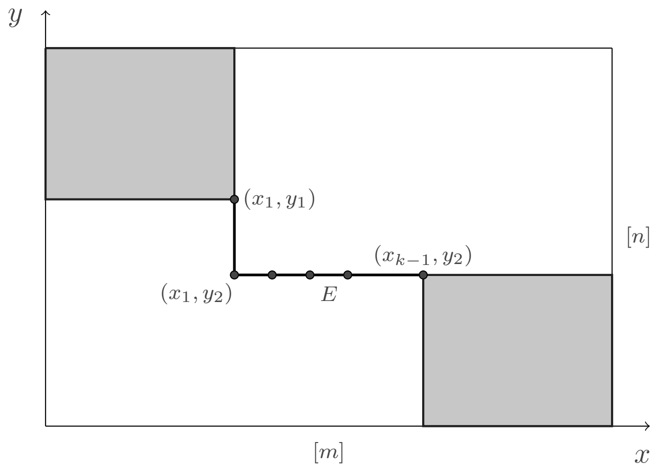

Figure 1: Only vertices that lie in these two shaded areas and the L-shaped path that connects these two areas can be adjacent to all vertices in E.

The following result gives an upper bound for the independence number of $\mathcal{L}_{m,n}$.

Proposition 5.2. *The hypergraph $\mathcal{L}_{m,n}$ satisfies $\alpha(\mathcal{L}_{m,n}) < m + (k-2)n$.*

Proof of Proposition 5.2. Let I be an independent in $\mathcal{L}_{m,n}$. Remove the topmost vertex of each column and the $k-2$ rightmost vertices of each row in I. It is easy to see that we removed at most $m + (k-2)n$ vertices from I, and I has no vertex left since otherwise I would contain an edge in $\mathcal{L}_{m,n}$. Therefore, $\alpha(\mathcal{L}_{m,n}) < m + (k-2)n$. ∎

Now we finish the proof of Theorem 1.10.

Proof of Theorem 1.10. First we prove the lower bound. Let $m = \lfloor \frac{t}{2} \rfloor$ and $n = \lfloor \frac{t-1}{2(k-2)} \rfloor$. By Propositions 5.1 and 5.2, the k-graph $\mathcal{L}_{m,n}$ is F^k-free and $\alpha(\mathcal{L}_{m,n}) \leq m + (k-2)n < t$. So,

$$r_k(F^k, t) > mn = \left\lfloor \frac{t}{2} \right\rfloor \left\lfloor \frac{t-1}{2(k-2)} \right\rfloor.$$

To prove the upper bound, let us show that $r_k(F^k, t) \leq r_k(S_t^k, t)$ first. Indeed, let $\mathcal{H}$ be a k-graph on $r_k(S_t^k, t)$ vertices. We may assume that $\mathcal{H}$ does not contain an independent set of size t. Then, there exist t distinct edges $E_1, \ldots, E_t$ and a vertex v in $\mathcal{H}$ such that $E_i \cap E_j = \{v\}$ for $1 \leq i < j \leq t$. Let S be a set that contains exactly one vertex from each $E_i \setminus \{v\}$ for $i \in [t]$. Then S has size t and hence contains an edge in $\mathcal{H}$, and it implies that $\mathcal{H}$ contains a copy of F^k. Therefore, $r_k(F^k, t) \leq r_k(S_t^k, t)$, and it follows from Theorem 1.9 that $r_k(F^k, t) \leq t(t-1) + 1$. ∎

6 Acknowledgement

We thank the referees for many helpful suggestions.

References

[1] M. Ajtai, P. Erdős, J. Komlós, and E. Szemerédi. On Turán's theorem for sparse graphs. *Combinatorica*, 1(4):313–317, 1981.

[2] M. Ajtai, J. Komlós, J. Pintz, J. Spencer, and E. Szemerédi. Extremal uncrowded hypergraphs. *J. Combin. Theory Ser. A*, 32(3):321–335, 1982.

[3] M. Ajtai, J. Komlós, and E. Szemerédi. A note on Ramsey numbers. *J. Combin. Theory Ser. A*, 29(3):354–360, 1980.

[4] N. Alon, K. E. Mellinger, D. Mubayi, and J. Verstraëte. The de Bruijn-Erdős theorem for hypergraphs. *Des. Codes Cryptogr.*, 65(3):233–245, 2012.

[5] P. Bennett and T. Bohman. A note on the random greedy independent set algorithm. *Random Structures Algorithms*, 49(3):479–502, 2016.

[6] C. Bonferroni. Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze*, 8:3–62, 1936.

[7] J. Cooper and D. Mubayi. Sparse hypergraphs with low independence number. *Combinatorica*, 37(1):31–40, 2017.

[8] M. Deza, P. Erdős, and P. Frankl. Intersection properties of systems of finite sets. *Proc. London Math. Soc. (3)*, 36(2):369–384, 1978.

[9] R. A. Duke, H. Lefmann, and V. Rödl. On uncrowded hypergraphs. In *Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993)*, volume 6, pages 209–212, 1995.

[10] A. Eustis and J. Verstraëte. On the independence number of Steiner systems. *Combin. Probab. Comput.*, 22(2):241–252, 2013.

[11] P. Frankl and Z. Füredi. Forbidding just one intersection. *J. Combin. Theory Ser. A*, 39(2):160–176, 1985.

[12] A. Frieze and M. Karoński. *Introduction to random graphs.* Cambridge University Press, Cambridge, 2016.

[13] A. Frieze and D. Mubayi. Coloring simple hypergraphs. *J. Combin. Theory Ser. B*, 103(6):767–794, 2013.

[14] J. H. Kim. The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$. *Random Structures Algorithms*, 7(3):173–207, 1995.

[15] A. Kostochka, D. Mubayi, and J. Verstraete. Hypergraph Ramsey numbers: triangles versus cliques. *J. Combin. Theory Ser. A*, 120(7):1491–1507, 2013.

[16] A. Kostochka, D. Mubayi, and J. Verstraëte. On independent sets in hypergraphs. *Random Structures Algorithms*, 44(2):224–239, 2014.

[17] P.-S. Loh. A note on embedding hypertrees. *Electron. J. Combin.*, 16(1):Note 18, 4, 2009.

[18] L. Lu and L. Székely. Using Lovász local lemma in the space of random injections. *Electron. J. Combin.*, 14(1):Research Paper 63, 13, 2007.

[19] D. Mubayi and O. Pikhurko. A new generalization of Mantel's theorem to k-graphs. *J. Combin. Theory Ser. B*, 97(4):669–678, 2007.

[20] J. Nešetřil and V. Rödl, editors. *Mathematics of Ramsey theory*, volume 5 of *Algorithms and Combinatorics.* Springer-Verlag, Berlin, 1990.

[21] V. Rödl and E. Šiňajová. Note on independent sets in Steiner systems. In *Proceedings of the Fifth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science (Poznań, 1991)*, volume 5, pages 183–190, 1994.

[22] J. B. Shearer. On the independence number of sparse graphs. *Random Structures Algorithms*, 7(3):269–271, 1995.

[23] J. Spencer. Turán's theorem for k-graphs. *Discrete Math.*, 2:183–186, 1972.

[24] F. Tian and Z.-L. Liu. Bounding the independence number in some (n,k,ℓ,λ)-hypergraphs. *Graphs Combin.*, 34(5):845–861, 2018.

[25] P. Turán. Eine Extremalaufgabe aus der Graphentheorie. *Mat. Fiz. Lapok*, 48:436–452, 1941.