Off-diagonal Ramsey numbers for slowly growing hypergraphs

Sam Mattheus* Dhruv Mubayi† Jiaxi Nie‡ Jacques Verstraëte §

January 12, 2025

Abstract

For a k-uniform hypergraph F and a positive integer n, the Ramsey number $r(F,n)$ denotes the minimum N such that every N-vertex F-free k-uniform hypergraph contains an independent set of n vertices. A hypergraph is *slowly growing* if there is an ordering $e_1, e_2, \dots, e_t$ of its edges such that $|e_i \setminus \bigcup_{j=1}^{i-1} e_j| \le 1$ for each $i \in \{2, \dots, t\}$. We prove that if $k \ge 3$ is fixed and F is any non k-partite slowly growing k-uniform hypergraph, then for $n \ge 2$,

$$r(F,n) = \Omega\Big(\frac{n^k}{(\log n)^{2k-2}}\Big).$$

In particular, we deduce that the off-diagonal Ramsey number $r(F_5,n)$ is of order $n^3/\mathrm{polylog}(n)$, where F_5 is the triple system $\{123, 124, 345\}$. This is the only 3-uniform Berge triangle for which the polynomial power of its off-diagonal Ramsey number was not previously known. Our constructions use pseudorandom graphs and hypergraph containers.

1 Introduction

A hypergraph is a pair (V,E) where V is a set, whose elements are called vertices, and E is a family of nonempty subsets of V, whose elements are called edges. A k-uniform hypergraph (k-graph for short) is a hypergraph whose edges are all of size k. An *independent set* of a hypergraph F is a subset of $V(F)$ which does not contain any edge of F.

Given a k-graph F, the *off-diagonal Ramsey number* $r(F,n)$ is the minimum integer such that every F-free k-graph on $r(F,n)$ vertices has an independent set of size n. Ajtai, Komlós and Szemerédi [1] proved the upper bound $r(K_3,n) = O(n^2/\log n)$, and Kim [12] proved the corresponding lower bound $r(K_3,n) = \Omega(n^2/\log n)$. The current state-of-the-art results are due to Fiz Pontiveros, Griffiths and Morris [8] and Bohman and Keevash [4], who determine $r(K_3,n)$ up to a small constant factor:

$$\left(\frac{1}{4} - o(1)\right)\frac{n^2}{\log n} \le r(K_3,n) \le (1+o(1))\,\frac{n^2}{\log n}.$$

For larger cliques, the current best general lower bounds are obtained by Bohman and Keevash [3] strengthening earlier bounds of Spencer [22, 23]. On the other hand, the current best upper bounds are proved by Li, Rousseau and Zang [15] by extending ideas of Shearer [21], which improve earlier bounds of

*Department of Mathematics, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium. sam.mattheus@vub.be. Research supported by a postdoctoral fellowship 1267923N from the Research Foundation Flanders (FWO).

†Department of Mathematics, Statistics and Computer Science, University of Illinois Chicago, Chicago, IL 60607, USA. mubayi@uic.edu. Research supported by NSF grants DMS-1952767, DMS-2153576 and a Simons Fellowship.

‡School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332 USA. jnie47@gatech.edu.

§Department of Mathematics, University of California, San Diego, La Jolla, CA 92093, USA. jacques@ucsd.edu. Research supported by NSF grant DMS-1952786.

Ajtai, Komlós and Szemerédi [1]. These bounds are as follows: for $s \geq 3$, there exists constant $c_1(s) > 0$ such that

$$c_1(s)\frac{n^{\frac{s+1}{2}}}{(\log n)^{\frac{s+1}{2}-\frac{1}{s-2}}} \leq r(K_s, n) \leq (1+o(1))\frac{n^{s-1}}{(\log n)^{s-2}}.$$

Recently, the first and fourth authors [16] determined the asymptotics of $r(K_4, n)$ up to a logarithmic factor by proving the following lower bounds.

Theorem 1.1 (Theorem 1, [16]). *As $n \to \infty$,*

$$r(K_4, n) = \Omega\left(\frac{n^3}{(\log n)^4}\right).$$

In this paper, we prove some hypergraph versions of these results. A *Berge triangle* is a hypergraph consisting of three distinct edges e_1, e_2 and e_3 such that there exists three distinct vertices x, y and z with the property that $\{x, y\} \subset e_1$, $\{y, z\} \subset e_2$, and $\{x, z\} \subset e_3$. It is easy to check that there are only four different 3-uniform Berge triangles: LC_3 (loose cycle of length 3), TP_3 (tight path on three edges and five vertices), F_5, and K_4^{3-} (3-uniform clique on four vertices minus an edge), as shown from left to right in Figure 1. It is natural to consider the problem of determining the off-diagonal Ramsey numbers for 3-uniform Berge triangles since they are in some sense the smallest non-trivial hypergraphs that provide a natural extension of $r(K_3, n)$.

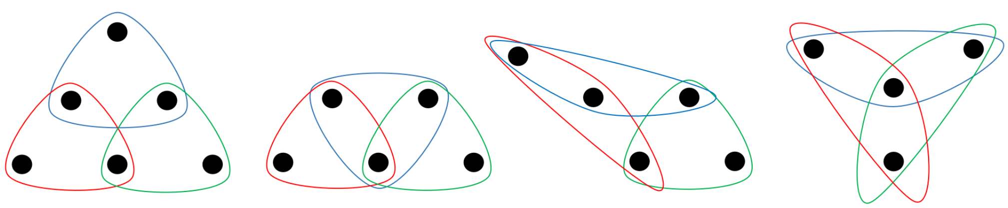

Figure 1: From left to right: LC_3, TP_3, F_5 and K_4^{3-}.

The off-diagonal Ramsey numbers for TP_3 and LC_3 have been determined up to a logarithmic factor: for TP_3, a result of Phelps and Rödl [19] shows that $c_1 n^2/\log n \leq r(TP_3, n) \leq c_2 n^2$; for LC_3, Kostochka, the second author and the fourth author [13] showed that $c_1 n^{3/2}/(\log n)^{3/4} \leq r(LC_3, n) \leq c_2 n^{3/2}$. It seems plausible to conjecture that for some constant c,

$$r(TP_3, n) \leq \frac{cn^2}{\log n} \quad \text{and} \quad r(LC_3, n) \leq \frac{cn^{\frac{3}{2}}}{(\log n)^{\frac{3}{4}}}.$$

It is conjectured explicitly in [13] that $r(LC_3, n) = o(n^{3/2})$ and the question of determining the order of magnitude of $r(TP_3, n)$ was posed in [6]. It was also shown in [6] that $r(TP_4, n)$ has order of magnitude n^2, leaving TP_3 as the only tight path for which the order of magnitude of $r(TP_s, n)$ remains open. We remark that if one can prove that every n-vertex TP_3-free 3-graph with average degree $d > 1$ has an independent set of size at least $\Omega(n\sqrt{\log d/d})$, then this implies that $r(TP_3, n) = \Theta(n^2/\log n)$.

The problem for K_4^{3-} is interesting in the sense that it is the smallest hypergraph whose off-diagonal Ramsey number is at least exponential: Erdős and Hajnal [7] proved $r(K_4^{3-}, n) = n^{O(n)}$ and Rödl (unpublished) proved $r(K_4^{3-}, n) \geq 2^{\Omega(n)}$. More recently, Fox and He [9] showed that $r(K_4^{3-}, n) = n^{\Theta(n)}$.

The problem for F_5, however, is not very well studied: a result of Kostochka, the second author and the fourth author [14] implies that $r(F_5, n) \leq c_1 n^3/\log n$, and the standard probabilistic deletion method shows that $r(F_5, n) \geq c_2 n^2/\log n$. In this paper, we fill this gap by showing that $r(F_5, n) = n^3/\mathrm{polylog}(n)$. This is a consequence of a more general theorem that we will prove.

Building upon techniques in [16], we prove lower bounds for the off-diagonal Ramsey numbers of a large family of hypergraphs. A k-graph F is *slowly growing* if its edges can be ordered as $e_1, \ldots, e_t$ such that

$$\forall\, i \in \{2, \ldots, t\},\ \left| e_i \setminus \bigcup_{j=1}^{i-1} e_j \right| \leq 1.$$

We use this terminology to describe the fact that at most one new vertex is added when we add a new edge in the ordering. Further, F is k-partite, or *degenerate*, if its vertices can be partitioned into k sets $V_1, \ldots, V_k$ such that each edge intersects each V_i, $1 \leq i \leq k$, in exactly one vertex. Otherwise, H is *non-degenerate*. The three hypergraphs TP_3, F_5 and K_4^{3-} in Figure 1 are slowly growing, whereas the first is not. The last two are non-degenerate.

In this paper, we obtain the following result for non-degenerate slowly growing hypergraphs.

Theorem 1.2. *For every $k \geq 3$, there exists a constant $c > 0$ such that for every slowly growing non-degenerate k-graph F and all integers $n \geq 2$*

$$r(F, n) \geq \frac{cn^k}{(\log n)^{2k-2}}.$$

The constant c here is independent of F because our construction simultaneously avoids all non-degenerate slowly growing F.

Theorem 1.2 is tight up to a logarithmic factor for the following family of hypergraphs which includes F_5. For $k \geq 3$, let F_{2k-1} be the k-graph on $2k-1$ vertices $v_1, \ldots, v_{k-1}, w_1, \ldots, w_k$ with k edges $\{v_1, \ldots, v_{k-1}, w_i\}$, $1 \leq i \leq k-1$, and $\{w_1, \ldots, w_k\}$. Further, let T_k be the k-graph obtained from F_{2k-1} by adding the edge $\{v_1, \ldots, v_{k-1}, w_k\}$. See Figure 2 for an illustration of F_7 and T_4. Note that T_2 is a (graph) triangle.

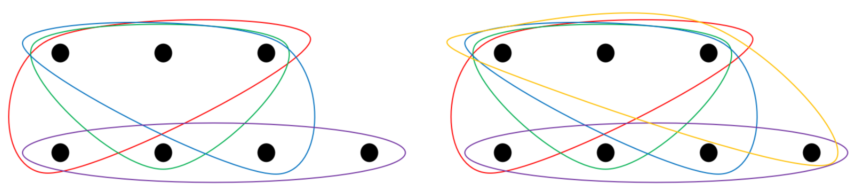

Figure 2: F_7 and T_4

The order of magnitude of $r(T_k, n)$ for $k \geq 3$ is determined by the upper bound result of Kostochka, the second author and the fourth author [14] together with the lower bound result of Bohman, the second author and Picollelli [5]. For $k = 2$, this theorem restates the known result [1, 4, 8, 12] that $r(K_3, n)$ has order of magnitude $n^2/\log n$.

Theorem 1.3 (Theorem 2, [14]; Theorem 1, [5])**.** *Let $k \geq 2$. Then there exist constants $c_1,\ c_2 > 0$ such that for all integers $n \geq 2$,*

$$\frac{c_1 n^k}{\log n} \leq r(T_k, n) \leq \frac{c_2 n^k}{\log n}.$$

Thus we have $r(F_{2k-1}, n) \leq r(T_k, n) \leq O(n^k/\log n)$. On the other hand, it is easy to check that F_{2k-1} is a slowly growing non-degenerate k-graph. Hence Theorem 1.2 together with Theorem 1.3 implies the following theorem.

Theorem 1.4. *Let $k \geq 3$. There exist constants $c_1, c_2 > 0$ such that for all integers $n \geq 2$,*

$$\frac{c_1 n^k}{(\log n)^{2k-2}} \leq r(F_{2k-1}, n) \leq \frac{c_2 n^k}{\log n}.$$

Theorem 1.4 determines $r(F_{2k-1}, n)$ up to a logarithmic factor. In particular, this determines $r(F_5, n)$ up to a polylogarithmic factor, and F_5 is the only 3-uniform Berge triangle for which the polynomial power of the off-diagonal Ramsey number was not previously known.

It would be interesting to determine its order of magnitude. We believe the current upper bounds are closer to the truth:

Conjecture 1. *There exists a constant $c > 0$ such that for $n \geq 2$,*

$$r(F_5, n) \geq \frac{cn^3}{\log n}.$$

2 The Construction

The proof of Theorem 1.2 uses the so-called random block construction, which first requires a pseudorandom bipartite graph. We build our construction using the following bipartite graph.

Definition 2.1. *For every prime power q and integer $m \geq 2$, let $\Gamma_{q,m}$ be the bipartite graph with two parts $X = \mathbb{F}_q^2$ and $Y = \mathbb{F}_q^m$, where two vertices $x = (x_0, x_1) \in X$ and $y = (y_0, \ldots, y_{m-1}) \in Y$ form an edge if and only if*

$$x_1 = \sum_{i=0}^{m-1} y_i x_0^i$$

One can view X as points on $\mathbb{F}_q^2$, and Y as one-variable polynomials defined on $\mathbb{F}_q$ of degree at most $m-1$. Now $\Gamma_{q,m}$ is simply the *incidence bipartite graph* of the points and the polynomials where a point $P \in X$ and a polynomial $F \in Y$ form an edge if and only if $P = (w, F(w))$ for some $w \in \mathbb{F}_q$.

For any vertex x of a graph G, we use $d(x)$ to denote the degree of x, that is, the number of neighbors of x in G. Further, for any set U of vertices, we use $d(U)$ to denote the number of common neighbors of vertices in U. When $U = \{x, y\}$, we use $d(x, y) = d(\{x, y\})$ for short. The following proposition collects some useful properties of $\Gamma_{q,m}$.

Proposition 2.2. *For every prime power q and integer $m \geq 2$, $\Gamma_{q,m}$ has the following properties:*

(i) *$\forall\, x \in X$, $d(x) = q^{m-1}$.*

(ii) *$\forall\, y \in Y$, $d(y) = q$.*

(iii) *$\forall\, y, y' \in Y$, if $y \neq y'$, then $d(y, y') \leq m-1$.*

(iv) *$\forall\, x, x' \in X$, let $x = (x_0, x_1)$ and $x' = (x_0', x_1')$. If $x_0 \neq x_0'$, then $d(x, x') = q^{m-2}$. If $x_0 = x_0'$ and $x_1 \neq x_1'$, then $d(x, x') = 0$.*

(v) *Let $U \subseteq X$ such that $1 \leq |U| \leq m$, then $d(U) \leq q^{m-|U|}$.*

Proof. (i) For every $x = (x_0, x_1) \in X$, to find a neighbor $y = (y_0, \ldots, y_{m-1})$ of x, one can choose y_i for $1 \leq i \leq m-1$ freely and then let $y_0 = x_1 - \sum_{i=1}^{m-1} y_i x^i$. Thus $d(x) = q^{m-1}$.

(ii) For every $y = (y_0, \ldots, y_{m-1})$, to find a neighbor $x = (x_0, x_1)$ of y, one can choose x_0 freely, and then let $x_1 = \sum_{i=0}^{m-1} y_i x_0^i$. Thus $d(y) = q$.

(iii) For every $y = (y_0, \ldots, y_{m-1}), y' = (y_0', \ldots, y_{m-1}') \in Y$, if $x = (x_0, x_1)$ is a common neighbor of y and y', then x_0 is a solution to the equation $\sum_{i=0}^{m-1} (y_i - y_i') x^i = 0$ where x is the only variable. By the Fundamental Theorem of Algebra for finite fields, such an equation has at most $m-1$ solutions. Since x_1 is determined by x_0, we conclude that $d(y, y') \leq m-1$.

(iv) For every $x = (x_0, x_1), x' = (x'_0, x'_1) \in X$, if $x_0 \neq x'_0$, then every common neighbor $y = (y_0, \ldots, y_{m-1})$ corresponds to a solution to a collection of two linear equations that are linearly independent. The solution space of such a collection of linear equations has rank $m-2$, which implies that the number of solutions is q^{m-2}. Thus in this case $d(x, x') = q^{m-2}$. On the other hand, if $x_0 = x'_0$ and $x_1 \neq x'_1$, then for every $y = (y_0, \ldots, y_n) \in Y$, $x_1 - \sum_{i=0}^{m-1} y_i x_0^i \neq x'_1 - \sum_{i=0}^{m-1} y_i x_0'^i$, which implies that the two equations cannot equal 0 at the same time. Thus $d(x, x') = 0$.

(v) Let $|U| = k$, and let $x^{(1)} = (x_0^{(1)}, x_1^{(1)}), \ldots, x^{(k)} = (x_0^{(k)}, x_1^{(k)})$ be the vertices in U. Then each common neighbor $y = (y_0, \ldots, y_{m-1})$ corresponds to a solution to the collection of k linear equations $\sum_{i=0}^{m-1} y_i x_0^{(t)\,i} = x_1^{(t)}$, $1 \leq t \leq k$. If there exists $1 \leq t_1 < t_2 \leq k$ such that $x_0^{(t_1)} = x_0^{(t_2)}$, then we must have $x_1^{(t_1)} \neq x_1^{(t_2)}$ since $x^{(t_1)}$ and $x^{(t_2)}$ are different. Then by the same argument as in (iv) we know that $d(U) = 0$. On the other hand, if all $x_0^{(t)}$ are distinct, then the solution space of the collection of linear equations has rank $m-k$, which implies that the number of solutions is q^{m-k}. Thus in this case $d(x, x') = q^{m-k}$.

□

For all $k \geq 3$, let $H_{q,k}$ be a k-uniform hypergraph on $X = X(\Gamma_{q,k-1})$ whose edges are all k-sets $\{x_1, \ldots, x_k\} \subseteq X$ such that there exists an element $y \in Y = Y(\Gamma_{q,k-1})$ such that $\{x_1, \ldots, x_k\} \subseteq N(y)$. By Proposition 2.2, $H_{q,k}$ is the union of q^{k-1} k-uniform cliques on q vertices such that each vertex is contained in q^{k-2} cliques and the vertex sets of every two cliques intersect in at most $k-2$ vertices. Let $H^*_{q,k}$ be the k-uniform hypergraph obtained by replacing each maximal clique of $H_{q,k}$ with a random complete k-partite k-graph on the same vertex set. More formally, for each $y \in Y$, we color the vertices in $N(y)$ with k colors $\{1, \ldots, k\}$ uniformly at random, and for each $1 \leq i \leq k$ we let $X_{y,i} \subseteq N(y)$ be the set of vertices with color i, and then we replace the clique on $N(y)$ with a complete k-partite k-graph on $N(y)$ with k-partition $X_{y,1} \sqcup \cdots \sqcup X_{y,k}$. It is easy to check the following proposition.

Proposition 2.3. *If F is a non-degenerate slowly growing k-graph, then $H^*_{q,k}$ is F-free.*

Proof. Consider an ordering $e_1, \ldots, e_t$ of the edges of F such that

$$\forall\, i \in \{2, \ldots, t\},\ \left| e_i \setminus \bigcup_{j=1}^{i-1} e_j \right| \leq 1.$$

Equivalently, we have

$$\forall\, i \in \{2, \ldots, t\},\ \left| e_i \cap \bigcup_{j=1}^{i-1} e_j \right| \geq k-1.$$

We claim that every copy of F in $H_{q,k}$ must be fully contained in one of the q^{k-1} k-uniform cliques of size q. Indeed, suppose that we want to build a copy of F in $H_{q,k}$ by consecutively picking the edges in the order given above. Then the fact that every two cliques of $H_{q,k}$ intersect in at most $k-2$ vertices shows that we must pick every edge in the clique containing the previous edges. Since $H^*_{q,k}$ is obtained from $H_{q,k}$ by replacing every clique by a complete k-partite k-graph and F itself is not k-partite, this proves the statement. □

We will fix an instance of $H^*_{q,k}$ with good *Balanced Supersaturation*, which means that each induced subgraph of $H^*_{q,k}$ on $q^{1+o(1)}$ vertices contains many edges that are evenly distributed. Using Balanced Supersaturation together with the Hypergraph Container Lemma [2, 20], we can find upper bounds on the number of independent sets in $H^*_{q,k}$ of size $t = (\log q)^2 q^{\frac{1}{k-1}}$.

We then take a random subset W of $V(H^*_{q,k})$ where each vertex is sampled independently with probability $p = \Theta(\frac{t}{q})$ as in [18]. Finally, our construction is obtained by arbitrarily deleting a vertex from each independent set of size t in $H^*_{q,k}[W]$.

We will give the details in the following sections.

3 Pseudorandomness of $\Gamma_{q,k-1}$

In this section we show the pseudorandomness of $\Gamma_{q,k-1}$, which will be useful later in showing balanced supersaturation of $H^*_{q,k}$.

Given an n-vertex graph G, let A_G be the adjacency matrix of G, which is the $n \times n$ symmetric matrix where

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

Let $\lambda_1(G) \geq \cdots \geq \lambda_n(G)$ denote the eigenvalues of A_G. If G is a bipartite graph with bipartition $V_1 \sqcup V_2$, we say G is (d_1, d_2)*-regular* if $d(v) = d_1$ for all $v \in V_1$ and $d(v) = d_2$ for all $v \in V_2$.

The seminal expander mixing lemma is an important tool that relates edge distribution to the second eigenvalue of a graph. Here we make use of the bipartite version.

Lemma 3.1 (Theorem 5.1, [10]). *Suppose that G is a (d_1, d_2)-regular bipartite graph with bipartition $V_1 \sqcup V_2$. Then for every $S \subset V_1$ and $T \subset V_2$, the number of edges between S and T, denoted by $e(S,T)$, satisfies*

$$\left| e(S,T) - \frac{d_2}{|V_1|}|S||T| \right| \leq \lambda_2(G)\sqrt{|S||T|}.$$

By Proposition 2.2, we know $\Gamma_{q,k-1}$ is (q^{k-2}, q)-regular. For convenience, from now on we let $n = |V(\Gamma_{q,k-1})| = q^2 + q^{k-1}$, $A = A_{\Gamma_{q,k-1}}$, $\lambda_i = \lambda_i(\Gamma_{q,k-1})$ for all $1 \leq i \leq n$, and let $d_1 = q^{k-2}$, $d_2 = q$.

Lemma 3.2. $\lambda_2 = q^{\frac{k}{2}-1}$.

Proof. Define the matrix

$$M = \left[\begin{array}{c|c} 0 & J \\ \hline J^t & 0 \end{array} \right],$$

where J is the $|X| \times |Y|$ all-one matrix. We will show that

$$A^3 = (q-1)q^{k-3}M + q^{k-2}A. \tag{1}$$

By definition, for any $x \in X$ and $y \in Y$, $A^3(x,y)$ is the number of walks of length three of the form $xy'x'y$ in $\Gamma_{q,k-1}$. There are two cases.

Case 1: $xy \in E(\Gamma_{q,k-1})$**.** When $x' = x$, the number of choices for y' is q^{k-2}. When $x' \neq x$, the number of choices for x' is $q-1$ and for each such x', by Proposition 2.2 (iv), the number of choices for y' is q^{k-3}. Thus in this case the number of walks $xy'x'y$ is $q^{k-2} + (q-1)q^{k-3}$.

Case 2: $xy \notin E(\Gamma_{q,k-1})$**.** Suppose $x = (x_0, x_1)$ and $x' = (x'_0, x'_1)$. If $x_0 = x'_0$, then $x_1 \neq x'_1$, and hence, by Proposition 2.2 (iv), x and x' have no common neighbor. When $x_0 \neq x'_0$ the number of choices for x' is $q-1$ and for each such x', the number of choices for y' is q^{k-3}, again by Proposition 2.2 (iv). Thus in this case the number of walks $xy'x'y$ is $(q-1)q^{k-3}$.

Combining the two cases above we obtain Equation (1). Next, let u_X be the characteristic vector of X, that is, $u_X(v) = 1$ for each $v \in X$ and $u_X(v) = 0$ otherwise. Similarly, let u_Y be the characteristic vector of Y. Let $a_1 = \sqrt{d_1}u_X + \sqrt{d_2}u_Y$ and let $a_n = \sqrt{d_1}u_X - \sqrt{d_2}u_Y$. It is easy to check that $\lambda_1 = -\lambda_n = \sqrt{d_1 d_2}$ and that a_1, a_n are eigenvectors corresponding to λ_1 and λ_n. Since A is symmetric, the spectral theorem implies that A has an orthornormal basis of eigenvectors. Hence, for each $1 < i < n$, there exists an eigenvector a_i corresponding to λ_i such that a_i is orthogonal to both a_1 and a_n. Thus a_i is orthogonal to u_X and u_Y, which implies that $M \cdot a_i = 0$. Multiplying both sides of Equation (1) by a_i, we obtain $\lambda_i^3 = q^{k-2}\lambda_i$. Because the rank of A is larger than 2, there exists at least one $\lambda_i \neq 0$, and hence $\lambda_i = \pm q^{\frac{k}{2}-1}$. Note that since $\Gamma_{q,k-1}$ is bipartite, we have $\lambda_i = \lambda_{n-i+1}$. Therefore, $\lambda_2 = q^{\frac{k}{2}-1}$. □

Let S be a subset of X with size $|S| = rq$. If we pick $y \in Y$ uniformly at random, then the expectation of $|N(y) \cap S|$ is r. Thus intuitively, the vertex set of a "typical" clique in $H_{q,k}$ intersects S in $\Theta(r)$ vertices. The following lemma shows that a substantial portion of all cliques are "typical".

Lemma 3.3. *Let S be a subset of X with size $|S| = rq$. For $0 < \delta < 1$, let*

$$Y_\delta = \left\{ y \in Y \;\middle|\; (1-\delta)r \le |N(y) \cap S| \le (1+\delta)r \right\}.$$

Then $|Y_\delta| \ge \left(1 - \frac{2}{\delta^2 r}\right) q^{k-1}$.

Proof. Let

$$Y_+ = \left\{ y \in Y \;\middle|\; |N(y) \cap S| > (1+\delta)r \right\} \text{ and } Y_- = \left\{ y \in Y \;\middle|\; |N(y) \cap S| < (1-\delta)r \right\}.$$

Apply Lemma 3.1 with $G = \Gamma_{q,k-1}$ and $T = Y_+$. Together with Lemma 3.2, we have

$$|e(S, Y_+) - \frac{q}{q^2} rq|Y_+|| \le q^{\frac{k}{2}-1} \sqrt{rq|Y_+|}.$$

By definition, $e(S, Y_+) \ge |Y_+|(1+\delta)r$. Thus $\delta r |Y_+| \le q^{\frac{k-1}{2}} \sqrt{r|Y_+|}$, which implies $|Y_+| \le \frac{q^{k-1}}{\delta^2 r}$. Similarly, we can show that $|Y_-| \le \frac{q^{k-1}}{\delta^2 r}$. Therefore,

$$|Y_\delta| = |Y| - |Y_+| - |Y_-| \ge \left(1 - \frac{2}{\delta^2 r}\right) q^{k-1}.$$

□

4 Balanced Supersaturation

In this section, we show that $H^*_{q,k}$ has balanced supersaturation with positive probability. We need to use the following concentration inequality.

Proposition 4.1 (Corollary 2.27, [11])**.** *Let $Z_1, \dots, Z_t$ be independent random variables, with Z_i taking values in a set Λ_i. Assume that a function $f : \Lambda_1 \times \cdots \times \Lambda_t \to \mathbb{R}$ satisfies the following Lipschitz condition for some numbers c_i:*

(L) If two vectors $z, z' \in \Lambda \times \cdots \times \Lambda_t$ differ only in the i^{th} coordinate, then $|f(z) - f(z')| \le c_i$.

Then, the random variable $X = f(Z_1, \dots, Z_t)$ satisfies, for any $\lambda \ge 0$,

$$\Pr(X \le \mathbb{E}(X) - \lambda) \le \exp\left(-\frac{\lambda^2}{2\sum_{i=1}^t c_i^2}\right)$$

Recall that $H^*_{q,k}$ is the k-uniform hypergraph obtained by replacing each maximal clique of $H_{q,k}$ with a random complete k-partite k-graph on the same vertex set. Concretely, for each $y \in Y$, we color the vertices in $N(y)$ with k colors $\{1, \dots, k\}$ uniformly at random, and for each $1 \le i \le k$ we let $X_{y,i} \subseteq N(y)$ be the set of vertices with color i, and then we replace the clique on $N(y)$ with a complete k-partite k-graph on $N(y)$ with k-partition $X_{y,1} \sqcup \cdots \sqcup X_{y,k}$. Note that the colorings for different cliques are independent.

Given a k-graph H, let $\Delta_i(H)$ denote the maximum integer such that there exists $S \subseteq V(H)$ such that $|S| = i$ and the number of edges containing S is $\Delta_i(H)$.

Lemma 4.2. *For q sufficiently large in terms of k, with positive probability, every $S \subseteq X$ with $|S| \geq 4kq$ satisfies the following. There exists a subgraph $H \subset H^*_{q,k}[S]$ such that, for all $1 \leq i \leq k$,*

$$\Delta_i(H) \leq \frac{6(16k)^{2k}|E(H)|}{|S|}\left(\frac{q^{\frac{1}{k-1}}}{|S|}\right)^{i-1}.$$

Proof. For a fixed $S \subseteq X$ with $|S| \geq 4kq$, let $r = |S|/q \geq 4k \geq 12$ and let

$$Y_{1/2} = \left\{y \in Y \;\middle|\; r/2 \leq |N_{\Gamma_{q,k-1}}(y) \cap S| \leq 3r/2\right\}.$$

By Lemma 3.3 we have $|Y_{1/2}| \geq q^{k-1}/3$.

Let H be a subgraph of $H^*_{q,k}[S]$ with edge set

$$E(H) = \left\{e \in E(H^*_{q,k}[S]) \;\middle|\; \exists y \in Y_{1/2} \text{ such that } e \in N(y)\right\}.$$

In other words, H contains only edges that are in the "typical" cliques. Define the random variable $Z = |E(H)|$. For all $y \in Y_{1/2}$ and $v \in N_{\Gamma_{q,k-1}}(y)$, let $A_{y,v}$ be the random variable with values in $\{1,\ldots,k\}$ such that $A_{y,v} = i$ if vertex v receives color i in the clique on $N_{\Gamma_{q,k-1}}(y)$. Let $B_1,\ldots,B_t$ be an arbitrary order of $A_{y,v}$ for $y \in Y_{1/2}$ and $v \in N_{\Gamma_{q,k-1}}(v)$. Clearly, Z is determined by $B_1,\ldots,B_t$, i.e., there exists a function $f : [k]^t \to \mathbb{N}$ such that $Z = f(B_1,\ldots,B_t)$. Observe that changing the color of a vertex v in a typical clique will only affect the number of edges containing v in that clique, which is at most $\binom{3r/2-1}{k-1} \leq (2r)^{k-1}$, since a typical clique has size at most $3r/2$. In other words, if two vectors $b, b' \in [k]^t$ differ in only one coordinate, then

$$|f(b) - f(b')| \leq (2r)^{k-1}.$$

Note that for any k vertices in a typical clique, the probability that they form an edge in H is $\frac{k!}{k^k}$. Hence, by linearity of expectation and the fact that a typical clique has size at least $r/2$ and that $r/2 - k \geq r/4$, we have

$$\mathbb{E}(Z) \geq |Y_{1/2}|\binom{r/2}{k}\frac{k!}{k^k} \geq \frac{q^{k-1}}{3}\frac{(r/2-k)^k}{k^k} \geq \frac{r^k q^{k-1}}{3(4k)^k}.$$

Thus by Proposition 4.1 with $\lambda = \frac{r^k q^{k-1}}{6(4k)^k}$, $c_i = (2r)^{k-1}$, and the fact that $t \leq |Y_{1/2}|(3r/2) \leq |Y|(3r/2) \leq 3rq^{k-1}/2$, we have

$$\Pr\left(Z \leq \frac{r^k q^{k-1}}{6(4k)^k}\right) \leq \exp\left(-\frac{(\frac{r^k q^{k-1}}{6(4k)^k})^2}{2(3rq^{k-1}/2)((2r)^{k-1})^2}\right) \leq \exp\left(-\frac{rq^{k-1}}{500(8k)^{2k}}\right).$$

Using the union bound, the probability that there exists an $S \subseteq X$ with $|S| = s = rq \geq 4kq$ such that $Z \leq \frac{r^k q^{k-1}}{6(4k)^k}$ is at most

$$\sum_{s=4kq}^{q^2}\binom{q^2}{s}\exp\left(-\frac{sq^{k-2}}{500(8k)^{2k}}\right) \leq \sum_{s=1}^{q^2}\exp\left(-\frac{sq^{k-2}}{1000(8k)^{2k}}\right) < 1$$

given that q is sufficiently large in terms of k.

Hence with positive probability, for every $S \subseteq X$ with $|S| = rq \geq 4kq$, the corresponding H satisfies $|E(H)| \geq \frac{r^k q^{k-1}}{6(8k)^{2k}}$. Let $J \subseteq S$ be such that $|J| = i$ and $1 \leq i \leq k-1$. Note that, by Proposition 2.2 (v), the number of y such that $J \subseteq N_{\Gamma_{q,k-1}}(y)$ is at most q^{k-1-i}, and for each such $y \in Y_{1/2}$ the number of edges in $N_{\Gamma_{q,k-1}}(y) \cap S$ containing J is at most $\binom{3r/2-i}{k-i} \leq (2r)^{k-i}$. Hence we have

$$\Delta_i(H) \leq (2r)^{k-i}q^{k-1-i}.$$

In addition, we know that $\Delta_k(H) \le 1$. By $|E(H)| \ge \frac{r^k q^{k-1}}{6(8k)^{2k}}$ and $|S| = rq$, we have

$$\frac{6(16k)^{2k}|E(H)|}{|S|}\left(\frac{q^{\frac{1}{k-1}}}{|S|}\right)^{i-1} \ge 2^{2k} r^{k-i} q^{k-1+\frac{i-1}{k-1}-i}.$$

Note that when $1 \le i \le k-1$, given that q is sufficiently large, we have

$$2^{2k} r^{k-i} q^{k-1+\frac{i-1}{k-1}-i} \ge (2r)^{k-i} q^{k-1-i} \ge \Delta_i(H),$$

and when $i = k$,

$$2^{2k} r^{k-i} q^{k-1+\frac{i-1}{k-1}-i} = 2^{2k} \ge \Delta_k(H).$$

Combining the inequalities above we have for all $1 \le i \le k$,

$$\Delta_i(H) \le \frac{6(16k)^{2k}|E(H)|}{|S|}\left(\frac{q^{\frac{1}{k-1}}}{|S|}\right)^{i-1},$$

concluding the proof. Note that a stronger bound actually holds for all $i \le k-1$, and the claimed bound only arises from the case $i = k$. $\square$

5 Counting independent sets

We make use of the hypergraph container method developed independently by Balogh, Morris and Samotij [2] and Saxton and Thomason [20]. Here we make use of the following simplified version of Theorem 1.5 in [17]:

Theorem 5.1 (Theorem 1.5, [17])**.** *For every integer $k \ge 2$, there exists a constant $\epsilon > 0$ such that the following holds. Let $B, L \ge 1$ be positive integers and let H be a k-graph satisfying*

$$\Delta_i(H) \le \frac{|E(H)|}{L}\left(\frac{B}{|V(H)|}\right)^{i-1}, \quad \forall 1 \le i \le k. \tag{2}$$

Then there exists a collection $\mathcal{C}$ of subsets of $V(H)$ such that:

(i) *For every independent set I of H, there exists $C \in \mathcal{C}$ such that $I \subset C$;*

(ii) *For every $C \in \mathcal{C}$, $|C| \le |V(H)| - \epsilon L$;*

(iii) *We have*

$$|\mathcal{C}| \le \exp\left(\frac{\log\left(\frac{|V(H)|}{B}\right) B}{\epsilon}\right).$$

Next, we use Theorem 5.1 together with Lemma 4.2 to count the number of independent sets of size $q^{\frac{1}{k-1}}(\log q)^2$ in $H^*_{q,k}$.

Theorem 5.2. *For every $k \ge 3$, there exists a constant $c' > 0$ such that, when q is sufficiently large, we can fix an instance of $H^*_{q,k}$ such that the number of independent sets of size $t = q^{\frac{1}{k-1}}(\log q)^2$ of $H^*_{q,k}$ is at most*

$$\left(\frac{c' q}{t}\right)^t.$$

Proof. By Lemma 4.2, we can fix an instance of $H^*_{q,k}$ such that for every $S \subset V(H^*_{q,k})$ with $|S| \ge 4kq$ there exists a subgraph H of $H^*_{q,k}[S]$ such that for all $1 \le i \le k$,

$$\Delta_i(H) \le \frac{6(16k)^{2k}|E(H)|}{|S|} \left(\frac{q^{\frac{1}{k-1}}}{|S|} \right)^{i-1}. \tag{3}$$

We will first prove the following claim.

Claim 5.3. *There exists a constant $\epsilon > 0$ such that for every $S \subset V(H^*_{q,k})$ with $|S| > 4kq$, there exists a collection $\mathcal{C}_S$ of at most*

$$\exp\left(\frac{\log q \cdot q^{\frac{1}{k-1}}}{\epsilon} \right)$$

subsets of S such that:

(i) *For every independent set I of $H^*_{q,k}[S]$, there exists $C \in \mathcal{C}_S$ such that $I \subset C$;*

(ii) *For every $C \in \mathcal{C}_S$, $|C| \le (1-\epsilon)|S|$.*

Proof. Fix an arbitrary $S \subset V(H^*_{q,k})$ with $|S| \ge 4kq$. By Lemma 4.2 there exists a subgraph H of $H^*_{q,k}[S]$ satisfying Equation (3). By Equation (3), it is easy to check that equation (2) holds for H, with $L = \frac{|S|}{6(16k)^{2k}}$ and $B = q^{\frac{1}{k-1}}$. Hence by Theorem 5.1, there exist a constant ϵ' (not depending on S) and a collection $\mathcal{C}_S$ of subsets of S such that

(i) For every independent set I of H, there exists $C \in \mathcal{C}_S$ such that $I \subset C$;

(ii) For every $C \in \mathcal{C}_S$, $|C| \le |V(H)| - \epsilon' L \le \left(1 - \frac{\epsilon'}{6(16k)^{2k}}\right)|S|$;

(iii) We have

$$|\mathcal{C}_S| \le \exp\left(\frac{\log\left(\frac{|V(H)|}{B}\right) B}{\epsilon'} \right) \le \exp\left(\frac{\log(q^2) q^{\frac{1}{k-1}}}{\epsilon'} \right).$$

Since H is a subgraph of $H^*_{q,k}[S]$, every independent set of $H^*_{q,k}[S]$ is also an independent set of H. Therefore, by taking ϵ sufficiently small with respect to ϵ' and k, we conclude that $\mathcal{C}_S$ has the desired properties. ☐

Now we apply Claim 5.3 iteratively as follows. Fix the constant ϵ guaranteed by Claim 5.3. Let $\mathcal{C}_0 = \{V(H^*_{q,k})\}$. Let $t_0 = |V(H^*_{q,k})| = q^2$ and let $t_i = (1-\epsilon)t_{i-1}$ for all $i \ge 1$. Let m be the smallest integer such that $t_m \le 4kq$. Clearly $m = O(\log q)$. Given a set of containers $\mathcal{C}_i$ such that every $C \in \mathcal{C}_i$ satisfies $|C| \le t_i$, we construct $\mathcal{C}_{i+1}$ as follows: for every $C \in \mathcal{C}_i$, if $|C| \le t_{i+1}$, then we put it into $\mathcal{C}_{i+1}$; otherwise, if $|C| > t_{i+1}$, by Claim 5.3, there exists a collection $\mathcal{C}'$ of containers for $H^*_{q,k}[C]$ such that every $C' \in \mathcal{C}'$ satisfies $|C'| < (1-\epsilon)|C| \le t_{i+1}$ — now we put every element of $\mathcal{C}'$ into $\mathcal{C}_{i+1}$. Let $\mathcal{C} = \mathcal{C}_m$. Note that

$$\frac{|\mathcal{C}_i|}{|\mathcal{C}_{i-1}|} \le \exp\left(\frac{\log q \cdot q^{\frac{1}{k-1}}}{\epsilon} \right).$$

Thus

$$|\mathcal{C}_m| = \prod_{i=1}^{m} \frac{|\mathcal{C}_i|}{|\mathcal{C}_{i-1}|} \le \exp\left(m \frac{\log q \cdot q^{\frac{1}{k-1}}}{\epsilon} \right).$$

As $m = O(\log q)$, we conclude that there exists a constant $c'' > 0$ such that

$$|\mathcal{C}| = |\mathcal{C}_m| \le \exp\left(c''(\log q)^2 q^{\frac{1}{k-1}} \right).$$

Also, by definition, we have $|C| \le 4kq$ for every $C \in \mathcal{C}$.

Recall that $t = (\log q)^2 q^{\frac{1}{k-1}}$ and let N_t be the number of independent set of H of size t. Since every independent set of H of size t is contained in some $C \in \mathcal{C}$, we have, for some constant $c' > 0$,

$$N_t \leq |\mathcal{C}| \binom{4kq}{t} \leq \left(\frac{c'q}{t}\right)^t.$$

□

6 Proof of Theorem 1.2

Proof of Theorem 1.2. For every sufficiently large prime power q, we let $t = (\log q)^2 q^{\frac{1}{k-1}}$. By Theorem 5.2 we can fix an instance of $H^*_{q,k}$ such that the number of independent sets of $H^*_{q,k}$ of size t is at most

$$\left(\frac{c'q}{t}\right)^t.$$

for some constant $c' > 0$. Let W be a random subset of $V(H^*_{q,k})$ where each vertex is sampled independently with probability $p = \frac{t}{c'q}$. Note that $p < 1$ as q is sufficiently large. Then the expected number of independent set of size t in $H^*_{q,k}[W]$ is at most

$$\left(\frac{c'q}{t}\right)^t p^t \leq 1.$$

Let $W' \subseteq W$ be obtained by arbitrarily deleting one vertex in each independent set of size t. Thus the expectation of $|W'|$ is at least

$$pq^2 - 1 = \frac{(\log q)^2}{c'} q^{\frac{k}{k-1}} - 1.$$

Hence there exists a choice W' with at least this many vertices. Let $H' = H^*_{q,k}[W']$. By definition of W', we have $\alpha(H') < t$. Moreover, by Proposition 2.3 we know that H' is F-free. Thus, we have

$$r(F, t) \geq \frac{(\log q)^2}{c'} q^{\frac{k}{k-1}}.$$

Recall that $t = (\log q)^2 q^{\frac{1}{k-1}}$. It is well-known that for every integer n there exists a prime q such that $n/2 \leq q \leq n$. Thus for every n sufficiently large, it is easy to find a prime q such that

$$(\log q)^2 q^{\frac{1}{k-1}} \leq n \leq 2(\log q)^2 q^{\frac{1}{k-1}}.$$

Therefore we conclude that there exists a constant $c > 0$ such that for all n sufficiently large,

$$r(F, n) \geq \frac{cn^k}{(\log n)^{2k-2}}.$$

□

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi. A note on Ramsey numbers. *Journal of Combinatorial Theory, Series A*, 29(3):354–360, 1980.

[2] J. Balogh, R. Morris, and W. Samotij. Independent sets in hypergraphs. *Journal of the American Mathematical Society*, 28(3):669–709, 2015.

[3] T. Bohman and P. Keevash. The early evolution of the H-free process. *Inventiones mathematicae*, 181(2):291–336, 2010.

[4] T. Bohman and P. Keevash. Dynamic concentration of the triangle-free process. *Random Structures & Algorithms*, 58(2):221–293, 2021.

[5] T. Bohman, D. Mubayi, and M. Picollelli. The independent neighborhoods process. *Israel Journal of Mathematics*, 214:333–357, 2016.

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

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

[8] G. Fiz Pontiveros, S. Griffiths, and R. Morris. The triangle-free process and the Ramsey number $R(3,k)$. *Memoirs of the American Mathematical Society*, 263(1274):v+125, 2020.

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

[10] W. H. Haemers. Interlacing eigenvalues and graphs. *Linear Algebra and its Applications*, 226/228:593–616, 1995.

[11] S. Janson, T. Luczak, and A. Rucinski. *Random graphs.* John Wiley & Sons, 2011.

[12] 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.

[13] A. Kostochka, D. Mubayi, and J. Verstraëte. Hypergraph Ramsey numbers: triangles versus cliques. *Journal of Combinatorial Theory, Series A*, 120(7):1491–1507, 2013.

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

[15] Y. Li, C. C. Rousseau, and W. Zang. Asymptotic upper bounds for Ramsey functions. *Graphs and Combinatorics*, 17:123–128, 2001.

[16] S. Mattheus and J. Verstraëte. The asymptotics of $r(4,t)$. *Annals of Mathematics*, 199(2):919–941, 2024.

[17] R. Morris, W. Samotij, and D. Saxton. An asymmetric container lemma and the structure of graphs with no induced 4-cycle. *Journal of the European Mathematical Society*, 26(5):1655–1711, 2024.

[18] D. Mubayi and J. Verstraëte. A note on pseudorandom Ramsey graphs. *Journal of the European Mathematical Society*, 26(1):153–161, 2024.

[19] K. T. Phelps and V. Rödl. Steiner triple systems with minimum independence number. *Ars Combinatoria*, 21:167–172, 1986.

[20] D. Saxton and A. Thomason. Hypergraph containers. *Inventiones mathematicae*, 201(3):925–992, 2015.

[21] J. B. Shearer. A note on the independence number of triangle-free graphs. *Discrete Mathematics*, 46(1):83–87, 1983.

[22] J. Spencer. Ramsey's theorem—a new lower bound. *Journal of Combinatorial Theory, Series A*, 18(1):108–115, 1975.

[23] J. Spencer. Asymptotic lower bounds for Ramsey functions. *Discrete Mathematics*, 20:69–76, 1977.