On generalized Ramsey numbers of Erdős and Rogers

Andrzej Dudek∗
Department of Mathematics
Western Michigan University
Kalamazoo, MI
andrzej.dudek@wmich.edu

Troy Retter
Department of Mathematics and
Computer Science
Emory University
Atlanta, GA
tretter@emory.edu

Vojtěch Rödl†
Department of Mathematics and
Computer Science
Emory University
Atlanta, GA
rodl@mathcs.emory.edu

January 16, 2014

Abstract

Extending the concept of Ramsey numbers, Erdős and Rogers introduced the following function. For given integers $2 \leq s < t$ let

$$f_{s,t}(n) = \min \{ \max \{|W| : W \subseteq V(G) \text{ and } G[W] \text{ contains no } K_s \} \},$$

where the minimum is taken over all $K_t$-free graphs $G$ of order $n$. In this paper, we show that for every $s \geq 3$ there exist constants $c_1 = c_1(s)$ and $c_2 = c_2(s)$ such that $f_{s,s+1}(n) \leq c_1(\log n)^c_2\sqrt{n}$. This result is best possible up to a polylogarithmic factor. We also show for all $t - 2 \geq s \geq 4$, there exists a constant $c_3 = c_3(s)$ such that $f_{s,t}(n) \leq c_3\sqrt{n}$. In doing so, we partially answer a question of Erdős by showing that

$$\lim_{n \to \infty} \frac{f_{s+1,s+2}(n)}{f_{s,s+2}(n)} = \infty$$

for any $s \geq 4$.

1 Introduction

In a graph $G$, a set $S \subseteq V(G)$ is independent if $G[S]$ does not contain a copy of $K_2$. More generally for any integer $s$, a set $S \subseteq V(G)$ can be called $s$-independent if $G[S]$ does not contain a copy of $K_s$. With this in mind, define the $s$-independence number of $G$, denoted by $\alpha_s(G)$, to be the size of the largest $s$-independent set in $G$. The classical Ramsey number

---

*Supported in part by Simons Foundation Grant #244712 and by a grant from the Faculty Research and Creative Activities Award (FRACAA), Western Michigan University.

†Supported in part by NSF grant DMS 0800070.
\(R(t, u)\) can be defined in this language as the smallest integer \(n\) such that every graph of order \(n\) contains either a copy of \(K_t\) or a 2-independent set of size \(u\). In other words, \(R(t, u)\) is the smallest integer \(n\) such that

\[
u \leq \min\{\alpha_2(G) : G \text{ is a } K_t\text{-free graph of order } n\}.
\]

Observe that if the right hand side of the above inequality is understood as a function of \(n\) and \(t\), then so is the classical Ramsey number.

A more general problem, first addressed by Hajnal (see, e.g., [10]), results by replacing the standard independence number by the \(s\)-independence number for some \(2 \leq s < t\).

Following this approach, in 1962 Erdős and Rogers [10] introduced the function

\[f_{s,t}(n) = \min\{\alpha_s(G) : G \text{ is a } K_t\text{-free graph of order } n\}.\]

The lower bound \(k \leq f_{s,t}(n)\) means that every \(K_t\)-free graph of order \(n\) contains a subset of \(k\) vertices with no copy of \(K_s\). The upper bound \(f_{s,t}(n) < \ell\) means that there exists a \(K_t\)-free graph of order \(n\) such that every subset of \(\ell\) vertices contains a copy of \(K_s\).

The case \(t = s + 1\) has received considerable attention over the last 50 years, in part due to the fact that it creates a general upper bound. For \(t' > t\) we have \(f_{s,t'}(n) \leq f_{s,t}(n)\), so it follows that \(f_{s,t}(n) \leq f_{s,s+1}(n)\) for all \(t \geq s + 1\). A first nontrivial upper bound for \(f_{s,s+1}(n)\) was established by Erdős and Rogers [10], which was subsequently addressed by Bollobás and Hind [4], Krivelevich [12, 13], Alon and Krivelevich [2], Dudek and Rödl [7], and most recently Wolfovitz [17]. The first nontrivial lower bound established by Bollobás and Hind [4] was later slightly improved by Krivelevich [13]. The most recent general bounds for \(s \geq 3\) were of the form:

\[
\Omega\left(\sqrt{\frac{n \log n}{\log \log n}}\right) = f_{s,s+1}(n) = O(n^{\frac{3}{2}}).
\]

The precise lower bound of (1) was first explicitly stated by Dudek and Mubayi [6], and was based upon their observation that the result of Krivelevich [13] could be slightly strengthened by incorporating a result of Shearer [14]. The upper bound of (1) appears in [7], where it was also conjectured that for all sufficiently large \(s\) the upper bound could be improved to show that

\[
f_{s,s+1}(n) = n^{\frac{1}{2} + o(1)}.
\]

Recently, Wolfovitz [17] showed that (2) holds when \(s = 3\). In this paper, we prove (2) for every \(s \geq 3\), establishing an upper bound that is tight up to a polylogarithmic factor. Our construction is based upon the projective plane random graph model that was first introduced in [7] and subsequently improved in [17]. Our more careful analysis of this model extends the ideas of Wolfovitz [17] and further build upon the ideas in [7], [12], and [13].

**Theorem 1.1** For every \(s \geq 3\) there is a constant \(c = c(s)\) such that

\[
f_{s,s+1}(n) \leq c(\log n)^{4s^2} \sqrt{n}.
\]
For the case $t = s + 2$, it follows from a result of Sudakov [15] (see also [7] for a simplified formula) that $f_{s,s+2}(n) = \Omega(n^{a_2})$, where $a_2 = \frac{1}{2} - \frac{1}{8(s-1)}$. On the other hand, clearly $f_{s,s+2}(n) \leq f_{s,s+1}(n)$. When $s \geq 4$, we establish an improved upper bound that omits the logarithmic factor.

**Theorem 1.2** For every $s \geq 4$ there is a constant $c = c(s)$ such that

$$f_{s,s+2}(n) \leq c\sqrt{n}.$$

This establishes the following corollary which provides the best known bounds on $f_{s,t}(n)$ for $t < 2s$.

**Corollary 1.3** For every $4 \leq s \leq t - 2$ there is a constant $c = c(s)$ such that

$$f_{s,t}(n) \leq f_{s,s+2}(n) \leq c\sqrt{n}.$$

When $t \geq 2s$, the upper bound $c(\log n)^{1/(s-1)n^{s/(t+1)}}$ of Krivelevich [12] remains best. For all values of $t > s + 1$, the best lower bounds follow from a recursive formula defined by Sudakov [15, 16]. We will return to these results concerning the general case in our concluding remarks. More related results are summarized in the survey [8].

Now that our two main results have been stated, we turn our attention towards an old question of Erdős [9], who asked if for fixed integers $s + 1 < t$,

$$\lim_{n \to \infty} \frac{f_{s+1,t}(n)}{f_{s,t}(n)} = \infty. \quad (3)$$

This central conjecture in the area is still wide open and asks for a rather precise estimation of $f_{s,t}(n)$. It is known due to Sudakov [16] that (3) holds for

$$(s, t) \in \{(2, 4), (2, 5), (2, 6), (2, 7), (2, 8), (3, 6)\}.$$

Observe that Theorem 1.2 together with the lower bound of [13] (and [7]) implies that for $s \geq 4$,

$$\frac{f_{s+1,s+2}(n)}{f_{s,s+2}(n)} \geq \frac{\Omega\left(\sqrt{\frac{n \log n}{\log \log n}}\right)}{O(\sqrt{n})} = \Omega\left(\sqrt{\frac{\log n}{\log \log n}}\right) \quad n \to \infty \to \infty.$$

That is, (3) holds for all pairs $(s, t) \in \{(4, 6), (5, 7), (6, 8), \ldots\}$.

In what follows, consider $s$ to be an arbitrary fixed integer and $n$ sufficiently large, i.e. $n \geq n_0(s)$. We will show that there exists a $K_{s+1}$-free graph of order $n$ such that every subset of $c(\log n)^{1/2s^2} \sqrt{n}$ vertices contains a copy of $K_s$ and that there exists a $K_{s+2}$-free graph of order $n$ such that every subset of $c\sqrt{n}$ vertices contains a copy of $K_s$. Indeed, this establishes Theorems 1.1 and 1.2 as stated (for all $n$), since the constant factors can subsequently be inflated to accommodate the finitely many cases where $n \leq n_0$. For simplicity, we do not round numbers that are supposed to be integers either up or down; this is justified since these rounding errors are negligible to the asymptomatic calculations we will make.
In Section 2, we begin our construction by considering the random hypergraph \( \mathbb{H} \) which is essentially the random hypergraph obtained from the affine plane by taking each hyperedge (line) with some uniform probability. We then use \( \mathbb{H} \) in Section 3 to construct a random graph \( G \) by replacing each hyperedge by a complete \( s \)-partite graph. In Section 4, the proof of Theorem 1.2 considers an induced subgraph of \( G \) whereas the proof of Theorem 1.1 considers yet another random subgraph of \( G \) which is analyzed by way of the Local Lemma.

Below we will use the standard notation to denote the neighborhood and degree of \( v \in G \) by \( N_G(v) \) and \( d_G(v) \) respectively.

## 2 The Hypergraph \( H \)

The **affine plane** of order \( q \) is an incidence structure on a set of \( q^2 \) points and a set of \( q^2 + q \) lines such that: any two points lie on a unique line; every line contains \( q \) points; and every point lies on \( q + 1 \) lines. It is well known that affine planes exist for all prime power orders. Clearly, an incidence structure can be viewed as a hypergraph with points corresponding to vertices and lines corresponding to hyperedges; we will use this terminology interchangeably.

In the affine plane, call lines \( L \) and \( L' \) parallel if \( L \cap L' = \emptyset \). In the affine plane there exist \( q + 1 \) sets of \( q \) pairwise parallel lines. (For more details see, e.g., [5].) Let \((V, L)\) be the hypergraph obtained by removing a parallel class of \( q \) lines from the affine plane or order \( q \).

The following lemma establishes some properties of this graph. We consider this graph in place of the projective or affine plane for numerical convenience.

**Lemma 2.1** For \( q \) prime, the \( q \)-uniform, \( q \)-regular hypergraph \((V, L)\) of order \( q^2 \) satisfies:

(P1) Any two vertices are contained in at most one hyperedge;

(P2) For every \( A \in \binom{V}{q} \), \(|\{L \in L : L \cap A \neq \emptyset\}| \geq \frac{q^2}{2} \).

**Proof.** By construction, \((V, L)\) is \( q \)-uniform, \( q \)-regular, and satisfies (P1). Consider \( A = \{v_1, v_2, \ldots, v_q\} \). Define \( d_+(v_i) = |\{L \in L : L \cap \{v_1, v_2, \ldots, v_i\} = \{v_i\}\}| \). Hence by property (P1), \( d_+(v_i) \geq q - i + 1 \). We now compute

\[
|\{L \in L : L \cap A \neq \emptyset\}| \geq \sum_{i=1}^{q} d_+(v_i) \geq \binom{q+1}{2} \geq \frac{q^2}{2}.
\]

\( \square \)

The objective of this section is to establish the existence of a certain hypergraph \((V, L') \subseteq (V, L)\) by considering a random sub-hypergraph of \((V, L)\). Preceding this, we introduce some terminology. Define

\[
L'_A = \{L \in L' : L \cap A \neq \emptyset\}, \quad \text{and} \quad L'_{B,\gamma} = \{L \in L' : |L \cap B| \geq \gamma\}.
\]

Call \( S \subseteq V \) \( L' \)-complete if every pair of points in \( S \) is contained in some common line in \( L' \). Let \( L(x, y) \) denote the unique line in \( L \) containing \( x \) and \( y \), provided such a line exists.
We will now distinguish 3 types of $L'$-dangerous subsets as depicted in Figure 1. The first two types have 5 vertices \{v_1, v_2, v_3, v_4, x\} and third type has 6 vertices \{v_1, v_2, v_3, v_4, y, z\}. All 3 types of dangerous sets must be $L'$-complete and have 4 points \{v_1, v_2, v_3, v_4\} in general position (i.e. no three points lie on a common line). Additionally we specify:

Type 1 $L'$-dangerous

The points \{v_1, v_2, v_3, v_4, x\} are in general position.

Type 2 $L'$-dangerous

The point x is contained in precisely one of the 6 lines $L(v_i, v_j)$ for 1 \( \leq i < j \leq 4 \). Up to relabeling, say $x \in L(v_2, v_3)$.

Type 3 $L'$-dangerous

The points y and z are each contained in exactly two of the lines $L(v_i, v_j)$ for 1 \( \leq i < j \leq 4 \). Up to relabeling, say $y \in L(v_1, v_3) \cap L(v_2, v_4)$ and $z \in L(v_1, v_2) \cap L(v_3, v_4)$.

All concepts above were defined relative to the subset $L' \subseteq L$. Obviously we can define the concepts $L$-complete, $L$-dangerous, $L_A$, and $L_{B,\gamma}$ related to the set $L$ analogously.

We are now ready to state the main result of this section.

Lemma 2.2 Let $q$ be a sufficiently large prime and $\alpha = (\log q)^2$. Then, there exists a $q$-uniform hypergraph $H = (V, L')$ of order $q^2$ such that:

(H1) Any two vertices are contained in at most one hyperedge;

(H2) For every $v \in V$, $d_H(v) \leq 2\alpha$;

(H3) $|\mathcal{D}| \leq 2\alpha^8 q$, where $\mathcal{D}$ is the set of $L'$-dangerous subsets;

(H4) For every $A \in \binom{V}{q}$, $|L'_A| \geq \frac{aq}{4}$;

(H5) For every integer $1 \leq \gamma \leq \frac{q}{16}$ and every $B \in \binom{V}{16\gamma q}$, $|L'_{B,\gamma}| \geq \frac{aq}{8}$.

5
Before proving the above lemma, we state a basic form of the Chernoff bound (as appearing in Corollary 2.3 of [11]) and mention what we will refer to as the union bound.

**Chernoff Bound** If $X \sim \text{Bi}(n,p)$ and $0 < \varepsilon \leq \frac{3}{2}$, then

$$\Pr \left( |X - E(X)| \geq \varepsilon \cdot E(X) \right) \leq 2 \exp \left\{ - \frac{E(X)\varepsilon^2}{3} \right\}.$$  

**Union Bound** If $E_i$ are events, then

$$\Pr \left( \bigcup_{i=1}^{k} E_i \right) \leq k \cdot \max \{ \Pr(E_i) : i \in [k] \}.$$  

**Proof of Lemma 2.2.** Take $(V, L)$ to be a hypergraph established by Lemma 2.1. Let $H = (V, L')$ be a random sub-hypergraph of $(V, L)$ where every line in $L$ is taken independently with probability $\frac{\alpha}{q} = (\log q)^2 q$. Since $H$ is a subgraph of $(V, L)$ any two vertices are in at most one line, so $H$ always satisfies (H1). We will show $H$ fails to satisfy (H2) and (H4) with probability at most $o(1)$ and that $H$ fails to satisfy (H3) with probability at most $\frac{1}{2}$. Together this implies $H$ satisfies (H1)-(H4) with probability at least $1 - \frac{1}{2} - o(1)$, establishing the existence of a hypergraph $H$ that satisfies (H1)-(H4). Finally, we use a counting argument to show that any such $H$ necessarily satisfies (H5).

**H2:** We first show that the probability that there exists a vertex of degree greater than $2\alpha$ is $o(1)$. Observe for fixed $v \in H$, $d_H(v) \sim \text{Bi}(q, \frac{\alpha}{q})$ and $E(d_H(v)) = \alpha$. So by the Chernoff bound with $\varepsilon = 1$,

$$\Pr \left( d_H(v) \geq 2\alpha \right) \leq \Pr \left( |d_H(v) - \alpha| \geq \alpha \right) \leq 2 \exp \left\{ - \frac{\alpha}{3} \right\}.$$  

Thus by the union bound the probability that there exists some $v \in V$ with $d_H(v) \geq 2\alpha$ is at most

$$q^2 \cdot 2 \exp \left\{ - \frac{\alpha}{3} \right\} = 2 \exp \left\{ 2 \log q - \frac{(\log q)^2}{3} \right\} = o(1).$$  

**H3:** In order to show $|D| > 4\alpha^3 q$ with probability at most $\frac{1}{2}$, we begin by counting the number of $L$-dangerous subsets of each type. Clearly the number of Type 1 $L$-dangerous subsets is at most $\binom{q}{3}$. To count the number of Type 2 $L$-dangerous subsets, first choose $\{v_1, v_2, v_3, v_4\}$ then $x$, observing $x$ must lie on one of the 6 lines which each have at most $q$ vertices. Thus there are at most $\binom{q}{4}(6q)$ configurations of this type. To count the number of Type 3 $L$-dangerous subsets, observe the lines $L(v_i, v_j)$ for $1 \leq i < j \leq 6$ intersect at most 3 points other than $v_1, v_2, v_3, v_4$. Hence there are at most $\binom{q}{4}(\binom{3}{2})$ subsets of this type in $L$. 

6
Since $L$-dangerous subsets of Type 1, Type 2, and Type 3 have 10, 8, and 7 lines respectively, an $L$-dangerous subset of each type will be $L'$-dangerous with respective probabilities $(\frac{\alpha}{q})^{10}, (\frac{\alpha}{q})^{8}$, and $(\frac{\alpha}{q})^{7}$. By the linearity of expectation, we now compute

$$E(|D|) \leq \left(\frac{q^2}{5}\right) \cdot \left(\frac{\alpha}{q}\right)^{10} + \left(\frac{q^2}{4}\right) \cdot (6q) \cdot \left(\frac{\alpha}{q}\right)^{8} + \left(\frac{q^2}{4}\right) \cdot \left(\frac{3}{2}\right) \cdot \left(\frac{\alpha}{q}\right)^{7} \leq \alpha^{10} + \frac{q\alpha^8}{4} + \frac{q\alpha^7}{8} \leq q\alpha^8.$$ 

Thus, the Markov inequality yields,

$$\Pr\left(|D| \geq 2\alpha^8q\right) \leq \Pr\left(|D| \geq 2E(|D|)\right) \leq \frac{1}{2}.$$

(H4): We will now prove that the probability that there exists $A \in \binom{V}{q}$ such that $|L_A'| < \frac{\alpha q}{4}$ is $o(1)$. Begin by considering any fixed $A \in \binom{V}{q}$. Then by Lemma 2.1, $|L_A| \geq \frac{q^2}{4}$, so we may fix $X \subseteq L_A$ with $|X| = \frac{q^2}{4}$. Let $X' = X \cap L'$. Since each line in $X$ appears in $X'$ independently with probability $\frac{\alpha}{q}$, $|X'| \sim \text{Bi}\left(\frac{q^2}{4}, \frac{\alpha}{q}\right)$ and $E(|X'|) = \frac{\alpha q}{2}$. Hence by the Chernoff bound with $\varepsilon = \frac{1}{2}$,

$$\Pr\left(|L_A'| < \frac{\alpha q}{4}\right) \leq \Pr\left(|X'| < \frac{\alpha q}{4}\right) \leq \Pr\left(|X'| - \frac{\alpha q}{2} \geq \frac{\alpha q}{4}\right) \leq 2\exp\left\{-\frac{\alpha q}{24}\right\}.$$ 

Consequently by the union bound, the probability that there exits some $A \subseteq V$, $|A| = q$, with $|L_A'| < \frac{\alpha q}{4}$ is at most

$$\binom{q^2}{q} \cdot 2\exp\left\{-\frac{\alpha q}{24}\right\} = q^{2q} \cdot 2\exp\left\{-\frac{(\log q)^2 q}{24}\right\} = \exp\left\{2q\log q - \frac{q(\log q)^2}{24}\right\} = o(1).$$

(H5): Finally, we will establish the following deterministic property: If $H$ satisfies (H2) and (H4), then $H$ also satisfies (H5).

Consider any integer $\gamma$ with $1 \leq \gamma \leq \frac{q}{10}$ and $B \in \binom{V}{16\gamma}$. Let $B = B_1 \cup B_2 \cup \ldots \cup B_{16\gamma}$ be a partition of $B$ into $16\gamma$ sets of size $q$. Consider the auxiliary bipartite graph $Aux$ with bipartition $\{B_1, B_2, \ldots, B_{16\gamma}\} \cup L'$. Join $B_i$ to $L \in L'$ if $B_i \cap L \neq \emptyset$. By property (H4) $d_{Aux}(B_i) \geq \frac{\alpha q}{4}$ for all $i \in [16\gamma]$, and thus the number of edges in $Aux$ satisfies

$$|e(Aux)| \geq \frac{\alpha q}{4} \cdot |\{B_1, B_2, \ldots, B_{16\gamma}\}| = 4\alpha q \gamma. \quad (4)$$

On the other hand, clearly $d_{Aux}(L) \leq |\{B_1, B_2, \ldots, B_{16\gamma}\}| = 16\gamma$ for all $L \in L'$ and by definition $d_{Aux}(L') \leq \gamma$ for all $L \in L' \setminus L_{B,\gamma}$. Also keeping in mind that by (H2)

$$|L' \setminus L_{B,\gamma}| \leq |L'| = \sum_{v \in V} d_{H(v)} \leq q^2 \cdot \frac{2\alpha q}{q} = 2\alpha q,$$

we compute

$$|e(Aux)| \leq |L'_{B,\gamma}| \cdot 16\gamma + |L' \setminus L_{B,\gamma}| \cdot \gamma \leq |L'_{B,\gamma}| \cdot 16\gamma + 2\alpha q \cdot \gamma. \quad (5)$$

Comparing (4) and (5), we obtain

$$4\alpha q \gamma \leq |e(Aux)| \leq |L'_{B,\gamma}| \cdot 16\gamma + 2\alpha q \gamma,$$

which yields $|L'_{B,\gamma}| \geq \frac{\alpha q}{8}$. 

\qed
3 The Graph $G$

Based upon the hypergraph $H$ established in the previous section, we will construct a graph $G$ with the following properties.

**Lemma 3.1** Let $q$ be a sufficiently large prime, $\alpha = (\log q)^2$, $\beta = (\log q)^{4s^2}$, and $s \geq 2$. Then, there exists a graph $G = (V, E)$ of order $q^2$ such that:

(G1) For every $C \in \binom{V}{16sq}$, $G[C]$ contains a copy of $K_s$;

(G2) For every $U \in \binom{V}{64sqq}$, $G[U]$ contains $\frac{\alpha \beta q}{s^2}$ edge disjoint copies of $K_s$;

(G3) Every edge $xy \in E$ is in at most $6^s \alpha^{2s-2}q$ copies of $K_{s+1}$;

(G4) If $s \geq 4$, then $G$ can be made $K_{s+2}$ free by removing $2^{0^8q}$ vertices.

**Proof.** Fix a hypergraph $H = (V, L')$ as established by Lemma 2.2. Construct the random graph $G$ as follows. For every $L \in L'$, let $\chi_L : L \rightarrow [s]$ be a random partition of the vertices of $L$ into $s$ classes, where for every $v \in L$, a class $\chi_L(v) \in [s]$ is assigned uniformly and independently at random. Then, let $xy \in E$ if $\{x, y\} \subseteq L$ for some $L \in L'$ and $\chi_L(x) \neq \chi_L(y)$. Thus for every $L \in L'$, $G[L]$ is a complete $s$-partite graph with vertex partition $L = \chi_L^{-1}(1) \cup \chi_L^{-1}(2) \cup \cdots \cup \chi_L^{-1}(s)$ (where the classes need not have the same size and the unlikely event that a class is empty is permitted). Observe that not only are $G[L]$ and $G[L']$ edge disjoint for distinct $L, L' \in L'$, but also that the partitions for $L$ and $L'$ were determined independently.

We will show $G$ does not satisfy (G1) and (G2) with probability at most $o(1)$ and that $G$ always satisfies (G3) and (G4). Hence the probability that $G$ satisfies properties (G1)-(G4) is at least $1 - o(1)$, implying the existence of a graph $G$ described in the lemma.

(G1): Consider any $C \in \binom{V}{16sq}$. We will bound the probability that $G[C] \not\supseteq K_s$. By (H1) with $\gamma = s$, the set of lines $L'_{C,s}$ that intersect $C$ in at least $s$ vertices has cardinality $|L'_{C,s}| \geq \frac{\alpha q}{s^2}$. For each $L \in L'_{C,s}$, let $X_L$ be the event $K_s \not\subseteq G[L \cap C]$. Since $|L \cap C| \geq s$ for all $L \in L'_{C,s}$, $\Pr(X_L) \leq 1 - \frac{s^2}{s^2}$. By independence,

$$\Pr \left( K_s \not\subseteq G[C] \right) \leq \Pr \left( \bigcap_{L \in L'_{C,s}} X_L \right) \leq \left( 1 - \frac{s^2}{s^2} \right)^{|L'_{C,s}|} \leq \left( 1 - \frac{s^2}{s^2} \right)^{\frac{\alpha q}{s^2}} \leq \exp \left\{ -\frac{s^2 \alpha q}{s^2} \right\}.$$ 

So by the union bound, the probability that there exists a subset of $16sq$ vertices in $G$ that contains no $K_s$ is at most

$$\left( \frac{q^2}{16sq} \right)^{exp \left\{ -\frac{s^2 \alpha q}{s^2} \right\}} \leq q^{16sq} \exp \left\{ -\frac{s^2 \alpha q}{s^2} \right\} = \exp \left\{ 16sq \log q - \frac{s^2 \alpha q (\log q)^2}{8s^2} \right\} = o(1),$$

where in the first inequality we used $\binom{a}{b} \leq \left( \frac{ae}{b} \right)^b$. 

8
(G2): For arbitrary $U \in \binom{V}{64s^2q}$, we will bound the probability that $G[U]$ does not contain $\frac{a^2\beta q}{16}$ edge disjoint copies of $K_s$. By (H5) with $\gamma = 4s\beta$, we may fix a subset $\mathcal{Z}_U \subseteq \mathcal{L}'_{U,4s\beta}$ of exactly $\frac{a^2\beta q}{8}$ lines with the property that each line has intersection at least $4s\beta$ with $U$. We will consider the lines in $\mathcal{Z}_U$ that contain the complete balanced $s$-partite graph on $2s\beta$ vertices, which we denote by $K_{2\beta,...,2\beta}$. Define $\mathcal{Z}'_U = \{L \in \mathcal{Z}_U : K_{2\beta,...,2\beta} \subseteq G[L \cap U]\}$. The graph $K_{2\beta,...,2\beta}$ certainly contains at least $\beta^2$ edge disjoint $K_s$ (Since we may choose a prime $\beta \leq p \leq 2\beta$ and it follows from [1] that we may then decompose $K_{p,...,p}$ into $p^2$ edge disjoint copies of $K_s$; this suffices for our purposes, but stronger results are known). Thus if we show $|\mathcal{Z}'_U| \geq \frac{a^2\beta q}{16}$ it will imply that $G[U]$ contains at least $|\mathcal{Z}'_U| \cdot \beta^2 \geq \frac{a^2\beta^2q}{16}$ edge disjoint copies of $K_s$.

For $L \in \mathcal{Z}_U$, let $Y_L$ be the event that $L \not\in \mathcal{Z}'_U$ and fix $L_{4s\beta} \subseteq L \cap U$ with $|L_{4s\beta}| = 4s\beta$. Now $Y_L$ will occur only if $| \chi_L^{-1}(i) \cap L_{4s\beta} | < 2\beta$ for some $i \in [s]$. Defining $X_i = | \chi_L^{-1}(i) \cap L_{4s\beta} |$, observe $X_i \sim Bi(4s\beta, \frac{1}{s})$ and $E(X_i) = 4\beta$. Chernoff's inequality reveals

$$\Pr \left( X_i < 2\beta \right) \leq \Pr \left( |X_i - E(X_i)| \geq \frac{E(X_i)}{2} \right) \leq 2 \exp \left\{ -\frac{\beta}{12} \right\} = 2 \exp \left\{ -\frac{\beta}{3} \right\}.$$  

By the union bound, $\Pr(Y_L) \leq \Pr \left( \bigcup_{i \in [s]} (X_i \leq 2\beta) \right) \leq s \cdot 2 \exp \left\{ -\frac{\beta}{3} \right\}$.

By independence, the probability that $Y_L$ occurs for at least $\frac{a^2\beta q}{16} = \frac{|\mathcal{Z}_U|}{2}$ of the lines in $\mathcal{Z}_U$ is at most

$$\binom{|\mathcal{Z}_U|}{\frac{|\mathcal{Z}_U|}{2}} \left( 2s \exp \left\{ -\frac{\beta}{3} \right\} \right)^{|\mathcal{Z}_U|/2} \leq 4^{|\mathcal{Z}_U|/2} \left( 2s \exp \left\{ -\frac{\beta}{3} \right\} \right)^{|\mathcal{Z}_U|/2} = \left( 8s \exp \left\{ -\frac{\beta}{3} \right\} \right)^{\frac{a^2\beta q}{16}}.$$  

That is, we have shown $|\mathcal{Z}'_U| < \frac{a^2\beta q}{16}$ with probability at most $\left( 8s \exp \left\{ -\frac{\beta}{3} \right\} \right)^{\frac{a^2\beta q}{16}}$ for fixed $U$. Thus by the union bound, the probability that there exits some $U \subseteq V$ with $|U| = 64s^2\beta q$ such that $|\mathcal{Z}'_U| < \frac{a^2\beta q}{16}$ is at most

$$\left( \frac{q^2}{64s^2\beta q} \right)^{\frac{a^2\beta q}{16}} \leq q^{64s^2\beta q} \left( \frac{a^2\beta q}{16} \right)^{\frac{a^2\beta q}{16}} \leq \exp \left\{ 64s^2\beta q \log q + \frac{a^2\beta q}{16} \log(8s) - \frac{a\beta q}{48} \right\} = o(1).$$

(G3): For any $xy \in E$, we will show the number of copies of $K_{s+1}$ that contain $xy$ is at most $6^s \alpha^{2s-2}$. Let $L \in \mathcal{L}'$ be the unique line such that $\{x, y\} \subseteq L$ as depicted in Figure 2. Let $N = (N_H(x) \cap N_H(y)) \setminus L$ be the set of all vertices not on $L$ that are collinear with both $x$ and $y$. Since $d_H(x), d_H(y) \leq 2\alpha$ by (H2), we infer that $|N| \leq 4\alpha^2$. Because $K_{s+1} \not\subseteq G[L]$, if a $K_{s+1}$ is to contain $x$ and $y$ it must contain at least one vertex $v \in N$. There are at most $|N| \leq 4\alpha^2$ choices for this vertex $v$. Once $v$ has been chosen, each of the remaining $s - 2$
vertices of the $K_{s+1}$ must lie in $N$ or in $L \cap N_H(v)$. Since $|N| + |L \cap N_H(v)| \leq 4\alpha^2 + 2\alpha$, the number of $K_{s+1}$ containing the edge $xy$ is at most $4\alpha^2(4\alpha^2 + 2\alpha)^s - 2 \leq 6\alpha^{2s-2}$.

$(G4)$: We will finally show that if $s \geq 4$, $G$ can be made $K_{s+2}$ free by removing at most $2\alpha^8q$ vertices. By (H3), all $L'$-dangerous sets can be destroyed by removing $2\alpha^8q$ vertices, so it suffices to show that every $K_{s+2}$ in $G$ contains a $L'$-dangerous subset.

Let $K$ be any copy of $K_{s+2}$ in $G$. By assumption $s \geq 4$, so $K$ must have at least 6 vertices, which clearly form a $L'$-complete set.

We first show that $K$ contains 4 vertices in general position. Suppose otherwise. Then there is some line $L \in L'$ that contains 3 vertices $\{p_1, p_2, p_3\}$ of $K$. Since $K_{s+1} \not\subseteq G[L]$, there must exist two vertices $a$ and $b$ in $K \setminus L$. Observe $\{a, b\}$ and any 2 vertices in $\{p_1, p_2, p_3\} \setminus L(a, b)$ are in general position.

Now fix 4 vertices $\{v_1, v_2, v_3, v_4\}$ of $K$ that are in general position and let $u_1, u_2$ be any two other vertices of $K$. Three cases are now considered. If either $u_1$ or $u_2$ do not lie on any of the 6 lines $L(v_i, v_j)$ for $1 \leq i < j \leq 4$, then there is a $L'$-dangerous subset of Type 1. If either $u_1$ or $u_2$ lie on exactly one line in $L(v_i, v_j)$ for $1 \leq i < j \leq 4$, then there is a $L'$-dangerous subset of Type 2. In the remaining case where both $u_1$ and $u_2$ each lie on at least 2 lines in $L(v_i, v_j)$ for $1 \leq i < j \leq 4$, then there is a $L'$-dangerous subset of Type 3. \qed

## 4 Proof of Theorem 1.1 and 1.2

Consider any sufficiently large integer $n$ and $s \geq 3$. By Bertrand’s postulate, we can find a prime $q$ such that $4n \leq q^2 \leq 16n$. Fix a graph $G$ procured by Lemma 3.1 of order $q^2$ and as before take

$$\alpha = (\log q)^2 \quad \text{and} \quad \beta = (\log q)^{4s^2}.$$  

Theorem 1.1 and Theorem 1.2 are now proved by considering different subgraphs of $G$ of order $n$.

**Proof of Theorem 1.2.** Consider the case where $s \geq 4$. To prove the theorem, we will show there exists a $K_{s+2}$-free induced subgraph of $G$ of order $n$ with the property that every subset of order $64s\sqrt{n}$ contains a copy of $K_s$. 

10
By (G1), every set of size $16sq$ (in $G$) contains a copy of $K_s$, so certainly every subset of size $64s\sqrt{n} \geq 16sq$ in any induced subgraph of $G$ must also contain a copy of $K_s$. Thus it will suffice to show that there is a $K_{s+2}$-free subset of $G$ of order $n$. But by (G4), we know that there is a set $R \subseteq V(G)$ of size $|R| = 2\alpha \delta q \leq n$ such that $G[V \setminus R]$ will be $K_{s+2}$-free. Finally since $|V \setminus R| \geq 4n - n \geq n$, the induced graph of $G$ on any $n$ vertices in $V \setminus R$ will have the desired properties.

Proof of Theorem 1.1. For $s \geq 3$, we will concentrate on constructing a $K_{s+1}$-free graph $G'$ on $q^2$ vertices with the property that every subset of size $64s\beta q$ vertices contains a copy of $K_s$. Since $\log(4n) \leq 2\log n$,

$$64s\beta q = 64s(\log q)^{4s^2} q \leq 64s(\log 4n)^{4s^2} 4\sqrt{n} \leq 2^{4s^2+8}\sqrt{s(\log n)^{4s^2}} \sqrt{n},$$

and so any induced subgraph of $G'$ of order $n$ will also be $K_{s+1}$-free and have the property that every set of order $2^{4s^2+8}\sqrt{s(\log n)^{4s^2}} \sqrt{n}$ contains a copy of $K_s$, exactly as desired.

Let $G'$ be a random subgraph of $G$ where each edge is taken with probability

$$\frac{1}{\gamma}, \quad \text{where } \gamma = (\log q)^8.$$

For a set $S \in \binom{V(G)}{s+1}$ that spans a copy of $K_{s+1}$ in $G$, let $A_S$ to be the event that all the edges of $S$ are in $G'$. Hence, $\bigcap A_S$ means that $K_{s+1} \not\subseteq G'$. For a set $U \in \binom{V(G)}{64s\beta q}$ let $K_U$ be a (fixed) set of

$$m = \frac{1}{16}\alpha \beta^2 q$$

equidistant copies $K_s$ contained in $U$, which are known to exist by (G2). Define $B_U$ to be the event that none of the $m$ edge disjoint $K_s$ appear in $G'$. Hence, $\bigcap B_U$ implies that for every $U \in \binom{V(G)}{64s\beta q}$ one of the disjoint copies of $K_s$ in $G[U]$ appears in $G'$. It will suffice to show that the probability that $(\bigcap A_S) \cap (\bigcap B_U)$ occurs is nonzero. In order to show this, we apply the Local Lemma (see, e.g., Lemma 5.1.1 in [3]).

Lovász Local Lemma Let $E_1, E_2, \ldots, E_k$ be events in an arbitrary probability space. A directed graph $D$ on the set of vertices $\{1, 2, \ldots, k\}$ is called a dependency digraph for the events $E_1, E_2, \ldots, E_k$ if for each $i$, $1 \leq i \leq k$, the event $E_i$ is mutually independent of all the events $\{E_j : (i,j) \notin D\}$. Suppose that $D$ is a dependency digraph for the above events and suppose there are real numbers $z_1, \ldots, z_k$ such that $0 \leq z_i < 1$ and $\Pr(E_i) \leq z_i \prod_{(i,j) \in D}(1 - z_j)$ for all $1 \leq i \leq k$. Then, $\Pr\left(\bigcap_{i=1}^{k} \overline{E_i}\right) > 0$.

Let $D$ be a dependency graph that corresponds to all events $A_S$ and $B_U$. Observe that $A_S$ depends only on the $\binom{s+1}{2}$ edges in $S$ and $B_U$ depends only on the $m(\frac{s}{2})$ edges of the $K_s$ in $K_U$. Also, observe that the number of events of the type $B_U$ is $\binom{q^2}{64s\beta q} \leq q^{64s\beta q}$. Thus by (G3), a fixed event $A_S$ depends on at most

$$d_{AA} = \binom{s+1}{2} 6^s \alpha^{2s-2}$$

11
other events $A_S$ and at most
\[ d_{AB} = q^{64s^3q} \]
events $B_U$. Similarly, a fixed event $B_U$ depends on at most
\[ d_{BA} = m \left( \frac{s}{2} \right) 6^s 2^{s-2} \]
events $A_S$ and at most
\[ d_{BB} = q^{64s^3q} \]
other events $B_{U'}$. Let
\[ x = \frac{1}{\alpha 2^{s^2}} \quad \text{and} \quad y = \frac{1}{(\log q)^4s^4q^{64s^3q}}. \]

To finish the proof, due to the Local Lemma it suffices to show that
\[ \left( \frac{1}{\gamma} \right)^{\left( \frac{s + 1}{2} \right)} \leq x (1 - x)^d_{AA} (1 - y)^d_{AB}, \]  
(6)
\[ \left( 1 - \left( \frac{1}{\gamma} \right)^{\left( \frac{s}{2} \right)} \right)^m \leq y (1 - x)^d_{BA} (1 - y)^d_{BB}. \]  
(7)

First we show that (6) holds. Using the fact that $e^{-2x} \leq 1 - x$ for $x$ sufficiently small (observe that $x \to 0$ with $q \to \infty$), a sufficient condition for (6) will be
\[ \left( \frac{1}{\gamma} \right)^{\left( \frac{s + 1}{2} \right)} \leq x e^{-2xd_{AA}} e^{-2yd_{AB}}, \]
and equivalently,
\[ \left( \frac{s + 1}{2} \right) \log (\gamma) \geq \log \left( \frac{1}{x} \right) + 2xd_{AA} + 2yd_{AB}. \]
The latter follows from the following three inequalities, which hold when $q$ is sufficiently large:
\[ \frac{2s^2}{2s^2 + 2s} \left( \frac{s + 1}{2} \right) \log (\gamma) \geq \log \left( \frac{1}{x} \right), \]
\[ \frac{s}{2s^2 + 2s} \left( \frac{s + 1}{2} \right) \log (\gamma) \geq 2xd_{AA}, \]
\[ \frac{s}{2s^2 + 2s} \left( \frac{s + 1}{2} \right) \log (\gamma) \geq 2yd_{AB}. \]

Similarly, using the facts that $e^{-2y} \leq 1 - y$ for $y$ sufficiently small and that $1 - \left( \frac{1}{\gamma} \right)^{\left( \frac{s}{2} \right)} \leq e^{-\left( \frac{1}{2} \right)^{\left( \frac{s}{2} \right)}}$, (7) will be satisfied if
\[ e^{-m\left( \frac{1}{\gamma} \right)^{\left( \frac{s}{2} \right)}} \leq ye^{-2xd_{BA}} e^{-2yd_{BB}}, \]
and equivalently,
\[ m\left(\frac{1}{\gamma}\right)^\left(\frac{1}{2}\right) \geq \log\left(\frac{1}{y}\right) + 2xd_{BA} + 2yd_{BB}. \]

As before the latter will follow from the following inequalities, which hold when \( q \) is sufficiently large:
\[ \frac{1}{3}m\left(\frac{1}{\gamma}\right)^\left(\frac{1}{2}\right) \geq \log\left(\frac{1}{y}\right), \]
\[ \frac{1}{3}m\left(\frac{1}{\gamma}\right)^\left(\frac{1}{2}\right) \geq 2xd_{BA}, \]
\[ \frac{1}{3}m\left(\frac{1}{\gamma}\right)^\left(\frac{1}{2}\right) \geq 2yd_{BB}. \]

This completes the proof of Theorem 1.1.

5 Concluding Remarks

We close this paper by discussing how the asymptotic behavior of \( f_{s,t}(n) \) changes for different values of \( 3 \leq s < t \).

If the difference between \( s \) and \( t \) is fixed, we make the following observation based upon the lower bound in Sudakov [15] (and Fact 3.5 in [7]) and Corollary 1.3.

**Observation 5.1** For any \( \epsilon > 0 \) and an integer \( k \geq 2 \) there is a constant \( s_0 = s_0(k, \epsilon) \) such that for all \( s \geq s_0 \),
\[ \Omega\left(n^{\frac{1}{2} - \epsilon}\right) = f_{s,s+k}(n) = O(\sqrt{n}). \]

In view of this observation and Theorem 1.2 we ask the following.

**Question 5.2** For any \( s \geq 3 \), is \( f_{s,s+2}(n) = o(\sqrt{n}) \)?

Another interesting question results from fixing the ratio between \( s \) and \( t \). The following is based upon [15] and [12] respectively.

**Observation 5.3** For any \( \epsilon > 0 \) and \( \lambda \geq 2 \) there is a constant \( s_0 = s_0(\lambda, \epsilon) \) such that for all \( s \geq s_0 \),
\[ \Omega\left(n^{\frac{1}{2\lambda} - \epsilon}\right) = f_{s,\lfloor \lambda s \rfloor}(n) = O\left(n^{\frac{1}{\lambda}}\right). \]

In particular, when \( \lambda = 3 \), we see \( \Omega(n^{1/6 - \epsilon}) = f_{s,\lfloor 3s \rfloor}(n) = O(n^{1/3}). \)

**Question 5.4** What is the asymptotic behavior of \( f_{s,\lfloor \lambda s \rfloor}(n) \)?

Recall that Erdős [9] asked if for fixed \( s + 2 \leq t \), \( \lim_{n \to \infty} \frac{f_{s+1,t}(n)}{f_{s,t}(n)} = \infty \). We ask a similar question, that if answered in the affirmative would imply an affirmative answer to the question of Erdős.

**Question 5.5** For all \( t > s \geq 3 \), is \( \lim_{n \to \infty} \frac{f_{s+1,t+1}(n)}{f_{s,t}(n)} = \infty \)?
6 Acknowledgment

We would like to thank Guy Wolfovitz for providing us with a manuscript of his work. We are also grateful to the referees for their valuable comments and suggestions.

References


