Induced subgraphs of K_r-free graphs and the Erdős–Rogers problem

Lior Gishboliner* Oliver Janzer† Benny Sudakov‡

Abstract

For two graphs F, H and a positive integer n, the function $f_{F,H}(n)$ denotes the largest m such that every H-free graph on n vertices contains an F-free induced subgraph on m vertices. This function has been extensively studied in the last 60 years when F and H are cliques and became known as the Erdős–Rogers function. Recently, Balogh, Chen and Luo, and Mubayi and Verstraëte initiated the systematic study of this function in the case where F is a general graph.

Answering, in a strong form, a question of Mubayi and Verstraëte, we prove that for every positive integer r and every K_{r-1}-free graph F, there exists some $\varepsilon_F > 0$ such that $f_{F,K_r}(n) = O(n^{1/2-\varepsilon_F})$. This result is tight in two ways. Firstly, it is no longer true if F contains K_{r-1} as a subgraph. Secondly, we show that for all $r \geq 4$ and $\varepsilon > 0$, there exists a K_{r-1}-free graph F for which $f_{F,K_r}(n) = \Omega(n^{1/2-\varepsilon})$. Along the way of proving this, we show in particular that for every graph F with minimum degree t, we have $f_{F,K_4}(n) = \Omega(n^{1/2-6/\sqrt{t}})$. This answers (in a strong form) another question of Mubayi and Verstraëte. Finally, we prove that there exist absolute constants $0 < c < C$ such that for each $r \geq 4$, if F is a bipartite graph with sufficiently large minimum degree, then $\Omega(n^{\frac{c}{\log r}}) \leq f_{F,K_r}(n) \leq O(n^{\frac{C}{\log r}})$. This shows that for graphs F with large minimum degree, the behaviour of $f_{F,K_r}(n)$ is drastically different from that of the corresponding off-diagonal Ramsey number $f_{K_2,K_r}(n)$.

1 Introduction

The Ramsey number $R(r,t)$ is the smallest n such that every n-vertex graph contains a clique of size r or an independent set of size t. The study of this function is one of the most important problems in discrete mathematics. The instances that have received the most attention are the "diagonal case" concerning $r = t$, and the case where r is fixed and $t \to \infty$ (which is often called the "off-diagonal case"). In this paper we will focus on the latter.

The first bound on this function was obtained by Erdős and Szekeres [14] in 1935, who proved that $R(r,t) = O(t^{r-1})$ for any fixed r and $t \to \infty$. Despite extensive research on the topic, the only (non-trivial) off-diagonal Ramsey number whose order of magnitude is known is $R(3,t)$. It was shown by Kim [20] in 1995 that $R(3,t) = \Omega(t^2/\log t)$, which matches an earlier upper bound by Ajtai, Komlós and Szemerédi [1]. Recently, a major breakthrough was obtained by Mattheus and Verstraëte [23], who proved that $R(4,t) \geq \Omega(t^3/(\log t)^4)$, matching the best known upper bound up to a polylogarithmic factor. Nevertheless, the problem of estimating $R(r,t)$ remains wide open for all $r \geq 5$, with the best bounds being

$$c_1(r)\frac{t^{\frac{r+1}{2}}}{(\log t)^{\frac{r+1}{2}-\frac{1}{r-2}}} \leq R(r,t) \leq c_2(r)\frac{t^{r-1}}{(\log t)^{r-2}},$$

*Department of Mathematics, University of Toronto, Canada. Email: lior.gishboliner@utoronto.ca

†Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, United Kingdom. Research supported by a fellowship at Trinity College. Email: oj224@cam.ac.uk

‡Department of Mathematics, ETH Zürich, Switzerland. Email: benjamin.sudakov@math.ethz.ch. Research supported in part by SNSF grant 200021-228014.

due to Bohman and Keevash [5] and Ajtai, Komlós and Szemerédi [1], respectively.

In 1962, Erdős and Rogers [13] considered the following generalization of the off-diagonal Ramsey problem. For positive integers $2 \le s < r$ and n, let $f_{s,r}(n)$ denote the largest m such that every K_r-free graph on n vertices contains a K_s-free induced subgraph on m vertices. Note that the Ramsey problem is recovered as the special case $s = 2$. The function $f_{s,r}$ has since become known as the Erdős–Rogers function and has attracted an extensive amount of research over the last 60 years (see, e.g. [13, 6, 21, 22, 2, 29, 28, 11, 12, 30, 9, 10, 16, 19, 25]).

In the last decade or so, there has been major progress towards finding the value of $f_{r-1,r}(n)$. Building on earlier work of Dudek and Rödl [11], Wolfovitz [30] proved that $f_{3,4}(n) \le n^{1/2+o(1)}$, matching the easy lower bound $f_{3,4}(n) \ge n^{1/2}$ up to the $o(1)$ term. Later, it was shown by Dudek, Retter and Rödl [10] that for all $r \ge 4$, we have $f_{r-1,r}(n) \le n^{1/2+o(1)}$, which is again tight up to the $o(1)$ term. Very recently, the upper bound was improved by Mubayi and Verstraëte [25], who showed that $f_{r-1,r}(n) = O(n^{1/2}\log n)$, coming close to the best known lower bound $f_{r-1,r}(n) = \Omega\left(\frac{n^{1/2}(\log n)^{1/2}}{(\log\log n)^{1/2}}\right)$, observed in [9].

While the $s = r-1$ case is more or less settled, the next case $s = r-2$ is already open in general. For $r = 4$, this problem is equivalent to determining the Ramsey numbers $R(4, k)$, and it follows from the recent breakthrough of Mattheus and Verstraëte [23] that $f_{2,4}(n) \le n^{1/3+o(1)}$, which is tight. Janzer and Sudakov [19] generalized this upper bound by proving that $f_{r-2,r}(n) \le n^{\frac{1}{2}-\frac{1}{8r-26}+o(1)}$ holds for all $r \ge 4$. It is unknown whether this is tight for $r > 4$; the best lower bound is $f_{r-2,r}(n) \ge n^{\frac{1}{2}-\frac{1}{6r-18}+o(1)}$, due to Sudakov [28].

Recently, Balogh, Chen and Luo [4] and Mubayi and Verstraëte [24] initiated a systematic study of the following generalization of the classical Erdős–Rogers function (see also [18] for an earlier paper in this direction). For graphs F and H and a positive integer n, we write $f_{F,H}(n)$ for the largest m such that every H-free graph on n vertices contains an F-free induced subgraph on m vertices. (Here H-free and F-free mean that they do not contain H or F as a not necessarily induced subgraph.) Both [4] and [24] are in fact mainly concerned with the case where H is a clique, and this will be the focus of our paper as well. Note that this problem still closely resembles the original Ramsey problem; the only difference is that we are looking for a large F-free induced subgraph rather than a large independent set.

Among other results, Mubayi and Verstraëte [24] proved that for every (non-empty) triangle-free graph F, we have $f_{F,K_3}(n) = n^{1/2+o(1)}$, thereby resolving the case where H is the triangle. Regarding the next case, namely that of $H = K_4$, they posed the following problem.

Problem 1.1 (Mubayi–Verstraëte [24])**.** *Is it true that for every triangle-free graph F, there exists $\varepsilon = \varepsilon_F > 0$ such that $f_{F,K_4}(n) = O(n^{1/2-\varepsilon})$?*

Our first result is an affirmative answer to Problem 1.1 in a more general form.

Theorem 1.2. *For every $r \ge 4$ and every K_{r-1}-free graph F, there exists $\varepsilon = \varepsilon_F > 0$ such that $f_{F,K_r}(n) = O(n^{1/2-\varepsilon})$.*

The assumption that F is K_{r-1}-free is necessary since if F contains K_{r-1}, then $f_{F,K_r}(n) \ge f_{K_{r-1},K_r}(n) \ge n^{1/2+o(1)}$ for any r. Mubayi and Verstraëte also conjectured that the $1/2$ in the exponent in Problem 1.1 cannot be replaced by anything smaller, and that this is witnessed by taking $F = K_{t,t}$ for large enough t.

Problem 1.3 (Mubayi–Verstraëte [24])**.** *Prove (or disprove) that for each $\varepsilon > 0$ there exists t such that $f_{K_{t,t},K_4}(n) = \Omega(n^{1/2-\varepsilon})$.*

We prove that this is indeed the case in the following more general form.

Theorem 1.4. *For every t and every graph F with minimum degree t, $f_{F,K_4}(n) = \Omega(n^{1/2-6/\sqrt{t}})$.*

In fact, using the same method, we prove that our Theorem 1.2 is tight for all r.

Theorem 1.5. *For every $r \geq 4$ and $\varepsilon > 0$ there is a K_{r-1}-free graph F such that $f_{F,K_r}(n) = \Omega(n^{1/2-\varepsilon})$.*

As a corollary of Theorem 1.4, we obtain the following result about graphs with large Turán number.

Corollary 1.6. *For every $\varepsilon > 0$ there exists some $\delta > 0$ such that if a bipartite graph F satisfies $ex(m,F) = \Omega(m^{2-\delta})$, then $f_{F,K_4}(n) = \Omega(n^{1/2-\varepsilon})$.*

This shows that for bipartite graphs F with large Turán number, the exponent in $f_{F,K_4}(n)$ is close to $1/2$. This complements a result of Balogh, Chen and Luo [4] which states that if $\mathrm{ex}(m,F) = O(m^{1+\alpha})$ for some $\alpha \in [0,1/2)$, then $f_{F,K_4}(n) \leq n^{\frac{1}{3-2\alpha}+o(1)}$.

Motivated by Theorem 1.4, it is natural to ask what happens if instead of K_4 one considers the case of a general clique K_r. Our methods allow us to address this question as well, and we obtain the following rather accurate estimates on $f_{F,K_r}(n)$ when F is a bipartite graph with large minimum degree.

Theorem 1.7. *For each $r \geq 4$ and $\varepsilon > 0$ there is t_0 such that for every $t \geq t_0$ and every graph F with minimum degree t, we have*

$$f_{F,K_r}(n) = \Omega(n^{\frac{1}{\lceil \log_2 r \rceil}-\varepsilon}).$$

Theorem 1.8. *There is an absolute constant $C > 0$ such that for every r and every bipartite graph F, we have*

$$f_{F,K_r}(n) = O(n^{\frac{C}{\log r}}).$$

From the above two theorems, we see that there are absolute constants $c, C > 0$ such that for every $r \geq 4$ and every bipartite graph F with large enough minimum degree (compared to r), we have

$$\Omega(n^{\frac{c}{\log r}}) \leq f_{F,K_r}(n) \leq O(n^{\frac{C}{\log r}}).$$

Note that this is in striking contrast with Ramsey numbers, for which we have $\Omega(n^{\frac{c}{r}}) \leq f_{K_2,K_r}(n) \leq O(n^{\frac{C}{r}})$. We also point out that both Theorem 1.7 and Theorem 1.8 use methods that are rather novel in the study of Erdős–Rogers functions.

As mentioned above, Mubayi and Verstraëte [24] proved that for all (non-empty) triangle-free graphs F, we have $f_{F,K_3}(n) = n^{1/2+o(1)}$. This shows that $f_{F,K_3}(n)$ is quite close to $f_{K_2,K_3}(n)$ for every triangle-free graph F. They asked to find an example where the two functions have different orders of magnitudes.

Problem 1.9 (Mubayi–Verstraëte [24])**.** *Find a triangle-free F for which $f_{F,K_3}(n)/f_{K_2,K_3}(n) \to \infty$.*

Note that by the celebrated result of Kim [20] on the Ramsey number $R(3,k)$, we have $f_{K_2,K_3}(n) = \Theta(\sqrt{n \log n})$. Problem 1.9 remains open, but in Subsection 3.3 we present a connection to the famous Zarankiewicz problem for 6-cycles, similar to the connection between Ramsey numbers and the Zarankiewicz problem discussed in [7].

Organization of the paper. In Section 2, we prove Theorems 1.4, 1.5 and 1.7 and Corollary 1.6. In Section 3, we prove Theorems 1.2 and 1.8. In this section we also discuss the problem of estimating $f_{F,K_3}(n)$ for an arbitrary triangle-free graph F, and reveal a connection to the Zarankiewicz problem for C_6. In Section 4, we give some concluding remarks.

In Section 2, logarithms are in base e, while in Section 3, logarithms are in base 2.

2 Lower bounds

In this section we prove Theorems 1.4, 1.5 and 1.7 and Corollary 1.6. We denote by $\alpha_F(G)$ the largest order of an F-free induced subgraph of G. The *s-domination number* $\gamma_s(F)$ of a graph F is the minimum k for which there is a set $A \subseteq V(F)$ with $|A| = k$ such that every $v \in V(F) \setminus A$ has at least s neighbours in A. We will need the following lemma, showing that graphs of large minimum degree have small s-domination number.

Lemma 2.1. *Let $t \geq 2$ and let F be a graph with minimum degree t. Let $\frac{6\log t}{t} \leq \delta \leq 1$ and set $s = \lfloor \frac{\delta t}{3} \rfloor$. Then $\gamma_s(F) \leq \delta \cdot v(F)$.*

Proof. TOPROVE 0 ∎

The following lemma, which we think is of independent interest, is a key for the proof of Theorems 1.4, 1.5 and 1.7. Here and below, for $X \subseteq V(G)$, we let $N(X)$ denote the common neighbourhood of X.

Lemma 2.2. *Let $0 < \delta < \beta < 1$, let F be a graph, let $s \geq 1$, and suppose that $\gamma_s(F) \leq \delta \cdot v(F)$. Let n be sufficiently large and let G be an n-vertex graph with $\alpha_F(G) < 0.5n^{\beta-2\delta}$. Then there are at least $0.5n^{(1-\beta+\delta)s}$ sets $X \subseteq V(G)$ of size s with $|N(X)| \geq n^{1-\beta}$.*

Proof. TOPROVE 1 ∎

2.1 Proof of Theorems 1.4 and 1.5 and Corollary 1.6

We will derive Theorems 1.4 and 1.5 from the following theorem.

Theorem 2.3. *Let $r \geq 4$ and let F be a graph which contains a copy of K_{r-2}. Let $\delta > 0$, and suppose that $\gamma_s(F) \leq \delta \cdot v(F)$ for $s = \lceil \frac{1}{\delta} \rceil$. Then $f_{F,K_r}(n) \geq 0.5n^{1/2-2\delta}$ holds for all sufficiently large n.*

Proof. TOPROVE 2 ∎

By combining Theorem 2.3 with Lemma 2.1, we get the following.

Theorem 2.4. *Let $r \geq 4$ and let F be a graph with minimum degree t which contains a copy of K_{r-2}. Then $f_{F,K_r}(n) \geq 0.5n^{1/2-6/\sqrt{t}}$ holds for all sufficiently large n.*

Proof. TOPROVE 3 ∎

Taking $r = 4$ in Theorem 2.4 immediately gives Theorem 1.4. Also, it is easy to see that there exists a graph F which has arbitrarily large minimum degree and contains K_{r-2} but not K_{r-1}. For example, we can take the complete $(r-2)$-partite graph with parts of size t, where t is sufficiently large. Hence, Theorem 2.4 implies Theorem 1.5.

To deduce Corollary 1.6, we will use the following result of Alon, Krivelevich and Sudakov.

Theorem 2.5 (Alon–Krivelevich–Sudakov [3]). *If F is a bipartite graph which does not contain a subgraph of minimum degree at least $t+1$, then $ex(n,F) = O(n^{2-\frac{1}{4t}})$.*

Proof. TOPROVE 4 ∎

2.2 Proof of Theorem 1.7

We will derive Theorem 1.7 from the following theorem.

Theorem 2.6. *Let F be a graph, let $\delta > 0$ and suppose that $\gamma_s(F) \leq \delta \cdot v(F)$ for $s = \lceil \frac{1}{\delta^3} \rceil$. Then for every $k \geq 1$, $f_{F,K_{2^k}}(n) \geq n^{\frac{1}{k} - 2^k \delta}$ for all sufficiently large n.*

Before proving Theorem 2.6, let us use it to prove Theorem 1.7.

Proof. TOPROVE 5 ∎

In the rest of this subsection, we prove Theorem 2.6. In the following lemma, $d(U,W)$ stands for the proportion of pairs $(u,w) \in U \times W$ for which uw is an edge.

Lemma 2.7. *Let $\delta, \varepsilon, \beta > 0$, Let $s \geq 2(\frac{\beta}{\varepsilon}+1)$ be an integer, and let F be a graph with $\gamma_s(F) \leq \delta \cdot v(F)$. Let n be sufficiently large and let G be an n-vertex graph with $\alpha_F(G) < 0.5 n^{\beta - 2\delta}$. Then there are $U, W \subseteq V(G)$ with $|U| \geq \Omega(n^{1-\beta-\frac{2\beta(\beta+\varepsilon)}{\varepsilon s}})$, $|W| \geq n^{1-\beta}$ and $d(U,W) \geq n^{-\varepsilon}$.*

Proof. TOPROVE 6 ∎

Proof. TOPROVE 7 ∎

3 Upper bound constructions

3.1 Proof of Theorem 1.8

It suffices to prove Theorem 1.8 for $F = K_{t,t}$ (since every bipartite graph is contained in $K_{t,t}$ for a sufficiently large t). Hence, Theorem 1.8 follows from the following result.

Theorem 3.1. *There is an absolute constant C such that for every $k \geq 2$ and every t we have $f_{K_{t,t},K_{2^k}}(n) = O(n^{C/k})$.*

For graphs G, H, the *lexicographic product* $G \cdot H$ is the graph obtained from G by substituting a copy of H for each vertex of G (and replacing edges of G with complete bipartite graphs). It is easy to see that $\omega(G \cdot H) = \omega(G) \cdot \omega(H)$ and $\chi(G \cdot H) \leq \chi(G) \cdot \chi(H)$.

Lemma 3.2. *For any positive integer t and graphs G and H, we have*

$$\alpha_{K_{t,t}}(G \cdot H) \leq \alpha(G)\alpha_{K_{t,t}}(H) + (t-1)\alpha_{K_{t,t}}(G).$$

In particular, $\alpha_{K_{t,t}}(G \cdot G) \leq t\alpha(G)\alpha_{K_{t,t}}(G)$.

Proof. TOPROVE 8 ∎

We construct K_{2^k}-free graphs with no large $K_{t,t}$-free set by induction on k. Roughly speaking, we start with a $K_{2^{k/2}}$-free graph G_0 with no large $K_{t,t}$-free set, take a union of it with a random graph on the same vertex set to obtain a graph H (where the random graph ensures that H has small independence number), and then consider $H \cdot H$. Then H has no large $K_{t,t}$-free set by Lemma 3.2. Unfortunately, H may contain a clique of size significantly greater than $2^{k/2}$, which means that $H \cdot H$ may contain a clique of size significantly greater than 2^k.

In order to overcome this issue, instead of considering the clique number, we consider the property of having no subgraph on $O(1)$ vertices with large chromatic number. This is more convenient because the chromatic number of the union of two graphs is at most the product of their chromatic numbers, whereas the clique number can be exponential in the clique numbers.

Definition 3.3. For an integer $r \geq 3$, let S_r be the set of all $\rho \geq 0$ with the property that for all positive integers t, s there is some $n_0 = n_0(\rho, r, t, s)$ such that for all $n \geq n_0$ there exists an n-vertex graph G in which every subgraph on s vertices is $(r-1)$-colourable and which has $\alpha_{K_{t,t}}(G) \leq n^{\rho}$.

Furthermore, let $\rho_r = \inf(S_r)$.

Note that $1 \in S_r$, so ρ_r is well-defined and $\rho_r \leq 1$.

Lemma 3.4. *Let H be a graph (on at least s vertices) in which every subgraph on s vertices is r-colourable. Then every subgraph of $H \cdot H$ on s vertices is r^2-colourable.*

Proof. TOPROVE 9 ■

We also need the following well-known properties of random graphs, which can be easily proved using the union bound.

Lemma 3.5. *Let s and r be fixed positive integers. Let $p = n^{-2/r}/\log n$. Then $G \sim G(n,p)$ satisfies the following properties.*

1. *Almost surely $\alpha(G) \leq n^{2/r}(\log n)^3$.*
2. *Almost surely every subgraph of G on at most s vertices has a vertex of degree at most $r-1$. Hence, every such subgraph is r-colorable.*

The following lemma establishes a recursive inequality for the numbers ρ_r, which we will then use to prove Theorem 3.1.

Lemma 3.6. *For every $1 \leq i \leq k/2$, we have*

$$\rho_{2^k} \leq \frac{1}{2}\rho_{2^i} + 2^{i-\lfloor k/2 \rfloor}.$$

Proof. TOPROVE 10 ■

Proof. TOPROVE 11 ■

3.2 The proof of Theorem 1.2

In this subsection we prove Theorem 1.2 in the following more precise form.

Theorem 3.7. *For every $r \geq 4$ and every K_{r-1}-free graph F on $s \geq 2$ vertices, we have $f_{F,K_r}(n) = O(n^{1/2-\frac{1}{8s-10}}(\log n)^3)$.*

Remark 3.8. In fact, our construction is such that every set of size roughly $n^{1/2-\frac{1}{8s-10}}(\log n)^3$ contains an *induced* copy of F.

The proof of Theorem 3.7 uses the method from [19] (which in turn built on [23]), where this result was proved in the special case $F = K_s$, $r = s + 2$. Similarly to those papers, the following graph provides the starting point in our construction.

Proposition 3.9 ([26] or [23])**.** *For every prime q, there is a bipartite graph K with vertex sets X and Y such that the following hold.*

1. $|X| = q^4 - q^3 + q^2$ *and* $|Y| = q^3 + 1$.
2. $d_K(x) = q + 1$ *for every* $x \in X$ *and* $d_K(y) = q^2$ *for every* $y \in Y$.
3. K *is* C_4*-free.*
4. K *does not contain the subdivision of* K_4 *as a subgraph with the part of size* 4 *embedded to* X.

Throughout this subsection, let $r \geq 4$ be a fixed positive integer and let F be a fixed K_{r-1}-free graph on s vertices. Let us identify the vertex set of F with $[s]$. Let q be a prime and let K be the graph provided by Proposition 3.9. We now construct a K_r-free graph H on vertex set X randomly as follows. For each $y \in Y$, partition $N_K(y)$ uniformly randomly as $A_1(y) \cup A_2(y) \cup \cdots \cup A_s(y)$ and place a complete bipartite graph between $A_i(y)$ and $A_j(y)$ whenever i and j are adjacent in F. In other words, we place a blow-up of F in $N_K(y)$ with parts $A_1(y), \ldots, A_s(y)$. The following lemma, proved in [19], combined with properties 3 and 4 of Proposition 3.9, shows that H is K_r-free with probability 1.

Lemma 3.10 ([19, Lemma 2.2])**.** *Assume that the edge set of a K_r is partitioned into cliques $C_1, \ldots, C_k$ of size at most $r - 2$. Then there exist four vertices such that all six edges between them belong to different cliques C_i.*

To see that Lemma 3.10 implies that H is K_r-free, assume that H does contain a copy of K_r on vertex set R. Note that by property 3 of Proposition 3.9, for any edge uv in the complete graph $H[R]$, there is a unique $y \in Y$ such that $u, v \in N_K(y)$. Hence, we can partition the edge set of $H[R]$ into cliques, one with vertex set $N_K(y) \cap R$ for each $y \in Y$ such that $|N_K(y) \cap R| \geq 2$. Moreover, any such clique has size at most $r-2$. (Indeed, F is K_{r-1}-free, so if $|N_K(y) \cap R| \geq r-1$, then $N_K(y) \cap R$ must contain distinct vertices $u \in A_i(y)$ and $v \in A_j(y)$ such that $ij \notin E(F)$ (or $i = j$), meaning that uv is not an edge in H.) Hence, by Lemma 3.10, there are four vertices in R such that for any two of them there is a different common neighbour in Y in the graph K, contradicting property 4 of Proposition 3.9.

Our key lemma, proved in Section 3.2.1, is as follows. Here and below we ignore floor and ceiling signs whenever they are not crucial.

Lemma 3.11. *Let q be a sufficiently large prime and let $t = q^{2-\frac{1}{s-1}}(\log q)^3$. Then with positive probability the number of sets $T \subset X$ of size t for which $H[T]$ is F-free is at most $(q^{\frac{1}{s-1}})^t$.*

It is easy to deduce Theorem 3.7 from this.

Proof. TOPROVE 12 ■

3.2.1 The number of F-free sets

In this subsection we prove Lemma 3.11. While the proof is very similar to that of Lemma 2.3 in [19], there are some small necessary changes, and we include a full proof for completeness. We will use the following lemma from [19].

Lemma 3.12 ([19, Lemma 2.4])**.** *Assume that q is sufficiently large. Then with positive probability, for every $U \subset X$ with $|U| \geq 500s^2q^2$ there exists some $\gamma \geq |U|/q^2$ such that the number of $y \in Y$ with $\gamma/(10s) \leq |A_i(y) \cap U| \leq \gamma$ for all $i \in [s]$ is at least $|U|q/(8(\log q)\gamma)$.*

Definition 3.13. Let us call an instance of H *nice* if it satisfies the conclusion of Lemma 3.12.

Lemma 3.11 can now be deduced from the following.

Lemma 3.14. *Let q be sufficiently large and let $t = q^{2-1/(s-1)}(\log q)^3$. If H is nice, then the number of sets $T \subset X$ of size t for which $H[T]$ is F-free is at most $(q^{1/(s-1)})^t$.*

In what follows, we will consider an s-uniform hypergraph on vertex set X whose hyperedges correspond to the copies of F in H. Then F-free subsets of X will correspond to independent sets in this hypergraph, so to prove Lemma 3.14, it suffices to bound the number of independent sets of certain size. This will be achieved using the hypergraph container method. For an s-uniform hypergraph $\mathcal{G}$ and some $\ell \in [s]$, we write $\Delta_\ell(\mathcal{G})$ for the maximum number of hyperedges in $\mathcal{G}$ containing the same set of ℓ vertices.

We use the following result from [19].

Lemma 3.15 ([19, Corollary 2.8])**.** *For every positive integer $s \geq 2$ and positive reals p and λ, the following holds. Suppose that $\mathcal{G}$ is an s-uniform hypergraph with at least two vertices such that $pv(\mathcal{G})$ and $v(\mathcal{G})/\lambda$ are integers, and for every $\ell \in [s]$,*

$$\Delta_\ell(\mathcal{G}) \leq \lambda \cdot p^{\ell-1}\frac{e(\mathcal{G})}{v(\mathcal{G})}.$$

Then there exists a collection $\mathcal{C}$ of at most $v(\mathcal{G})^{spv(\mathcal{G})}$ sets of size at most $(1-\delta\lambda^{-1})v(\mathcal{G})$ such that for every independent set I in $\mathcal{G}$, there exists some $R \in \mathcal{C}$ with $I \subset R$, where $\delta = 2^{-s(s+1)}$.

Let $\mathcal{H}$ be the s-uniform hypergraph on vertex set X in which s vertices form a hyperedge if they induce a copy of F in H. The next lemma shows that if H is nice, then a suitable subgraph of $\mathcal{H}$ (chosen with the help of Lemma 3.12) satisfies the codegree conditions in Lemma 3.15 with small values of λ and p.

Lemma 3.16. *Assume that H is nice. Then for each $U \subset X$ of size at least $500s^2q^2$ there exists a subgraph $\mathcal{G}$ of $\mathcal{H}[U]$ (on vertex set U) which satisfies*

$$\Delta_\ell(\mathcal{G}) \leq \lambda \cdot p^{\ell-1}\frac{e(\mathcal{G})}{v(\mathcal{G})} \tag{1}$$

for every $\ell \in [s]$ with $\lambda = O_s(\log q)$ and $p \leq |U|^{-1}q^{2-1/(s-1)}$.

Proof. TOPROVE 13 ■

Combining Lemma 3.15 and Lemma 3.16, we prove the following result.

Lemma 3.17. *Let q be sufficiently large and assume that H is nice. Let U be a subset of X of size at least $500s^2q^2$. Now there exists a collection $\mathcal{C}$ of at most $(q^4)^{sq^{2-1/(s-1)}}$ sets of size at most $(1-\Omega_s((\log q)^{-1}))|U|$ such that for any F-free (in H) set $T \subset U$ there exists some $R \in \mathcal{C}$ with $T \subset R$.*

Proof. TOPROVE 14 ■

Corollary 3.18. *Let q be sufficiently large and assume that H is nice. Then there is a collection $\mathcal{C}$ of at most $(q^4)^{O_s(q^{2-1/(s-1)}(\log q)^2)}$ sets of size at most $500s^2q^2$ such that for any F-free (in H) set $T \subset X$ there exists some $R \in \mathcal{C}$ such that $T \subset R$.*

Proof. TOPROVE 15 ■

Corollary 3.18 implies that if q is sufficiently large and H is nice, then the number of F-free sets of size $t = q^{2-1/(s-1)}(\log q)^3$ in H is at most

$$(q^4)^{O_s(q^{2-1/(s-1)}(\log q)^2)}\binom{500s^2q^2}{t} \leq (q^4)^{O_s(q^{2-1/(s-1)}(\log q)^2)}(q^{1/(s-1)}/\log q)^t \leq (q^{1/(s-1)})^t,$$

proving Lemma 3.14.

3.3 F-free induced subgraphs in triangle-free graphs

In this subsection, we observe a connection between Problem 1.9 and the Zarankiewicz problem for 6-cycles.

Let $z(n,m,\{C_4,C_6\})$ denote the maximum number of edges in a bipartite graph with $n+m$ vertices which does not contain C_4 or C_6 as a subgraph. An old result of de Caen and Székely [8] states that $z(n,m,\{C_4,C_6\}) = O(n^{2/3}m^{2/3})$ for $n^{1/2} \leq m \leq n^2$. They observed that there are matching constructions for $m=n$, $m=n^{7/8}$, $m=n^{4/5}$ and $m=n^{1/2}$, but that there is some function $h(n)\to\infty$ such that $z(n,m,\{C_4,C_6\}) = o(n^{2/3}m^{2/3})$ holds for $\omega(n^{1/2}) \leq m \leq n^{1/2}h(n)$. We note that $h(n)$ comes from an application of the Ruzsa–Szemerédi $(6,3)$-theorem [27] and is of order $e^{\Theta(\log^*(n))}$, where $\log^*(n)$ is the iterated logarithm function.

Roughly speaking, we prove that if $z(n,m,\{C_4,C_6\}) = \Theta(n^{2/3}m^{2/3})$ for $m \approx n^{1/2}(\log n)^{3/2}$, then $f_{F,K_3}(n) = \Theta_F(\sqrt{n\log n})$ for every triangle-free graph F. Note that this would also give a new proof of $R(3,t) = \Theta(t^2/\log t)$.

Proposition 3.19. *For every triangle-free graph F, if c_F is sufficiently large, then the following holds. Let $m = c_F n^{1/2}(\log n)^{3/2}$. Assume that there is a $\{C_4,C_6\}$-free biregular bipartite graph with $n+m$ vertices and $\Omega((nm)^{2/3})$ edges. Then $f_{F,K_3}(n) \leq c_F\sqrt{n\log n}$.*

Remark 3.20. The biregularity assumption can be relaxed. Furthermore, any C_6-free graph can be made C_4-free by discarding at most half of its edges [17], so the same conclusion holds assuming the existence of a suitable C_6-free graph.

Proof. TOPROVE 16 ■

4 Concluding remarks

4.1 Remark about improving some lower bounds in [24]

We outline an argument which is implicit in [28] and can be used to improve some of the lower bounds for $f_{F,K_4}(n)$ proved in [24]. The improvement comes from the fact that the proof in [24] uses that a K_4-free graph with average degree d has independence number at least $\sqrt{d}$; this follows by considering a vertex of degree at least d and using the fact that a triangle-free graph with m vertices has independence number at least $\sqrt{m}$. The following proposition gives a better bound in the relevant range of d.

Proposition 4.1 ([28])**.** *Every n-vertex K_4-free graph with average degree $d \geq n^{2/3}$ contains an independent set of size $\Omega(\frac{d}{n^{1/3}})$.*

Note that the bound $\frac{d}{n^{1/3}}$ beats the bound $\sqrt{d}$ whenever $d \gg n^{2/3}$. When trying to prove a lower bound of the form $f_{F,K_4}(n) \geq n^{1/3+\varepsilon}$, one can assume that the average degree d of the host graph G is at most $n^{2/3+2\varepsilon}$ (because otherwise $\alpha(G) \geq \sqrt{d} \geq n^{1/3+\varepsilon}$). This is part of the proof in [24]. By instead using Proposition 4.1, one obtains the stronger $d \leq n^{2/3+\varepsilon}$, which immediately leads to improved bounds (with the rest of the proof in [24] remaining the same). For example, one can improve the constant $\frac{1}{100}$ in the bound $f_{C_k,K_4}(n) \geq n^{\frac{1}{3}+\frac{1}{100k}}$ proved in [24]. Since we think that this may be useful in future works on this topic, we decided to include Proposition 4.1 and its proof. The proof uses the dependent random choice method [15].

Proof. TOPROVE 17 ∎

4.2 Open problems

- We proved that $n^{1/2-O(1/\sqrt{t})} \leq f_{K_{t,t},K_4}(n) \leq n^{1/2-\Omega(1/t)}$, with the lower bound coming from Theorem 1.4 and the upper bound from Theorem 3.7. It might be interesting to determine for this problem the correct order of magnitude of the error term in the exponent.

- Another natural question is to estimate $f_{K_{2,t},K_4}(n)$. By an argument along the lines of the proof of Theorem 1.4, using also Proposition 4.1, one can show that $f_{K_{2,t},K_4}(n) \geq n^{\frac{8}{21}-o_t(1)}$. We believe that it would be interesting to decide whether $f_{K_{2,t},K_4}(n) \leq O(n^{1/2-c})$ for some $c > 0$ which is independent of t.

- The construction of Mattheus and Verstraëte [23] shows that the bound in Proposition 4.1 is tight (up to polylogarithmic terms) for $d = \Theta(n^{2/3})$. Here an interesting problem is to prove a tight bound for the size of the largest independent set one can guarantee in every n-vertex K_4-free graph with average degree $d = \Theta(n^{\alpha})$ for $\frac{2}{3} < \alpha < 1$. In particular, is there any $\alpha > \frac{2}{3}$ for which the bound given by Proposition 4.1 is tight?

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] N. Alon and M. Krivelevich. Constructive bounds for a Ramsey-type problem. *Graphs and Combinatorics*, 13:217–225, 1997.

[3] N. Alon, M. Krivelevich, and B. Sudakov. Turán numbers of bipartite graphs and related Ramsey-type questions. *Combinatorics, Probability and Computing*, 12(5-6):477–494, 2003.

[4] J. Balogh, C. Chen, and H. Luo. On the maximum F-free induced subgraphs in K_t-free graphs. *arXiv preprint arXiv:2406.13780*, 2024.

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

[6] B. Bollobás and H. Hind. Graphs without large triangle free subgraphs. *Discrete mathematics*, 87(2):119–131, 1991.

[7] D. Conlon, S. Mattheus, D. Mubayi, and J. Verstraëte. Ramsey numbers and the Zarankiewicz problem. *Bulletin of the London Mathematical Society*, 2024.

[8] D. De Caen and L. A. Székely. On dense bipartite graphs of girth eight and upper bounds for certain configurations in planar point–line systems. *Journal of Combinatorial Theory, Series A*, 77(2):268–278, 1997.

[9] A. Dudek and D. Mubayi. On generalized Ramsey numbers for 3-uniform hypergraphs. *Journal of Graph Theory*, 76(3):217–223, 2014.

[10] A. Dudek, T. Retter, and V. Rödl. On generalized Ramsey numbers of Erdős and Rogers. *Journal of Combinatorial Theory, Series B*, 109:213–227, 2014.

[11] A. Dudek and V. Rödl. On K_s-free subgraphs in K_{s+k}-free graphs and vertex Folkman numbers. *Combinatorica*, 31:39–53, 2011.

[12] A. Dudek and V. Rödl. On the function of Erdős and Rogers. *Ramsey Theory: Yesterday, Today, and Tomorrow*, pages 63–76, 2011.

[13] P. Erdős and C. A. Rogers. The construction of certain graphs. *Canadian Journal of Mathematics*, 14:702–707, 1962.

[14] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio mathematica*, 2:463–470, 1935.

[15] J. Fox and B. Sudakov. Dependent random choice. *Random Structures and Algorithms*, 38:1–32, 2011.

[16] W. Gowers and O. Janzer. Improved bounds for the Erdős-Rogers function. *Advances in Combinatorics*, 2020.

[17] E. Győri. C_6-free bipartite graphs and product representation of squares. *Discrete Mathematics*, 165:371–375, 1997.

[18] X. Hu and Q. Lin. Ramsey numbers and a general Erdős-Rogers function. *Discrete Mathematics*, 347(12):114203, 2024.

[19] O. Janzer and B. Sudakov. Improved bounds for the Erdős-Rogers $(s, s+2)$-problem. *arXiv preprint arXiv:2307.05441*, 2023.

[20] 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.

[21] M. Krivelevich. K_s-free graphs without large K_r-free subgraphs. *Combinatorics, Probability and Computing*, 3(3):349–354, 1994.

[22] M. Krivelevich. Bounding Ramsey numbers through large deviation inequalities. *Random Structures & Algorithms*, 7(2):145–155, 1995.

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

[24] D. Mubayi and J. Verstraëte. Erdős-Rogers functions for arbitrary pairs of graphs. *arXiv preprint arXiv:2407.03121*, 2024.

[25] D. Mubayi and J. Verstraëte. On the order of Erdős-Rogers functions. *arXiv preprint arXiv:2401.02548*, 2024.

[26] M. E. O'Nan. Automorphisms of unitary block designs. *J. Algebra*, 20(3):495–511, 1972.

[27] I. Z. Ruzsa and E. Szemerédi. Triple systems with no six points carrying three triangles. *Combinatorics (Keszthely, 1976), Coll. Math. Soc. J. Bolyai*, 18(939-945):2, 1978.

[28] B. Sudakov. Large K_r-free subgraphs in K_s-free graphs and some other Ramsey-type problems. *Random Structures & Algorithms*, 26(3):253–265, 2005.

[29] B. Sudakov. A new lower bound for a Ramsey-type problem. *Combinatorica*, 25(4):487–498, 2005.

[30] G. Wolfovitz. K_4-free graphs without large induced triangle-free subgraphs. *Combinatorica*, 33:623–631, 2013.