Large K_r-Free Subgraphs in K_s-Free Graphs and Some Other Ramsey-Type Problems

Benny Sudakov*
Department of Mathematics, Princeton University, Princeton, New Jersey 08544; e-mail: bsudakov@math.princeton.edu

Received 4 November 2003; accepted 8 May 2004
DOI 10.1002/rsa.20035
Published online 30 July 2004 in Wiley InterScience (www.interscience.wiley.com).

ABSTRACT: In this paper we present three Ramsey-type results, which we derive from a simple and yet powerful lemma, proved using probabilistic arguments. Let $3 \leq r < s$ be fixed integers and let G be a graph on n vertices not containing a complete graph K_s on s vertices. More than 40 years ago Erdős and Rogers posed the problem of estimating the maximum size of a subset of G without a copy of the complete graph K_r. Our first result provides a new lower bound for this problem, which improves previous results of various researchers. It also allows us to solve some special cases of a closely related question posed by Erdős. For two graphs G and H, the Ramsey number $R(G, H)$ is the minimum integer N such that any red-blue coloring of the edges of the complete graph K_N, contains either a red copy of G or a blue copy of H. The *book* with n pages is the graph B_n consisting of n triangles sharing one edge. Here we study the book-complete graph Ramsey numbers and show that $R(B_n, K_n) \leq O(n^3/\log^{3/2} n)$, improving the bound of Li and Rousseau. Finally, motivated by a question of Erdős, Hajnal, Simonovits, Sós, and Szemerédi, we obtain for all $0 < \delta < 2/3$ an estimate on the number of edges in a K_4-free graph of order n which has no independent set of size $n^{1-\delta}$. © 2004 Wiley Periodicals, Inc. Random Struct. Alg., 25, 253–265, 2005

1. INTRODUCTION

For two graphs G and H, the Ramsey number $R(G, H)$ is the minimum integer N such that any red-blue coloring of the edges of the complete graph K_N on N vertices, contains

*Research supported in part by NSF Grants DMS-0355497 and DMS-0106589, and by an Alfred P. Sloan fellowship.

either a red copy of G or a blue copy of H. The problem of determining or accurately estimating Ramsey numbers is one of the central problems in modern Combinatorics, and it has received a considerable attention (see, e.g., [13], [6]). The classical Ramsey numbers are those of the complete graphs and are denoted by $R(s, t) = R(K_s, K_t)$.

A more general function than $R(s, t)$ was first considered (for a special case) by Erdős and Gallai in [9]. Suppose $2 \le r < s < n$ are integers and let G be a K_s-free graph. Denote by $f_r(G)$ the maximum cardinality of a subset of vertices of G that contains no copy of K_r, and define

$$f_{r,s}(n) = \min f_r(G),$$

where the minimum is taken over all K_s-free graphs on n vertices. It is easy to see that for $r = 2$, we have that $f_{2,s}(n) < t$ if and only if the Ramsey number $R(s, t)$ satisfies $R(s, t) > n$. Therefore, the problem of determining the function $f_{r,s}(n)$ extends that of determining Ramsey numbers.

Erdős and Rogers [11] considered the case of fixed s, $r = s - 1$ and n tending to infinity. They proved that there exist a constant $\epsilon(s)$ and a K_s-free graph G of order n, such that every induced subgraph of G of order $n^{1-\epsilon(s)}$ contains a copy of K_{s-1}, i.e., $f_{s-1,s}(n) \le n^{1-\epsilon(s)}$. The next step came about 30 years later when Bollobás and Hind [5] improved the result of Erdős and Rogers. They also obtained the following first lower bound, $f_{r,s}(n) \ge n^{1/(s-r+1)}$. Later, their bounds were again improved by Krivelevich, who in [16], [17] showed that

$$c_1 n^{1/(s-r+1)} (\log\log n)^{1-1/(s-r+1)} \le f_{r,s}(n) \le c_2 n^{r/(s+1)} (\log n)^{1/(r-1)},$$

where c_1, c_2 are positive constants depending only on r and s. Most of these results were obtained using the probabilistic method. Recently, two explicit constructions which provide an upper bound on $f_{r,s}(n)$ were obtained by Alon and Krivelevich in [1].

As one can see, the upper bound on $f_{r,s}(n)$ has attracted a lot of attention and been improved considerably during the last forty years. On the other hand, not much progress was made on obtaining good lower bounds. As pointed by Bollobás and Hind [5], no essentially nontrivial lower bound was known. Recently the first such bound was obtained in [22], where it was proved that $f_{r,s}(n) \ge \Omega(n^{b_s})$, with b_s satisfying $1/b_{s+1} = 1 + \frac{1}{r-1}\sum_{i=0}^{r-2} 1/b_{s-i}$ if $s \ge r \ge 3$ and $b_2 = \cdots = b_r = 1$. In [22] it was also proved that this recursion implies that $f_{r,s}(n) \ge \Omega\big(n^{r/(2s)+O(1/s^2)}\big)$. But despite this improvement, there was still a large gap between lower and upper bounds for $f_{r,s}(n)$.

In this paper we will obtain a new lower bound which slightly reduces this gap. Let $r \ge 3$ and let a_s be the sequence which satisfies

$$a_3 = \cdots = a_r = 1, \qquad a_{r+1} = \frac{3r-4}{5r-6}, \qquad \text{and}$$

$$\frac{1}{a_{s+1}} = 1 + \frac{1}{r-1}\sum_{i=0}^{r-2} \frac{1}{a_{s-i}} \qquad \forall s \ge r+1.$$

We show that if $s > r + 1$ then $f_{r,s}(n) \ge \Omega(n^{a_s})$. This improves the above mentioned result from [22], since it is easy to see from the definitions that $a_s > b_s$ for all $s \ge r + 1$. Our bound can be further improved by a polylogarithmic factor using rather standard techniques. We will illustrate this by proving that $f_{3,5} \ge \Omega\left(n^{5/12}\frac{\log^{1/3} n}{(\log\log n)^{1/12}}\right)$. Also, using the new bound, we can make some progress on the following question of Erdős. In [8] he asked to estimate $f_{r,s}(n)$ as accurate as possible and, in particular to show that

$\lim_{n\to\infty} f_{r+1,s}(n)/f_{r,s}(n) = \infty$ for all $r+1 < s$. Our result, together with known upper bounds on $f_{r,s}(n)$, proves the above conjecture of Erdős for $r = 2, 4 \le s \le 8$ and $r = 3$, $s = 6$.

One of the main tools which we use to obtain an improved bound on $f_{r,s}(n)$ is a simple and yet powerful lemma, whose proof is based on probabilistic arguments. We believe that this lemma is of independent interest and might have additional applications. The two Ramsey-type results, which we discuss next, were both obtained using this lemma.

The *book* with n pages is the graph B_n consisting of n triangles sharing one edge. Ramsey problems involving books and their generalizations have been studied extensively by various researchers (see, e.g., [18] and its references). One of the problems, which was considered by Li and Rousseau, is to estimate the book-complete graph Ramsey numbers. In [18] they proved that for all sufficiently large n these numbers satisfy $\frac{n^3}{44\log^2 n} < R(B_n, K_n) < \frac{n^3}{\log(n/e)}$. An interesting question which was left open is to determine the asymptotic behavior of $R(B_n, K_n)$. Here we make some progress in this direction. Using our main lemma, we can show that

$$R(B_n, K_n) \le O\left(\frac{n^3}{\log^{3/2} n}\right),$$

thus reducing the gap between the upper and lower bounds.

Let H be a fixed *forbidden* graph and let f be a function of n. Denote by $\mathbf{RT}(n, H, f(n))$ the maximum number of edges in an H-free graph G on n vertices which has no independent set of size $f(n)$. The problem of estimating this parameter was motivated by the classical Ramsey and Turán theorems and attracted a lot of attention during the last 30 years (see, e.g., the excellent survey [20] of Simonovits and Sós).

One of the first and probably one of the most celebrated results in Ramsey-Turán theory claims that $\mathbf{RT}(n, K_4, o(n)) = (1+o(1))n^2/8$. The upper bound was obtained by Szemerédi and the lower bound was proved by a surprising construction of Bollobás and Erdős. Despite the fact that this construction has independence number $o(n)$, it is not difficult to check that it is still rather large. So replacing $o(n)$ by slightly smaller functions, like $n/\log n$ or even $n^{1-\delta}$ for some fixed $\delta > 0$, perhaps one could get smaller upper bounds on the number of edges. This natural question was posed in [10] and [20]. The first results in this direction were obtained in [21], where it was proved that $\mathbf{RT}(n, K_4, n^{1-o(1)}) = o(n^2)$ together with the estimates for $\mathbf{RT}(n, K_4, n^{1-\delta})$ for a few fixed values of δ. In this paper we will further extend these results using our main lemma. We show that if $0 < \delta \le 2/3$ and r is an integer such that $1/(r+1) < \delta \le 1/r$, then

$$\mathbf{RT}(n, K_4, n^{1-\delta}) \le \begin{cases} n^{2-\delta/(r+1)} & \text{for } \dfrac{r+1}{r^2+r+1} \le \delta, \\ n^{2-((r+2)\delta-1)/(2r+1)} & \text{for } \delta \le \dfrac{r+1}{r^2+r+1}. \end{cases}$$

The rest of this paper is organized as follows. In the next section we prove our key lemma together with some of its immediate applications. First we obtain an improved bound on $f_{3,5}(n)$. We consider this case separately to illustrate our main ideas and also to show how our general bound for $f_{r,s}(n)$ can be improved by a polylogarithmic factor. Next we consider the book-complete graph Ramsey numbers and obtain an improvement of the result of Li and Rousseau. We conclude Section 2 with the proof of upper bounds on the Ramsey-Turán numbers of K_4. In Section 3 we study the general case of the Erdős-Rogers problem and show how to find large K_r-free subsets in K_s-free graphs. The final section of the paper contains some concluding remarks.

The most part of our notation is standard. Given a graph $G = (V, E)$ and a vertex $v \in V$, let $N(v)$ be the set of vertices of G adjacent to v and let $d(v) = |N(v)|$ be the degree of v. For a subset $W \subset V$ we denote by $N^*(W)$ the set of all common neighbors in G of vertices from W. Moreover, for a pair of vertices $u, v \in G$ the size of their common neighborhood is denoted by $codeg(u, v) = |N^*(u, v)|$. Throughout the paper we make no attempts to optimize various absolute constants. To simplify the presentation, we often omit floor and ceiling signs whenever these are not crucial. All logarithms are in the natural base e.

2. THE KEY LEMMA AND ITS APPLICATIONS

We start this section by proving our main lemma, which we think is of independent interest. Its proof is probabilistic and uses a variant of a process which may be called *dependent random choice*. This process and its extensions were recently developed by various researches and successfully applied to different Ramsey and Turán type problems (see, e.g., [21], [15], [3], [14], [12], and their references). After proving the key lemma, in the next three subsections we discuss its application to various Ramsey-type problems.

Lemma 2.1. *Let $i \geq 2$ be a positive integer and let $G = (V, E)$ be a graph of order n with average degree d such that, for every edge $e = (v, w)$ of G, the number of common neighbors of v and w is at most a. Then G contains an induced subgraph with at least $\frac{d^i}{2n^{i-1}}$ vertices and average degree at most $\frac{2a^i}{d^{i-1}}$.*

Proof. Consider a subset $I \subset V$ consisting of i (not necessarily distinct) vertices of G which we chose uniformly at random. Denote

$$U = \{v \in V \mid I \subseteq N(v)\}.$$

Let $X = |U|$ and let Y be the random variable counting the number of edges of G spanned by U. We estimate the expectations of X and Y.

$$\begin{aligned}\mathbb{E}[X] &= \sum_{v \in V} \left(\frac{|N(v)|}{n}\right)^i = \frac{1}{n^i} \sum_{v \in V} (d(v))^i \geq \frac{1}{n^i} n \left(\frac{\sum_{v \in V} d(v)}{n}\right)^i \\ &= \frac{1}{n^{i-1}} \left(\frac{2|E(G)|}{n}\right)^i = \frac{d^i}{n^{i-1}},\end{aligned}$$

where the inequality follows from the convexity of $f(x) = x^i$. In order to estimate the expected value of Y, observe that for a fixed edge $e = (v, w)$ of G the probability that both v and w are in U is precisely $\left(\frac{codeg(v,w)}{n}\right)^i \leq (a/n)^i$. As the number of edges of G is $nd/2$ it follows that

$$\mathbb{E}[Y] = \sum_{e \in E(G),\, e=(u,w)} \Pr[v, w \in U] \leq \frac{nd}{2} \left(\frac{a}{n}\right)^i = \frac{da^i}{2n^{i-1}}.$$

In particular, since Y is nonnegative in the case $a = 0$ we get that Y is identically zero.

Let Z be the random variable which we define as follows. If $a = 0$, then $Z = X - \frac{d^i}{2n^{i-1}}$; otherwise $Z = X - \frac{d^{i-1}}{a^i}Y - \frac{d^i}{2n^{i-1}}$. By linearity of expectation it is easy to see that in both cases $\mathbb{E}[Z] \geq 0$. Hence there exists a choice of I for which $Z \geq 0$. Note that in this case $X - \frac{d^i}{2n^{i-1}} \geq Z \geq 0$ and $Y \leq \frac{a^i}{d^{i-1}}(X - Z) \leq \frac{a^i}{d^{i-1}}X$. Choose such an I and consider the subgraph of G induced by the set U. Then this graph has $X \geq \frac{d^i}{2n^{i-1}}$ vertices and average degree at most $2Y/X \leq \frac{2a^i}{d^{i-1}}$. ∎

Note that the assertion of this lemma holds trivially for $i = 1$. To see this, it is enough to consider the neighborhood of a vertex of maximum degree. Also the result of Lemma 2.1 is interesting only for $a < d^2/n$. Otherwise a stronger statement can be obtained using a simple averaging argument. Finally, we do not know if our result is tight and it would be very interesting to improve it.

2.1. Triangle-Free Subsets in K_5-Free Graphs

Here we show how using our Lemma 2.1 one can obtain an improved lower bound on $f_{3,5}(n)$. We consider this case separately to illustrate some ideas in the proof of our more general result which we present in Section 3 and also to indicate how the bound for $f_{r,s}(n)$ can be improved by a polylogarithmic factor.

Theorem 2.2. *Every K_5-free graph of order n contains a triangle-free subset of vertices of size*

$$\Omega\left(n^{5/12}\frac{\log^{1/3} n}{(\log\log n)^{1/12}}\right).$$

To prove this result we need a lower bound on the size of the maximum independent set in 3-uniform hypergraphs in which every pair of vertices is contained in only few edges. Such hypergraphs are usually called *uncrowded*. The following proposition is a corollary of more general result on the size of an independent set in uncrowded hypergraphs, obtained by Duke, Lefmann, and Rödl [7].

Proposition 2.3. *Let H be a 3-uniform hypergraph on n vertices and with m edges. Let $(m/n)^{1/2} \leq t$ and suppose there exist an $\epsilon > 0$ such that the number of edges containing any fixed pair of vertices of H is at most $t^{1-\epsilon}$. Then H contains an independent set of size*

$$\alpha(H) \geq \Omega\left(\frac{n}{t}\log^{1/2} t\right).$$

We also need the result of Shearer [19] on the size of an independent set in K_s-free graphs.

Proposition 2.4. *Let G be a K_s-free ($s \geq 4$ fixed) graph on n vertices with average degree at most d. Then the independence number of G has size at least*

$$\alpha(G) \geq \Omega\left(\frac{n}{d}\left(\frac{\log d}{\log\log d}\right)\right).$$

Having finished all the necessary preparations, we are now ready to complete the proof of our first result.

Proof of Theorem 2.2. Let G be a K_5-free graph of order n. If it contains an edge (v, w) such that the number of common neighbors of v and w is at least $n^{5/12}\frac{\log^{1/3} n}{(\log\log n)^{1/12}}$, then this set is triangle-free and we are done. Thus we can assume that every edge of G is contained in at most $a = n^{5/12}\frac{\log^{1/3} n}{(\log\log n)^{1/12}}$ triangles. If the average degree d of G is at least $d \geq n^{3/4}(\log\log n)^{1/4}$ then, using Lemma 2.1 for $i = 2$, we conclude that G contains an induced subgraph with at least $\frac{d^2}{2n} \geq \frac{1}{2}n^{1/2}(\log\log n)^{1/2}$ vertices and average degree at most $\frac{2a^2}{d} \leq 2n^{1/12}\log^{2/3} n(\log\log n)^{-5/12}$. Since this subgraph is also K_5-free, using Proposition 2.4, we obtain that it contains an independent set of size

$$\Omega\left(\frac{n^{1/2}(\log\log n)^{1/2}}{n^{1/12}\log^{2/3} n(\log\log n)^{-5/12}}\left(\frac{\log n}{\log\log n}\right)\right) = \Omega\left(n^{5/12}\frac{\log^{1/3} n}{(\log\log n)^{1/12}}\right).$$

Next suppose that the average degree d of G is at most $d \leq n^{3/4}(\log\log n)^{1/4}$ and every edge of G is contained in at most $n^{5/12}\frac{\log^{1/3} n}{(\log\log n)^{1/12}}$ triangles. Let H be a 3-uniform hypergraph whose vertices are the vertices of G and whose edges are the triangles of G. Denote by

$$t = n^{7/12}\log^{1/6} n(\log\log n)^{1/12}.$$

Then an easy calculation shows that the number m of the triangles in graph G is at most

$$\begin{aligned} m &\leq \frac{1}{3}|E(G)| \max_{(u,v)\in E(G)} codeg(u, v) \leq \frac{1}{6}nd \max_{(u,v)\in E(G)} codeg(u, v) \\ &\leq \frac{n}{6}\left(n^{3/4}(\log\log n)^{1/4}\right)\left(n^{5/12}\frac{\log^{1/3} n}{(\log\log n)^{1/12}}\right) = \frac{n}{6}t^2. \end{aligned}$$

As the number of triangles containing any fixed pair of vertices of G is at most $n^{5/12}\frac{\log^{1/3} n}{(\log\log n)^{1/12}} \ll t^{6/7}$, we can apply Proposition 2.3 to conclude that the hypergraph H contains an independent set of size

$$\Omega\left(\frac{n}{t}\log^{1/2} t\right) = \Omega\left(\frac{n}{n^{7/12}\log^{1/6} n(\log\log n)^{1/12}}\log^{1/2} n\right) = \Omega\left(n^{5/12}\frac{\log^{1/3} n}{(\log\log n)^{1/12}}\right).$$

Since an independent set in H corresponds to an induced triangle-free subgraph of G, this completes the proof of the theorem. ∎

2.2. Book-Complete Graph Ramsey Numbers

Recall that the book with n pages is the graph B_n consisting of n triangles sharing one edge. We shall prove the following theorem which improves the upper bound of Li and Rousseau [18] on the book-complete graph Ramsey numbers.

Theorem 2.5. *For all sufficiently large n,* $R(B_n, K_n) \leq \frac{200n^3}{\log^{3/2} n}$.

In the proof of this result, in addition to Lemma 2.1, we will use the following well known bound ([4], Lemma 12.16) on the independence number of a graph containing few triangles (see also [2] for a more general result).

Proposition 2.6. *Let G be a graph on n vertices with average degree at most d and let m be the number of triangles in G. Then G contains an independent set of size at least*

$$\frac{4n}{39d}(\log d - 1/2\log(m/n)).$$

Proof of Theorem 2.5. Let G be graph of order $N = \frac{200n^3}{\log^{3/2} n}$ not containing B_n. Denote by d the average degree of G. By definition, every edge of G is contained in at most $a = n - 1$ triangles. Therefore the total number m of triangles in G is bounded by

$$m < |E(G)| \max_{(u,v)\in E(G)} codeg(u, v) < Ndn.$$

If $d \le \frac{10n^2}{\log^{1/2} n}$ then, by Proposition 2.6, G contains an independent set of size

$$\frac{4N}{39d}(\log d - 1/2\log(m/N)) \ge \frac{2N}{39d}\log(d/n) \ge (40/39 + o(1))n > n,$$

and we are done.

Thus we can assume that the average degree of G is at least $\frac{10n^2}{\log^{1/2} n}$. Then, using Lemma 2.1 for $i = 2$, we obtain that G contains an induced subgraph with at least $n' \ge \frac{d^2}{2N} \ge \frac{1}{4}n\log^{1/2} n$ vertices and average degree at most $d' \le \frac{2a^2}{d} \le \frac{1}{5}\log^{1/2} n$. Now to complete the proof note that, by Turán's theorem, this subgraph contains an independent set of size $n'/(d' + 1) \ge (5/4 + o(1))n > n$. ∎

2.3. Ramsey-Turán Numbers of K_4

Here we present an application of Lemma 2.1 to a problem which arises in Ramsey-Turán theory. For all $0 < \delta \le 2/3$, we obtain the following estimate on the number of edges in a K_4-free graph of order n which contains no independent set of size $n^{1-\delta}$.

Theorem 2.7. *Let* $0 < \delta \le 2/3$ *and let r be an integer such that* $1/(r + 1) < \delta \le 1/r$. *Then*

$$\mathbf{RT}\big(n, K_4, n^{1-\delta}\big) \le \begin{cases} n^{2-\delta/(r+1)} & \textit{for } \dfrac{r+1}{r^2+r+1} \le \delta, \\ n^{2-((r+2)\delta-1)/(2r+1)} & \textit{for } \delta \le \dfrac{r+1}{r^2+r+1}. \end{cases}$$

Proof. First consider the case when $(r + 1)/(r^2 + r + 1) \le \delta \le 1/r$. Let G be a K_4-free graph on n vertices with at least $n^{2-\delta/(r+1)}$ edges. It is enough to show that $\alpha(G) \ge n^{1-\delta}$. Since G is a K_4-free graph, note that the common neighborhood of any pair of adjacent vertices forms an independent set. Therefore, we can assume that any two adjacent vertices have at most $a = n^{1-\delta}$ common neighbors. Using Lemma 2.1 for $i = r + 1$ together with the fact that the average degree of G is $d \ge 2n^{1-\delta/(r+1)}$, we conclude that G contains an

induced subgraph on $n' \geq d^{r+1}/(2n^r) \geq 2^r n^{1-\delta} \geq 2n^{1-\delta}$ vertices with average degree at most
$$d' \leq 2a^{r+1}/d^r \leq 2^{-(r-1)} n^{1-\delta((r^2+r+1)/(r+1))} \leq 1.$$
Therefore, by Turán's theorem it contains an independent set of size at least $n'/(d'+1) \geq n^{1-\delta}$ and we are done.

Now consider the case when $\frac{1}{r+1} < \delta \leq \frac{r+1}{r^2+r+1}$. Suppose G is a K_4-free graph on n vertices with at least $n^{2-((r+2)\delta-1)/(2r+1)}$ edges. Similarly as before we can assume that any two adjacent vertices of G have at most $a = n^{1-\delta}$ common neighbors. The average degree of G is $d \geq 2n^{1-((r+2)\delta-1)/(2r+1)}$. Again using Lemma 2.1 for $i = r+1$ we obtain that G contains an induced subgraph on
$$n' \geq \frac{d^{r+1}}{2n^r} \geq 2^r n^{1-\frac{(r+1)(r+2)\delta-(r+1)}{2r+1}} \geq 2n^{1-\frac{(r+1)(r+2)\delta-(r+1)}{2r+1}}$$
vertices with average degree at most
$$d' \leq \frac{2a^{r+1}}{d^r} \leq 2^{-(r-1)} n^{\frac{r+1-(r^2+r+1)\delta}{2r+1}} \leq n^{\frac{r+1-(r^2+r+1)\delta}{2r+1}}.$$
Therefore, by Turán's theorem, it contains an independent set of size at least $n'/(d'+1) \geq n^{1-\delta}$. This completes the proof. ∎

Remark. Note that, since G is K_4-free graph we can estimate the size of the independent set in the above proof by using Proposition 2.4 instead of Turán's theorem. When $1/(r+1) < \delta < ((r+1)/(r^2+r+1))$ this indeed gives an improvement by polylogarithmic factor. We leave the details of this computations to the interested reader.

3. LARGE K_r-FREE SUBGRAPHS IN K_s-FREE GRAPHS

In this section we obtain a new lower bound on the maximum size of a K_r-free subgraph of a K_s-free graph for all $r \geq 3$ and $s > r+1$. This bound can be easily improved by a polylogarithmic factor using Proposition 2.4 together with results on uncrowded hypergraphs. In the previous section we show how this can be done for $f_{3,5}(n)$, but the computations for the general case are much more involved. Therefore, in order to clarify the presentation, we omit them here and will only concentrate on obtaining the best possible exponent of n in our result.

Theorem 3.1. *Let $r \geq 3$ be a fixed integer and let $a_k(r)$ satisfies the following recurrence relation*
$$a_i = 1, \qquad -(r-3) \leq i \leq 0, \qquad a_1 = \frac{3r-4}{5r-6}, \qquad \textit{and}$$
$$\frac{1}{a_{k+1}} = 1 + \frac{1}{r-1}\sum_{j=0}^{r-2} \frac{1}{a_{k-j}}, \qquad \forall k \geq 1.$$
Then $f_{r,r+k}(n) \geq \Omega(n^{a_k})$ for all $k \geq 2$.

We need the following well-known Turán-type estimate on the size of a maximum independent set in r-uniform hypergraphs. We include its short proof for the sake of completeness.

Lemma 3.2. *Let $r \geq 3$ and let $H = (V, E)$ be an r-uniform hypergraph on n vertices with m edges. If $m < n/r$ edges, then $\alpha(H) > n/2$, and if $m \geq n/r$, then H contains an independent set of size*

$$\alpha(H) > \frac{1}{3}\frac{n}{(m/n)^{1/(r-1)}}.$$

Proof. If $m < n/r$, then, by deleting one vertex from every edge of H, we obtain an independent set of size $n - m > n/2$. Now suppose that $m \geq n/r$. Choose a random subset V_0 of V by taking each $v \in V$ into V_0 independently with probability $p = (n/(rm))^{1/(r-1)}$. Define random variables X, Y by letting X be the number of vertices in V_0 and letting Y be the number of edges spanned by V_0. Then by linearity of expectation there exists a set V_0, for which

$$X - Y \geq np - mp^r = \frac{r-1}{r^{r/(r-1)}}\frac{n}{(m/n)^{1/(r-1)}} > \frac{1}{3}\frac{n}{(m/n)^{1/(r-1)}}.$$

Fix such a set V_0 and for every edge spanned by V_0 delete from V_0 an arbitrary vertex of this edge. This produces an independent set of the size guaranteed by assertion of the lemma. ∎

Lemma 3.3. *Given graph a $G = (V, E)$ and a positive integer $r \geq 3$, let h_i be the maximum number of common neighbors of a subset I of vertices of G, where I is a clique of size $2 \leq i \leq r-1$. Then the number of copies in G of the complete graph K_r is bounded by $|E(G)| \cdot \prod_{i=2}^{r-1} h_i$.*

Proof. For every edge (v_1, v_2) of G the number of copies of K_r which contain (v_1, v_2) can be very roughly estimated as follows. Let $\{v_1, v_2, \ldots, v_r\}$ form a clique. Note that for every $2 \leq i \leq r-1$ the vertex v_{i+1} is a common neighbor of the i-clique $\{v_1, v_2, \ldots, v_i\}$ and therefore can be chosen in at most h_i ways. This implies that altogether there are at most $\prod_{i=2}^{r-1} h_i$ r-cliques which contain e and the total number of r-cliques in G is bounded by $|E(G)| \cdot \prod_{i=2}^{r-1} h_i$. ∎

Lemma 3.4. *Let $r \geq 3$ be an integer and let G be a K_{r+2}-free graph such that the largest K_r-free subset of G has size at most h. Then for every $2 \leq t \leq r$ G contains at most*

$$O\left(h^{t-2+\frac{11r-12}{3r-4}}\right)$$

copies of the complete graph K_t.

Proof. Let h_i be the maximum number of common neighbors of a subset I of vertices of G, where I is a clique of size $2 \leq i \leq r-1$. Clearly by definition $h_i \leq h_2$ for every $i \geq 2$. Since G is a K_{r+2}-free graph, we note that the common neighborhood of any pair of adjacent vertices is K_r-free. Therefore, such pairs have at most h common neighbors and so

$h_2 \leq h$. It is easy to see that it is enough to prove that the number of edges in G is bounded by $O\big(h^{(11r-12)/(3r-4)}\big)$. Indeed, together with Lemma 3.3 this implies that the number of t-cliques in G is at most

$$|E(G)| \cdot \prod_{i=2}^{t-1} h_i \leq |E(G)| \cdot h^{t-2} = O\Big(h^{t-2+\frac{11r-12}{3r-4}}\Big).$$

Suppose, for the sake of contradiction, that G has $e > (3h)^{(11r-12)/(3r-4)}$ edges. If the number of vertices of G is $n \leq e^{(6r-6)/(11r-12)}$ then the average degree of G is at least $d = 2e/n \geq 2e^{(5r-6)/(11r-12)}$. Therefore, using Lemma 2.1 for $i = 2$, we conclude that G contains an induced subgraph on $n' \geq d^2/(2n) \geq 2e^{(4r-6)/(11r-12)}$ vertices with average degree at most $d' \leq 2h^2/d \leq e^{(r-2)/(11r-12)}$. Thus by Turán's theorem it contains an independent set of size at least $n'/(d'+1) > e^{(3r-4)/(11r-12)} > h$, a contradiction.

Now we can assume that $n \geq e^{(6r-6)/(11r-12)}$. Let H be an r-uniform hypergraph whose vertices are the vertices of G and whose edges are all copies of K_r contained in the graph G. Clearly, by definition, an independent set in H corresponds to an induced K_r-free subgraph of G. Denote by m the number of edges in H. Then by Lemma 3.3 we have that $m \leq h^{r-2}|E(G)| \leq e^{1+((3r-4)(r-2))/(11r-12)} = e^{(3r^2+r-4)/(11r-12)}$. Now, using Lemma 3.2, we obtain that H contains an independent set (i.e., K_r-free subset of G) of size

$$\alpha(H) > \frac{n}{3(m/n)^{1/(r-1)}} = \frac{n^{r/(r-1)}}{3m^{1/(r-1)}} \geq \frac{1}{3}e^{\frac{6r^2-6r}{(r-1)(11r-12)}-\frac{3r^2+r-4}{(r-1)(11r-12)}}$$
$$= \frac{1}{3}e^{\frac{3r^2-5r+4}{(r-1)(11r-12)}} = \frac{1}{3}e^{\frac{3r-4}{11r-12}} > h.$$

This contradiction completes the proof of the lemma. ∎

Proof of Theorem 3.1. Let a_k be a sequence which satisfies

$$a_i = 1, \qquad -(r-3) \leq i \leq 0, \qquad a_1 = \frac{3r-4}{5r-6}, \qquad \textit{and}$$
$$\frac{1}{a_{k+1}} = 1 + \frac{1}{r-1}\sum_{j=0}^{r-2}\frac{1}{a_{k-j}}, \qquad \forall k \geq 1, \tag{1}$$

and let G be a graph on n vertices, not containing a clique of size $r+k$. We prove that G contains a K_r-free subset of size $\Omega\big(n^{a_k}\big)$ by induction on k.

If $k = 2$, then G is K_{r+2}-free and $a_2 = \frac{3r-4}{6r-6}$. Denote by h the size of the largest K_r-free subset of G. Let H be a r-uniform hypergraph whose vertices are the vertices of G and whose edges are all copies of K_r contained in graph G. Clearly, by definition, an independent set in H corresponds to an induced K_r-free subgraph of G and has size at most h. By Lemmas 3.4 and 3.2 we have that the number of edges of H is $m \leq O\big(h^{r-2+(11r-12)/(3r-4)}\big)$ and it contains an independent set of size $h \geq \alpha(H) \geq \frac{n}{3(m/n)^{1/(r-1)}}$. This implies that $h \geq \Omega\big(n^{(3r-4)/(6r-6)}\big)$ and completes the proof for $k = 2$.

Next suppose that our statement is true for all $k' \leq k$ and let G be a graph not containing a clique of size $r+(k+1)$. For very $2 \leq i \leq r-1$, let h_i be the maximum number of common neighbors of an i-clique of G, and let d be the maximum degree of G. First, consider the case when $k \geq r$. Suppose that $2 \leq i \leq r-1$ and let I be an i-clique of G. By definition, the subgraph of G induced by the set $N^*(I)$ contains no clique of size $r+(k+1-i)$. Therefore, if $|N(I)| \geq n^{a_{k+1}/a_{k+1-i}}$, then by the induction hypothesis $G[N^*(I)]$ contains a K_r-free set of size

$$\Omega\big(|N^*(I)|^{a_{k+1-i}}\big) = \Omega(n^{a_{k+1}}),$$

and we are done. Hence we can assume that $d \leq n^{a_{k+1}/a_k}$ and $h_i \leq n^{a_{k+1}/a_{k+1-i}}$ for every $2 \leq i \leq r-1$.

Let H be a r-uniform hypergraph whose vertices are the vertices of G and whose edges are all copies of K_r contained in graph G. Then by Lemma 3.3 we can bound the number of edges of H by

$$m \leq |E(G)| \cdot \prod_{i=2}^{r-1} h_i \leq nd \cdot \prod_{i=2}^{r-1} h_i \leq n^{1+\sum_{i=1}^{r-1} \frac{a_{k+1}}{a_{k+1-i}}}.$$

We may also assume that $m \geq n/r$, since otherwise H contains an independent set of size $\Omega(n)$. Thus, we can apply Lemma 3.2 to show that the hypergraph H contains an independent set of size

$$\Omega\left(\frac{n}{(m/n)^{1/(r-1)}}\right) = \Omega\left(n^{1-\frac{a_{k+1}}{r-1}\sum_{i=1}^{r-1}\frac{1}{a_{k+1-i}}}\right) = \Omega\left(n^{1-\frac{a_{k+1}}{r-1}\sum_{j=0}^{r-2}\frac{1}{a_{k-j}}}\right).$$

This completes the proof of the first case, since from the recurrence relation (1) it is easy to see that

$$a_{k+1} = 1 - \frac{a_{k+1}}{r-1}\sum_{j=0}^{r-2}\frac{1}{a_{k-j}}.$$

Next consider the case when $2 \leq k \leq r-1$. Then, similarly as above, by the induction hypothesis we can assume that $|N^*(I)| \leq n^{a_{k+1}/a_{k+1-i}}$ for every clique I in G now only of size $1 \leq i \leq k-2$. This implies that the maximum degree of G is $d \leq n^{a_{k+1}/a_k}$ and the maximum number of common neighbors of every clique of size i is $h_i \leq n^{a_{k+1}/a_{k+1-i}}$ for all $2 \leq i \leq k-2$. Let J be a clique of size $k-1$ in G. By definition, an induced subgraph $G[N^*(J)]$ contains no clique of size $r+2$ and its largest K_r-free subgraph has size $h \leq n^{a_{k+1}}$. Hence, by Lemma 3.4, it contains at most

$$O\left(n^{a_{k+1}\left(\frac{11r-12}{3r-4}+r-k-1\right)}\right) = O\left(n^{a_{k+1}\left(1/a_2+1/a_1+\sum_{i=-r+k+2}^{0}1/a_i\right)}\right)$$

copies of the complete graph K_{r-k+1}. Then, similarly as in proof of Lemma 3.3, we can bound the total number of r-cliques in G by

$$\begin{aligned} m &\leq \left(nd \cdot \prod_{i=2}^{k-2} h_i\right) O\left(n^{a_{k+1}\left(1/a_2+1/a_1+\sum_{i=-r+k+2}^{0}1/a_i\right)}\right) \\ &\leq O\left(n^{1+\frac{a_{k+1}}{a_k}+\sum_{i=2}^{k-2}\frac{a_{k+1}}{a_{k+1-i}}} \cdot n^{\sum_{i=k-1}^{r-1}\frac{a_{k+1}}{a_{k+1-i}}}\right) \\ &= O\left(n^{1+\sum_{i=1}^{r-1}\frac{a_{k+1}}{a_{k+1-i}}}\right). \end{aligned}$$

Now we can finish the induction step in the same way as in the first case. ∎

Using Theorem 3.1 together with some arguments from [22] one can show that $1/a_k$ is asymptotically $\frac{2}{r}k + c(r)$, for some constant c depending only on r. Therefore, when r and k are large, this theorem gives only slightly improvement of the result in [22]. On the other hand, for small r and k this improvement is more substantial. In particular, Theorem 3.1 together with the known upper bound on $f_{r,s}(n)$ immediately implies an affirmative answer to the following cases of the Erdős conjecture mentioned in the introduction.

Corollary 3.5. *If* $4 \leq s \leq 8$, *then* $\lim_{n\to\infty} \frac{f_{3,s}(n)}{f_{2,s}(n)} = \infty$. *We also have* $\lim_{n\to\infty} \frac{f_{4,6}(n)}{f_{3,6}(n)} = \infty$.

Proof. Computing the values of the sequence $a_s(r)$ from Theorem 3.1 we obtain that $f_{3,5} \geq n^{5/12}$, $f_{3,6} \geq n^{10/31}$, $f_{3,7} \geq n^{4/15}$, $f_{3,8} \geq n^{40/177}$ and $f_{4,6} \geq n^{4/9}$. This, together with the estimate $f_{3,4} \geq n^{1/2}$ and the upper bound $f_{r,s}(n) \leq n^{r/(s+1)}(\log n)^{1/(r-1)}$ implies the corollary. ∎

4. CONCLUDING REMARKS

- In this paper we obtained an improved lower bound on $f_{r,s}(n)$ for all $s > r+1$. Nevertheless, it is easy to see that the gap between our result and the known upper bounds is still relatively large and it would be very interesting to close it. In case when $r = s-1$ the exponent in the current lower bound is $1/2$ and our methods are not sufficient to improve it. On the other hand, the best upper bound for this case has the form $n^{1-\epsilon(s)}$ where $\epsilon(s)$ tends to zero when s tends to infinity. It is an intriguing open question to decide if for every $0 < \delta < 1$, the value of $f_{s-1,s}(n)$ is greater than $n^{1-\delta}$ for sufficiently large s. Even to obtain a lower bound for $f_{s-1,s}(n)$ with exponent strictly larger than $1/2$ would be extremely interesting.
- Another question which remains open is to understand the asymptotics of the book-complete graph Ramsey numbers $R(B_n, K_n)$. We think that the new upper bound for these numbers which we obtain here gives some indication that the current lower bound for $R(B_n, K_n)$ determines its correct behavior. Therefore, we conjecture that $R(B_n, K_n) = \Theta(n^3/\log^2 n)$.
- The upper bound on the $\mathbf{RT}(n, K_4, n^{1-\delta})$ which we obtain in Theorem 2.7 is probably far from being tight and one should try to improve it. Also it would be interesting to construct dense K_4-free graphs which have small independence number. This will give lower bounds on the Ramsey-Turán numbers of K_4. One construction, which unfortunately gives a weak result, can be obtained as follows. Let G be a K_4-free graph of order t which has $\Omega(t^{8/5})$ edges and contains no independent set of size $O(t^{2/5}\log t)$. Such graph can be obtained by taking an appropriate subgraph of the random graph $G(t,p), p = ct^{-2/5}$ for some small constant c (see [17] for details). Replace each vertex of G by an independent set of size k and connect two vertices in the new graph if and only if the corresponding vertices of G are connected by an edge. It is easy to see that this gives a K_4-free graph of order kt with $\Omega(k^2 t^{8/5})$ edges whose independence number is $O(kt^{2/5}\log t)$. Taking $n = t^{3/(5\delta)}(\log t)^{-1/\delta}$ and $k = n/t$, we obtain a construction of a K_4-free graph of order n with $n^{2-2\delta/3-o(1)}$ edges and with no independent set of size $O(n^{1-\delta})$.

ACKNOWLEDGMENT

We thank Peter Keevash and anonymous referees for helpful comments.

REFERENCES

[1] N. Alon and M. Krivelevich, Constructive bounds for a Ramsey-type problem, Graphs Combin 13 (1997), 217–225.

[2] N. Alon, M. Krivelevich, and B. Sudakov, Coloring graphs with sparse neighborhoods, J Combin Theory Ser B 77 (1999), 73–82.

[3] N. Alon, M. Krivelevich, and B. Sudakov, Turán numbers of bipartite graphs and related Ramsey-type questions, Combin Probab Comput 12 (2003), 477–494.

[4] B. Bollobás, Random graphs, 2nd edition, Cambridge Studies in Advanced Mathematics 73, Cambridge University Press, Cambridge, 2001.

[5] B. Bollobás and H. R. Hind, Graphs without large triangle free subgraphs, Discrete Math 87(2) (1991), 119–131.

[6] F. Chung and R. Graham, Erdős on graphs. His legacy of unsolved problems, A K Peters, Wellesley, MA, 1998.

[7] R. Duke, H. Lefmann, and V. Rödl, On uncrowded hypergraphs, Random Structures Algorithms 6 (1995), 209–212.

[8] P. Erdős, "Some of my recent problems in combinatorial number theory, geometry and combinatorics," Graph theory, combinatorics, and algorithms, Vol. 1, 2 (Kalamazoo, MI, 1992), 335–349. Wiley, New York, 1995.

[9] P. Erdős and T. Gallai, On the minimal number of vertices representing the edges of a graph, Publ Math Inst Hungar Acad Sci 6 (1961), 181–203.

[10] P. Erdős, A. Hajnal, M. Simonovits, V. T. Sós, and E. Szemerédi, Turán-Ramsey theorems and simple asymptotically extremal structures, Combinatorica 13 (1993), 31–56.

[11] P. Erdős and C. A. Rogers, The construction of certain graphs, Canad J Math 14 (1962), 702–707.

[12] W. T. Gowers, A new proof of Szemerédi's theorem for arithmetic progressions of length four, Geom Funct Anal 8 (1998), 529–551.

[13] R. Graham, B. Rothschild and J. Spencer, Ramsey theory, 2nd edition Wiley, New York, 1990.

[14] A. Kostochka and V. Rödl, On graphs with small Ramsey numbers, J Graph Theory 37 (2001), 198–204.

[15] A. Kostochka and B. Sudakov, On Ramsey numbers of sparse graphs, Combinatorics, Probability and Computing 12 (2003), 627–641.

[16] M. Krivelevich, K^s-free graphs without large K^r-free subgraphs, Combin Probab Comput 3 (1994), 349–354.

[17] M. Krivelevich, Bounding Ramsey numbers through large deviation inequalities, Random Structures Algorithms 7 (1995), 145–155.

[18] Y. Li and C. Rousseau, On book-complete graph Ramsey numbers, J Combin Theory Ser B 68 (1996), 36–44.

[19] J. Shearer, On the independence number of sparse graphs, Random Structures Algorithms 7 (1995), 269–271.

[20] M. Simonovits and V. Sós, Ramsey-Turán theory, Discrete Math 229 (2001), 293–340.

[21] B. Sudakov, Few remarks on the Ramsey-Turán-type problems, J Combin Theory Ser B 88 (2003), 99–106.

[22] B. Sudakov, A new lower bound for a Ramsey-type problem, Combinatorica, to appear.