Multicolor Ramsey numbers via pseudorandom graphs

Xiaoyu He*
Department of Mathematics
Stanford University
Stanford, U.S.A.
alkjash@stanford.edu

Yuval Wigderson†
Department of Mathematics
Stanford University
Stanford, U.S.A.
yuvalwig@stanford.edu

Submitted: Oct 18, 2019; Accepted: Jan 23, 2020; Published: Feb 7, 2020

Abstract

A weakly optimal K_s-free (n, d, λ)-graph is a d-regular K_s-free graph on n vertices with $d = \Theta(n^{1-\alpha})$ and spectral expansion $\lambda = \Theta(n^{1-(s-1)\alpha})$, for some fixed $\alpha > 0$. Such a graph is called optimal if additionally $\alpha = \frac{1}{2s-3}$. We prove that if $s_1, \ldots, s_k \geqslant 3$ are fixed positive integers and weakly optimal K_{s_i}-free pseudorandom graphs exist for each $1 \leqslant i \leqslant k$, then the multicolor Ramsey numbers satisfy

$$\Omega\Big(\frac{t^{S+1}}{\log^{2S} t}\Big) \leqslant r(s_1, \ldots, s_k, t) \leqslant O\Big(\frac{t^{S+1}}{\log^{S} t}\Big),$$

as $t \to \infty$, where $S = \sum_{i=1}^{k}(s_i - 2)$. This generalizes previous results of Mubayi and Verstraëte, who proved the case $k = 1$, and Alon and Rödl, who proved the case $s_1 = \cdots = s_k = 3$. Both previous results used the existence of optimal rather than weakly optimal K_{s_i}-free graphs.

Mathematics Subject Classifications: 05C55, 05D10

1 Introduction

The central object of study in Ramsey theory is the Ramsey number $r(s_1, \ldots, s_k)$, which is defined to be the smallest posititive integer N such that in any k-coloring of the complete graph K_N, there is a monochromatic K_{s_i} of some color $i \in \{1, \ldots, k\}$.

In the case $k = 2$, the order of growth of $r(3, t)$ as $t \to \infty$ was determined to be

$$r(3, t) = \Theta\Big(\frac{t^2}{\log t}\Big)$$

*Supported by NSF Graduate Research Fellowship DGE-1656518.
†Supported by NSF Graduate Research Fellowship DGE-1656518.

by Ajtai, Komlós, and Szemerédi [1] and Kim [8]. It is one of the central open problems in Ramsey theory to generalize these bounds and determine the growth rates of $r(s,t)$ for all fixed $s \geqslant 3$ and $t \to \infty$. Unfortunately, when $s \geqslant 4$ even the polynomial order of $r(s,t)$ is not known, and the best known bounds are

$$\Omega\Big(\frac{t^{\frac{s+1}{2}}}{(\log t)^{\frac{s+1}{2}-\frac{1}{s-2}}}\Big) \leqslant r(s,t) \leqslant O\Big(\frac{t^{s-1}}{\log^{s-2} t}\Big).$$

The lower bound is due to Bohman and Keevash [7], while the upper bound is again due to Ajtai, Komlós, and Szemerédi [1].

Recently, Mubayi and Verstraëte [10] connected the growth rate of $r(s,t)$ to a problem in the theory of pseudorandom graphs. Recall that an (n,d,λ)-*graph* is a d-regular graph on n vertices such that all of its nontrivial eigenvalues have absolute value at most λ.

Definition 1. A *family of weakly optimal K_s-free (n,d,λ)-graphs* is a collection of K_s-free (n_i,d_i,λ_i)-graphs for which $d_i = \Theta(n_i^{1-\alpha})$ and $\lambda_i = \Theta(n_i^{1-(s-1)\alpha})$ as $n_i \to \infty$, for some fixed $\alpha > 0$. We call α the *parameter of weak optimality.* If, moreover, $\lambda_i = \Theta(\sqrt{d_i})$ (so that $\alpha = \frac{1}{2s-3}$), then this family is said to be *optimal.*

Note that α and the implicit constants may not depend on i. Informally, we say that weakly optimal K_s-free (n,d,λ)-graphs exist if there exists a family of weakly optimal K_s-free (n,d,λ)-graphs, for some fixed $\alpha > 0$. Note that the t-blowup of an (n,d,λ)-graph is an $(nt,dt,\lambda t)$-graph with the same clique number; thus, the existence of optimal K_s-free (n,d,λ)-graphs implies the existence of weakly optimal K_s-free (n,d,λ)-graphs for all $0 < \alpha \leqslant \frac{1}{2s-3}$ (this fact was observed already by Krivelevich, Sudakov, and Szabó [9] when $s=3$). Because of this, the existence of weakly optimal K_s-free (n,d,λ)-graphs is indeed weaker than the existence of optimal ones.

Sudakov, Szabó, and Vu [11] conjectured the existence of optimal K_s-free (n,d,λ)-graphs for all $s \geqslant 3$ and all n; such graphs where constructed by Alon [2] in the case $s=3$ but the conjecture remains open for $s \geqslant 4$ (see [6] for the best known construction for $s \geqslant 5$, which agrees with Alon's bound for $s=4$). Conditional on this conjecture, Mubayi and Verstraëte showed that $r(s,t)$ grows like t^{s-1} up to polylogarithmic factors.

Theorem 2. *(Mubayi and Verstraëte [10].) If optimal K_s-free (n,d,λ)-graphs exist for all n, then*

$$\Omega\Big(\frac{t^{s-1}}{\log^{2s-4} t}\Big) \leqslant r(s,t) \leqslant O\Big(\frac{t^{s-1}}{\log^{s-2} t}\Big),$$

where the implicit constants may depend only on s.

Theorem 2 relies heavily on a lemma of Alon and Rödl [4], which was originally used to prove the following bound on the multicolor Ramsey number $r_k(s,t) \coloneqq r(s,\dots,s,t)$ where s appears k times.

Theorem 3. *(Alon and Rödl [4].) For all $k \geqslant 1$,*

$$\Omega\Big(\frac{t^{k+1}}{\log^{2k} t}\Big) \leqslant r_k(3,t) \leqslant O\Big(\frac{t^{k+1}}{\log^{k} t}\Big),$$

where the implicit constants may depend only on k.

Note that Theorem 3 depends on the existence of optimal K_3-free (n, d, λ)-graphs, which were constructed by Alon [2].

Our main result is the following natural common generalization of Theorems 2 and 3, which also replaces the assumption of optimality by that of weak optimality.

Theorem 4. *If $s_1, \ldots, s_k \geqslant 3$, $S = \sum_{i=1}^{k}(s_i - 2)$, and for each $1 \leqslant i \leqslant k$ there exist weakly optimal K_{s_i}-free (n, d, λ)-graphs for all n, then*

$$\Omega\Big(\frac{t^{S+1}}{\log^{2S} t}\Big) \leqslant r(s_1, \ldots, s_k, t) \leqslant O\Big(\frac{t^{S+1}}{\log^{S} t}\Big), \tag{1}$$

where the implicit constants may depend only on S and the weak optimality parameters $\alpha_1, \ldots, \alpha_k$.

Like Theorems 2 and 3, Theorem 4 is a consequence of a lemma of Alon and Rödl [4] which shows that an (n, d, λ)-graph has few independent sets of order just over n/d. We will need the following slightly stronger version, which is proved in exactly the same way.

Lemma 5. *If G is an (n, d, λ)-graph and $t \geqslant \frac{2n \log^2 n}{d}$, then the number of t-tuples $(v_1, \ldots, v_t) \in V(G)^t$ of vertices of G, no pair of which are adjacent, is at most*

$$\Big(\frac{4en\lambda}{d}\Big)^t.$$

In the next section we prove the lower bound in Theorem 4. The proofs of Lemma 5 and the upper bound in Theorem 4 are relatively standard and are confined to the appendix.

2 The Proof

The main difficulty in applying Lemma 5 to construct Ramsey graphs is rescaling a given (n, d, λ)-graph to have the appropriate number of vertices. The proofs of Theorems 2 and 3 each provide half the picture. In the proof of Theorem 2, a K_s-free (n, d, λ)-graph is scaled down to a smaller K_s-free graph with no independent sets of size t by sampling a random induced subgraph. In the proof of Theorem 3, a K_3-free (n, d, λ)-graph is scaled up to a larger K_3-free graph with few independent sets by performing a balanced blowup.

The natural common generalization of these two constructions is a random blowup; using random blowups, we will be able to scale the weakly optimal K_s-free (n, d, λ)-graphs to K_s-free graphs of any size with few independent sets. Define $i_t(G)$ to be the number of independent sets of order t in G.

Lemma 6. *If there exists a K_s-free (n, d, λ)-graph G and $t \geqslant \frac{2n \log^2 n}{d}$, then for every N there exists a K_s-free graph $G(N)$ on N vertices with*

$$i_t(G(N)) \leqslant \Big(\frac{2e^2\lambda N}{n \log^2 n}\Big)^t.$$

Proof. We will define $G(N)$ as follows. Pick a uniform random map $f : [N] \to G$, and let $G(N)$ be the graph on $[N]$ whose edges are exactly the pairs (i, j) that map to edges in G. Since G is K_s-free, so is $G(N)$. It suffices to prove the desired upper bound on $\mathbb{E}[i_t(G(N))]$.

By Lemma 5 (proved in Appendix A) and linearity of expectation,

$$\begin{aligned}\mathbb{E}[i_t(G(N))] &= \binom{N}{t} \Pr[f([t]) \text{ is an independent set}] \\ &= \binom{N}{t} \frac{\left(\frac{4e\lambda n}{d}\right)^t}{n^t},\end{aligned}$$

since $f([t])$ is a uniform random t-tuple in $V(G)^t$. Bounding $\binom{N}{t} \leqslant \left(\frac{eN}{t}\right)^t$, we find that with positive probability,

$$i_t(G(N)) \leqslant \Big(\frac{eN}{t}\Big)^t \Big(\frac{4e\lambda}{d}\Big)^t \leqslant \Big(\frac{2e^2\lambda N}{n \log^2 n}\Big)^t$$

since $t \geqslant \frac{2n \log^2 n}{d}$. □

We are ready to prove the main result. The upper bound is proved in Appendix B.

Proof of the lower bound in Theorem 4. Henceforth all implicit constants are allowed to depend on $S = \sum_{i=1}^{k}(s_i - 2)$ and on the weak optimality parameters $\alpha_1, \ldots, \alpha_k$. Let G_i be a weakly optimal K_{s_i}-free (n_i, d_i, λ_i)-graph, where $d_i = \Theta(n_i^{1-\alpha_i})$ and $\lambda_i = \Theta(n_i^{1-(s_i-1)\alpha_i})$. As these are assumed to exist for all n_i, we pick

$$n_i = \Theta\Big(\Big(\frac{t}{\log^2 t}\Big)^{1/\alpha_i}\Big)$$

so that with $d_i = \Theta(n_i^{1-\alpha_i})$, the bound $t \geqslant \frac{2n_i \log^2 n_i}{d_i}$ holds. Take

$$N = \Theta\Big(\frac{t^{S+1}}{\log^{2S} t}\Big),$$

the implicit constant to be chosen later. Rescaling each G_i to a $G_i(N)$ on N vertices satisfying Lemma 6, we get k graphs $G_i(N)$ on the same vertex set $[N]$ such that $G_i(N)$ is K_{s_i}-free and

$$i_t(G_i(N)) \leqslant \Big(\frac{2e^2\lambda_i N}{n_i \log^2 n_i}\Big)^t. \tag{2}$$

We define a random $(k+1)$-coloring of $\binom{[N]}{2}$ so that in each of the first k colors, the edges form a subgraph of $G_i(N)$. To do so, simply take a uniform random vertex permutation of $G_i(N)$ as the edges in the i-th color; when multiple colors are given to the same edge, break ties arbitrarily. All remaining edges are given color $k + 1$.

This $(k + 1)$-colored graph has no monochromatic K_{s_i} in any of the first k colors. It remains to show that with positive probability, it has no K_t in the last color. Indeed, the probability that a given set I of order t induces a K_t in the last color is exactly the product

$$\prod_{i=1}^{k} \frac{i_t(G_i(N))}{\binom{N}{t}},$$

since I must be an independent set in each of the first k colors. By (2), we have that

$$\begin{aligned}\prod_{i=1}^{k} \frac{i_t(G_i(N))}{\binom{N}{t}} &\leqslant \prod_{i=1}^{k} \Big(\frac{2e^2\lambda_i N}{n_i \log^2 n_i}\Big)^t \Big/ \Big(\frac{N}{t}\Big)^t \\ &\leqslant \prod_{i=1}^{k} (C\lambda_i/d_i)^t\end{aligned}$$

for an absolute constant $C > 0$. With our choices of λ_i and d_i,

$$\frac{\lambda_i}{d_i} = \Theta\left(n_i^{-\alpha_i(s_i-2)}\right) = \Theta\left(\left(\frac{t}{\log^2 t}\right)^{-(s_i-2)}\right).$$

By taking a union bound over all I, the probability that there exists a K_t in the last color is at most

$$\binom{N}{t} \prod_{i=1}^{k} O\Big(\Big(\frac{t}{\log^2 t}\Big)^{-(s_i-2)}\Big)^t \leqslant O\Big(\frac{N}{t}\Big(\frac{t}{\log^2 t}\Big)^{-S}\Big)^t < 1$$

for the appropriate choice of the constant in the definition of N. This completes the proof. □

Acknowledgements. The authors would like to thank Ryan Alweiss and Jacob Fox for helpful discussions on this problem. We are also grateful to Anurag Bishnoi for bringing reference [9] to our attention.

References

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

[2] N. Alon, Explicit Ramsey graphs and orthonormal labelings, *Electron. J. Combin.* **1** (1994), #R12.

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

[4] N. Alon and V. Rödl, Sharp bounds for some multicolor Ramsey numbers, *Combinatorica* **25** (2005), 125–141.

[5] N. Alon and J. Spencer, *The Probabilistic Method*, third edition, Wiley, New York, 2008.

[6] A. Bishnoi, F. Ihringer, and V. Pepe, A construction for clique-free pseudorandom graphs, preprint, arXiv:1905.04677.

[7] T. Bohman and P. Keevash, The early evolution of the H-free process, *Invent. Math.* **181** (2010), 291–336.

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

[9] M. Krivelevich, B. Sudakov, and T. Szabó, Triangle factors in sparse pseudo-random graphs, *Combinatorica* **24** (2004), 403–426.

[10] D. Mubayi and J. Verstraëte, A note on pseudorandom Ramsey graphs, preprint, arXiv:1909.01461.

[11] B. Sudakov, T. Szabó, and V. Vu, A generalization of Turán's theorem, *J. Graph Theory* **49** (2005), 187–195.

A Proof of Lemma 5

We give a short proof of Lemma 5 using the Expander Mixing Lemma (see e.g. [5, Corollary 9.2.5]).

Lemma 7. *(Expander Mixing Lemma.) If G is an (n, d, λ)-graph and $S, T \subseteq V(G)$, then*

$$|e(S,T) - \frac{d}{n}|S||T|| < \lambda\sqrt{|S||T|}.$$

Here $e(S, T)$ denotes the number of ordered pairs $(s, t) \in S \times T$ which are edges of G.

Proof of Lemma 5. We count the number of ways to pick $v_1, \dots, v_t$ one-by-one. Let S_k be the set of all vertices with no edges to $v_1, \dots, v_{k-1}$ (including $v_1, \dots, v_{k-1}$), and let $T_k = \{v \in S_k : |N(v) \cap S_k| < \frac{d}{2n}|S_k|\}$. Thus, S_k is the set of all valid candidates for v_k, and T_k is the subset of valid candidates for which S_{k+1} is not much smaller than S_k. In particular, every time we choose $v_k \in S_k \backslash T_k$, we find that

$$|S_{k+1}| \leqslant (1 - \frac{d}{2n})|S_k| < e^{-\frac{d}{2n}}|S_k|,$$

so since $|S_0| = n$, the total number of k for which v_k can be chosen from $S_k \backslash T_k$ is bounded by $t' = \frac{2n}{d}\log n$.

On the other hand, by the definition of T_k we have $e(S_k, T_k) < \frac{d}{2n}|S_k||T_k|$, and so applying Lemma 7 we get

$$\frac{d}{2n}|S_k||T_k| < \lambda\sqrt{|S_k||T_k|}.$$

In particular, since $T_k \subseteq S_k$, we have

$$|T_k| < \frac{2n\lambda}{d}.$$

Thus, the total number of sequences $v_1, \dots, v_t$ where all pairs are not adjacent is bounded by

$$\binom{t}{t'} n^{t'} \left(\frac{2n\lambda}{d}\right)^t,$$

since we can choose the t' steps on which $v_k \in S_k \backslash T_k$ in $\binom{t}{t'}$ ways, the number of such choices is bounded by n on each step, and in all the other steps the number of choices for v_k is at most $|T_k| < \frac{2n\lambda}{d}$. Bounding $\binom{t}{t'} < 2^t$ and $n^{t'} < n^{t/\log n} = e^t$, we obtain a bound of

$$\left(\frac{4en\lambda}{d}\right)^t,$$

as claimed. □

B The upper bound in Theorem 4

Alon and Rödl [4] proved the upper bound in (1) when $s_1 = s_2 = \cdots = s_k = 3$, and our proof is a generalization of theirs.

Proof of the upper bound in Theorem 4. We fix k and induct on S. The base case $S = 1$ is just $r(2, 2, \dots, 2, 3, t) = O(t^2/\log t)$ for any number of 2's, by Ajtai, Komlós and Szemerédi [1]. Assume by induction that there exist absolute constants $C_{S'} > 0$ for all $S' < S$ such that for all vectors $(s_1, \dots, s_k)$ with $s_i \geqslant 2$ and $\sum_{i=1}^{k}(s_i - 2) = S'$,

$$r(s_1, \dots, s_k, t) \leqslant n_{S'} := \frac{C_{S'} t^{S'+1}}{\log^{S'} t}.$$

Now let $n_S = C_S t^{S+1}/\log^S t$ for some C_S to be determined, and suppose we are given a $(k+1)$-coloring of K_{n_S} such that there is no monochromatic K_{s_i} of color i, nor a monochromatic K_t of color $k+1$. Define T to be the spanning subgraph of K_{n_S} obtained by taking only the edges of the first k colors. If D is the maximum degree in T, then

$$D < kn_{S-1}, \tag{3}$$

If (3) is false, then there is a vertex $v \in V(T)$ and some color $i \leqslant k$ such that v is incident to at least

$$n_{S-1} \geqslant r(s_1, \dots, s_i - 1, \dots, s_k, t)$$

edges of color i. The induced subgraph on the set of vertices connected to v by color i must not contain a monochromatic clique K_{s_j} of any color $j \neq i$, so there will be a $K_{s_i - 1}$ of color i inside. But then this forms a K_{s_i} of color i together with v, which is a contradiction. This proves inequality (3).

Next, let D' denote the maximum number of edges in some neighborhood $N_T(v)$ of a vertex in T. We show

$$D' < k^2 D n_{S-2}. \tag{4}$$

Suppose otherwise, and let v be the vertex with the most edges in its neighborhood. If $u \in N_T(v)$, define $d_v(u)$ as the number of common neighbors $w \in N_T(v) \cap N_T(u)$ for which either uv, uw, vw are all the same color, or uw and vw are different colors. Each edge $uw \in N_T(v)$ contributes either once or twice to the sum of the $d_v(u)$, so

$$\sum_{u \in N_T(v)} d_v(u) \geqslant k^2 D n_{S-2}.$$

In particular, there is some u for which $d_v(u) \geqslant k^2 n_{S-2}$. We can categorize the vertices w of $N_T(v)$ counted in $d_v(u)$ by the pair of colors of uw and vw, and find that there exists colors i, j (not necessarily different) and a set W of n_{S-2} vertices such that for every $w \in W$, uw is of color i and vw is of color j. If $i \neq j$, this implies a contradiction from the fact that

$$|W| \geqslant n_{S-2} \geqslant r(s_1, \ldots, s_i - 1, \ldots, s_j - 1, \ldots, s_k, t).$$

Otherwise, if $i = j$, then by the definition of $d_v(u)$ it must be that uv is of color i as well, and so we also get a contradiction since

$$|W| \geqslant n_{S-2} \geqslant r(s_1, \ldots, s_i - 2, \ldots, s_k, t).$$

This proves (4). It is a corollary of a result of Alon, Krivelevich, and Sudakov [3] that if a graph has maximum degree D and every neighborhood has at most $D' = \frac{D^2}{f}$ edges, then its independence number is at least $\Omega(\frac{n \log f}{D})$. In particular, we see that the independence number of T is at least

$$\Omega\Big(\frac{n_S \log t}{D}\Big),$$

since (4) implies $D' = O(D^2 \log t / t)$. On the other hand, an independent set in T forms a monochromatic clique in K_{n_S} of color $k + 1$, so

$$t > \Omega\Big(\frac{n_S \log t}{D}\Big),$$

which shows that

$$n_S < O\Big(\frac{Dt}{\log t}\Big) = O\Big(\frac{C_{S-1} t^{S+1}}{\log^S t}\Big).$$

Picking C_S sufficiently large in terms of C_{S-1}, this gives the desired contradiction. $\square$