1 Introduction

Ramsey's theorem, proved by Ramsey in 1930, says that for every positive integer t, there exists some positive integer N so that any two-coloring of $E(K_N)$ contains a monochromatic copy of K_t. Since every graph is contained in a complete graph, an equivalent but superficially stronger statement is the following: for every graph H, there exists a graph G so that every two-coloring of $E(G)$ contains a monochromatic copy of H. In case G satisfies this property, we say that G is *Ramsey for H*.

Broadly speaking, the field of graph Ramsey theory is concerned with the question "which graphs G are Ramsey for a given graph H?". In general, this question is extremely difficult, so one studies various natural subquestions. For example, it is natural to ask for the *smallest* graph G which is Ramsey for H. The most well-studied version of this question is to determine the *Ramsey number* $r(H)$ of H, which is defined as the least number of vertices in a graph G which is Ramsey for H. Equivalently, $r(H)$ is the least N so that K_N is Ramsey for H.

However, rather than asking for few vertices, one can ask for few *edges*. Specifically, one can define the *size Ramsey number* $\hat{r}(H)$ as

$$\hat{r}(H) = \min\{e(G) : G \text{ is Ramsey for } H\}.$$

This quantity was defined in 1978 by Erdős, Faudree, Rousseau, and Schelp. One simple observation they made is that if H has no isolated vertices, then

$$\frac{r(H)}{2} \leq \hat{r}(H) \leq \binom{r(H)}{2}. \tag{1}$$

Indeed, for the upper bound, one notes that if $N = r(H)$, then K_N is Ramsey for H and has $\binom{N}{2}$ edges. For the lower bound, suppose that G is Ramsey for H and has $\hat{r}(H)$ edges. Since H has no isolated vertices, we may assume that G also does not, and thus G has at most $2e(G) = 2\hat{r}(H)$ vertices, showing that $r(H) \leq 2\hat{r}(H)$.

Thanks to (1), we know that $\hat{r}(H)$ is somewhere between a linear and a quadratic function of $r(H)$. And interestingly, it turns out that there are situations where one of the two bounds is close to optimal, and other situations where the other is close to optimal. A nice pair of illustrative examples is given by stars and double stars. Let us define the *starburst graph* $S_n^{(k)}$ to consist of a clique K_k with n pendant vertices coming off of each of the vertices of the clique. Thus, $S_n^{(1)}$ is simply the star $K_{1,n}$, and $S_n^{(2)}$ is the double star, consisting of two stars joined along a common edge. It is well-known and simple to prove that any tree has Ramsey number which is linear in its number of vertices, so $r(S_n^{(i)}) = O(n)$ for $i = 1, 2$. This implies, by (1), that $\Omega(n) \leq \hat{r}(S_n^{(i)}) \leq O(n^2)$ for $i = 1, 2$. And surprisingly, the answers in the two cases are very different!

Proposition 1.1. $\hat{r}(S_n^{(1)}) = 2n - 1$.

Proof. Consider $S_{2n-1}^{(1)}$, which has $2n - 1$ edges. In any two-coloring of $E(S_{2n-1}^{(1)})$, at least n edges receive the same color, and they form a monochromatic copy of $S_n^{(1)}$.

For the lower bound, if G is a graph with at most $2n-2$ edges, then we may 2-color its edges so that there are fewer than n edges in either color. As $S_n^{(1)}$ has n edges, there can be no monochromatic copy of $S_n^{(1)}$ in this coloring. $\square$

Proposition 1.2. $\hat{r}(S_n^{(2)}) > n^2$.

Proof. Let G be a graph with at most n^2 edges. This implies that G has at most $2n$ vertices of degree at least n; call the set of these high-degree vertices U. Then color all edges within U red, and all other edges blue. There can be no red copy of $S_n^{(2)}$, since $S_n^{(2)}$ has $2n+2$ vertices and U has at most $2n$ vertices. Additionally, any blue copy of $S_n^{(2)}$ would have to include a blue edge between two vertices of degree at least n, and there are no such blue edges. This shows that G is not Ramsey for H. $\square$

As these examples show, either of the bounds in (1) can be close to optimal, and the value of $\hat{r}(H)$ really depends on the structure of H; superficially, $S_n^{(1)}$ and $S_n^{(2)}$ are fairly similar graphs (e.g. they are both trees of bounded diameter), yet their size Ramsey numbers are very different.

An even more surprising result, due to Chvátal, is that the upper bound in (1) is *exactly* tight when H is a clique.

Theorem 1.3 (Chvátal)**.** *For any positive integer t,*

$$\hat{r}(K_t) = \binom{r(K_t)}{2}.$$

Proof (Folklore?) Let G be a graph which is Ramsey for K_t. We first claim that $\chi(G) \geq r(K_t)$. Indeed, if $\chi(G) < r(K_t)$, then there is a graph homomorphism $\varphi : G \to K_{r(K_t)-1}$. By the definition of the Ramsey number, we know that there exists a 2-coloring of $E(K_{r(K_t)-1})$ with no monochromatic K_t. By pulling this coloring back along the homomorphism φ (i.e. by coloring an edge (u, v) of G according to the color of the edge $(\varphi(u), \varphi(v))$), we obtain a 2-coloring of $E(G)$ with no monochromatic K_t, a contradiction.

Now, it remains to observe that G has at least $\binom{\chi(G)}{2}$ edges. Indeed, suppose this is not the case, and fix a proper coloring of G with $\chi(G)$ colors. By the pigeonhole principle, there exist two colors such that there is no edge of G connecting them. But this means that the union of these two color classes is an independent set of G, so we can combine these two colors to obtain a proper coloring of G with $\chi(G) - 1$ colors, a contradiction. Putting these two observations together yields the desired result. $\square$

Theorem 1.3 is very nice, in that it allows us to reduce the study of the size Ramsey number $\hat{r}(K_t)$ to the study of the ordinary Ramsey number $r(K_t)$. Unfortunately, the determination of $r(K_t)$ is a notoriously difficult problem. The bounds

$$\sqrt{2}^t \leq r(K_t) \leq 4^t$$

were proved over 75 years ago by Erdős and Erdős–Szekeres, respectively, and despite decades of effort, there has been no improvement to either of the exponential constants $\sqrt{2}$ and 4.

Because of Theorem 1.3, all of the study of size Ramsey numbers concerns itself with graphs which are not complete. In their original 1978 paper, Erdős, Faudree, Rousseau, and Schelp proved a number of bounds on size Ramsey numbers of various classes of graphs, and they ended the paper with four questions: to determine (up to a constant factor) the size Ramsey numbers of four classes of graphs.

1. Complete bipartite graphs, $K_{s,t}$.
2. Book graphs, $B_n^{(k)}$. This graph consists of n copies of K_{k+1} glued along a common K_k.
3. Starburst graphs, $S_n^{(k)}$, defined above.
4. Path graphs, P_n.

Let me begin with the final question. As P_n is a tree, we know that $r(P_n) = \Theta(n)$ (and the precise value of $r(P_n)$ has been known for decades, thanks to a 1967 result of Gerencsér and Gyárfás). By (1), this shows that $\hat{r}(P_n)$ is somewhere between linear and quadratic in n. In 1983, Beck proved that the lower bound is essentially tight, namely that $\hat{r}(P_n) = \Theta(n)$. The key idea in Beck's proof is to take G to be a good expander on $\Theta(n)$ vertices, which we can take to have $\Theta(n)$ edges. If we two-color the edges of G, then one of the monochromatic subgraphs contains another good expander, and it is not too hard to show that good expanders contain long paths, which together implies that G is Ramsey for P_n. In fact, one can even take G to be an Erdős–Rényi random graph $G(N, p)$, where $N = \Theta(n)$ and $p = \Theta(1/n)$, and this graph will be Ramsey for H with high probability.

Almost all of the subsequent work on size Ramsey numbers begins with Beck's result. There were a sequence of results proving that certain special sparse graphs have linear size Ramsey numbers, including cycles and bounded-degree trees. More generally, suppose we have a graph H with n vertices and maximum degree Δ. A foundational result of Chvátal, Rödl, Szemerédi, and Trotter says that $r(H) = O_\Delta(n)$, that is, that bounded-degree graphs have linear Ramsey numbers. Because of this and (1), we know that $\Omega(n) \le \hat{r}(H) \le O_\Delta(n^2)$. It is natural to wonder whether the upper bound can be improved for all such H, as it is for the special sparse graphs mentioned above. This is indeed true, since Kohayakawa, Rödl, Schacht, and Szemerédi proved that $\hat{r}(H) \le n^{2-1/\Delta+o(1)}$ for all n-vertex H with maximum degree Δ. Like Beck, they prove this by taking G to be an appropriate Erdős–Rényi random graph, namely $G(N, p)$ where $N = \Theta(n)$ and $p = n^{-1/\Delta+o(1)}$. However, proving that this graph is Ramsey for H with high probability is quite involved, and uses the techniques of sparse regularity, which is an important tool in many results in the area.

In the other direction, Rödl and Szemerédi proved that the lower bound $\hat{r}(H) = \Omega(n)$ is not always tight, by showing that there exists an n-vertex graph H with maximum degree 3 and $\hat{r}(H) \ge n(\log n)^c$ for some absolute constant $c > 0$. It remains a major open problem to close the gap between the lower and upper bounds; in particular, it is unknown whether one has $\hat{r}(H) \ge n^{1+c}$ for some bounded-degree H and some positive $c > 0$. Just two weeks ago, major progress was made by Tikhomirov, who modified the Rödl–Szemerédi construction to find a graph H with maximum degree 3 and and $\hat{r}(H) \ge ne^{\Omega(\sqrt{\log n})}$.

Over the past 40 years, there was very little progress on any of the remaining three questions asked by Erdős, Faudree, Rousseau, and Schelp. For those three classes of graphs—complete bipartite, books, and starbursts—the techniques they developed allowed Erdős, Faudree, Rousseau, and Schelp to prove some bounds, but there remained a gap in all cases. Additionally, it seems that they raised these specific questions for exactly this reason, in order to see how their techniques could be improved to go beyond the barriers they encountered.

In recent work with David Conlon and Jacob Fox, we determine up to a constant factor $\hat{r}(B_n^{(k)})$ and $\hat{r}(S_n^{(k)})$ when $n \gg k$, and make significant progress on determining $\hat{r}(K_{s,t})$. We use a variety of different tools and techniques, including a number of relatively intricate probabilistic constructions. In the following sections, I will discuss these results.

2 Complete bipartite graphs

Erdős, Faudree, Rousseau, and Schelp proved that if $s \le t$, then

$$\Omega(st2^s) \le \hat{r}(K_{s,t}) \le O(s^2t2^s),$$

where the upper bound holds for all $s \le t$ and the lower bound holds whenever $t \gg s$. Later, Erdős and Rousseau actually proved that the lower bound holds for all $s \le t$, so there remains only a $\Theta(s)$ gap between the lower and upper bounds. When t is very large with respect to s, then Pikhurko proved that the upper bound is tight, and even determined the correct asymptotic constant. However, his proof gives no information on how large t must be as a function of s, and thus says nothing about $\hat{r}(K_{s,t})$ for any given s, t.

Our first result improves the Erdős–Rousseau lower bound. It gives a power improvement on s whenever $t \ge (1+\delta)s$ for any $\delta > 0$, and it matches the upper bound once $t = \Omega(s \log s)$.

Theorem 2.1 (Conlon–Fox–W.)**.** *For all $t \ge s$,*

$$\hat{r}(K_{s,t}) = \Omega(s^{2-\frac{s}{t}}t2^s)$$

The lower bound argument of Erdős and Rousseau is very simple to describe, and it is also at the heart of our improved lower bound. They first prove the following lemma.

Lemma 2.2 (Erdős–Rousseau)**.** *If $s \le t$ are integers and G is a graph with m edges, then G has at most*

$$\left(\frac{100m}{st}\right)^t$$

copies of $K_{s,t}$.

Once they have this lemma, the lower bound is straightforward. In a random coloring of G, every copy of $K_{s,t}$ is monochromatic with probability 2^{1-st}. If we choose $m = st2^s/200$ and fix a graph G with m edges, then in a random coloring of G, the expected number of monochromatic $K_{s,t}$ is at most

$$2^{1-st}\left(\frac{100m}{st}\right)^t = 2\left(\frac{100m}{st2^s}\right)^t = 2^{1-t} < 1.$$

Therefore, any graph with $st2^s/200$ edges can be colored to have no monochromatic $K_{s,t}$.

For the upper bound, let $S = 2s^2$ and $T = 10t2^s$. We claim that $K_{S,T}$ is Ramsey for $K_{s,t}$. To see this, suppose we are given a two-coloring of $E(K_{S,T})$. Without loss of generality, at least $T/2$ vertices on the right are incident to at least $S/2$ red edges. Now pick a random subset of size s from the left. The expected number of red common neighbors it has is exactly

$$\frac{1}{\binom{S}{s}} \sum_v \binom{\deg_R(v)}{s} \geq \frac{1}{\binom{S}{s}} \cdot \frac{T}{2} \cdot \binom{S/2}{s} = \frac{T}{2} \cdot \frac{\binom{S/2}{s}}{\binom{S}{s}},$$

where the sum is over all vertices on the right. By our choice of $S = 2s^2$, one can check that $\binom{S/2}{s}/\binom{S}{s} \geq \frac{1}{5} \cdot 2^{-s}$. So an average s-set on the left has at least $T/(10 \cdot 2^s) = t$ red neighbors on the right, and so we find a red $K_{s,t}$.

Recall that for the lower bound argument, Erdős and Rousseau colored the edges independently at random. But if we examine the upper bound, we see that this doesn't really seem like the right thing to do. Indeed, this upper bound argument is tight when all vertices on the right of $K_{S,T}$ have equal red and blue degrees. However, if we color the edges independently at random, then the red degree of a vertex on the right is distributed as $\mathrm{Bin}(S, \frac{1}{2})$. This random variable has standard deviation on the order of $\sqrt{S} = \Theta(s)$. Moreover, since we have $\exp(s)$ many vertices on the right, we will expect a large number of vertices to be off from the mean by a noticeable number of standard deviations. So one should expect a random coloring to not be optimal in case the host graph really is $K_{S,T}$, which was the graph we used for the upper bound.

Instead, a better coloring to use would be *hypergeometric* in nature. We want every vertex on the right to have equal red and blue degree, but otherwise we want the edges to behave randomly. So we should make every vertex on the right select a uniformly random subset of size $S/2$ from the left, make that subset its red neighborhood, and its complement its blue neighborhood.

More generally, if we wish to color an arbitrary graph G, we should try to color hypergeometrically from the low-degree to the high-degree vertices. Namely, a natural thing to try is the following. Let us order the vertices of G according to their degree, let A consist of the $\approx s^2$ highest-degree vertices in G, and let B be the remaining vertices. We want to pretend that G looks like a complete bipartite graph between A and B. In particular, we expect edges inside B not to matter too much, so we color all those edges independently and uniformly at random. All remaining edges are incident to A, and those we color hypergeometrically: for every vertex v, it has some number k of neighbors which precede it in the ordering and which are in A. It picks a random subset of exactly $k/2$ of these, colors the edges to those vertices red, and colors the remaining $k/2$ edges blue.

As it turns out, this doesn't *quite* work; it does improve on the Erdős–Rousseau bound, but the exponent on s is $2 - \frac{s}{t} + o(1)$, where the $o(1)$ tends to 0 as $s \to \infty$.

The problem is that we actually had a fair amount of wiggle room in our choice of S and T above. For example, we can multiply S by an absolute constant factor c, multiply T by roughly $1/c$, and as long as we are careful, we can still get a complete bipartite graph that is Ramsey for $K_{s,t}$, and which has essentially the same number of edges as

$K_{S,T}$. So we can't just "model" G as $K_{S,T}$; if we try to do that, we might be off by a big constant factor on the "correct" size for A. So the right thing to do is to hedge our bets. We pick sets $A_1, A_2, \ldots, A_m$, where A_1 consists of the two highest-degree vertex in G, A_2 consists of the next four highest-degree vertices, and in general A_i will consist of the 2^i highest-degree vertices that have not yet been included in some A_j. We then color all edges hypergeometrically independently into each A_i, making all these choices independently. Since the A_i's cover all possible "scales", the issue encountered above goes away, and one can show that this coloring has no monochromatic $K_{s,t}$ with high probability, as long as the number of edges in G is at most $O(s^{2-\frac{s}{t}}t2^s)$.

3 Book graphs

Recall that the book graph $B_n^{(k)}$ consists of n copies of K_{k+1} glued along a common K_k. Why should one care about book graphs? As it turns out, there is good reason to care: book graphs play a surprisingly central role in a number of other Ramsey-theoretic questions. For example, all known techniques for upper-bounding $r(K_t)$ use (implicitly or explicitly) bounds on $r(B_n^{(k)})$ somewhere. In fact, Ramsey's original proof of Ramsey's theorem proved the finiteness of $r(K_t)$ by inductively proving the finiteness of $r(B_n^{(k)})$ for appropriate choices of k and n. Additionally, books show up in other Ramsey-theoretic problems; for example, a result of Sudakov on the Ramsey number of a graph with a given number of edges works by iteratively finding large monochromatic books in a coloring of K_N.

Erdős, Faudree, Rousseau, and Schelp proved that for $n \gg k \geq 2$,

$$\Omega(k^2n^2) \leq \hat{r}(B_n^{(k)}) \leq O(16^k n^2).$$

The lower bound follows from a simple explicit coloring, and the upper bound follows from a result on the ordinary Ramsey number of $B_n^{(k)}$. Indeed, Erdős, Faudree, Rousseau, and Schelp proved that

$$(2^k - o(1))n \leq r(B_n^{(k)}) \leq 4^k n, \tag{2}$$

where the $o(1)$ term tends to 0 as n tends to infinity, for fixed k. Here, the lower bound follows immediately from a random coloring, and the upper bound follows by a simple neighborhood-chasing argument. Closing the gaps here remained a major open problem, until it was essentially fully resolved by Conlon in 2019.

Theorem 3.1 (Conlon). *If k is fixed, then*

$$r(B_n^{(k)}) = (2^k + o(1))n$$

as $n \to \infty$.

Conlon's original proof used Szemerédi's regularity lemma and gave very poor quantitative estimates on the $o(1)$ term. In follow-up work, Conlon, Fox, and I simplified Conlon's proof, got a much better quantitative dependence, and proved some generalizations which I won't discuss here. But suffice to say that these works give powerful techniques for finding monochromatic books in colorings of graphs, and I will return to this momentarily.

As an immediate corollary of Conlon's theorem, we see that $\hat{r}(B_n^{(k)}) \leq \binom{r(B_n^{(k)})}{2} = O(4^k n^2)$ for $n \gg k$. But even with this improvement, there remains a massive gap between the lower and upper bounds in (2), between quadratic in k and exponential in k behavior.

As it turns out, the truth is essentially halfway between these two bounds.

Theorem 3.2 (Conlon–Fox–W.). *If n is sufficiently large with respect to $k \geq 2$, then*

$$\hat{r}(B_n^{(k)}) = \Theta(k 2^k n^2).$$

Let's begin with the upper bound, which is technically harder but provides good intuition for the lower bound. We wish to find a graph G which is Ramsey for $B_n^{(k)}$. I think a natural choice is $B_N^{(K)}$ for appropriate parameters $K > k, N > n$, and we hope to embed the common K_k of $B_n^{(k)}$ in the common K_K of $B_N^{(K)}$. For this to work, how big must we take K and N?

First, we claim that we need $K > (k-1)n$. Indeed, suppose that $K \leq (k-1)n$. Then we can use the *Turán coloring* on the K_K of $B_N^{(K)}$. Namely, we split this K_K into $(k-1)$ sets, each of size n, and color each set red and all edges between distinct sets blue. Additionally, we color all edges between the K_K and the remaining N vertices blue. Inside the K_K, there is no red $B_n^{(k)}$, since the red graph has connected components with only n vertices, so $B_n^{(k)}$ can't fit. Additionally, the blue graph on the K_K has chromatic number $k-1$, so it contains no K_k. Thus, we can't possibly find a blue $B_n^{(k)}$ in the entire coloring. So we need $K > (k-1)n$.

On the other hand, suppose that we color the edges uniformly at random. Then a simple application of the Chernoff bound and the union bound shows that no monochromatic K_k will have many more than $2^{-k}(K+N)$ monochromatic extensions to a K_{k+1}. Since we hope to find n monochromatic extensions, we see that we need to take N at least roughly $2^k n$ for this to have any hope of working.

As it turns out, these are essentially the only two constraints! This is the content of the following result, which implies the upper bound in Theorem 3.2.

Lemma 3.3 (Conlon–Fox–W.). *Suppose n is sufficiently large with respect to k. If $K = 2kn$ and $N = 2^{k+1}n$, then $B_N^{(K)}$ is Ramsey for $B_n^{(k)}$.*

The proof is quite involved, and requires the techniques developed to find monochromatic books in the earlier works of Conlon and Conlon–Fox–W. But morally, the content of these techniques is that once one is past the "Turán barrier" and the "random barrier", then any coloring contains a monochromatic book.

This upper bound also hints at how to prove the lower bound in Theorem 3.2. Namely, we are given a graph G with few edges, and we wish to color it so that it there is no monochromatic $B_n^{(k)}$. If we take as our model that G should be "similar to" $B_N^{(K)}$, then we want to use both the Turán coloring and the random coloring. We want to use the Turán coloring within the "high-degree part" of G, and the random coloring everywhere else. So a natural thing to try is to pull out the $(k-1)n$ highest-degree vertices of G, further partition them into $k-1$ sets and put the Turán coloring between these sets, and then to color all remaining edges uniformly at random. As it turns out, this doesn't quite work, but something similar to it does.

Namely, suppose we are given a graph G with at most $k2^kn^2/10000$ edges. Let $s = k/3$. Let V_0 consist of the $n/10$ highest-degree vertices in G, let V_1 consist of the $n/10$ next-highest-degree vertices, and so on until V_s. Finally, let U consist of all the remaining vertices. We color each V_i red, and color all edges between V_i and V_j for $i \neq j$ blue. Inside U, we color each edge red or blue uniformly at random. However, for all edges between V_i and U, we color them red with probability $p_i = \frac{1}{2}(i/s)^{1/k}$ and blue with probability $1 - p_i$. Basically, for smaller i, the vertices in V_i have higher degree, and are thus more "dangerous" for red books. So we wish to counteract this by coloring them red with probability that is smaller than $\frac{1}{2}$, and it turns out that the choice of p_i above is a good one.

As it turns out, this coloring works. The proof is not hard, but it is fairly technically involved; there are many different "types" of books one needs to worry about, and for each one, one needs to analyze the probability that it is monochromatic. So I will not discuss the proof any further.

4 Starburst graphs

As we saw earlier, there is a notable difference between $\hat{r}(S_n^{(1)})$ and $\hat{r}(S_n^{(2)})$, coming from the presence of adjacent high-degree vertices. Given this, it is natural to ask about $\hat{r}(S_n^{(k)})$ for $k \geq 3$ and $n \gg k$.

Erdős, Faudree, Rousseau, and Schelp proved that for $k \geq 2$,

$$\Omega(k^3n^2) \leq \hat{r}(S_n^{(k)}) \leq O(k^4n^2) \tag{3}$$

when n is sufficiently large with respect to k, and they asked to close the $\Theta(k)$ gap. The lower bound is again proved via a simple explicit coloring.

For the upper bound on $\hat{r}(S_n^{(k)})$, Erdős, Faudree, Rousseau, and Schelp determined (up to a constant factor) the ordinary Ramsey number $r(S_n^{(k)})$, namely $r(S_n^{(k)}) = \Theta(k^2n)$ for n sufficiently large with respect to k. By (1), this implies that $\hat{r}(S_n^{(k)}) \leq \binom{r(S_n^{(k)})}{2} = O(k^4n^2)$. The proof that $r(S_n^{(k)}) = O(k^2n)$ is fairly straightforward, and uses a neighborhood-chasing argument akin to the ordinary Erdős–Szekeres proof which bounds $r(K_t)$.

Our final result closes the $\Theta(k)$ gap in (3), establishing that the lower bound is asymptotically tight.

Theorem 4.1 (Conlon–Fox–W.). *If n is sufficiently large with respect to $k \geq 2$, then*

$$\hat{r}(S_n^{(k)}) = \Theta(k^3n^2).$$

The idea of the proof is as follows. As mentioned above, if $N = \Theta(k^2n)$, then K_N is Ramsey for $S_n^{(k)}$. Rather than using K_N, we use the random graph $G(N, p)$ where $p = \Theta(1/k)$, which has $\Theta(k^3n^2)$ edges with high probability. The key idea is that, since the edges of $G(N,p)$ are very well-distributed with high probability, the same neighborhood-chasing argument that worked in K_N can be made to work in $G(N,p)$, establishing the upper bound.

This proof works once $n \geq k^{Ck}$ for some constant $C > 0$. Moreover, the proof really breaks down below this threshold, since $G(N,p)$ will not even contain a *single* copy of K_k if $p = \Theta(1/k)$ and $N < k^{ck}$ for some small $c > 0$. However, by using a more involved random model, one can prove that $\hat{r}(S_n^{(k)}) = \Theta(k^3 n^2)$ already once $n \geq 2^{Ck}$. Namely, rather than considering $G(N,p)$, we consider the following random graph model, which we call $G(N,p,\omega)$, and which is designed to artificially inflate the clique number of $G(N,p)$ without affecting much else. We form $G(N,p,\omega)$ by starting with an empty graph on N vertices, repeatedly adding the edges of a uniformly random clique K_ω, and continuing in this way until we've added at least $p\binom{N}{2}$ edges. One can show that if ω is not too large, then this graph looks similar to $G(N,p)$ with respect to the "global" distribution of edges (e.g. any two large subsets have edge density roughly p between them). However, the "local" distribution of edges is much denser, as $G(N,p,\omega)$ has enormous cliques. Because of this, one can use a similar proof technique to show that with high probability, $G(N,p,\omega)$ is Ramsey for $S_n^{(k)}$ if $N = \Theta(kn), p = \Theta(1/k)$, and $\omega = N^{1/8}$, and that this already holds once $n \geq 2^{Ck}$.