Improving $R(3,k)$ in just two bites

Zion Hefty* Paul Horn† Dylan King‡ Florian Pfender§

December 3, 2025

Abstract

We present a random construction proving that the extreme off-diagonal Ramsey numbers satisfy $R(3,k) \geq \left(\frac{1}{2}+o(1)\right)\frac{k^2}{\log k}$ (conjectured to be asymptotically tight), improving the previously best bound $R(3,k) \geq \left(\frac{1}{3}+o(1)\right)\frac{k^2}{\log k}$. In contrast to all previous constructions achieving the correct order of magnitude, we do not use a nibble argument.

1 Introduction

The Ramsey number $R(\ell,k)$ is the least N such that every N-vertex graph contains either a clique of size ℓ or an independent set of size k. In 1930 Ramsey [20] showed the existence of $R(\ell,k)$, and in the intervening century much effort has been dedicated to understanding their asymptotic behavior. The best current bounds on the diagonal Ramsey numbers $R(k,k)$ (which are known exactly only up to $k=4$, with $43 \leq R(5,5) \leq 46$ and the upper bound only very recently announced by Angeltveit and McKay [4]) are

$$2^{k/2+o(1)} \leq R(k,k) \leq (4-c)^k,$$

for some $c>0$. The lower bound is one of the earliest instances of Erdős' probabilistic method [11], and the upper bound was the first exponential improvement over a result of Erdős and Szekeres [12], only recently announced in 2023 by Campos, Griffiths, Morris and Sahasrabudhe [6], subsequently optimized to a value of $c \approx 0.2$ (for sufficiently large k) by Gupta, Ndiaye, Norin and Wei [15].

The second most studied family of graph Ramsey numbers are the extreme off-diagonal numbers $R(\ell,k)$ for ℓ fixed and k tending to infinity. The case $\ell=3$ yields the first non-trivial asymptotic question. The exciting history of the problem until the early 2000s is described by Spencer [23] in a chapter of the book "Ramsey Theory" [22], while the more recent history is very well captured by Campos, Jenssen, Michelen and Sahasrabudhe [8]. We briefly summarize the results determining the order of magnitude. The upper bound of $R(3,k) = O(\frac{k^2}{\log k})$ goes back to work of Ajtai, Komlós and Szemeŕedi [1,2] in the early 1980s, and in 1983 Shearer [21] reduced the implied constant to $R(3,k) \leq (1+o(1))\frac{k^2}{\log k}$. The lower bound was improved in several steps until Kim [16] proved in 1995 that $R(3,k) \geq c\frac{k^2}{\log k}$, with a constant $c \approx \frac{1}{160}$. In his paper, which led to him winning the Fulkerson Prize, Kim uses a probabilistic method commonly known as a "nibble", where small batches of edges are added randomly for many steps. This carefully avoids the appearance of triangles while keeping the edge distribution sufficiently random to destroy large independent sets. Spencer [23] writes in 2011:

"The value of a constant c so that $R(3,k) \sim c\frac{k^2}{\log k}$ remains open to this day, but this problem seems beyond our reach."

*Department of Mathematics, University of Denver, Denver, USA. Email: Zion.Hefty@du.edu

†Department of Mathematics, University of Denver, Denver, USA and Department of Mathematics and Applied Mathematics, University of Johannesburg, Johannesburg, South Africa. Email: Paul.Horn@du.edu

‡California Institute of Technology, Pasadena, USA. Email: dking@caltech.edu

§Department of Mathematical and Statistical Sciences, University of Colorado Denver, Denver, USA. E-mail: Florian.Pfender@ucdenver.edu. Research is partially supported by NSF FRG DMS-2152498.

Nevertheless, only a short time later in 2013, two groups announced that they had greatly improved on the constant in the lower bound, showing that $R(3,k) \geq \left(\frac{1}{4} + o(1)\right) \frac{k^2}{\log k}$. Bohman and Keevash [5] and Fiz Pontiveros, Griffiths and Morris [13] very carefully analyzed the triangle-free process, in which an edge, chosen uniformly at random from those which would not create a triangle, is added to the graph (and this operation is repeated until no such edges exist). Both of these papers go through serious technical efforts to control the randomness in the process, and in particular show that with high probability the triangle-free process terminates with independence number yielding this bound. Fiz Pontiveros, Griffiths and Morris went so far as to conjecture that $c = \frac{1}{4}$ is the correct constant for $R(3,k)$.

In another breakthrough 12 years later, Campos, Jenssen, Michelen and Sahasrabudhe [8] showed that the story of the lower bound was not finished, and in fact $R(3,k) \geq \left(\frac{1}{3} + o(1)\right) \frac{k^2}{\log k}$, now conjecturing that $c = \frac{1}{2}$ should be the correct constant.

Conjecture 1.1 (Campos, Jenssen, Michelen and Sahasrabudhe [8])**.**

$$R(3,k) = \left(\frac{1}{2} + o(1)\right) \frac{k^2}{\log k}.$$

In this paper, we prove the lower bound in Conjecture 1.1.

Theorem 1.2.

$$R(3,k) \geq \left(\frac{1}{2} + o(1)\right) \frac{k^2}{\log k}.$$

As all other previous results, we prove the following theorem concerning independence numbers, which easily implies Theorem 1.2.

Theorem 1.3. *For all $\varepsilon > 0$, there exists an $n_0 = n_0(\varepsilon)$ so that for $n \geq n_0$ there exists a triangle-free graph G on n vertices with independence number*

$$\alpha(G) < (1+\varepsilon)\sqrt{n \log n}.$$

While inspired by ideas in [8], the construction in Theorem 1.3 does not require a nibble, and thus circumvents many of the technical difficulties faced by the previous groups. Spencer states in the epilogue of his book chapter:

"Is the story of $R(3,k)$ over? I think not. I think there is plenty of room for a consolidation of the results. My dream is a ten-page paper which gives $R(3,k) = \Theta(k^2/\log k)$."

If one were interested only in the correct order of magnitude, our construction could achieve Spencer's dream, by combining slightly weaker proofs of the ingredients of Theorem 1.3 with Shearer's one page proof of the upper bound.

Our method seems quite versatile and has already seen some applications to related problems. In Section 5 we give a direct application to hypergraph Ramsey numbers of stars, suggested to us by Mubayi. Most recently, Campos, Jenssen, Michelen and Sahasrabudhe together with the fourth author [7] have built on our method to find the first polynomial improvements over the edge deletion threshold for the cycle-complete Ramsey numbers $r(C_\ell, K_k)$, for all odd $\ell \geq 9$ as k tends to infinity.

2 The Construction of G

Before formally presenting our construction witnessing Theorem 1.3, we provide some heuristic justification. The edge deletion method is a typical starting point in modern combinatorics when one wishes to construct a large graph G which avoids a forbidden subgraph H, but carries certain other properties held by the random graph (in our case, G should have no large independent sets). Implementing the method first requires one to calculate the edge deletion threshold p_H, so that (in expectation) the number of copies of H in the random graph $G(n, p_H)$ is of the same order as $p_H n^2$, the expected number of edges. If $p \ll p_H$ and one samples G from $G(n,p)$ before deleting an edge from every copy of H, only $o(p_H n^2)$ edges are lost. Some further effort can show that the resulting graph will have small independence number, because $G(n,p)$ had small independence number and only a small number of edges were removed. For many problems, this approach gives the best known constructions. To outperform such examples, one may try to construct an H-free graph with edge density larger than p_H, while still being careful to obtain some of the pseudorandomness that the random graph enjoys.

Most recently, Campos, Jenssen, Michelen and Sahasrabudhe [8] constructed a triangle-free graph with edge density $(1+o(1))\sqrt{\frac{2\log n}{3n}}$. To do so, they let $N = n/\log^2(n)$ and $p = \sqrt{\frac{\log(n)}{6n}}$ before sampling from G_1 from $G(N,p)$. Noting that $p = o(\sqrt{\frac{1}{N}})$ they can delete few edges to make the resulting graph (still on N vertices) triangle free, before blowing up each vertex by a factor $\log^2(n)$ to obtain a new graph on n vertices, still triangle-free and with edge density $\sqrt{\frac{\log n}{6n}}$, well above the edge deletion threshold. Unfortunately the blow-up operation also inflates independent sets. In [8] it is shown that if one then runs a tailored nibble (loosely speaking, a guided triangle-free process), one will eventually double the density of the blow-up graph and regain the necessary pseudo-randomness, destroying the (structured) large independent sets. Our construction replaces the nibble in the second stage of [8] by a second copy of a blow-up of the smaller random graph. We recover the necessary pseudorandom properties by randomly overlaying the two blow-ups; each destroys the large independent sets in the other.

We note that a similar idea of stacking triangle-free graphs to reduce the independence number was used before by Alon and Rödl [3] where they find bounds for multi-color Ramsey numbers $R(3,\ldots,3,k)$ by combining several copies of a Ramsey graph for $R(3,k)$. A second idea we utilize in our construction is that of sampling a subgraph from a larger graph to reduce the relative size of the independence number. Mubayi and Verstraëte use this idea in [19] to deduce lower bounds on Ramsey numbers from the existence of graphs with good pseudorandom properties (there measured by eigenvalues). A similar idea is used by Mattheus and Verstraëte [17] in their celebrated construction providing (up to logarithmic factors) a tight lower bound on $R(4,k)$.

Let us now formally define our construction of the desired n vertex graph G. For the sake of readability, we omit floors and ceilings whenever they are inconsequential. All of our logarithms are natural with base e. For any n and ε, set

$$N = \frac{n}{\log^2(n)}, \qquad p = \beta\sqrt{\frac{\log n}{n}}, \qquad k = \kappa\sqrt{n\log n},$$

where ultimately

$$\beta = \frac{1}{2}, \qquad \kappa = 1+\varepsilon.$$

The construction then consists of the following steps:

1. Sample two (independent) copies of the random graph $G(N,p)$, G_R and G_B, on vertex sets $V_R = \{r_1,\ldots,r_N\}$ and $V_B = \{b_1,\ldots,b_N\}$, with $p = \beta\sqrt{\frac{\log n}{n}}$. We will refer to edges in G_R and G_B as red and blue edges respectively.

2. Let $G_1 = G_R * G_B$ be the co-normal product, with vertex set $V_R \times V_B$ and $((r_i,b_j),(r_k,b_\ell)) \in E(G_1)$ if $(r_i,r_k) \in E(G_R)$ or $(b_j,b_\ell) \in E(G_B)$. See Figure 1 for a small example of such a product G_1. We call an edge $((r_i,b_j),(r_k,b_\ell)) \in E(G_1)$ red if $(r_i,r_k) \in E(G_R)$ and blue if $(b_j,b_\ell) \in E(G_B)$, noting that it is possible for a single edge to be colored both red and blue.

3. For any vertex $r_i \in V_R$, let $N(r_i)$ denote the neighborhood of r_i in V_R, and $N^+(r_i) := N(r_i) \cap \{r_{i+1}\ldots,r_N\}$ the 'right' neighbors which lie above r_i in the natural ordering of $V(G_R)$. Denote the sets $X_{r_i} = N(r_i)\times V_B$ and $X^+_{r_i} = N^+(r_i)\times V_B$; similarly define $X_{b_i} = V_R\times N(b_i)$ and $X^+_{b_i} = V_R\times N^+(b_i)$ for $b_i \in V_B$.

   We create a triangle-free graph $G_2 \subseteq G_1$ by deleting all red edges in

$$\bigcup_{i=1}^{N}\left(\binom{X^+_{r_i}}{2} \cup \binom{X_{b_i}}{2}\right),$$

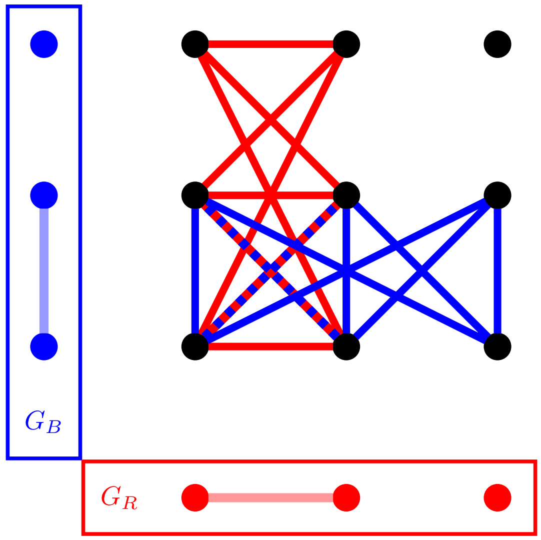

Figure 1: A small example of G_1

and all blue edges in

$$\bigcup_{i=1}^{N}\left(\binom{X_{r_i}}{2}\cup\binom{X_{b_i}^{+}}{2}\right).$$

Indeed, suppose that $(r_{i_1}, b_{j_1}), (r_{i_2}, b_{j_2}), (r_{i_3}, b_{j_3})$ form a triangle in G_1. If this triangle is monochromatic red, then necessarily $r_{i_1}, r_{i_2}, r_{i_3}$ are distinct, and by relabeling we assume that $r_{i_1} < r_{i_2} < r_{i_3}$. Then the red edge $((r_{i_2}, b_{j_2}), (r_{i_3}, b_{j_3}))$ is removed as it belongs to $\binom{X_{r_{i_1}}^{+}}{2}$ (and an analogous argument holds for monochromatic blue triangles). Suppose then that $(r_{i_1}, b_{j_1}), (r_{i_2}, b_{j_2}), (r_{i_3}, b_{j_3})$ is a triangle with two red edges and (after relabeling) $((r_{i_1}, b_{j_1}), (r_{i_2}, b_{j_2}))$ is blue. Then $((r_{i_1}, b_{j_1}), (r_{i_2}, b_{j_2})) \in \binom{X_{r_3}}{2}$ and this blue edge would be removed. The case of two blue edges and one red edge is identical.

4. Let $V(G) = \{v_1, \ldots, v_n\}$, choose an injection $\varphi : V(G) \rightarrowtail V_R \times V_B$ uniformly at random, and include an edge (v, v') in $E(G)$ if and only if $(\varphi(v), \varphi(v')) \in E(G_2)$. It will be useful in the future to denote the graph induced by G_2 on $\varphi(V(G))$ as G_3 so that G is the pre-image of G_3.

   In the course of analyzing the construction we will utilize the parameter hierarchy

$$0 < \varepsilon_2 \ll \varepsilon_1 \ll \varepsilon \ll 1.$$

   First we will use ε_2 in Lemma 3.1 to quantify standard concentration results for degrees and codegrees in random graphs, and ε_1 will be used in Lemmas 3.2, 3.3, 3.4, and 3.5 to quantify properties of k-sets in G. Finally in Section 4 we will show that $\alpha(G) < (1+\varepsilon)\sqrt{n \log n}$ with high probability.

3 Preliminaries

In this section, we introduce some notation and prove several structural lemmas that enable us to analyze the independent sets in G. For $r_i \in V_R$ let $F(r_i) = \{(r_i, b_j) \in V(G_3)\}$ and for $b_i \in V_B$ let $F(b_i) = \{(r_j, b_i) \in V(G_3)\}$ denote the respective fibers in $V(G_3)$. Let

$$N_3(r_i) = \bigcup_{r_j \in N(r_i)} F(r_j), \qquad N_3^{+}(r_i) = \bigcup_{r_j \in N^{+}(r_i)} F(r_j),$$

$$N_3(b_i) = \bigcup_{b_j \in N(b_i)} F(b_j), \qquad N_3^{+}(b_i) = \bigcup_{b_j \in N^{+}(b_i)} F(b_j).$$

For any set $X \subseteq V_R \times V_B$ denote the projections onto V_R and V_B by

$$\pi_R(X) = \{r_i \in V_R : F(r_i) \cap X \neq \emptyset\},$$
$$\pi_B(X) = \{b_i \in V_B : F(b_i) \cap X \neq \emptyset\}.$$

For any set $X \subseteq V(G)$ we slightly abuse the notation, letting $\pi_R(X) = \pi_R(\varphi(X))$ and $\pi_B(X) = \pi_B(\varphi(X))$.

The following routine lemma shows that the degrees and neighborhoods in the model behave as expected with high probability.

Lemma 3.1. *Let $C = 3\sqrt{20}$ and $\mathcal{D}$ be the event that the following hold for all distinct $v, w \in (V_R \cup V_B)$.*

$$\left|F(v) - \log^2(n)\right| \leq \varepsilon_2 \log^2(n) \tag{1}$$
$$\left||N(v)| - pN\right| \leq \varepsilon_2 pN \tag{2}$$
$$|N(v) \cap N(w)| \leq C \log n \quad \text{if } v, w \in V_R \text{ or } v, w \in V_B \tag{3}$$
$$\left||N_3(v)| - pn\right| \leq \varepsilon_2 pn \tag{4}$$
$$|N_3(v) \cap N_3(w)| \leq C \log^3(n) \tag{5}$$
$$|\pi_B(N_3(v)) \cap \pi_B(N_3(w))| \leq C \log^3(n) \quad \text{if } v, w \in V_R \tag{6}$$
$$|\pi_R(N_3(v)) \cap \pi_R(N_3(w))| \leq C \log^3(n) \quad \text{if } v, w \in V_B \tag{7}$$

Then $\mathbb{P}(\mathcal{D}) = 1 - o(1)$.

Recall here that $1/n \ll \varepsilon_2 \ll \varepsilon$ and therefore the conclusion that $\mathbb{P}(\mathcal{D}) = 1 - o(1)$ is statement on asymptotic behavior as n grows and ε_2 remains fixed.

Proof of Lemma 3.1. In our construction, G_3 is chosen specifically to have exactly n vertices; we first analyze (1) – (7) under a different model using independent vertex inclusion. To that end, let G_4 be the random graph obtained from G_1 by taking an induced graph on a ρ-random subset of G_1 with vertex-inclusion probability $\rho = \frac{n}{N^2}$. We use all of the same neighborhood notation ($N(\cdot), F(\cdot)$, etc.) for this model, and will use later that, when conditioned on $|V(G_4)| = n$, G_4 is distributed identically to G_3.

Although the conclusion of the lemma takes all of the constants in the error to be the same (either ε_2 or C), due to dependencies between the involved quantities it is easier to prove these statements with slightly different constants. Thus we prove (1) and (2) with $\varepsilon_2' = \varepsilon_2/3$, and we prove (3) with $C_3 = \sqrt{20}$. Of course since $\varepsilon_2' < \varepsilon_2$ and $C_3 < C$, the claimed bounds will still hold.

Let $A_1, \ldots, A_7$ be the events that each of (1) – (7) hold for G_4 with the updated constants described in the previous paragraph, and define $\mathcal{D}' = \bigcap_{i=1}^7 A_i$. We utilize the expansion

$$\mathbb{P}({\mathcal{D}'}^c) \leq \mathbb{P}(A_1^c)+\mathbb{P}(A_2^c)+\mathbb{P}(A_3^c)+\mathbb{P}(A_4^c \mid A_1\cap A_2)+\mathbb{P}(A_5^c \mid A_1\cap A_3)+\mathbb{P}(A_6^c \mid A_1\cap A_2\cap A_3)+\mathbb{P}(A_7^c \mid A_1\cap A_2\cap A_3)$$

and analyze each of the terms on the right individually, starting with $\mathbb{P}(A_1^c)$.

Since G_4 includes each vertex of G_1 independently, it is straightforward to check the expectation $\mathbb{E}(F(v)) = \frac{n}{N^2} \cdot N = \log^2(n)$. Applying the Chernoff bound Theorem A.2 (here and for the remaining applications we are interested in random variables which are the sum of independent Bernoulli variables) with $t = c_1 \log^{3/2}(n)$ shows that, for a fixed $v \in V_R \cup V_B$,

$$\mathbb{P}(|F(v) - \log^2(n)| \geq c_1 \log^{3/2}(n)) \leq 2 \exp\left(-\frac{c_1^2 \log^3(n)}{4(\log^2(n))}\right) = O\left(\frac{1}{N^4}\right)$$

where the last step holds for some c_1 a large constant. A union bound over v shows that $\mathbb{P}(A_1^c) = O(\frac{1}{N^3})$.

A similar strategy is effective for A_2. Again for a fixed v we can calculate $\mathbb{E}(N(v)) = \beta\sqrt{\frac{n}{\log^3(n)}}$, and applying Theorem A.2 now with $t = c_2\sqrt{\frac{n}{\log^4(n)}}$ shows

$$\mathbb{P}\left(\left|N(v) - \beta\sqrt{\frac{n}{\log^3(n)}}\right| \geq c_2\sqrt{\frac{n}{\log^4(n)}}\right) \leq 2\exp\left(-\frac{c_2^2\sqrt{n}}{4(\log^{5/2}(n))}\right) = O\left(\frac{1}{N^4}\right)$$

for some c_2 a large constant, and a union bound over v shows $\mathbb{P}(A_2^c) = O(\frac{1}{N^3})$.

For A_3 the expectation is much smaller since we assume $u \neq w$; now $\mathbb{E}(|N(v) \cap N(w)|) = Np^2 = o(1)$. Then applying Theorem A.2 with $t = C_3 \log n$ we see that for a fixed pair v, w

$$\mathbb{P}\left(|N(v) \cap N(w)| \geq C_3 \log n\right) \leq 2 \exp\left(-\frac{C_3^2 \log n}{4}\right) = O\left(\frac{1}{N^5}\right)$$

for our choice of $C_3 = \sqrt{20}$. A union bound over v and w shows that $\mathbb{P}(A_3^c) = O(\frac{1}{N^3})$.

For A_4, since $(1+\varepsilon_2')^2 = (1+\varepsilon_2/3)^2 \leq 1+\varepsilon_2$, it follows that $A_1 \cap A_2 \subseteq A_4$ and therefore $\mathbb{P}(A_4^c \mid A_1 \cap A_2) = 0$. This deduction is one of the reasons why we prove (1) and (2) with modified constants.

The event A_5 considers two types of pairs v, w; ones where they both belong to V_R (or both to V_B) and pairs where, say, $v \in V_R$ and $w \in V_B$. In the first case, since $(1+\varepsilon_2')(C_3) \leq C$, (5) holds on $A_1 \cap A_3$. In the second case $N_3(v) \cap N_3(w)$ is a $(\frac{n}{N^2})$-random subset of the box $N(v) \times N(w)$ which has size $(1+o(1))p^2N^2$ on A_2. Computing the conditional expectation $\mathbb{E}(|N_3(v) \cap N_3(w)| \mid A_1 \cap A_3) = (1+o(1))\beta^2 \log n$ and applying Theorem $A.2$ with $t = c_5 \log n$ shows that

$$\mathbb{P}\left(|N_3(v) \cap N_3(w)| \geq (2\beta^2 + c_5) \log n\right) \leq 2 \exp\left(-\frac{c_5^2 \log n}{4}\right) = O\left(\frac{1}{N^5}\right)$$

for some sufficiently large fixed c_5 – again, $c_5 = \sqrt{20}$ suffices. Observe then that $2\beta^2 + \sqrt{20} \leq 3\sqrt{20} = C$. A union bound over the pairs $v \in V_R, w \in V_B$ suffices to conclude that $\mathbb{P}(A_5^c \mid A_1 \cap A_3) = O(\frac{1}{N^3})$.

Finally we turn to $\mathbb{P}(A_6^c \mid A_1 \cap A_2 \cap A_3)$ (the analysis for A_7 is entirely analogous). For $v, w \in V_R$, we decompose

$$\pi_B(N_3(v)) \cap \pi_B(N_3(w)) \subseteq \Bigg(\pi_B(N_3(v) \cap N_3(w)) \cup \Bigg(\pi_B(N_3(v) \setminus N_3(w)) \cap \pi_B(N_3(w) \setminus N_3(v))\Bigg)\Bigg). \tag{8}$$

If $A_1 \cap A_3$ occurs, then deterministically

$$|\pi_B(N_3(v) \cap N_3(w))| \leq (1+\varepsilon_2')C_3 \log^3(n) < 2\sqrt{20} \log^3(n).$$

For the second term, let $Z_{v,w} = |\pi_B(N_3(v) \setminus N_3(w)) \cap \pi_B(N_3(w) \setminus N_3(v))|$. To compute $\mathbb{E}(Z_{v,w})$, consider a fixed element $b \in V_B$, which lies in the above set if there are $x \in N(v) \setminus N(w)$ and $y \in N(w) \setminus N(v)$ with $(x, b) \in V(G_3)$ and $(y, b) \in V(G_3)$. Using now that $N(w) \setminus N(v)$ and $N(v) \setminus N(w)$ are disjoint, this occurs with probability $(1-(1-n/N^2)^{|N(v)\setminus N(w)|})(1-(1-n/N^2)^{|N(w)\setminus N(v)|})$. Having conditioned on A_2 and A_3, the intersection $N(v) \cap N(w)$ is small and therefore it still holds that $|N(v) \setminus N(w)| = (1+o(1))pN$, and likewise for $|N(w) \setminus N(v)|$. Using the large n approximation for $(1-n/N^2)^{(1+o(1))pN} = (1+o(1))(1-\frac{pn}{N})$, we see

$$\mathbb{E}(Z_{v,w} \mid A_1 \cap A_2 \cap A_3) = (1+o(1))\beta^2 N \frac{\log^5(n)}{n} \leq \log^3(n),$$

for n sufficiently large. Furthermore, this intersection is binomially distributed, as for different b these events are independent.

Now, applying Theorem $A.2$ with $t = c_6 \log^2(n)$ implies

$$\mathbb{P}\left(Z_{v,w} \geq \log^3(n) + c_6 \log^2(n)\right) \leq 2 \exp\left(-\frac{c_6^2 \log n}{4}\right) = O\left(\frac{1}{N^5}\right)$$

for some sufficiently large $c_6 > 0$. Combining the bounds obtained on $Z_{v,w}$ and $|\pi_B(N_3(v) \cap N_3(w))|$ we see that

$$|\pi_B(N_3(v)) \cap \pi_B(N_3(w))| \leq |\pi_B(N_3(v) \cap (N_3(w))| + Z_{v,w} \leq 3\sqrt{20} \log^3(n) = C \log^3(n)$$

with high probability – and as before, $\mathbb{P}(A_6^c \mid A_1 \cap A_2 \cap A_3) = O(\frac{1}{N^3})$ by the union bound. Likewise, by symmetry $\mathbb{P}(A_7^c \mid A_1 \cap A_2 \cap A_3) = O(\frac{1}{N^3})$.

Summing the bounds obtained above for each A_i shows that $\mathbb{P}(\mathcal{D}'^c) = O(\frac{1}{N^3})$.

We are now ready to transfer this analysis back to our model where $V(G_2)$ is selected randomly from all n-vertex subsets. As remarked at the start of the proof, $\mathbb{P}(\mathcal{D}'^c \mid v(G') = n) = \mathbb{P}(\mathcal{D}^c)$. One can also observe that $\mathbb{P}(|V(G_3)| = n) = \Omega(\frac{1}{N^2})$, since $|V(G_3)|$ as a random variable takes on N^2+1 discrete values and n is the most probable. Putting these together,

$$\mathbb{P}(\mathcal{D}^c) \leq \frac{\mathbb{P}(\mathcal{D}'^c)}{\mathbb{P}(|V(G_3)| = n)} = O\left(\frac{1}{N}\right)$$

and we are done. $\square$

Given $I \subseteq V(G)$ with $|I| = k$ and a vertex $v \in V_R \cup V_B$, we let

$$X_v(I) = \varphi(I) \cap X_v, \qquad X_v^+(I) = \varphi(I) \cap X_v^+.$$

For a set I, we call the sets of pairs

$$C(I) = \bigcup_{v \in V_R \cup V_B} \binom{X_v(I)}{2}, \qquad C^+(I) = \bigcup_{v \in V_R \cup V_B} \binom{X_v^+(I)}{2}$$

'closed' and 'closed$^+$' pairs, while the pairs in $O(I) = \binom{\varphi(I)}{2} \setminus C(I)$ and $O^+(I) = \binom{\varphi(I)}{2} \setminus C^+(I)$ are 'open' and 'open$^+$', respectively. By definition, we have $C^+(I) \subseteq C(I)$ and $O(I) \subseteq O^+(I)$. As G_2 is triangle-free, all edges in G_2 are in $O^+(I)$. The main challenge then is to understand the set $C^+(I)$ for various I, which we do through a series of lemmas. Due to the greater symmetry it is easier to consider $C(I)$, though, and all lemmas in this section contend with $C(I)$.

For a set I, we will partition the vertices of $V_R \cup V_B$ by the size of $|X_v(I)|$. Consider the cutoffs

$$t_1 = \frac{\sqrt{n \log n}}{\log \log n}, \qquad t_2 = n^{1/4+\varepsilon}, \qquad t_3 = n^{2\varepsilon},$$

and partition $V_R \cup V_B$ by size as

$$\begin{aligned} H_I &= \{v \in V_R \cup V_B : |X_v(I)| > t_1\}, \\ L_I &= \{v \in V_R \cup V_B : t_2 < |X_v(I)| \leq t_1\}, \\ M_I &= \{v \in V_R \cup V_B : t_3 < |X_v(I)| < t_2\}, \\ S_I &= \{v \in V_R \cup V_B : |X_v(I)| < t_3\}. \end{aligned}$$

The next four lemmas address the size of $\binom{X_v(I)}{2}$ over each of the four parts. For each of L_I, M_I, and S_I the total contribution of closed pairs will be $o(k^2)$, and we address these first.

Lemma 3.2. *Conditioned on the event $\mathcal{D}$ defined in Lemma 3.1, for all $I \subseteq V(G)$ with $|I| = k$,*

$$\left| \bigcup_{v \in L_I} \binom{X_v(I)}{2} \right| \leq \varepsilon_1 k^2.$$

assuming n is sufficiently large.

Proof. Suppose $\mathcal{D}$ holds and fix an $I \subseteq V(G)$ with $|I| = k$. Consider any subset $L \subseteq L_I$. By inclusion-exclusion and (5) of $\mathcal{D}$,

$$k = |I| \geq \sum_{v \in L} |X_v(I)| - \sum_{\{v,w\} \in \binom{L}{2}} |X_v(I) \cap X_w(I)| \geq |L| n^{1/4+\varepsilon} - \binom{|L|}{2} 3\sqrt{20} \log^3(n).$$

This would yield a contradiction if $|L| = \lfloor n^{\frac{1-\varepsilon}{4}} \rfloor$. From here, we observe that $|L_I| < n^{\frac{1-\varepsilon}{4}}$, and in turn that $\sum_{v \in L_I} |X_v(I)| \leq (1 + o(1))k$. Thus,

$$\sum_{v \in L_I} \binom{|X_v(I)|}{2} \leq \sum_{v \in L_I} \frac{|X_v(I)| t_1}{2} = o(k^2)$$

since $t_1 = o(k)$, and the result follows for n sufficiently large. $\square$

Let us now look at the sets $X_v(I)$ of the next smaller size – those with $v \in M_I$. The following is a weaker version of a lemma in [8], showing that the contribution of closed pairs from M_I is $o(k^2)$, with high probability.

Lemma 3.3. *Let $\mathcal{M}$ be the event that for all $I \subseteq V(G)$ with $|I| = k$,*

$$\left| \bigcup_{v \in M_I} \binom{X_v(I)}{2} \right| \leq \varepsilon_1 k^2. \tag{9}$$

Then $\mathbb{P}(\mathcal{M}) = 1 - o(1)$.

Proof. We will show that in order for (9) to fail, there would be an unusually dense subgraph of moderate size in G, which is very unlikely to happen.

Fix an $I \subseteq V(G)$ with $|I| = k$. Let

$$Y_I = \sum_{v \in M_I} |X_v(I)|.$$

We prove that with high probability for all I, $Y_I \leq kn^{1/4-\varepsilon}$. From this the lemma immediately follows, as

$$\sum_{v \in M_I} \binom{|X_v(I)|}{2} \leq \sum_{v \in M_I} |X_v(I)| t_2 = o(k^2).$$

Suppose that $Y_I > kn^{1/4-\varepsilon}$. Let $B \subseteq M_I$ be minimal so that $Y = \sum_{v \in B} |X_v(I)| > kn^{1/4-\varepsilon}$. Note that $Y < 2kn^{1/4-\varepsilon}$ (since each $|X_v(I)| \leq n^{1/4+\varepsilon}$) and that $|B| \leq n^{-2\varepsilon} Y \leq 2kn^{1/4-3\varepsilon}$. Let $B_R = B \cap V_R$ and $B_B = B \cap V_B$. Then Y is the number of edges in the union of two bipartite graphs Γ_R and Γ_B on $V(\Gamma_R) = (B_R, \varphi(I))$, $V(\Gamma_B) = (B_B, \varphi(I))$ where

$$E(\Gamma_R) = \{v(w, w') : vw \in E(G_R)\}, \; E(\Gamma_B) = \{v(w, w') : vw' \in E(G_B)\}.$$

Assuming that $\mathcal{D}$ of Lemma 3.1 occurs, every edge $vw \in E(G_R)$ is causing at most $|F(w)| + |F(v)| < 3\log^2(n)$ edges $v(w, w'), w(v, v') \in E(\Gamma_R)$, and similarly for every edge $v'w' \in E(G_B)$ causing edges in $E(\Gamma_B)$. That is, the bipartite subgraphs of G_R on $(B_R, \pi_R(I))$ and of G_B on $(B_B, \pi_R(I))$ must together contain at least $\frac{2kn^{1/4-\varepsilon}}{6\log^2(n)}$ edges in G_R and G_B.

Now, consider a fixed $\tilde{B} = \tilde{B}_R \cup \tilde{B}_B \subseteq (V_R \cup V_B)$ of size $2kn^{1/4-3\varepsilon}$ and $I \subseteq V(G)$ of size k. Suppose $Z \sim Bin(|\tilde{B}|k, p)$. The probability that there are at least $\frac{kn^{1/4-\varepsilon}}{3\log^2(n)}$ edges in the union of the two bipartite subgraphs on $(\tilde{B}_R, \pi_R(I))$ in G_R and on $(\tilde{B}_B, \pi_B(I))$ in G_B is at most (for sufficiently large n)

$$\begin{aligned}
\mathbb{P}\left(Z \geq \frac{kn^{1/4-\varepsilon}}{3\log^2(n)}\right) &\leq \mathbb{P}\left(Z \geq \mathbb{E}(Z) + \frac{kn^{1/4-\varepsilon}}{6\log^2(n)}\right) \\
&\leq \exp\left(-\frac{kn^{1/4-\varepsilon}}{12\log^2(n)}\right) \\
&\leq \exp\left(-n^{3/4-2\varepsilon}\right),
\end{aligned}$$

where here we used that $\frac{kn^{1/4-\varepsilon}}{3\log^2(n)} \geq n^{\varepsilon}\mathbb{E}(Z) \geq 3\mathbb{E}(Z)$, along with Corollary A.3. Now, observing that

$$\binom{n}{k}\binom{2N}{2kn^{1/4-3\varepsilon}} \leq \exp\left(k\log n + 2kn^{1/4-3\varepsilon}\log n\right) = \exp\left(o(n^{3/4-2\varepsilon})\right),$$

and that $B \subseteq \tilde{B}$ for some $\tilde{B}$, a union bound over all possible choices of $\tilde{B}$ and I shows that with high probability none of the potential choices of $(B, \varphi(I))$ yield graphs with as many edges as required above.

Letting $\mathcal{O}$ denote the existence of such an overfull $(B, \varphi(I))$, we conclude that the probability that the high probability conclusion of the lemma is violated is at most $\mathbb{P}(\mathcal{O}^c \cup \mathcal{D}^c) = o(1)$. $\square$

Next, we consider the sets $X_v(I)$ with $v \in S_I$.

Lemma 3.4. *Let $\mathcal{S}$ be the event that for all $I \subseteq V(G)$ with $|I| = k$,*

$$\left| \bigcup_{v \in S_I} \binom{X_v(I)}{2} \right| \le \varepsilon_1 k^2.$$

Then $\mathbb{P}(\mathcal{S}) = 1 - o(1)$.

Proof. Fixing a set $I \subseteq V(G)$ with $|I| = k$ we consider

$$\left| \bigcup_{v \in S_I} \binom{X_v(I)}{2} \right| \le \left| \bigcup_{v \in S_I \cap (\pi_R(I) \cup \pi_B(I))} \binom{X_v(I)}{2} \right| + \left| \bigcup_{v \in S_I \setminus (\pi_R(I) \cup \pi_B(I))} \binom{X_v(I)}{2} \right|.$$

Since $|\pi_R(I) \cup \pi_B(I)| \le 2|I| = 2k$, the first term satisfies

$$\left| \bigcup_{v \in S_I \cap (\pi_R(I) \cup \pi_B(I))} \binom{X_v(I)}{2} \right| \le \binom{n^{2\varepsilon}}{2}(2k) \le n^{4\varepsilon} k \le \frac{\varepsilon_1}{2} k^2. \tag{10}$$

The second term requires some probabilistic analysis. Set

$$Z := \left| \bigcup_{v \in S_I \setminus (\pi_R(I) \cup \pi_B(I))} \binom{X_v(I)}{2} \right|.$$

For a pair $e \in \binom{F(w) \cap \varphi(I)}{2}$ for some $w \in V_R \cup V_B$, $e \in \binom{X_v(I)}{2}$ for any v with $vw \in E(G_R) \cup E(G_B)$. Assuming $\mathcal{D}$, we have $|F(w) \cap \varphi(I)| < 2\log^2(n)$ for all w. As the sets $\binom{F(w)}{2}$ are disjoint, and every vertex in $\varphi(I)$ is in exactly two fibers, there are fewer than $\frac{2k}{2\log^2 n} 2\log^4(n) = 2k\log^2(n)$ such pairs.

For a fixed pair $e \in \binom{\varphi(I)}{2} \setminus \bigcup_w \binom{F(w)}{2}$, the probability that e lies in the union $\bigcup_v \binom{X_v(I)}{2}$ (regardless of whether or not $v \in S_I$) is at most

$$2\big(1 - (1-p^2)^N\big) \;\le\; 2p^2 N \;=\; \frac{2\beta^2}{\log n}.$$

Whence with $\beta = \frac{1}{2}$,

$$\mathbb{E}(Z) \le \frac{2\beta^2}{\log n}\binom{k}{2} + 2k\log^2(n) = \frac{1}{2\log n}\binom{k}{2} + 2k\log^2(n).$$

To apply McDiarmid's inequality we will argue that Z is controlled by a family of independent random variables

$$Y_v := \begin{cases} N(v) \cap (\pi_R(I) \cup \pi_B(I)), & \text{if } v \in S_I, \\ \emptyset, & \text{if } v \notin S_I. \end{cases}$$

indexed by $v \in (V_R \setminus \pi_R(I)) \cup (V_B \setminus \pi_B(I))$. For fixed φ, these variables depend only on $N(v)$, which is determined by the choices of G_R and G_B. They are mutually independent because we are only considering those where $v \notin \pi_R(I) \cup \pi_B(I)$. As the union over such v, Z is a deterministic function of the Y_v and changing the value of a single such Y_v only alters the set $\binom{X_v(I)}{2}$ therefore changing the value of Z by at most

$$\binom{n^{2\varepsilon}}{2} \;\le\; \frac{n^{4\varepsilon}}{2}.$$

As there are $|S_I \setminus \pi_R(I) \cup \pi_B(I)| \le 2N$ random variables in our family, McDiarmid's Inequality (Theorem A.4) yields, for any $t > 0$,

$$\mathbb{P}\left(Z \ge \mathbb{E}(Z) + t\right) \le \exp\left(-\frac{t^2}{N\, n^{8\varepsilon}}\right).$$

Applying this with $t = \frac{1}{\sqrt{\log n}}\binom{k}{2}$ and recalling $k = \kappa\sqrt{n\log n}$, we obtain

$$
\begin{aligned}
\mathbb{P}\left(Z \geq \frac{2}{\sqrt{\log n}}\binom{k}{2}\right) &\leq \mathbb{P}\left(Z \geq \mathbb{E}(Z) + \frac{1}{\sqrt{\log n}}\binom{k}{2}\right) \\
&\leq \exp\left(-\frac{\left(\frac{1}{\sqrt{\log n}}\binom{k}{2}\right)^2}{N\, n^{8\varepsilon}}\right) \\
&= \exp\left(-\Omega\left(n^{1-8\varepsilon}\log^3(n)\right)\right) \\
&= o\left(\binom{n}{k}^{-1}\right).
\end{aligned}
$$

Therefore, with probability at least $1 - o\left(\binom{n}{k}^{-1}\right)$,

$$Z \leq \frac{2}{\sqrt{\log n}}\binom{k}{2} \leq \frac{\varepsilon_1}{2}k^2$$

for all sufficiently large n. Combining this with (10), we find that, for a fixed I, the whole quantity

$$\left|\bigcup_{v\in S_I}\binom{X_v(I)}{2}\right| \leq \varepsilon_1 k^2$$

with probability $1-o\left(\binom{n}{k}^{-1}\right)$. A union bound over the $\binom{n}{k}$ choices of I shows that the desired bound holds simultaneously for all I. ☐

Finally, we turn to the largest of the $X_v(I)$ – those with $v \in H_I$. Here is where the biggest potential contributions to the closed pairs come from.

Lemma 3.5. *If $\mathcal{D}$ of Lemma 3.1 occurs, and n is sufficiently large, then all $I \subseteq V(G)$ with $|I| = k$ satisfy*

$$\sum_{v\in H_I\cap V_R}\binom{|\pi_R(X_v(I))|}{2} \leq \varepsilon_1 k^2, \tag{11}$$

$$\sum_{v\in H_I\cap V_B}\binom{|\pi_B(X_v(I))|}{2} \leq \varepsilon_1 k^2, \tag{12}$$

$$\sum_{v\in H_I\cap V_R}\binom{|\pi_B(X_v(I))|}{2} \leq (1+\varepsilon_1)\min\left\{\binom{k-|\pi_R(I)|}{2}, \binom{pn}{2} + \binom{k-|\pi_R(I)|-pn}{2}\right\}, \tag{13}$$

$$\sum_{v\in H_I\cap V_B}\binom{|\pi_R(X_v(I))|}{2} \leq (1+\varepsilon_1)\min\left\{\binom{k-|\pi_B(I)|}{2}, \binom{pn}{2} + \binom{k-|\pi_B(I)|-pn}{2}\right\}. \tag{14}$$

Proof. Assuming $\mathcal{D}$, an analogous proof to that of Lemma 3.2 implies that

$$\sum_{v\in H_I}|X_v(I)| \leq (1+o(1))\left|\bigcup_{v\in H_I}X_v(I)\right|. \tag{15}$$

Indeed, for any $H \subseteq H_I$

$$k \geq \left|\bigcup_{v\in H_I}X_v(I)\right| \geq \sum_{v\in H}|X_v(I)| - \sum_{\{u,v\}\in\binom{H}{2}}|X_v(I)\cap X_u(I)| \geq |H|\frac{\sqrt{n\log n}}{\log\log n} - \binom{|H|}{2}C\log^3(n).$$

This is a contradiction for any $|H| = (1+o(1))k\frac{\log\log n}{\sqrt{n\log n}}$, implying that

$$|H_I| \leq (1+o(1))k\frac{\log\log n}{\sqrt{n\log n}} < 2\log\log n \tag{16}$$

and hence (15).

From this and the fact (from $\mathcal{D}$) that all vertices in V_R and V_B have degree at most $(1+\varepsilon_2)pN \leq (\beta+\varepsilon_2)\sqrt{\frac{n}{\log^3(n)}}$ we have

$$\sum_{v\in H_I\cap V_R} |\pi_R(X_v(I))| \leq |H_I|(1+\varepsilon_2)pN = o(k) \qquad \text{and} \qquad \sum_{v\in H_I\cap V_B} |\pi_B(X_v(I))| = o(k). \tag{17}$$

This already implies (11) and (12). Now, if $u,v \in H_I\cap V_R$, part (6) of $\mathcal{D}$ shows that $\pi_B(X_u(I))$ and $\pi_B(X_v(I))$ have small intersection, from which one gets with (16) that

$$\sum_{v\in H_I\cap V_R} |\pi_B(X_v(I))| = \left|\bigcup_{v\in H_I\cap V_R} \pi_B(X_v(I))\right| + o(k).$$

Order the vertices of $H_I\cap V_R$ as $v_1,\ldots,v_t$, and set $x_{v_i} = |\pi_B(X_{v_i}(I))|$. We may assume that x_{v_1} is the maximum of the x_{v_i}. Note that $x_{v_1} \leq |N_3(v_1)| \leq (1+\varepsilon_2)pn$. By the convexity of the binomial coefficients, we have for sufficiently large n

$$\begin{aligned}
\sum\binom{x_{v_i}}{2} &\leq \binom{\sum x_{v_i}}{2}\\
&\leq \binom{(1+\varepsilon_2)\left|\bigcup \pi_B(X_{v_i}(I))\right|}{2}\\
&\leq (1+3\varepsilon_2)\binom{\left|\bigcup \pi_B(X_{v_i}(I))\right|}{2},
\end{aligned}$$

and for $\sum x_{v_i} > (1+\varepsilon_2)pn$,

$$\begin{aligned}
\sum\binom{x_{v_i}}{2} &\leq \binom{x_{v_1}}{2} + \sum_{i\geq 2}\binom{x_{v_i}}{2}\\
&\leq \binom{(1+\varepsilon_2)pn}{2} + \binom{\sum x_{v_i} - (1+\varepsilon_2)pn}{2}\\
&\leq \binom{(1+\varepsilon_2)pn}{2} + \binom{(1+\varepsilon_2)\left(\left|\bigcup \pi_B(X_{v_i}(I))\right| - pn\right)}{2}\\
&\leq (1+3\varepsilon_2)\left(\binom{pn}{2} + \binom{\left|\bigcup \pi_B(X_{v_i}(I))\right| - pn}{2}\right).
\end{aligned}$$

Finally, to complete the proof of (13), remember that ε_2 is sufficiently small compared to ε_1, and bound

$$\left|\bigcup \pi_B(X_{v_i}(I))\right| \leq \left|\bigcup X_{v_i}(I)\right| \leq k - |\pi_R(I)| + o(k).$$

This last inequality follows from (17), observing that

$$\begin{aligned}
|\pi_R(I)| &\leq \left|\pi_R\left(\bigcup X_{v_i}(I)\right)\right| + \left|\pi_R\left(\varphi(I)\setminus\left(\bigcup X_{v_i}(I)\right)\right)\right|\\
&\leq \left|\pi_R\left(\bigcup X_{v_i}(I)\right)\right| + \left|\varphi(I)\setminus\left(\bigcup X_{v_i}(I)\right)\right|\\
&= o(k) + k - \left|\bigcup X_{v_i}(I)\right|.
\end{aligned}$$

The proof of (14) works analogously, considering $v \in H_I\cap V_B$. $\square$

4 Independent sets in G and the proof of Theorem 1.3

We now turn to the crucial part of Theorem 1.3, showing that vertex sets of size k are not independent. This is first and only time that it is important to remove edges from monochromatic triangles by considering $C^+(I)$. Up until this time, we could have omitted all edges in monochromatic triangles by using $C(I)$ instead.

For what follows, fix a k-set $I \subseteq V(G)$. We will begin by revealing the edges in G_R and G_B incident to vertices not in $\pi_R(I) \cup \pi_B(I)$ and a small selection of edges within $\pi_R(I) \cup \pi_B(I)$. We ensure that most open pairs in $\pi_R(I) \cup \pi_B(I)$ remain to be considered while most closed pairs are already revealed. As all open pairs must be non-edges, this allows us to bound the probability that the set is independent at the end of the process. Consider the following procedure.

1. First, let $F_0 = (V_R \cup V_B) \setminus (\pi_R(I) \cup \pi_B(I))$. For all vertices $v \in F_0$, reveal all neighbors and non-neighbors in $\pi_R(I)$ and $\pi_B(I)$ in G_R and G_B respectively. These, in turn, reveal edges within G_2.

2. Let $F_1 = \{v \in \pi_R(I) \cup \pi_B(I) : |F(v) \cap I| > \log n\}$. For all vertices in F_1, reveal all neighbors and non-neighbors in $\pi_R(I)$ and $\pi_B(I)$ in G_R and G_B respectively. These, in turn, reveal edges within G_2.

3. Let $F_2 = \{v \in \pi_R(I) \cup \pi_B(I) : |N(v) \cap F_1| > \frac{t_2}{\log n}\}$. Note that
$$(\pi_R(I) \cup \pi_B(I)) \cap H_I \subseteq F_2 \subseteq L_I \cup H_I.$$
Indeed, for $v \in \pi_R(I) \cup \pi_B(I)$ with $v \notin F_2$ then
$$|X_v(I)| < (1+\varepsilon_2)(\log^2 n)\frac{t_2}{\log n} + (\log n)(1+\varepsilon_2)pN \ll t_1,$$
so $v \notin H_I$, proving the first inclusion. The second inclusion follows as $|N(v) \cap F_1| > \frac{t_2}{\log n}$ along with the definition of F_1 implies that $|X_v(I)| > t_2$.

   For all vertices in F_2, reveal all neighbors and non-neighbors in $\pi_R(I)$ and $\pi_B(I)$ in G_R and G_B respectively. These, in turn, reveal edges within G_2.

4. Reveal all pairs in $X_v(I)$ for $v \in F = F_0 \cup F_1 \cup F_2$.

Let G_2^I be the subgraph of G_2 containing exactly all the edges revealed so far. To reduce notation, let $\mathcal{R} = \mathcal{D} \cap \mathcal{S} \cap \mathcal{M}$ in the following. Let
$$f(\ell_R, \ell_B) = \binom{\ell_R}{2} + \binom{\ell_B}{2} - \min\left\{\binom{k-\ell_R}{2}, \binom{pn}{2} + \binom{k-\ell_R-pn}{2}\right\} - \min\left\{\binom{k-\ell_B}{2}, \binom{pn}{2} + \binom{k-\ell_B-pn}{2}\right\}.$$

For integers $\ell_R, \ell_B \in [k]$ with $\ell_R \cdot \ell_B \geq k$, let $\mathcal{S}_{I,\ell_R,\ell_B}$ denote the event that $|\pi_R(I)| = \ell_R$ and $|\pi_B(I)| = \ell_B$. By Lemmas 3.2-3.5, if $\mathcal{R} \cap \mathcal{S}_{I,\ell_R,\ell_B}$ holds, then $\pi_R(I) \cup \pi_B(I)$ contains at least $f(\ell_R, \ell_B) - 9\varepsilon_1 k^2$ open pairs. As $|F_1| \leq 2\frac{k}{\log n}$ and $|F_2| \leq |L_I \cup H_I| < n^{\frac{1}{4}}$ by the proof of Lemma 3.2, the procedure reveals at most $k|F_1 \cup F_2| < \varepsilon_1 k^2$ further pairs in $O(I)$. Thus at this point, there are at least $f(\ell_R, \ell_B) - 10\varepsilon_1 k^2$ open pairs unrevealed.

We observe

Lemma 4.1. *Suppose $I \subseteq V(G)$ with $|I| = k$. Let $\mathcal{B}_I$ be the event that I is independent in G. Then, assuming n is sufficiently large,*
$$\mathbb{P}(\mathcal{B}_I \cap \mathcal{R} \mid \mathcal{S}_{I,\ell_R,\ell_B}) \leq (1-p)^{f(\ell_R,\ell_B) - \varepsilon^3 k^2} \leq \exp(-p(f(\ell_R,\ell_B) - \varepsilon^3 k^2)).$$

Proof. Let $E_{I,R}$ and $E_{I,B}$ be the sets of pairs not revealed so far in G_R and G_B, and $E_I = E_{I,R} \cup E_{I,B}$. The key observation regards how I can be independent at the end. Recall that edges in $O^+(I)$ would yield edges in I, and thus $\pi_R(O^+(I))$ and $\pi_B(O^+(I))$ can not contain edges if I is independent. Thus, if I is independent, all edges in E_I must be in $\pi_R(C^+(I)) \cup \pi_B(C^+(I))$.

Let

$$T_R = E_{I,R} \setminus \bigcup_{i=1}^{N} \binom{N^+(r_i)}{2},$$
$$T_B = E_{I,B} \setminus \bigcup_{i=1}^{N} \binom{N^+(b_i)}{2},$$

Let $U_R = T_R \cap E(G_R)$, $U_B = T_B \cap E(G_R)$. These are the edges in E_I which are not closed$^+$ by vertices of the same color, and thus must be closed by vertices of the opposite color. For $0 \le u_R, u_B \le k^2$, denote the event that $|U_R| = u_R$ and $|U_B| = u_B$ by $\mathcal{U}_{I,u_R,u_B}$. We have

$$\mathbb{P}(\mathcal{B}_I \cap \mathcal{R} \mid \mathcal{S}_{I,\ell_R,\ell_B}) = \sum_{u_R}\sum_{u_B} \mathbb{P}(\mathcal{B}_I \cap \mathcal{R} \cap \mathcal{U}_{I,u_R,u_B} \mid \mathcal{S}_{I,\ell_R,\ell_B}). \tag{18}$$

Assume by symmetry that $u_R \ge u_B$. Then

$$\begin{aligned}
\mathbb{P}(\mathcal{B}_I \cap \mathcal{R} \cap \mathcal{U}_{I,u_R,u_B} \mid \mathcal{S}_{I,\ell_R,\ell_B}) &\le \binom{3\varepsilon_1 k^2}{u_R} p^{u_R} \binom{|T_B|}{u_B} p^{u_B} (1-p)^{|T_R|+|T_B|-u_R-u_B} && (19)\\
&\le \binom{3\varepsilon_1 k^2}{u_R} p^{u_R} \binom{k^2}{u_B} p^{u_B} (1-p)^{f(\ell_R,\ell_B)-15\varepsilon_1 k^2} && (20)\\
&\le \left(\frac{3e\varepsilon_1 pk^2}{u_R}\right)^{u_R} \left(\frac{epk^2}{u_B}\right)^{u_B} (1-p)^{f(\ell_R,\ell_B)-15\varepsilon_1 k^2} \\
&\le \left(\frac{3e\sqrt{\varepsilon_1} pk^2}{u_R}\right)^{u_R} \left(\frac{e\sqrt{\varepsilon_1} pk^2}{u_B}\right)^{u_B} (1-p)^{f(\ell_R,\ell_B)-15\varepsilon_1 k^2} && (21)\\
&\le \exp\Big(4\sqrt{\varepsilon_1} pk^2 - p(f(\ell_R,\ell_B) - 10\varepsilon_1 k^2)\Big). && (22)
\end{aligned}$$

To see this, reveal first the pairs in $E_{I,B}$ in lexicographic order, and then the pairs in $E_{I,R}$. It is critically important that whenever we reach a pair in E_I in this order, it is already determined if the pair is in $T_R \cup T_B$. For (19), note that there are $\binom{|T_B|}{u_B}$ choices for U_B, and for each of these choices each of these pairs is an edge of G_B with probability p. For all other pairs in T_B, the pair must be a non-edge. Then, the edges in U_R have to be chosen from the pairs in

$$\bigcup_{b_i \in \pi_B(I)\setminus F} \pi_R\left(\binom{X_{b_i}(I)}{2}\right),$$

a set that is completely determined once we have revealed all of G_B. As all vertices in $\pi_B(I) \setminus F$ are in $L_I \cup M_I \cup S_I$, Lemmas 3.2-3.4 bound the number of such pairs to $3\varepsilon_1 k^2$ if $\mathcal{R}$ holds. Finally, all other pairs in $T_R \cup T_B$ must be non-edges.

For (20), we observe that trivially $|T_B| \le k^2$, and that all open pairs in E_I are non-edges in T_I. Inequality (21) is due to $u_R \ge u_B$, and (22) follows from the fact that $(C/x)^x$ is maximized for $x = C/e$. Therefore by (18),

$$\begin{aligned}
\mathbb{P}(\mathcal{B}_I \cap \mathcal{R} \mid \mathcal{S}_{I,\ell_R,\ell_B}) &\le k^4 \exp\Big(8\sqrt{\varepsilon_1} pk^2 - p(f(\ell_r,\ell_B) - 10\varepsilon_1 k^2)\Big)\\
&= \exp\Big(-p(f(\ell_r,\ell_B) - 8\sqrt{\varepsilon_1}k^2 - 10\varepsilon_1 k^2) + 4\log k\Big),
\end{aligned}$$

which, for sufficiently small ε_1 and large enough k implies the result. □

We next resolve the dependence on $\mathcal{S}_{I,\ell_R,\ell_B}$.

Lemma 4.2. *Suppose $I \subseteq V(G)$ with $|I| = k$. Then*

$$\mathbb{P}(\mathcal{B}_I \cap \mathcal{R}) = o(1)\binom{n}{k}^{-1}.$$

Proof. Note that since φ is chosen uniformly at random, $\varphi(I)$ is a uniformly chosen k-subset of $V_R \times V_B$. For any $\ell_R, \ell_B \in [k]$ with $\ell_R \cdot \ell_B \geq k$, and writing $\ell_R = x_R k$, $\ell_B = x_B k$, we have

$$\begin{aligned}
\mathbb{P}(\mathcal{S}_{I,\ell_R,\ell_B}) &= \mathbb{P}(|\pi_R(I)| = \ell_R \wedge |\pi_B(I)| = \ell_B) \\
&\leq \mathbb{P}(|\pi_R(I)| \leq \ell_R \wedge |\pi_B(I)| \leq \ell_B) \\
&\leq \binom{N}{\ell_R}\binom{N}{\ell_B}\frac{\binom{\ell_R\ell_B}{k}}{\binom{N^2}{k}} \\
&\leq e^k \binom{N}{\ell_R}\binom{N}{\ell_B}\left(\frac{\ell_R}{N}\right)^k\left(\frac{\ell_B}{N}\right)^k \\
&\leq N^{(x_R+x_B-2)k} k^{(-x_R-x_B+2)k} e^{(1+x_R+x_B)k} x_R^{(1-x_R)k} x_B^{(1-x_B)k} \\
&= n^{-\frac{1}{2}(2-x_R-x_B)k+o(k)} \\
&= \exp\left(-\frac{2-x_R-x_B}{2}(1+o(1))k\log n\right).
\end{aligned}$$

Therefore, using that $\binom{n}{k} = n^{\frac{1}{2}k+o(k)}$,

$$\begin{aligned}
\mathbb{P}(\mathcal{B}_I \cap \mathcal{R})\binom{n}{k} &\leq k^2 \max_{\ell_R,\ell_B} \mathbb{P}(\mathcal{B}_I \cap \mathcal{R} \mid \mathcal{S}_{I,\ell_R,\ell_B})\mathbb{P}(\mathcal{S}_{I,\ell_R,\ell_B})\binom{n}{k} \\
&\leq \max_{\ell_R,\ell_B} \exp\left(-\frac{1-x_R-x_B}{2}(1+o(1))k\log n\right)\mathbb{P}(\mathcal{B}_I \cap \mathcal{R} \mid \mathcal{S}_{I,\ell_R,\ell_B}). \qquad (23)
\end{aligned}$$

We consider three cases based on $x_R + x_B$. If $x_R + x_B \leq 1 - \frac{\varepsilon}{2}$ at the maximum, then we use the trivial bound $\mathbb{P}(\mathcal{B}_I \cap \mathcal{R} \mid \mathcal{S}_{I,\ell_R,\ell_B}) \leq 1$ to show that

$$\begin{aligned}
(23) &\leq \exp\left(-\frac{\varepsilon}{4}k\log n + o(k\log n)\right) \\
&= o(1).
\end{aligned}$$

If $x_R + x_B \geq 1 + \frac{\varepsilon}{2}$ at the maximum, then by Lemma 4.1 and the values $\beta = \frac{1}{2}$ and $\kappa = 1 + \varepsilon$,

$$\begin{aligned}
(23) &\leq \exp\left(-\frac{1-x_R-x_B}{2}(1+o(1))k\log n - p\left(\binom{\ell_R}{2} + \binom{\ell_B}{2} - \binom{k-\ell_R}{2} - \binom{k-\ell_B}{2} - \varepsilon^3 k^2\right)\right) \\
&\leq \exp\left(\left(\frac{x_R+x_B-1}{2} - \frac{\beta(-2\kappa + 2\kappa(x_R+x_B) - 2\varepsilon^3\kappa)}{2}\right)k\log n + o(k\log n)\right) \\
&= \exp\left(\left(-\frac{\varepsilon(x_R+x_B-1-\varepsilon^3-\varepsilon^4)}{2}\right)k\log n + o(k\log n)\right) \\
&\leq \exp\left(-\left(\frac{\varepsilon^2(1-2\varepsilon^2-2\varepsilon^3)}{4}\right)k\log n + o(k\log n)\right) \\
&= o(1).
\end{aligned}$$

Finally, if $1 - \frac{\varepsilon}{2} < x_R + x_B < 1 + \frac{\varepsilon}{2}$ at the maximum, then we may assume by symmetry that $x_B \leq x_R$ and thus $x_B \leq \frac{x_R+x_B}{2} < \frac{1}{2} + \frac{\varepsilon}{4}$, which implies that $k - \ell_B > pn$ and

$$\binom{k-\ell_B}{2} > \binom{pn}{2} + \binom{k-\ell_B-pn}{2}.$$

Therefore, by Lemma 4.1 and the values $\beta = \frac{1}{2}$ and $\kappa = 1 + \varepsilon$,

$$
\begin{aligned}
(23) &\leq \exp\left(-\frac{1 - x_R - x_B}{2}(1 + o(1))k \log n - p\left(\binom{\ell_R}{2} + \binom{\ell_B}{2} - \binom{k - \ell_R}{2} - \binom{pn}{2} - \binom{k - \ell_B - pn}{2} - \varepsilon^3 k^2\right)\right) \\
&= \exp\left(\left(-\frac{1 - x_R - x_B}{2} - \frac{\beta}{2\kappa}\left(-2\kappa^2 + 2\kappa^2(x_R + x_B) + 2\kappa\beta - 2x_B\kappa\beta - 2\beta^2 - 2\varepsilon^3\kappa^2\right)\right) k \log n + o(k \log n)\right) \\
&= \exp\left(\frac{1}{8(1 + \varepsilon)}\left(2x_B - 1 - (4\varepsilon^2 + 2\varepsilon)(x_R + x_B - 1) - 2\varepsilon x_R + 8\varepsilon^3(1 + \varepsilon)^2\right) k \log n + o(k \log n)\right) \\
&\leq \exp\left(\frac{1}{8(1 + \varepsilon)}\left(2(1 - \varepsilon)x_B - 1 - (4\varepsilon^2 + 2\varepsilon)(x_R + x_B - 1) + 9\varepsilon^3\right) k \log n + o(k \log n)\right) \\
&\leq \exp\left(\frac{1}{8(1 + \varepsilon)}\left(2(1 - \varepsilon)\frac{2 + \varepsilon}{4} - 1 + 2\varepsilon^3 + \varepsilon^2 + 9\varepsilon^3\right) k \log n + o(k \log n)\right) \\
&= \exp\left(\frac{\varepsilon(-1 + \varepsilon + 22\varepsilon^2)}{16(1 + \varepsilon)} k \log n + o(k \log n)\right) \\
&= o(1).
\end{aligned}
$$

□

We are finally ready to prove Theorem 1.3.

Proof of Theorem 1.3. Now we estimate the probability that G has an independent set of size k. This is

$$
\begin{aligned}
\mathbb{P}\left(\bigcup_{\substack{I \subseteq V(G) \\ |I| = k}} \mathcal{B}_I\right) &= \mathbb{P}\left(\bigcup_{\substack{I \subseteq V(G) \\ |I| = k}} \mathcal{B}_I \cap \mathcal{R}\right) + \mathbb{P}\left(\bigcup_{\substack{I \subseteq V(G) \\ |I| = k}} \mathcal{B}_I \cap \mathcal{R}^c\right) \\
&\leq \mathbb{P}\left(\bigcup_{\substack{I \subseteq V(G) \\ |I| = k}} \mathcal{B}_I \cap \mathcal{R}\right) + \mathbb{P}(\mathcal{R}^c) \\
&\leq \sum_{\substack{I \subseteq V(G) \\ |I| = k}} \mathbb{P}(B_I \cap \mathcal{R}) + o(1) \\
&= o(1)
\end{aligned}
$$

by Lemma 4.2. Thus with probability $1 - o(1)$, G has $\alpha(G) < k = (1 + \varepsilon)\sqrt{n \log n}$. □

5 The hypergraph Ramsey number $R\left(S_4^{(3)}, S_k^{(3)}\right)$

In this section, we show that our method can also be applied in other settings, and we apply it to hypergraphs. Ramsey numbers for hypergraphs are defined analogously to Ramsey numbers for graphs. Define the 3-uniform star hypergraph $S_k^{(3)}$ as the hypergraph on k vertices containing exactly all edges incident to a given central vertex. In particular, the hypergraph on four vertices with three edges is $S_4^{(3)} = K_4^{(3)-}$, the complete hypergraph minus one edge. The link of a vertex v in a 3-uniform hypergraph is the graph induced by the edges containing v.

Therefore, if we want to find a 2-coloring of the complete hypergraph containing no red $S_4^{(3)}$ and no blue $S_k^{(3)}$, we are looking for a (red) hypergraph in which no link of any vertex contains a triangle or an independent set of size $k - 1$. In other words, we are looking for a hypergraph in which every vertex link is a Ramsey graph for $R(3, k - 1)$. This immediately implies that $R\left(S_4^{(3)}, S_k^{(3)}\right) \leq R(3, k - 1) + 1$.

Mubayi suggested to us that with only minor modifications, our proof for the lower bound for $R(3,k)$ also gives a lower bound for $R\left(S_4^{(3)}, S_k^{(3)}\right)$, improving the implied constant in a recent paper by Mubayi and Spanier [18], which proves a conjecture by Conlon, Fox, He, Mubayi, Suk and Verstraëte [9] using a nibble argument. We outline the proof below, leaving the details to the reader.

Theorem 5.1. $\left(\frac{1}{2}-o(1)\right)\frac{k^2}{\log k} \leq R\left(S_4^{(3)}, S_k^{(3)}\right) \leq (1+o(1))\frac{k^2}{\log k}$.

Proof sketch. In a similar way as we have constructed G, we construct a random hypergraph H as follows:

1. Independent $H(N,p)$ random 3-graphs H_R and H_B are generated on vertex sets $V_R = \{r_1, \ldots, r_N\}$ and $V_B = \{b_1, \ldots, b_N\}$, where $p = \beta\sqrt{\frac{\log n}{n}}$. Edges in H_R and H_B are referred to as red and blue edges respectively.

2. The 3-graph $H_1 = H_R * H_B$ on vertex set $V_R \times V_B$ is constructed. This is similar to the co-normal product of H_R and H_B: $(r,b)(r',b')(r'',b'') \in E(H_1)$ if at least one of $rr'r'' \in E(H_R)$ and $bb'b'' \in E(H_B)$ while $|\{r,r',r'',b,b',b''\}| = 6$ (this last condition differs from the co-normal product, but it simplifies our ensuing analysis). If $rr'r'' \in E(H_R)$, the resulting edge is colored red. If $bb'b'' \in E(H_B)$, the resulting edge is colored blue. Edges may, potentially, be colored both red and blue.

3. Let $V(H) = \{v_1, \ldots, v_n\}$. Choose an injection $\varphi : V(H) \rightarrowtail V_R \times V_B$ uniformly at random. The graph induced by H_1 on $\varphi(V(H))$ is denoted H_2, which has vertex set $V(H_2) = \varphi(V(H))$.

4. Create the final $S_4^{(3)}$-free graph H as follows:

   - First, red edges in H_2 are inspected in lexicographical order (i.e., edges are ordered first by their lowest red coordinate, then by their middle coordinate, and then by the highest red coordinate.) If the pre-image (under φ) of a considered edge does not create a monochromatic red $S_4^{(3)}$ with those edges in H previously added, it is added to H. Otherwise, it is not.
   - Next, the same is done with blue edges.
   - Then, from any non-monochromatic $S_4^{(3)}$ with two blue edges that may remain in G, the red edge is removed.
   - Finally, from any non-monochromatic $S_4^{(3)}$ with two red edges that may remain in G, the blue edge is removed.

This procedure results in an $S_4^{(3)}$-free 3-graph H isomorphic to a subgraph of H_2 – the first two steps result in no monochromatic $S_4^{(3)}$ being created, while the last two steps destroy the remaining copies of $S_4^{(3)}$.

The link of any vertex $(v,v') \in V(H_2)$ is exactly the same as the graph G_2 above (when we replace n by $n+1$, and delete horizontal and vertical edges, which get deleted with high probability when we construct G anyways). So considering pairs $((v,v'), I)$ instead of I in $V(G)$ yields the exact same computations as in Section 3. The extra factor of n when counting these pairs can always be swallowed up in the error terms.

The only addition is that some edges in the link of (v,v') may be omitted or deleted due to $S_4^{(3)}$ in H_2 not centered at (v,v'), and these are not captured in our analysis of the sets $X_w(I)$ inside the link of (v,v'). For this to happen, we would need three vertices $x,y,z \in V_R$ (or V_B) with edges $vxy, xyz \in E(H_R)$ (or $v'xy, xyz \in E(H_B)$) in which case the pair yz would be closed in the link of (v,v'). For every k-set $\varphi(I) \subseteq V(G_2)$, we expect at most

$$\pi_R(I)^2 N p^2 \leq \frac{k^2}{4}\frac{n}{\log^2(n)}\frac{\log n}{n} = \frac{k^2}{4\log n}$$

such closed pairs yz in $\varphi(I)$ coming from H_R, and the same bound from H_B. A Chernoff bound then shows that with high probability the number of these pairs is $o(k^2)$ for all $((v,v'), I)$.

Before the proof analogous to Lemma 4.1, we first reveal all edges of H_R and H_B not containing v and v', then we proceed as in the graph case with edges in the link not contained in $\pi_R(I) \cup \pi_B(I)$, and then iteratively reveal all edges in closed pairs in the link until the process stops.

The proof of Lemma 4.1 and the remaining computations are the same, concluding the proof sketch. $\square$

6 Conclusion

In any n-vertex triangle free graph G we have that $\alpha(G) \geq \Delta(G) \geq \frac{2e(G)}{n}$. Our construction has $\alpha(G) = (1+o(1))\frac{2e(G)}{n} = (1+o(1))\sqrt{n\log n}$; furthermore these independent sets appear both as the neighborhoods of single vertices, as well as 'mixed' throughout G, resembling those in a random graph of the same density. This is one of the main supporting heuristics for Conjecture 1.1, and we agree that Shearer's [21] upper bound on $R(3,k)$ can likely be improved. Improving Shearer's Theorem is related to the well-studied question of whether there is an efficient algorithm for finding large independent sets in the random graph $G(n,p)$. For most given densities, known efficient algorithms can find only independent sets of about half the size of the independence number.

Shearer's algorithmic lower bound on the independence number is $\alpha(G) \geq (1+o(1))n\frac{\log d}{d}$, where d is the average degree of a triangle-free graph. This is half of our upper bound when using the average degree in our construction, matching the gap that persists for random graphs. For the best general lower bound on the independence number, one then uses $\alpha(G) \geq d$ and optimizes the average degree.

This gives further support to Conjecture 1.1: if one believes that maximum Ramsey graphs have certain pseudo-random properties, it is unsurprising that a lower bound implied by a bound from a greedy algorithm based on the average degree would be smaller than the actual independence number by a factor of $\sqrt{\frac{1}{2}}$. A conjecture by Davies, Jenssen, Perkins, and Roberts [10] relating maximum and average size of independence sets in triangle-free graphs would also imply the upper bound in Conjecture 1.1.

7 Acknowledgments

An early version of the construction of G was first considered by the fourth author after listening to Marcus Michelen's beautiful talk about their new lower bound at the "Rocky Mountain Summer Workshop 2025" at Colorado State University, supported by NSF grant CCF-2309707, followed by stimulating discussions over the next days.

The authors then started seriously working on this project at the "Graduate Workshop in Combinatorics" which was held a couple weeks later at Iowa State University in Ames, Iowa, and was supported by NSF grant DMS-2152490 and the Combinatorics Foundation.

We would like to thank David Conlon, Dhruv Mubayi, Will Perkins and Yuval Wigderson for helpful comments.

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi. A note on Ramsey numbers. *J. Combin. Theory Ser. A*, 29(3):354–360, 1980. `doi:10.1016/0097-3165(80)90030-8`.

[2] M. Ajtai, J. Komlós, and E. Szemerédi. A dense infinite Sidon sequence. *European J. Combin.*, 2(1):1–11, 1981. `doi:10.1016/S0195-6698(81)80014-5`.

[3] N. Alon and V. Rödl. Sharp bounds for some multicolor Ramsey numbers. *Combinatorica*, 25(2):125–141, 2005. `doi:10.1007/s00493-005-0011-9`.

[4] V. Angeltveit and B. D. McKay. $R(5,5) \leq 46$, 2025. URL: `https://arxiv.org/abs/2409.15709`, `arXiv:2409.15709`.

[5] T. Bohman and P. Keevash. Dynamic concentration of the triangle-free process. *Random Structures Algorithms*, 58(2):221–293, 2021. `doi:10.1002/rsa.20973`.

[6] M. Campos, S. Griffiths, R. Morris, and J. Sahasrabudhe. An exponential improvement for diagonal Ramsey. *Annals of Mathematics*, to appear, 2025. URL: `https://arxiv.org/abs/2303.09521`, `arXiv:2303.09521`.

[7] M. Campos, M. Jenssen, M. Michelen, F. Pfender, and J. Sahasrabudhe. A polynomial improvement for the odd cycle-complete ramsey numbers, 2025. URL: https://arxiv.org/abs/2511.10641, arXiv:2511.10641.

[8] M. Campos, M. Jenssen, M. Michelen, and J. Sahasrabudhe. A new lower bound for the Ramsey numbers $R(3, k)$, 2025. URL: https://arxiv.org/abs/2505.13371, arXiv:2505.13371.

[9] D. Conlon, J. Fox, X. He, D. Mubayi, A. Suk, and J. Verstraëte. Hypergraph Ramsey numbers of cliques versus stars. *Random Structures Algorithms*, 63(3):610–623, 2023. doi:10.1002/rsa.21155.

[10] E. Davies, M. Jenssen, W. Perkins, and B. Roberts. On the average size of independent sets in triangle-free graphs. *Proc. Amer. Math. Soc.*, 146(1):111–124, 2018. doi:10.1090/proc/13728.

[11] P. Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.*, 53:292–294, 1947. doi:10.1090/S0002-9904-1947-08785-1.

[12] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935. URL: http://www.numdam.org/item?id=CM_1935__2__463_0.

[13] G. Fiz Pontiveros, S. Griffiths, and R. Morris. The triangle-free process and the Ramsey number $R(3, k)$. *Mem. Amer. Math. Soc.*, 263(1274):v+125, 2020. doi:10.1090/memo/1274.

[14] A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, Cambridge, 2016. doi:10.1017/CBO9781316339831.

[15] P. Gupta, N. Ndiaye, S. Norin, and L. Wei. Optimizing the CGMS upper bound on Ramsey numbers, 2024. URL: https://arxiv.org/abs/2407.19026, arXiv:2407.19026.

[16] 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. doi:10.1002/rsa.3240070302.

[17] S. Mattheus and J. Verstraëte. The asymptotics of $r(4, t)$. *Ann. of Math. (2)*, 199(2):919–941, 2024. doi:10.4007/annals.2024.199.2.8.

[18] D. Mubayi and N. Spanier. K_4^--free triple systems without large stars in the complement, 2025. URL: https://arxiv.org/abs/2504.06076, arXiv:2504.06076.

[19] D. Mubayi and J. Verstraëte. A note on pseudorandom Ramsey graphs. *J. Eur. Math. Soc. (JEMS)*, 26(1):153–161, 2024. doi:10.4171/jems/1359.

[20] F. P. Ramsey. On a Problem of Formal Logic. *Proc. London Math. Soc. (2)*, 30(4):264–286, 1929. doi:10.1112/plms/s2-30.1.264.

[21] J. B. Shearer. A note on the independence number of triangle-free graphs. *Discrete Math.*, 46(1):83–87, 1983. doi:10.1016/0012-365X(83)90273-X.

[22] A. Soifer, editor. *Ramsey theory*, volume 285 of *Progress in Mathematics*. Birkhäuser/Springer, New York, 2011. Yesterday, today, and tomorrow, Papers from the workshop held at Rutgers University, Piscataway, NJ, May 27–29, 2009. doi:10.1007/978-0-8176-8092-3.

[23] J. Spencer. *Eighty Years of Ramsey $R(3, k)$... and Counting!*, pages 27–39. Birkhäuser Boston, Boston, MA, 2011. doi:10.1007/978-0-8176-8092-3_2.

A Inequalities

Here we briefly record a few standard inequalities we use, in the notation of Frieze and Karoński [14].

Lemma A.1.

1. $1 + x \le e^x, \ \forall x.$

2. $\left(\frac{n}{k}\right)^k \le \binom{n}{k} \le \left(\frac{en}{k}\right)^k, \ \forall n \ge k.$

Theorem A.2 (Chernoff/Hoeffding Inequality)**.**
Suppose that $S_n = X_1 + X_2 + \ldots + X_n$ where $0 \le X_i \le 1$ are independent random variables, $\mu_i = \mathbb{E}(X_i)$, $\mu = \sum \mu_i$. Then for $t \ge 0$,

$$\mathbb{P}\left(S_n \ge \mu + t\right) \le \exp\left(-\frac{t^2}{2(\mu + t/3)}\right),$$

and for $0 \le t \le \mu$

$$\mathbb{P}\left(S_n \le \mu - t\right) \le \exp\left(-\frac{t^2}{2(\mu - t/3)}\right).$$

Corollary A.3. *If $t \ge \frac{3}{2}\mu$, then*

$$\mathbb{P}(S_n \ge \mu + t) \le \exp\left(-\frac{t}{2}\right).$$

Theorem A.4 (McDiarmid's Inequality)**.**
Suppose that $Z = Z(W_1, W_2, \ldots, W_n)$ is a random variable that depends on n independent random variables $W_1, W_2, \ldots, W_n$. Further, suppose that

$$|Z(W_1, W_2, \ldots, W_i, \ldots, W_n) - Z(W_1, W_2, \ldots, W_i', \ldots, W_n)| \le c_i$$

for all $i = 1, 2, \ldots, n$ and $W_1, W_2, \ldots, W_n, W_i'$. Then for all $t > 0$ we have

$$\mathbb{P}(Z \ge \mathbb{E}(Z) + t) \le \exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right), \text{ and}$$
$$\mathbb{P}(Z \le \mathbb{E}(Z) - t) \le \exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right).$$