# RAMSEY GAMES AGAINST A ONE-ARMED BANDIT

EHUD FRIEDGUT, YOSHIHARU KOHAYAKAWA, VOJTĚCH RÖDL,
ANDRZEJ RUCIŃSKI, AND PRASAD TETALI

ABSTRACT. We study the following one-person game against a random graph: the Player's goal is to 2-colour a random sequence of edges $e_1, e_2, \ldots$ of a complete graph on $n$ vertices, avoiding a monochromatic triangle for as long as possible. The game is over when a monochromatic triangle is created. The online version of the game requires that the Player should colour each edge when it comes, before looking at the next edge.

While it is not hard to prove that the expected length of this game is about $n^{4/3}$, the proof of the upper bound suggests the following relaxation: instead of colouring online, the random graph is generated in only two rounds, and the Player colours the edges of the first round before the edges of the second round are generated. Given the size of the first round, how many edges can there be in the second round if the Player is to win with reasonable probability? In the extreme case, when the first round consists of a random graph with $cn^{3/2}$ edges, where $c$ is a positive constant, we show that the Player can win only if *constantly* many edges are generated in the second round.

The analysis of the two-round version of the game is based on a delicate lemma concerning edge-coloured random graphs.

## 1. INTRODUCTION

In this paper, we study *Ramsey one-person triangle avoidance games against a random graph*. In all versions of our game, the goal of the player, called throughout *the Painter*, is to colour randomly generated edges, using a given number of colours, without creating a monochromatic triangle. Informally, the *length* of the game is the maximum number of edges on which the Painter can win with large probability. We distinguish three versions of the game: *online*, *two-round*, and *offline*.

*Date*: Copy produced on June 20, 2002.

1991 *Mathematics Subject Classification.* 05C80, 91A43, Secondary: 05C55, 91A60.

*Key words and phrases.* Random graphs, Ramsey games, thesholds.

The second author was partially supported by MCT/CNPq through ProNEx Programme (Proc. CNPq 664107/1997–4) and by CNPq (Proc. 300334/93–1 and 468516/2000–0). The third author was partially supported by NSF Grant 0071261. The collaboration between the second and third authors is supported by a CNPq/NSF cooperative grant. The last author was partially supported by the NSF grant DMS-0100298.

1.1. **The three games.** In the *online* version of our game, the Painter receives a random sequence of edges, and she has to colour each edge as it comes, avoiding monochromatic triangles as long as possible.

It turns out that, when the number of colours available to the Painter is $r = 2$, this game is quite easy to analyse; indeed, a fairly simple 'local' argument shows that this game typically lasts for $\Theta(n^{4/3})$ steps.

We observe that the length of the game is not so sharply determined. When the Painter has seen only $o(n^{4/3})$ edges, she is essentially certain to be safe; she is most unlikely to survive up to $\omega n^{4/3}$ edges, for any $\omega$ with $\omega = \omega(n) \to \infty$ as $n \to \infty$; but, for any $c > 0$, she has probability bounded away from 0 and 1 of surviving up to $cn^{4/3}$ edges.

What happens if the Painter has a *mercy period*? That is, suppose she is allowed *not* to colour the first, say, $N_0$ edges online, but may wait to see *all of them* before committing herself to the colour of those edges. Clearly, she should be able to survive longer, because her initial colouring can be cleverer.

Instead of switching back to the online version after the mercy period of $N_0$ edges, let us simply suppose that another $N_1$ random edges are now presented to the Painter. If she is able to extend her initial colouring of the $N_0$ edges to the union of the $N_0$ and the $N_1$ edges still avoiding triangles, she wins; otherwise, she loses. This is the *two-round* version of our game.

We shall analyse the two-round game with $N_0 = cn^{3/2}$, for $c$ a small positive constant. (We shall see in a moment why this is the interesting choice for $N_0$.) The result here is that, with such a value of $N_0$, if the Painter has $r = 2$ colours available, then typically she can win only if $N_1 = O(1)$: for any $N_1 = N_1(n)$ that tends to infinity, the probability that the Painter wins tends to 0.

We shall also show that, if $N_0 = cn^{3/2}$ and $r = 3$ colours are available, then the breakpoint for $N_1$ is of order $n$: if $N_1 = o(n)$, then the Painter typically wins; if $N_1 = \omega n$, where $\omega = \omega(n)$ tends to infinity, then the Painter typically loses. If $N_1 = \Theta(n)$, then she wins with probability bounded away from 0 and 1. Note that, as in the online game, the transition is not sharp.

We finally turn to the *offline* version of our game. In the offline game, an $n$-vertex graph $G(n, N)$ with a fixed number of edges $N = N(n)$ is generated uniformly at random and is presented to the Painter; her task is to colour the edges of $G(n, N)$ avoiding monochromatic triangles.

The reader familiar with the theory of random graphs will immediately see that the analysis of this game amounts exactly to determining the threshold for the Ramsey property $\mathcal{R}_r$, which consists of all graphs such that every $r$-colouring of their edges results in a monochromatic triangle.

The threshold for the property $\mathcal{R}_r$ has been investigated for some time now. In particular, it has been proved that there exist constants $c_r$ and $C_r$

such that

$$\lim_{n\to\infty} \mathbb{P}[G(n,N) \in \mathcal{R}_r] = \begin{cases} 1 & \text{if } N > C_r n^{3/2} \\ 0 & \text{if } N < c_r n^{3/2}. \end{cases} \tag{1}$$

The above threshold of $n^{3/2}$ was first proved for $r = 2$ in [19] (see also [8]), and then validated for arbitrary $r \geq 2$ in [20]. It is interesting to note that the dependence on $r$ is not present in the exponent of $n$, but only in the constants.

Returning to the case $r = 2$, the proof in [19] leaves a wide gap between $c_2$ and $C_2$. Using a criterion from [10] for the existence of sharp thresholds, that gap was recently annihilated in [9]. In the usual parlance, the Ramsey property $\mathcal{R}_2$ admits a *sharp threshold*. We shall come back to this in Section 2. Note that the existence of a sharp threshold for $\mathcal{R}_2$ means that the offline game for $r = 2$ colours has a sharply determined value, unlike the online and the two-round games.

The result in (1) makes the two-round game with $N_0$ of order $n^{3/2}$ quite interesting. Note that the game would be trivial if $N_0 > C_r n^{3/2}$, for in that case the chances of the Painter surviving even the first round are rather slim. On the other hand, if $N_0 = cn^{3/2}$ with $c < c_r$, then the Painter may be happy, because she has a very good chance of being able to colour the edges of the first round. However, our result (Theorem 6 below) shows that, at least for $r = 2$, her happiness is short-lived: *regardless of how small $c$ is*, the fact that she has to colour those $N_0$ edges *before seeing the final graph* makes it very unlikely that she will succeed in colouring even as few as, say, $N_1 = \log\log\log n$ further edges! This is in contrast to the fact that, if $N_0 + N_1 < c_r n^{3/2}$, then, typically, $G(n, N_0 + N_1)$ may be 2-edge-coloured avoiding monochromatic triangles (see (1)).

We mention in passing that, in Section 2, we shall attempt to explain how the two-round version of our game naturally arose from a technical part of the proof of the existence of a sharp threshold for $\mathcal{R}_2$.

Interestingly enough, both the two-round and online versions of our Ramsey game are much more sensitive to the number of colours available to the Painter than the offline one is. Recall that, for example, in the two-round game, while for $r = 2$ colours the survival time is $O(1)$, for $r = 3$ colours the survival time is of order $n$. For the online, 3-colour game, we have a lower bound of roughly $n^{7/5}$ for its length, which is much larger than for two colours. At this point, we do not have a matching upper bound for the 3-colour game.

1.2. **Organization.** This paper is organized as follows. In Section 2 we briefly discuss the proof of the existence of a sharp threshold for the property $\mathcal{R}_2$, and its relation to this paper. In Section 3, we state and prove our results on Ramsey games. As it turns out, most of the work will go into proving a certain delicate lemma (Lemma 4); Section 4 will be entirely

devoted to proving this lemma. In Section 5, we briefly discuss some open problems.

1.3. **Notation and terminology.** Our notation is fairly standard. We let $[n] = \{1, \ldots, n\}$. We write $[X]^k$ for the family of $k$-element subsets of $X$. Let $G$ be a graph. If $x$ is a vertex of $G$, then $N_G(x) = N(G; x)$ denotes the neighbourhood of $x$ in $G$. If $U \subseteq V(G)$, we write $G[U]$ for the subgraph of $G$ induced by $U$. Thus, $E(G[U]) = E(G) \cap [U]^2$. We put $e_G(U)$ for $|E(G[U])|$, the number of edges induced by $U$ in $G$. If $U$, $W \subseteq V(G)$ are disjoint, then $G[U, W]$ stands for the bipartite graph with vertex classes $U$ and $W$ naturally induced by $U$ and $W$ in $G$. We write $E_G(U, W)$ for the set of edges of $G$ with one endvertex in $U$ and the other in $W$. That is $E_G(U, W) = E(G[U, W])$. We let $e_G(U, W) = |E_G(U, W)|$. We sometimes identify a graph with its edge set.

A colouring of the edges of a graph will be called *proper* if it does not contain a monochromatic triangle.

We shall consider the three standard models for random graphs: the binomial random graphs $G(n, p)$, the uniform random graphs $G(n, N)$, and the random graph processes $G_0 \subseteq \cdots \subseteq G_{\binom{n}{2}}$, which naturally correspond to random permutations $e_1, \ldots, e_{\binom{n}{2}}$ of the edges of the complete graph on $[n]$. In particular, if we stop a random graph process at time $t$, we obtain a random graph in the uniform model $G(n, t)$. When convenient, we switch between these models without discussion. A rule of thumb is that most results can be "translated" back and forth between the binomial and uniform models, provided $N \sim p\binom{n}{2}$. For more on the (asymptotic) equivalence of these models and on random graphs in general, we refer the reader to [6] and [12].

Most of the time, we have a parameter $n$ and we are interested in the case in which $n \to \infty$. As usual, we use the term 'asymptotically almost surely' (a.a.s.) to mean 'with probability tending to 1 as $n \to \infty$'. Therefore, we often tacitly assume that $n$ is large enough (for instance, for some inequalities to hold). Also, we drop the $\lfloor\ \rfloor$ and $\lceil\ \rceil$ from our formulae when they are not important. This will simplify the exposition considerably.

Our logarithms are natural logarithms.

*Acknowledgements.* We thank Svante Janson for his interest and suggestions. Part of this work was done when some of the authors were at the meeting *Combinatorics, Probability, and Computing*, at the Mathematisches Forschungs Institut Oberwolfach.

## 2. Sharp thresholds

Recall that $\mathcal{R}_2$ consists of all graphs such that every 2-colouring of their edges results in a monochromatic triangle. The main result in [9] is the following theorem, asserting that property $\mathcal{R}_2$ admits a sharp threshold.

**Theorem 1.** *There exists a function $b = b(n)$ such that for all $\varepsilon > 0$ we have*

$$\lim_{n\to\infty} \mathbb{P}[G(n,p) \in \mathcal{R}_2] = \begin{cases} 1 & \text{if } p > (1+\varepsilon)b(n)/\sqrt{n} \\ 0 & \text{if } p < (1-\varepsilon)b(n)/\sqrt{n}. \end{cases}$$

As the proof of Theorem 1 has some implications with respect to the two-round version of our game, we give a very brief outline of it here. (The remainder of this paper does *not* depend on this section.)

The proof of Theorem 1 is based on a consequence of Friedgut's criterion from [10], which says that in order for $\mathcal{R}_2$ to have a sharp threshold it is sufficient that for any choice of the constants $c$, $\alpha$, and $\xi > 0$, and for any balanced graph $M$ with average degree 4 there is a graph property $\mathcal{G}$ such that $G(n,p) \in \mathcal{G}$ a.a.s., where $p = c/\sqrt{n}$, and such that, for all $n$-vertex graphs $G \in \mathcal{G}$ with sufficiently large $n$,

(*) if $\mathbb{P}[G \cup M^* \in \mathcal{R}_2] > \alpha$, then $\mathbb{P}[G \cup G(n,\xi p) \in \mathcal{R}_2] \to 1$ as $n \to \infty$, where $M^*$ is a random copy of $M$ in the complete graph $K_n$.

It can be proved in at least two ways that a graph with average degree 4 cannot be a member of $\mathcal{R}_2$ (see [9, 18]) and hence $M \notin \mathcal{R}_2$. Thus, if $G \notin \mathcal{R}_2$ as well, then the assumption of the above implication indicates a very complex structure of $G$, which in a sense must be close to possessing property $\mathcal{R}_2$. This closeness leads, after $50^+$ pages of proof involving a special, tailor-cut regularity lemma, to the conclusion that, for some $\lambda > 0$, there exists a family $\mathcal{K}$ of subgraphs of $G$ such that

($i$) $|\mathcal{K}| = 2^{o(n^{3/2})}$,
($ii$) for all $K \in \mathcal{K}$, we have $|E(K)| > \lambda c n^{3/2}$,

and, most importantly,

($iii$) for every proper 2-colouring of the edges of $G$, there is a $K \in \mathcal{K}$ that is monochromatic.

Now, all we need is for every $K$ to have a triangle $T = T(K)$ in $G(n,\xi p)$ such that each of the edges of $T$ connects the endpoints of a path of length two in $K$. Then, if $K$ is monochromatic, there is no way to extend the colouring of $G$ to a proper colouring of $G \cup G(n,\xi p)$. To establish the existence of $T = T(K)$ above, we define for a graph $F$ the *base graph* $\mathrm{Base}(F)$ on the vertex set $V(F)$ as follows: a pair $e$ of vertices of $F$ is an edge of $\mathrm{Base}(F)$ if it forms a triangle with two edges of $F$.

Since $G$ is essentially a pseudorandom graph resembling $G(n, c/\sqrt{n})$, Lemma 2 below, proved in [9], which we believe is of independent interest, implies that there are plenty of triangles in the base of each $K \in \mathcal{K}$. Given two real numbers $0 < \lambda < 1$ and $0 < a < 1/6$, we say that a graph $G$ has property $\mathcal{T}(\lambda, a)$ if, for any subgraph $F$ of $G$ with at least $\lambda|E(G)|$ edges, the graph $\mathrm{Base}(F)$ contains at least $a|V(G)|^3$ triangles.

**Lemma 2.** *For any $\lambda > 0$, there exists $a > 0$ such that if $np^2 \geq \lambda$ then, with probability $1 - o(1)$, the random graph $G(n,p)$ has property $\mathcal{T}(\lambda, a)$.*

Finally, a standard application of Janson's inequality [11] yields for any given $K$ that with probability $1 - 2^{-\Theta(n^{3/2})}$ there is a triangle $T(K)$ in $\mathrm{Base}(K) \cap G(n, \xi p)$. This, together with $(i)$ and $(iii)$ above completes the proof of (*).

The convergence

$$\mathbb{P}[G \cup G(n, \xi p) \in \mathcal{R}_2] \to 1$$

may be viewed as a statement about a two-round Ramsey game against a random graph. This explains why Lemma 2 and a related result, Lemma 4 given below, are useful tools in studying such games.

## 3. Ramsey triangle avoidance games

This section contains our main results. In Section 3.1 below, we prove a relatively simple result for the online version of the Ramsey game. In Section 3.2, we state two results on the two-round game, together with their proofs, pending a major technical lemma, Lemma 4, which will be proved in Section 4.

### 3.1. **The online game.**

Consider the random graph process

$$e_1, \ldots, e_{\binom{n}{2}}$$

on $n$ vertices (see Section 1.3), revealing its edges one by one, and the following one-person game related to it. The Painter's task is to 2-colour the edges online, with colours red and blue, say, and not to create a monochromatic triangle for as long as she can. Here *online* means that the Painter has to decide on the colour of $e_i$ before $e_{i+1}$ is generated. The game is over when a monochromatic triangle is created (this does happen if $n \geq 6$). For how long can she stay in the game with probability approaching 1 as $n \to \infty$?

More formally, for each strategy $\pi$ of the Painter, let $X_\pi$ be the length of the game, *i.e.*, $X_\pi$ is the largest index $N$ such that $e_1, \ldots, e_N$ contains no monochromatic triangle. The goal of the game is to choose a strategy that maximizes $\mathbb{E}(X_\pi)$. By a *strategy* we mean a complete set of rules assigning a colour to the next edge.

Since $e_1, \ldots, e_N$ is exactly the random graph $G(n, N)$, clearly, by (1), an upper bound on the expected duration of the game is $C_2 n^{3/2}$. We shall soon show that the threshold for the length of the online game is in fact much smaller.

A crucial tool in our investigation is the base graph $\mathrm{Base}(F)$ of a graph $F$. Recall that a pair $e$ of vertices of $F$ is an edge in $\mathrm{Base}(F)$ if it forms a triangle with two edges of $F$. Now consider an online 2-colouring of the edges $e_1, \ldots, e_{\binom{n}{2}}$. Let $R_i$ and $B_i$ be the sets of the edges $e_j$ $(j \leq i)$ coloured so far with colours red and blue, respectively. The Painter is stuck if and only if

$$e_{i+1} \in \mathrm{Base}(R_i) \cap \mathrm{Base}(B_i).$$

Similarly, the game is bound to be over by time $j$ if for some $i < j$, the intersection of the set of edges $\{e_{i+1}, \ldots, e_j\}$ with $\mathrm{Base}(R_i)$ (or with $\mathrm{Base}(B_i)$) contains a triangle.

Let us begin with a trivial lower bound. As long as $N = o(n^{6/5})$, there is a.a.s. no copy of a diamond in $e_1, \ldots, e_N$, that is, a $K_4$ with an edge omitted, and the Painter never gets stuck until the $N$th edge, since

$$e_{i+1} \notin \mathrm{Base}(R_i) \cap \mathrm{Base}(B_i)$$

for each $i < N$.

But she can do better: as long as $N = o(n^{4/3})$ a.a.s. neither a copy of $K_4$ nor a copy of the pyramid graph (see Figure 1) has emerged, and so $\mathrm{Base}(G(n, N))$ does not contain a triangle. Hence the greedy strategy $\pi_0$ of colouring each edge red, unless such a move would create a red $K_3$, is successful. Indeed, the base of $R_N$, which is a subgraph of $\mathrm{Base}(G(n, N))$, is triangle-free, and no blue triangle will be created. We have just proved that a.a.s. $X_{\pi_0} > n^{4/3}/\omega$ for any $\omega = \omega(n)$ with $\omega \to \infty$.

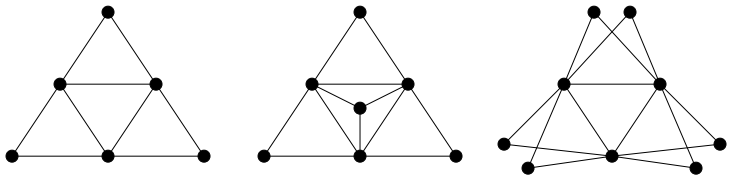

FIGURE 1. The pyramid, the enhanced pyramid, and the double pyramid

Quite surprisingly, $n^{4/3}$ is the right threshold for the length of the random online triangle avoidance game.

**Theorem 3.** *For every $\omega$ with $\omega = \omega(n) \to \infty$ as $n \to \infty$ and for every strategy $\pi$ of the Painter, a.a.s. $X_\pi < \omega n^{4/3}$.*

*Proof.* We relax the online regime by allowing a *mercy period* of $N_0 = n^{4/3}$ steps; *i.e.*, only after that many edges are generated, do they have to be coloured by the Painter. We shall argue that no matter how they are coloured, the Painter loses a.a.s. after another $(\omega - 1)n^{4/3}$ edges. This is nothing else but a two-round exposure technique (see [12]), with a colouring in between.

The aim is to show that a.a.s. every subgraph of $G(n, N_0)$ with at least half of the edges (the majority colour) has at least $\Omega(n^2)$ triangles in its base. Then, by the second moment method, any $\omega n^{4/3}$ additional random edges will a.a.s. contain at least one of these triangles and the game will be over.

A.a.s. the number of copies of $K_{2,3}$ in $G(n, N_0)$, denoted by $X(K_{2,3})$, is smaller than, say, $n \log n$. Let $R$ be the majority colour in $G(n, N_0)$ and let $d_1, \ldots, d_n$ be the degree sequence in $R$. Then the number of triangles in

the base of $R$ is at least

$$\sum_{i=1}^{n} \binom{d_i}{3} - X(K_{2,3}) \geq n\binom{N_0/n}{3} - X(K_{2,3}) = \Omega(n^2),$$

where the last equation holds a.a.s. □

The reader will have no difficulty in verifying that the proofs above show that, for any $c > 0$, the Painter has probability bounded away from 0 and 1 of surviving up to $cn^{4/3}$ edges.

With three colours at hand the Painter can survive quite a bit longer. The painter may use the natural greedy strategy, say $\pi_1$, which colours an edge red if possible, otherwise blue if possible, and finally yellow if none of red and blue will do. Using $\pi_1$, the painter gets stuck only if a copy of the double pyramid, or of the enhanced pyramid (see Figure 1), or of $K_5$ with an edge omitted has emerged. Since such subgraphs a.a.s. do not appear in the random graph with $o(n^{7/5})$ edges, we obtain that $\mathbb{P}(X_{\pi_1} > n^{7/5}/\omega) \to 1$. However, we do not know whether this is the right threshold.

3.2. **The two-round game.** The proof of Theorem 3 bridges the online and the two-round game, which we now define formally. Given two integers $N_0$ and $N_1$, first, $N_0$ edges are generated randomly (let $G_0 = G(n, N_0)$ be the graph defined by these edges), and the Painter colours them *properly* with $r$ colours (if possible), that is, avoiding monochromatic triangles. Then a second round of $N_1$ random edges (graph $G_1$) is thrown in and the Painter wins if her colouring of $G_0$ has an extension to a proper colouring of $G_0 \cup G_1$ (if $G_0$ cannot be properly coloured, the Painter already lost at that stage).

For a given strategy $\pi$ of the Painter, let $P_\pi(r, N_0, N_1)$ be the probability that the Painter wins the game when using $\pi$. Here, by a *strategy*, we mean an assignment of an $r$-edge-colouring to every graph with $N_0$ edges. Thus, the goal of the Painter is to find a strategy $\pi$ that maximizes $P_\pi(r, N_0, N_1)$. Let

$$P^*(r, N_0, N_1) = \max_\pi P_\pi(r, N_0, N_1).$$

The above proof of Theorem 3 gives that $P^*(2, n^{4/3}, \omega n^{4/3}) = o(1)$, because the base of the majority colour in $G_0 = G(n, n^{4/3})$ contains $\Theta(n^2)$ triangles. This simple argument may be extended to a wider range of $N_0$ (for $r = 2$), namely,

$$n^{7/6} \leq N_0 \leq cn^{3/2},$$

where here and below $c$ is always assumed to be so small that

$$\mathbb{P}(G(n, cn^{3/2}) \in \mathcal{R}_2) = o(1).$$

Indeed, along similar lines, one may easily prove that there are $\Theta(N_0^3/n^2)$ triangles in the base graph of the majority colour of $G_0 = G(n, N_0)$, and consequently, arguing as in the proof of Theorem 3, we have

$$P^*(2, N_0, \omega n^{8/3}/N_0) = o(1).$$

In particular, in the most extreme case, we obtain that the number of triangles in the base graph of the majority colour of $G_0 = G(n, cn^{3/2})$ is $\Theta(n^{5/2})$, and thus $P^*(2, cn^{3/2}, \omega n^{7/6}) = o(1)$.

Using Lemma 2 with $\lambda = \min\{c^2, 1/2\}$ and with $F$ consisting of the edges coloured by the majority colour, we can improve the above and conclude that in fact a.a.s., for any 2-edge-colouring of $G(n, cn^{3/2})$, the number of triangles in the base graph of the majority colour is of maximal possible order $\Theta(n^3)$. By the second moment method this yields that

$$P^*(2, cn^{3/2}, \omega n) = o(1)$$

(a triangle is needed in the base of one colour). To improve this any further we need to find a stronger tool.

As we have just seen, Lemma 2, put in the colouring context, says that the base of the majority colour is rich in triangles. A powerful extension of that result, Lemma 4 below, which is the main technical result of this paper, implies that the same is true for the intersection of the bases of two colours, provided the colouring is proper. This result not only settles the two-round game for $r = 2$, but also for $r = 3$, the other interesting case.

**Lemma 4.** *For all $c > 0$, there are constants $\alpha$ and $\beta > 0$ such that with $p = c/\sqrt{n}$ the random graph $G = G(n,p)$ a.a.s. satisfies the following property. For every proper 3-colouring $E(G) = R \cup B \cup Y$ of the edges of $G$, there are two colours, say $R$ and $B$, such that, letting*

$$I = \mathrm{Base}(R) \cap \mathrm{Base}(B), \tag{2}$$

*we have*

*(i)* $|E(I)| \geq \alpha n^2$

*and, in fact,*

*(ii) the graph $I$ contains at least $\beta n^3$ triangles.*

**Remark 5.** *Note that Lemma 2 is not a special case of the seemingly stronger Lemma 4. In Lemma 2 the parameter $\lambda$ may be arbitrarily small, while in Lemma 4 the assumption that the colouring must be proper sets an implicit lower bound on the size of the second largest colour (see (44)). In particular, it seems that Lemma 4 cannot replace Lemma 2 in the proof of Theorem 1 in [9].*

*Nevertheless, the proof of Lemma 4, besides some new ideas, uses the ideas from the proof of Lemma 2. The relation between the proofs of Lemmas 2 and 4 is briefly discussed in Section 4.2.*

Taking Lemma 2 for granted, we are ready to state and prove our results for the two-round game.

**Theorem 6.** *For any $\omega$ with $\omega = \omega(n) \to \infty$ as $n \to \infty$, we have*

$$P^*(2, cn^{3/2}, \omega) = o(1).$$

*Proof.* This follows immediately from Lemma 4($i$) (with $Y = \emptyset$), since a.a.s. at least one of the $\omega$ new edges will fall into $I$. □

Finally, let us consider the two-round game for $r > 2$ and $N_0 = cn^{3/2}$. If $r \geq 4$, then a trivial conservative strategy is to use only two colours for $G_0$ and then another two colours for $G_1$. This, together with the threshold result from [20] settles the game at $\Theta(n^{3/2})$, *i.e.*, there exist constants $b$ and $B > 0$ for which we have

$$\lim_{n\to\infty} P^*(r, cn^{3/2}, N_1) = \begin{cases} 1 & \text{if } N_1 < bn^{3/2} \\ 0 & \text{if } N_1 > Bn^{3/2}. \end{cases}$$

In the remaining case $r = 3$, we have the following, complete solution.

**Theorem 7.** *We have*

$$\lim_{n\to\infty} P^*(3, cn^{3/2}, N_1) = \begin{cases} 1 & \text{if } N_1 \ll n \\ 0 & \text{if } N_1 \gg n. \end{cases}$$

*Proof.* The '1-statement', that is, the fact that $P^*(3, cn^{3/2}, N_1) \to 1$ as $n \to \infty$, follows by using only two colours for $G_0$, and then exclusively the third colour for $G_1$. As long as $N_1 = o(n)$, a.a.s. there will not be any triangles in $G_1$. The 0-statement follows from the second moment method and Lemma 4($ii$). □

We remark that the proof above shows that $P^*(3, cn^{3/2}, c'n)$ is bounded away from 0 and 1 for any $c' > 0$.

## 4. Pairs at distance two in red and blue

Our aim in this section is to prove Lemma 4. On the way we shall come close to proving Lemma 2 (see Section 4.2). However, we refer the reader to [9] for the complete proof of Lemma 2.

Let us introduce a piece of terminology. Let $\varrho$ and $d$ be positive reals. We say that a graph $H$ is $(\varrho, d)$-*dense* if, for all $U \subseteq V(H)$ with $|U| \geq \varrho|V(H)|$, we have

$$e_H(U) \geq d\binom{|U|}{2}.$$

An easy averaging argument shows that, when verifying this property, it suffices to check only all subsets $U$ of cardinality $|U| = \lceil \varrho|V(H)| \rceil$.

We now state a strengthening of Lemma 4, which, together with an easy property of $(\varrho, d)$-dense graphs, implies Lemma 4.

**Lemma 8.** *For all $c > 0$, there is a constant $d > 0$ such that, for any $\varrho > 0$, there is a constant $\sigma_1 > 0$ such that for $p = c/\sqrt{n}$ the random graph $G = G(n, p)$ a.a.s. satisfies the following property. For every proper 3-colouring $E(G) = R \cup B \cup Y$ of the edges of $G$, there are two colours, say $R$ and $B$, for which, letting*

$$I = \mathrm{Base}(R) \cap \mathrm{Base}(B),$$

*there exists a set $U \subseteq V(G)$ such that*

(*i*) $|U| \geq \sigma_1 n$ *and*
(*ii*) $I[U]$ *is* $(\varrho, d)$*-dense.*

Graphs that are $(\varrho, d)$-dense for small $\varrho$ are 'rich in subgraphs'; in particular, the following result holds [21].

**Lemma 9.** *For all* $d > 0$ *there exist* $\varrho$, $n_0$, *and* $c_0 > 0$ *such that every* $(\varrho, d)$*-dense graph on* $n \geq n_0$ *vertices contains at least* $c_0 n^3$ *triangles.*

With Lemmas 8 and 9 at hand, it is easy to prove Lemma 4.

*Proof of Lemma* 4. Given $c > 0$, Lemma 8 gives us a constant $d > 0$. By Lemma 9 we obtain constants $\varrho$, $n_0$, and $c_0 > 0$ such that any $u$-vertex $(\varrho, d)$-dense graph with $u \geq n_0$ contains at least $c_0 u^3$ triangles. We now feed this constant $\varrho > 0$ to Lemma 8, to get a constant $\sigma_1 > 0$.

We claim that the choice of

$$\alpha = \frac{1}{3} d \sigma_1^2 \tag{3}$$

and

$$\beta = c_0 \sigma_1^3, \tag{4}$$

will do in Lemma 4.

To prove our claim, suppose we have a proper 3-colouring of the edges of $G = G(n,p)$. By Lemma 8, a.a.s. there are two colours, say $R$ and $B$, and a set $U \subseteq V(G)$, $|U| \geq \sigma_1 n$, for which $I[U] = (\mathrm{Base}(R) \cap \mathrm{Base}(B))[U]$ is $(\varrho, d)$-dense. Then

$$|E(I)| \geq |E(I[U])| \geq d \binom{|U|}{2} \geq \frac{1}{3} d |U|^2 \geq \left( \frac{1}{3} d \sigma_1^2 \right) n^2 = \alpha n^2.$$

Moreover, by the choice of $\varrho > 0$ and $c_0$, we deduce from Lemma 9 that the number of triangles in the $(\varrho, d)$-dense graph $I[U]$ is at least

$$c_0 |U|^3 \geq c_0 \sigma_1^3 n^3 = \beta n^3.$$

Lemma 4 is proved. □

Let us now briefly comment on the organization of this section. The proof of Lemma 8 fills the remainder of this section, and will require considerable work. This proof is based on some technical, auxiliary results presented in Section 4.1 and proved in Section 4.3. The actual proof of Lemma 8 is given in Section 4.2.

4.1. **Auxiliary lemmas.** In this section, we state three lemmas that will be crucial in the proof of Lemma 8.

Throughout this section, we assume that

$$p = \frac{c}{\sqrt{n}},$$

where $0 < c \leq 1/5$ is an arbitrary constant that we fix once and for all.

Let $H$ be a graph. A pair of disjoint sets $U$, $V \subseteq V(H)$ will be called $(p, \lambda; H, \varepsilon)$-*lower semi-regular*, or $(p, \lambda; H, \varepsilon)$-*semi-regular* for short, if, for all $U' \subseteq U$, $V' \subseteq V$, with $|U'| \geq \varepsilon|U|$ and $|V'| \geq \varepsilon|V|$, we have

$$\frac{e_H(U', V')}{p|U'||V'|} \geq \lambda.$$

Our first auxiliary lemma, Lemma 10, concerns the existence of a certain substructure in a properly edge-coloured $G(n,p)$. A tripartite graph on vertex sets $Z$, $U$, and $W$ with no edge between $Z$ and $W$ will be called an *amalgam* (of two bipartite graphs). The amalgam guaranteed by Lemma 10 in a properly edge-coloured $G(n,p)$ satisfies certain properties with respect to its size, colouring and structure (see Figure 2).

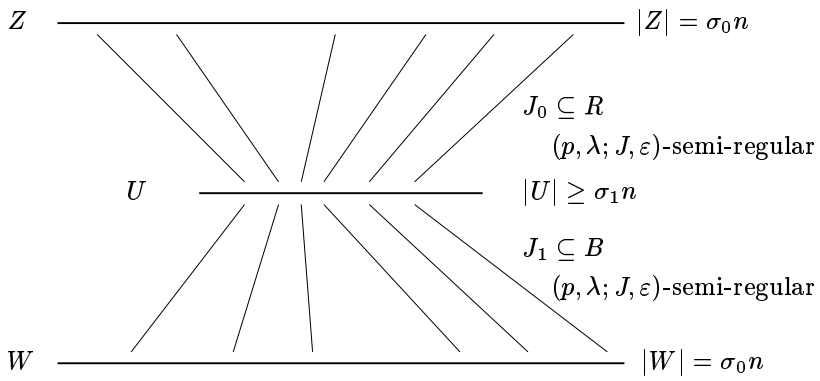


FIGURE 2. The amalgam given by Lemma 10

**Lemma 10.** *Let $\sigma_0 = (1/200)c^2$ and $\lambda = (1/30)c^2$. For any $\varepsilon > 0$, there is a constant $\sigma_1 > 0$ for which the following assertion holds a.a.s. for $G = G(n,p)$. Suppose*

$$E(G) = R \cup B \cup Y \tag{5}$$

*is a proper colouring of the edges of $G$. Then there exist two colours, say $R$ and $B$, and a subgraph $J$ of $G$ as follows:*

- *(i) the vertex set of $J$ is the union of three pairwise disjoint sets $Z$, $U$, and $W$, with $|Z| = |W| = \sigma_0 n$ and $|U| \geq \sigma_1 n$,*
- *(ii) we have $E_J(Z,U) \subseteq R$ and $E_J(U,W) \subseteq B$,*
- *(iii) the pairs $(Z,U)$ and $(U,W)$ are $(p, \lambda; J, \varepsilon)$-semi-regular.*

For convenience, let $J_0 = J[Z,U]$ and $J_1 = J[U,W]$ be the bipartite graphs induced by the pairs $(Z,U)$ and $(U,W)$ in $J$. Basically, our second lemma concerns the graph $J_0$ in Lemma 10.

Recall that, for a graph $F$, we write $\mathrm{Base}(F)$ for the graph whose edges are the pairs of vertices $\{x,y\} \subseteq V(F)$ that are joined by a path of length 2 in $F$. To investigate the 'density' of $\mathrm{Base}(F)$, we shall concentrate on some suitable bipartite subgraphs of $F$. Suppose $F_0$ is a bipartite subgraph of $F$, with vertex classes $U$ and $W$. We shall write $B_U(F_0)$ for the subgraph $\mathrm{Base}(F_0)[U]$

of $\mathrm{Base}(F_0)$ induced by the vertex class $U$. For convenience, we shall often simply write $B_U(F_0)$ instead of $E(B_U(F_0))$. Explicitly, we have

$$B_U(F_0) = \big\{\{u,u'\} \subseteq U\colon \exists w \in W \text{ with } \{u,w\},\ \{u',w\} \in F_0\big\}.$$

Note that, clearly, $B_U(F_0) \subseteq \mathrm{Base}(F)$.

**Lemma 11.** *For all positive constants $\sigma_0$, $\lambda$, and $\varrho$, there exists a constant $\varepsilon > 0$ such that a.a.s. $G = G(n,p)$ satisfies the following property. For all bipartite subgraphs $J_0 = (Z,U;E)$ of $G$ with*

*(i) $|Z| = \sigma_0 n$ and $u = |U| \geq n/\log n$, and such that*
*(ii) $(Z,U)$ is $(p,\lambda;J_0,\varepsilon)$-semi-regular,*

*the graph $\mathrm{Base}(J_0)[U] = B_U(J_0)$ is $(\varrho,d')$-dense, where*

$$d' = \min\left\{\frac{1}{4}\lambda^2 c^2 \sigma_0, \frac{1}{16}\lambda^4\right\}. \tag{6}$$

As an immediate consequence of Lemmas 10 and 11 we conclude that a.a.s. $G = G(n,p)$ contains an amalgam $J$ that satisfies $(i)$–$(iii)$ of Lemma 10 and such that $\mathrm{Base}(J_0)[U] = B_U(J_0)$ is $(\varrho,d')$-dense. Suppose that $I[U] = (\mathrm{Base}(R) \cap \mathrm{Base}(B))[U]$ is *not* $(\varrho,d)$-dense for some $d$ to be chosen later (cf. (24)). Then, because

$$I[U] = (\mathrm{Base}(R) \cap \mathrm{Base}(B))[U] \supseteq (\mathrm{Base}(J_0) \cap \mathrm{Base}(J_1))[U],$$

the graph $(\mathrm{Base}(J_0) \cap \mathrm{Base}(J_1))[U]$ is not $(\varrho,d)$-dense either. This means that there exists a set $U^* \subseteq U$, with $u^* = |U^*| \geq \varrho|U|$, such that

$$|(\mathrm{Base}(J_0) \cap \mathrm{Base}(J_1))[U^*]| < d\binom{u^*}{2}.$$

The subset $U^*$ generates a subamalgam $H$ of $J$, which, as our next result shows, is very rare among all amalgams of that size. Consequently, it will follow that a.a.s. $G(n,p)$ does not contain such an amalgam $H$, contradicting the assumption that $I[U]$ is not $(\varrho,d)$-dense.

Now come the details. Let $Z$, $U^*$, and $W$ be three pairwise disjoint sets of vertices. Suppose we have constants

$$\sigma_0,\ \lambda,\ \tau, \text{ and } \varepsilon' > 0$$

and integers

$$n,\ T \text{ and } T'.$$

Suppose

$$|Z| = |W| = \sigma_0 n \text{ and } |U^*| \geq n/\log n.$$

We now let

$$\mathcal{B}(Z,U^*,W;\sigma_0,\lambda,\tau,\varepsilon';n,T,T') \tag{7}$$

be the family of amalgams $H$ on $Z \cup U^* \cup W$ satisfying the following conditions:

(*i*) $E(H) = E_H(Z, U^*) \cup E_H(U^*, W)$, where

$$|E_H(Z, U^*)| = T' \quad \text{and} \quad |E_H(U^*, W)| = T,$$

and

$$\Delta(H) \le 2c\sqrt{n}, \tag{8}$$

(*ii*) the pairs $(Z, U^*)$ and $(U^*, W)$ are $(p, \lambda; H, \varepsilon')$-semi-regular,

(*iii*) letting $H_0 = H[Z, U^*]$ and $H_1 = H[U^*, W]$, we have

$$|B_{U^*}(H_0)| \ge \tau \binom{|U^*|}{2},$$

and

(*iv*) we have

$$|B_{U^*}(H_0) \cap B_{U^*}(H_1)| < d\binom{|U^*|}{2}, \tag{9}$$

where $d$ is as in (24).

We may now state our third and final auxiliary lemma.

**Lemma 12.** *For all positive constants $\sigma_0$, $\lambda$, $\tau$, and $\alpha \le 1$, there exist constants $\varepsilon' > 0$ and $n_0$ such that, with the notation as above, for all $n \ge n_0$ and all $T$ and $T'$, we have*

$$|\mathcal{B}(Z, U^*, W; \sigma_0, \lambda, \tau, \varepsilon'; n, T, T')| \le \alpha^T \binom{|Z||U^*|}{T'} \binom{|W||U^*|}{T}.$$

We observe that Lemma 12 above may be interpreted as a statement concerning the *uniform* space of amalgams of bipartite graphs on vertex sets $Z$, $U^*$, and $W$ and number of edges $T'$ and $T$. Indeed, Lemma 12 tells us that the probability that a random element from this space should be an element of the family in (7) is at most $\alpha^T$ for any given $\alpha > 0$, if $n$ is sufficiently large and $\varepsilon'$ is sufficiently small.

Hence, one may say that the members $H$ in the family in (7) are *extremely rare*, as long as $T$ is large (which will turn out to be our case). We now use this fact to show that the random graph $G(n, p)$ is quite unlikely to contain such rare graphs $H$.

Given $\sigma_0$, $\lambda$, $\tau$, $\varepsilon' > 0$, and $n$, let us call an amalgam

$$H = (Z, U^*, W; E_H(Z, U^*), E_H(U^*, W))$$

*rare* if $|Z| = |W| = \sigma_0 n$,

$$u^* = |U^*| \ge n/\log n, \tag{10}$$

$$\Delta(H) \le 2c\sqrt{n}, \tag{11}$$

and (*ii*)–(*iv*) from the definition of the family in (7) hold. Note that if $H$ is rare, then, setting $|E_H(Z, U^*)| = T'$ and $|E_H(U^*, W)| = T$, we have

$$T,\ T' \ge \lambda p|Z||U^*| = \lambda p|W||U^*| = \lambda\sigma_0 pn|U^*| = \lambda\sigma_0 pnu^*, \tag{12}$$

and hence, as (10) holds, we have

$$T,\ T' \gg n\log n. \tag{13}$$

On the other hand, we observe that, by (11), we have $T' \le 2c\sqrt{n}|U^*|$, and hence, using (12), we have

$$\frac{T}{T'} \ge \frac{\lambda p|U^*||W|}{2c\sqrt{n}|U^*|} = \frac{\lambda(c/\sqrt{n})\sigma_0 n}{2c\sqrt{n}} = \frac{1}{2}\lambda\sigma_0. \tag{14}$$

in every rare amalgam. Finally, we invoke Lemma 12 with the constants $\sigma_0$, $\lambda$, and $\tau$ above and

$$\alpha = \left(\frac{\lambda}{2e}\right)^{4/\lambda\sigma_0}. \tag{15}$$

Lemma 12 then gives us a constant $\varepsilon' = \varepsilon'(\mathrm{L12}) > 0$.

After having done this preliminary work, we state and prove the following crucial claim.

**Claim 13.** *A.a.s., $G = G(n,p)$ contains no rare amalgam with parameters $\sigma_0$, $\lambda$, $\tau$, and $\varepsilon'$.*

*Proof.* We shall deduce this claim from Lemma 12. We start by setting

$$\vartheta_0 = \alpha^{\lambda\sigma_0/4} \qquad \text{and} \qquad \vartheta_1 = \alpha^{1/2}. \tag{16}$$

Note that (16) gives that

$$\vartheta_1 < \vartheta_0 = \frac{\lambda}{2e}. \tag{17}$$

Since $T \ge (\lambda\sigma_0/2)T'$ (see (14)), we have

$$\alpha^T = \alpha^{T/2}\alpha^{T/2} = \vartheta_0^{(2/\lambda\sigma_0)T}\vartheta_1^T \le \vartheta_0^{T'}\vartheta_1^T. \tag{18}$$

We are now ready to show that a.a.s. no rare amalgam occurs as a subgraph in $G = G(n,p)$. We shall use the first moment method. The expected number of such subgraphs is

$$\sum |\mathcal{B}(Z,U^*,W;\sigma_0,\lambda,\tau,\varepsilon';n,T,T')|p^{T+T'}, \tag{19}$$

where the sum is over all choices of $Z$, $U^*$, $W \subseteq V(G)$ and over all choices of $T$ and $T'$. We use Lemma 12, (12), (13), (17), and (18), to estimate (19). Below, we write $\sum_{u^*}$ for the sum over $n/\log n \le u^* \le n$ and $\sum_{T,T'}$ for the

sum over $T, T' \geq \lambda\sigma_0 pnu^*$ (see (12)). The quantity in (19) is at most

$$\begin{aligned}
&\sum_{u^*}\sum_{T,T'} n^{u^*} n^{2\sigma_0 n}\alpha^T \binom{\sigma_0 nu^*}{T'}\binom{\sigma_0 nu^*}{T} p^{T'+T} \\
&\qquad \leq \sum_{u^*}\sum_{T,T'} n^{3n}\vartheta_0^{T'}\vartheta_1^{T}\left(\frac{e\sigma_0 nu^*}{T'}\right)^{T'}\left(\frac{e\sigma_0 nu^*}{T}\right)^{T} p^{T'+T} \\
&\qquad = \sum_{u^*}\sum_{T,T'} n^{3n}\left(\vartheta_0\frac{e\sigma_0 pnu^*}{T'}\right)^{T'}\left(\vartheta_1\frac{e\sigma_0 pnu^*}{T}\right)^{T} \\
&\qquad \leq \sum_{u^*}\sum_{T,T'} n^{3n}\left(\vartheta_0\frac{e}{\lambda}\right)^{T'}\left(\vartheta_1\frac{e}{\lambda}\right)^{T} \\
&\qquad \leq n^{4n}\sum_{T,T'}\left(\frac{1}{2}\right)^{T+T'} \\
&\qquad \leq n^{4n}(n^2)^2 2^{-\omega n\log n}
\end{aligned}$$

for some $\omega = \omega(n)$ with $\omega \to \infty$ as $n \to \infty$ (see (13)). We conclude that the expectation in (19) is $o(1)$, and Claim 13 follows by Markov's inequality. □

4.2. **Proof of Lemma 8.** We shall use Lemmas 10, 11, and 12 to prove Lemma 8.

Let $c > 0$ be given. A moment's thought shows that the almost sure property asserted for $G = G(n,p)$ is increasing, and therefore we may suppose that $c \leq 1/5$ (for details concerning this argument, for instance see [12, Lemma 1.10]). To define the constant $d$ whose existence is guaranteed in Lemma 8, we proceed as follows. Let

$$\sigma_0 = \frac{1}{200}c^2, \tag{20}$$

$$\lambda = \frac{1}{30}c^2, \tag{21}$$

$$\tag{22}$$

and

$$\tau = \min\left\{\frac{1}{4}\lambda^2 c^2\sigma_0, \frac{1}{16}\lambda^4\right\}. \tag{23}$$

We then set

$$d = \frac{1}{800}\lambda^2 c^2\sigma_0\tau^2, \tag{24}$$

and claim that this choice for $d$ will do.

To prove our claim above, let an arbitrary $0 < \varrho \leq 1$ be given. We have to prove the existence of a suitable constant $\sigma_1 > 0$. To this end, we first invoke Lemma 11 with the constants $\sigma_0$, $\lambda$, and $\varrho$ above, to obtain a

constant $\varepsilon(\mathrm{L11}) > 0$. Moreover, we invoke Lemma 12 with the constants $\sigma_0$, $\lambda$, and $\tau$ above and

$$\alpha = \left(\frac{\lambda}{2e}\right)^{4/\lambda\sigma_0}. \tag{25}$$

Lemma 12 then gives us a constant $\varepsilon'(\mathrm{L12}) > 0$.

We now let

$$\varepsilon = \min\{\varepsilon(\mathrm{L11}), \varrho\varepsilon'(\mathrm{L12})\}. \tag{26}$$

Finally, we invoke Lemma 10 for the value of $\varepsilon$ in (26) to obtain

$$\sigma_1 = \sigma_1(\mathrm{L10}) > 0. \tag{27}$$

We claim that $\sigma_1$ chosen above will do in Lemma 8. In the remainder of this proof, we verify this claim.

For later convenience, let

$$\varepsilon' = \frac{\varepsilon}{\varrho}. \tag{28}$$

Observe that, by the choice of $\varepsilon'$ in (28), Lemma 12 tells us that the family

$$\mathcal{B}(Z, U^*, W; \sigma_0, \lambda, \tau, \varepsilon'; n, T, T') \tag{29}$$

is 'small'. In fact, we have already proved that, with this choice of $\varepsilon' \leq \varepsilon'(\mathrm{L12})$, a.a.s. $G(n,p)$ contains no *rare* amalgam with parameters $\sigma_0$, $\lambda$, $\tau$, and $\varepsilon'$ (see Claim 13).

We now proceed to prove that the value of $\sigma_1$ given in (27) will do in Lemma 8. To that end, suppose that we have a colouring $E(G) = R \cup B \cup Y$ of $G = G(n,p)$ with no monochromatic triangles.

It follows easily from the Chernoff bound that a.a.s.

$$\Delta(G) \leq 2c\sqrt{n}, \tag{30}$$

and hence we shall assume that (30) holds for $G$. By Claim 13, we may also suppose that $G$ contains no rare amalgam with parameters $\sigma_0$, $\lambda$, $\tau$, and $\varepsilon'$. Finally, we may suppose that $G$ satisfies the almost sure properties given in Lemmas 10 and 11. We shall now show that $G$ must then contain the required set $U$.

Since $G$ satisfies the conclusion of Lemma 10, there is an amalgam $J \subseteq G$ as in the statement of that lemma. We follow the notation of Lemma 10, so that, *e.g.*, $V(J) = Z \cup U \cup W$ (see Figure 2 as well).

Clearly, we have $|U| \geq \sigma_1 n$, and hence condition $(i)$ in Lemma 8 holds. Hence, the proof of Lemma 8 will be finished if we prove the following fact.

**Fact 14.** *The graph $I[U]$ is $(\varrho, d)$-dense.*

*Proof.* Our proof strategy has already been unveiled just after the statement of Lemma 11. Here we give all the details..

Let $J_0 = J[Z,U]$ and $J_1 = J[U,W]$. Since $G$ satisfies the conclusion of Lemma 11, the graph $\mathrm{Base}(J_0)[U] = B_U(J_0)$ is $(\varrho, \tau)$-dense.

We now claim that $(\mathrm{Base}(J_0)\cap\mathrm{Base}(J_1))[U] = B_U(J_0)\cap B_U(J_1)$ is $(\varrho,d)$-dense. Note that, then, $I[U]$ must be $(\varrho,d)$-dense, because

$$\begin{aligned} I[U] &= (\mathrm{Base}(R)\cap\mathrm{Base}(B))[U] \\ &\supseteq (\mathrm{Base}(J_0)\cap\mathrm{Base}(J_1))[U] = B_U(J_0)\cap B_U(J_1). \end{aligned}$$

Suppose for a contradiction that $B_U(J_0)\cap B_U(J_1)$ is not $(\varrho,d)$-dense, and let $U^*\subseteq U$ be a witness to this fact, that is, suppose that

$$|U^*| \geq \varrho|U| \tag{31}$$

and

$$|(B_U(J_0)\cap B_U(J_1))\cap[U^*]^2| < d\binom{|U^*|}{2}. \tag{32}$$

We let $H$ be the subgraph of $J$ induced by $Z\cup U^*\cup W$. Also, let $H_0 = H[Z,U^*]$ and $H_1 = H[U^*,W]$, and let $T' = |E(H_0)|$ and $T = |E(H_1)|$.

Since $(Z,U)$ and $(U,W)$ are $(p,\lambda;J,\varepsilon)$-semi-regular, by the choice of $\varepsilon'$ (see (28)) and (31), we have that $(Z,U^*)$ and $(U^*,W)$ are $(p,\lambda;H,\varepsilon')$-semi-regular.

We already know that $B_U(J_0)$ is $(\varrho,\tau)$-dense, so that

$$|B_{U^*}(H_0)| = |B_U(J_0)\cap[U^*]^2| \geq \tau\binom{|U^*|}{2}. \tag{33}$$

However, we may deduce from (32) that

$$|B_{U^*}(H_0)\cap B_{U^*}(H_1)| = |(B_U(J_0)\cap B_U(J_1))\cap[U^*]^2| < d\binom{|U^*|}{2}.$$

The crucial observation now is that the graph $H$ is a *rare amalgam* with parameters $\sigma_0$, $\lambda$, $\tau$, and $\varepsilon'$. Indeed, it only remains to check (10) and (11). Both are clear: we have $|U^*| = \varrho|U| \geq \varrho\sigma_1 n \gg n/\log n$, and hence (10) does hold. Inequality (11) follows trivially, as $\Delta(H)\leq\Delta(G)\leq 2c\sqrt{n}$ (see (30)).

We have reached a contradiction, for we supposed that $G$ contains no rare amalgam with the parameters above. This contradiction completes the proof of Fact 14. □

As observed before, Fact 14 completes the proof of Lemma 8.

*Comments on the proof of Lemma* 2. We observed in the beginning of Section 4 that we would come close to proving Lemma 2. We may now expand on this remark a little (we shall be sketchy).

Suppose $p = c/\sqrt{n}$, where $c > 0$ is an arbitrary constant, and consider $G = G(n,p)$. Furthermore, suppose $F\subseteq E(G)$ has cardinality at least $\lambda|E(G)|$ for some fixed $\lambda > 0$. We wish to show that $\mathrm{Base}(F)$ contains $cn^3$ triangles, for some constant $c > 0$.

Recall that Lemma 10 asserts the existence of a certain substructure in a $G(n,p)$ (an 'amalgam' of two bipartite graphs) whose edges have been properly 3 coloured (see Figure 2). In the case in which we have a subset of edges $F$ of $G = G(n,p)$ as above, one may prove the existence, within $F$,

of the 'upper half' of the amalgam in Figure 2. In fact, using the notation introduced immediately after the statement of Lemma 10, one may prove the existence of the graph $J_0$. (The proof is roughly speaking 'contained' in the proof of Lemma 10.)

With a suitable graph $J_0$ at hand, one may apply Lemma 11, and deduce that Base($F$) contains a $(\varrho, d)$-dense, induced subgraph with at least $\sigma n$ vertices, where $\varrho$, $d$, and $\sigma$ are certain positive constants. It now follows from Lemma 9 that Base($F$) must contain $cn^3$ triangles, for some constant $c$, as required in Lemma 2.

4.3. **Proof of the auxiliary lemmas.** The three auxiliary lemmas stated in Section 4.1 are proved in the first three subsections of this section. However, the hardest of the lemmas to prove, Lemma 12, will be deduced from a new auxiliary lemma (Lemma 23), which is proved in the last subsection of this section.

4.3.1. *Proof of Lemma* 10. We start with two basic lemmas.

**Lemma 15.** *Let $G$ be a graph on $k \geq 1$ vertices and at least $\gamma\binom{k}{2}$ edges, where $0 < \gamma \leq 1$. Then there is a set of vertices $Y \subseteq V(G)$ with*

$$|Y| \geq \frac{1}{2}\gamma k + 1 \tag{34}$$

*and so that the graph $G[Y]$ induced by $Y$ in $G$ has minimum degree*

$$\delta(G[Y]) \geq \frac{1}{2}\gamma k. \tag{35}$$

*In particular, the number of vertices of degree at least $\gamma k/2$ in $G$ is at least $\gamma k/2 + 1$.*

*Proof.* The result holds if $k \leq 2/\gamma$, and hence we suppose $k > 2/\gamma$. Pick a subset $Y \subseteq V(G)$ with $|Y| \geq 2$ that maximizes the ratio $|E(G[Y])|/(|Y|-1)$. Put $\ell = |E(G[Y])|$ and $y = |Y|$. By the choice of $Y$, we have

$$\frac{\ell}{y-1} \geq \frac{\gamma}{k-1}\binom{k}{2} = \frac{1}{2}\gamma k. \tag{36}$$

This implies that $\ell \geq (\gamma/2)k > 1$, and hence $y > 2$. Let $v$ be a vertex in $Y$, and write $d$ for its degree in $G[Y]$. Then, by the choice of $Y$, considering the subgraph $G[Y \setminus \{y\}]$ (recall $|Y| = y > 2$ and hence $|Y \setminus \{y\}| \geq 2$), we see that

$$\frac{\ell}{y-1} \geq \frac{|E(G[Y \setminus \{y\}])|}{y-2} = \frac{\ell - d}{y-2}.$$

This implies that $d \geq \ell/(y-1) \geq \gamma k/2$ (recall (36)). Therefore, (35) holds. Clearly, inequality (35) implies (34). □

**Lemma 16.** *Suppose $k \geq 1$ and*

$$E(K_k) = R \cup B \cup Y \cup S \tag{37}$$

*is a colouring of the edges of the complete graph $K_k$ with colours red, blue, yellow, and sienna, with*

$$|R|,\ |B| \geq \gamma \binom{k}{2}$$

*and*

$$|S| \leq \frac{1}{8}\gamma^2 \binom{k}{2}. \tag{38}$$

*Then there is a vertex $x$ in $K_k$ and distinct colours $C_1$, $C_2 \in \{R, B, Y\}$ such that $x$ is incident to at least $\gamma k/8$ edges of colour $C_1$ and at least $\gamma k/8$ edges of colour $C_2$.*

*Proof.* For a colour $C$ and vertex $x$, let $d(C, x)$ be the number of edges of colour $C$ incident to $x$ in our colouring (37). For $C = R$ and $B$, let

$$U_C = \left\{ u \in V(K_k) \colon d(C, x) \geq \frac{1}{2}\gamma k \right\}. \tag{39}$$

By Lemma 15, we have $|U_R|$, $|U_B| \geq \gamma k/2 + 1 \geq \gamma k/2$. Clearly, we are done if $U_R \cap U_B \neq \emptyset$. Hence we suppose that $U_R \cap U_B = \emptyset$. Consider the $|U_R||U_B|$ edges in $E(U_R, U_B) = E_{K_k}(U_R, U_B)$. There is a colour $C \in \{R, B, Y\}$ for which we have

$$|E(U_R, U_B) \cap C| \geq \frac{1}{3}\left(|U_R||U_B| - |T|\right).$$

Without loss of generality, we may assume that $C \neq R$. By averaging, we may deduce that there is a vertex $x \in U_R$ for which we have

$$\begin{aligned} d(C, x) \geq \frac{|E(U_R, U_B) \cap C|}{|U_R|} &\geq \frac{1}{3}\left(|U_B| - \frac{|T|}{|U_R|}\right) \\ &\geq \frac{1}{3}\left(\frac{1}{2}\gamma k - \frac{(1/16)\gamma^2 k^2}{\gamma k/2}\right) = \frac{1}{8}\gamma k, \end{aligned}$$

and we are again done: $x$ is incident to at least $\gamma k/2 \geq \gamma k/8$ edges of colour $R$ and to at least $\gamma k/8$ edges of colour $C \neq R$. □

We now recall Szemerédi's well-known regularity lemma [23] (see also [12, Section 8.3]). Since we are dealing with sparse graphs (that is, graphs with a subquadratic number of edges), we shall make use of a suitable variant. We need to introduce some definitions.

A graph $H$ is called $(p; b, \beta)$-*bounded* if for all pairs of disjoint subsets $U$, $V \subseteq V(H)$ with $|U|$, $|V| > \beta|V(H)|$ we have $e_H(U, V) \leq bp|U||V|$. A pair $U$, $V \subseteq V(H)$, $U \cap V = \emptyset$, is called $(p; H, \varepsilon)$-*regular* if for all $U' \subseteq U$ and $V' \subseteq V$, with $|U'| \geq \varepsilon|U|$ and $|V'| \geq \varepsilon|V|$, we have

$$\left| \frac{e_H(U', V')}{p|U'||V'|} - \frac{e_H(U, V)}{p|U||V|} \right| \leq \varepsilon.$$

Observe that we have been dealing with a 'one-sided' version of regularity, namely, 'semi-regularity'. Note that, clearly, if a pair is $(p; H, \varepsilon)$-regular, then it is also $(p, \lambda; H, \varepsilon)$-semi-regular for $\lambda = e_H(U, V)/p|U||V| - \varepsilon$.

When dealing with regular pairs, the following simple observation is often useful.

**Lemma 17.** *If $(U,V)$ is $(p,\lambda;H,\varepsilon)$-semi-regular and $W \subseteq V$ satisfies $|W| \geq \varepsilon|V|$, then more than $(1-\varepsilon)|U|$ vertices of $U$ have each at least $\lambda p|W|$ neighbours in $W$.*

Let $|V(H)| = n$. A partition $V(H) = V_1 \cup \cdots \cup V_k$ is $(p;H,\varepsilon,k)$-*regular* if all but at most $\varepsilon\binom{k}{2}$ pairs $(V_i,V_j)$ are $(p;H,\varepsilon)$-regular, and, for all $i = 1,\ldots,k$, $\lfloor n/k \rfloor \leq |V_i| \leq \lceil n/k \rceil$. The following sparse version of Szemerédi's regularity lemma will be important in what follows (for general discussions on this lemma, see, *e.g.*, [12, 15, 16]).

**Lemma 18.** *For all $\varepsilon > 0$, $b \geq 1$, and all integers $m$ and $r \geq 1$, there exist $\beta > 0$ and $K \geq m$ for which the following holds. For all $p$ and all $r$-tuples of $(p;b,\beta)$-bounded graphs $(H_1,\ldots,H_r)$, with all the $H_i$ on the same vertex set $V$ with $|V| \geq m$, there exists a partition $V = V_1 \cup \cdots \cup V_k$ of $V$ with $m \leq k \leq K$ that is $(p;H_i,\varepsilon,k)$-regular for all $1 \leq i \leq r$.*

We are now ready to prove Lemma 10.

*Proof of Lemma* 10. Let $\varepsilon > 0$ be given. Since the smaller $\varepsilon$ is, the stronger is the conclusion in Lemma 10, we may suppose that

$$\varepsilon \leq \frac{1}{30}. \tag{40}$$

Put

$$\delta = \min\left\{\varepsilon^2, \frac{1}{10^2}\lambda, \frac{1}{2\times 10^4}c^4\right\} \quad \text{and} \quad m = \frac{16}{3\lambda^2}. \tag{41}$$

Let

$$\beta = \beta(\delta,m) \quad \text{and} \quad K = K(\delta,m) \tag{42}$$

be the constants given by Lemma 18, for the constants $\varepsilon = \delta$, $b = 2$, $m$, and $r = 3$ as above. We now let

$$\sigma_1 = \frac{1}{K(\delta,m)}, \tag{43}$$

and claim that this choice for $\sigma_1$ will do. We now proceed to prove this claim. Thus, let the graph $G$ and the edge-colouring $E(G) = R \cup B \cup Y$ be as in the statement of Lemma 10.

The number of triangles in $G$ is a.a.s. $(1/6+o(1))(np)^3 = (1/6+o(1))c^3n^{3/2}$. Also, the expected number of edges $e = \{a,b\}$ together with two further vertices $x$ and $y$ $(x \neq y)$ with $x,\, y \in N(G;a) \cap N(G;b)$ (that is, the number of diamonds) is

$$\binom{n}{2}\binom{n-2}{2}p^5 \sim \left(\frac{1}{4}+o(1)\right)c^5n^{3/2}.$$

Standard variance calculations show that, by Chebyshev's inequality, the number of diamonds as above is a.a.s. $(1/4+o(1))c^5n^{3/2}$. Therefore, the

number of triangles in $G$ that are 'solitary', that is, that share no edge with any other triangle, is a.a.s.

$$\frac{1}{6}c^3n^{3/2} - \frac{1}{4}c^5n^{3/2} \geq \frac{1}{7}c^3n^{3/2},$$

where we used that $c \leq 1/5$. We may hence suppose that $G$ contains at least $c^3n^{3/2}/7$ solitary triangles.

By adjusting the notation, we may suppose that

$$|R| \geq |B| \geq |Y|.$$

Since we are supposing that the colouring (5) contains no monochromatic triangle, we may deduce that $|B \cup Y| \geq c^3n^{3/2}/7$; indeed, each solitary triangle must contain either a blue edge or a yellow edge, for otherwise it would be entirely red. In particular, we have

$$|R|,\ |B| \geq \frac{1}{14}c^3n^{3/2}. \tag{44}$$

By Chernoff's inequality, we may also suppose that, for each pair $U$, $W \subseteq V(G)$ with $U \cap W = \emptyset$ and $|U|$, $|W| \geq n/\log n$, we have

$$\frac{1}{2}|U||W|p \leq e_G(U,W) = |E_G(U,W)| \leq 2|U||W|p, \tag{45}$$

and

$$e_G(U) = |E(G[U])| \leq |U|^2 p. \tag{46}$$

Moreover, we may suppose that $G$ and all its subgraphs are $(p;2,\beta)$-bounded.

Let us now apply Lemma 18 to obtain a partition $V(G) = V_1 \cup \cdots \cup V_k$ of $V(G)$ with $m \leq k \leq K$ that is $(p;F,\delta,k)$-regular for $F = R$, $F = B$, and $F = Y$. We now state and prove a claim that will be used in what follows.

**Claim 19.** *Suppose $F \subseteq E(G)$ is a subset of edges of $G = G(n,p)$, with $|F| \geq \lambda cn^{3/2}$. Let $V_1, \ldots, V_k$ be a $(p;F,\delta,k)$-regular partition of $V(G)$, where $m \leq k \leq K$. Suppose for simplicity that $|V_1| = \cdots = |V_k|$. Let $\Phi$ be the set of pairs $\{i,j\}$ $(1 \leq i < j \leq k)$ for which $(V_i,V_j)$ is $(p;F,\delta)$-regular and*

$$e_F(V_i,V_j) \geq \frac{1}{2}\lambda p\left(\frac{n}{k}\right)^2. \tag{47}$$

*Then*

$$|\Phi| \geq \frac{5}{8}\lambda\binom{k}{2}. \tag{48}$$

*Proof.* Let $\ell = x\binom{k}{2} = |\Phi|$ and recall that we are assuming that $|V_i| = n/k$ $(1 \leq i \leq k)$. Using that $G$ satisfies (45) and (46) for all 'large' $U$ and $W \subseteq V(G)$, we deduce that

$$\lambda cn^{3/2} \leq |F| \leq 2\ell\left(\frac{n}{k}\right)^2 p + \left(\binom{k}{2} - \ell\right)\frac{\lambda p}{2}\left(\frac{n}{k}\right)^2 + k\left(\frac{n}{k}\right)^2 p,$$

and so

$$\lambda \le x + \frac{1}{4}(1-x)\lambda + \frac{1}{k}.$$

Solving the last inequality for $x$, we obtain $x \ge 3\lambda/4$, because $1/k \le 1/m = 3\lambda^2/16$. Hence, at least

$$\frac{3}{4}\lambda\binom{k}{2} - \delta\binom{k}{2} \ge \frac{5}{8}\lambda\binom{k}{2}$$

pairs $(V_i, V_j)$ are $(p; F, \delta)$-regular, each satisfying (47); that is, the set $\Phi$ satisfies (48), as required. □

For $F = R$, $B$, and $Y$, let $\Phi_F$ be the set of pairs $\{i, j\}$ $(1 \le i < j \le k)$ such that $(V_i, V_j)$ is $(p; F, \delta)$-regular and

$$e_F(V_i, V_j) \ge \frac{1}{2}\left(\frac{1}{14}c^2\right)p\left(\frac{n}{k}\right)^2 = \frac{1}{28}c^2 p\left(\frac{n}{k}\right)^2. \tag{49}$$

We apply Claim 19 with $F = R$ and $B$ and $\lambda = c^2/14$ (see (44)). That claim tells us that

$$|\Phi_R|,\ |\Phi_B| \ge \frac{1}{25}c^2\binom{k}{2}. \tag{50}$$

Since

$$3 \times \frac{1}{28}c^2 p\left(\frac{n}{k}\right)^2 \le \frac{1}{2}p\left(\frac{n}{k}\right)^2,$$

the fact that (45) holds for all 'large' $U$ and $W \subseteq V(G)$ implies that if a pair $\{i, j\}$ $(1 \le i < j \le k)$ is not a member of $\Phi_R \cup \Phi_B \cup \Phi_Y$, then $(V_i, V_j)$ is *not* $(p; F, \delta)$-regular for some $F \in \{R, B, Y\}$. Put

$$\Phi_S = [k]^2 \setminus (\Phi_R \cup \Phi_B \cup \Phi_Y).$$

The observations above, the fact that $V_1, \ldots, V_k$ is a $(p; F, \delta, k)$-regular partition of $V(G)$ for $F \in \{R, B, Y\}$, and our choice of $\delta$ (see (41)) imply that

$$|\Phi_S| \le 3\delta\binom{k}{2} \le \frac{3}{2 \times 10^4}c^4\binom{k}{2} < \frac{1}{8}\left(\frac{1}{25}c^2\right)^2\binom{k}{2}. \tag{51}$$

We now apply Lemma 16 to the following colouring of the edges of the complete graph $K_k$ on $[k] = \{1, \ldots, k\}$:

$$E(K_k) = \Phi_R \cup \Phi_B \cup \Phi_Y \cup \Phi_S.$$

In view of (50) and (51), we may apply Lemma 16 with $\gamma = c^2/25$. We then obtain $i_0 \in [k] = V(K_k)$ and two disjoint sets $J'_C \subseteq [k]$ ($C = R$ and $B$, say) with

$$|J'_C| \ge \frac{1}{8}\left(\frac{1}{25}c^2\right)k = \frac{1}{200}c^2 k, \tag{52}$$

such that, for all $j \in J'_C$, we have that

$$(V_{i_0}, V_j) \text{ is } (p; C, \delta)\text{-regular} \tag{53}$$

and

$$e_C(V_{i_0}, V_j) \ge \frac{1}{28}c^2 p\left(\frac{n}{k}\right)^2. \tag{54}$$

We let $J_C \subseteq J'_C$ ($C = R$ and $B$) be an arbitrary subset of $J'_C$ with $|J_C| = (c^2/200)k$. Moreover, let

$$U = V_{i_0}, \quad Z = \bigcup_{j \in J_R} V_j, \quad \text{and} \quad W = \bigcup_{j \in J_B} V_j,$$

and let

$$J = R[Z, U] \cup B[U, W]. \tag{55}$$

**Claim 20.** *We claim that $J$, $Z$, $U$, and $W$ satisfy $(i)$, $(ii)$, and $(iii)$ of Lemma 10.*

*Proof.* We start by noticing that

$$|Z| = |W| = \frac{1}{200}c^2 k \times \frac{n}{k} = \frac{1}{200}c^2 n = \sigma_0 n.$$

Moreover,

$$|U| = \frac{n}{k} \geq \frac{n}{K(\delta, m)} = \sigma_1 n.$$

Property $(ii)$ asserted in Lemma 10 is clear from the definition of $J$ (see (55)). We are now left with proving $(iii)$, that is, that $(Z, U)$ and $(U, W)$ are $(p, \lambda; J, \varepsilon)$-semi-regular. We deal with $(Z, U)$, as the other case is analogous.

Let $U' \subseteq U$ and $Z' \subseteq Z$ be such that $|U'| \geq \varepsilon|U|$ and $|Z'| \geq \varepsilon|Z|$. We have to show that

$$\frac{e_J(U', Z')}{p|U'||Z'|} \geq \frac{1}{30}c^2 = \lambda. \tag{56}$$

Put $Z'_j = Z' \cap V_j$ ($j \in J_R$). Clearly, we have $e_J(U', Z') = \sum_{j \in J_R} e_J(U', Z'_j)$. Therefore, we have

$$\begin{aligned} \frac{e_J(U', Z')}{p|U'||Z'|} &= \sum_{j \in J_R} \frac{e_J(U', Z'_j)}{p|U'||Z'_j|} \times \frac{|U'||Z'_j|}{|U'||Z'|} \\ &= \left( \sum_{j \in J'_R} + \sum_{j \in J''_R} \right) \frac{e_J(U', Z'_j)}{p|U'||Z'_j|} \times \frac{|Z'_j|}{|Z'|}, \end{aligned} \tag{57}$$

where $J'_R = \{j \in J_R \colon |Z'_j|/|V_j| < \delta\}$, and $J''_R = J_R \setminus J'_R$. We now claim that

$$\sum_{j \in J''_R} \frac{|Z'_j|}{|Z'|} \geq 1 - \varepsilon. \tag{58}$$

Indeed,

$$\sum_{j \in J'_R} |Z'_j| \leq \delta \sum_{j \in J'_R} |V_j| \leq \delta |Z|.$$

Moreover, $|Z'| \geq \varepsilon|Z|$, and hence

$$\sum_{j \in J'_R} |Z'_j| \leq \delta \left( \frac{1}{\varepsilon}|Z'| \right) \leq \varepsilon |Z'|, \tag{59}$$

as $\delta \leq \varepsilon^2$ (see (41)). It follows from (59) that (58) does indeed hold.

We now go back to (57). Clearly, the right-hand side of (57) is, by (53), (54), and (58), at least

$$\sum_{j\in J_R''} \frac{e_J(U',Z_j')}{p|U'||Z_j'|}\times\frac{|Z_j'|}{|Z'|} \geq \left(\frac{1}{28}c^2-\delta\right)\sum_{j\in J_R''}\frac{|Z_j'|}{|Z'|}$$
$$\geq \left(\frac{1}{28}c^2-\delta\right)(1-\varepsilon) \geq \frac{1}{30}c^2.$$

We have thus concluded that (56) does indeed hold, as required. This concludes the proof that $(Z,U)$ is $(p,\lambda;J,\varepsilon)$-semi-regular, and the proof of Claim 20 is complete. □

Claim 20 completes the proof of Lemma 10. □

4.3.2. *Proof of Lemma* 11. We start by stating the following upper tail estimate from [21]. We give this result in a very general form. Given a finite set $\Gamma$ and $0<p<1$, we denote by $\Gamma_p$ a random binomial subset of $\Gamma$. Note that $([n]^2)_p$ is nothing else but the random graph $G(n,p)$. Let $\mathcal{S}\subseteq[\Gamma]^s$ and set $\mu=\mathbb{E}|\{S\in\mathcal{S}\colon S\subseteq\Gamma_p\}|$.

**Lemma 21.** *For all $q$ and $t>0$, with probability at least*

$$1-\exp\left\{-\frac{qt}{s(\mu+t)}\right\}$$

*there is $E_0\subseteq\Gamma_p$ with $|E_0|=q$ such that $\Gamma_p\setminus E_0$ contains fewer than $\mu+t$ sets from $\mathcal{S}$.*

The above lemma implies the following result.

**Corollary 22.** *For every $c>0$, a.a.s., for each pair $U$, $W\subseteq[n]$ with $U\cap W=\emptyset$, there exists $E_0=E_0(U,W)\subseteq E(G(n,c/\sqrt{n}))$ with $|E_0|=n\log n$ such that the bipartite subgraph of $G(n,c/\sqrt{n})\setminus E_0$ spanned between the sets $U$ and $W$ contains no more than*

$$2\binom{|U|}{2}\binom{|W|}{2}\frac{c^4}{n^2}$$

*copies of the 4-cycle $C_4$.*

*Proof.* Apply Lemma 21 to $\Gamma=U\times W$ and $\mathcal{S}$, the family of the edge sets of all 4-cycles in $\Gamma$, with $s=4$, $p=c/\sqrt{n}$, $q=n\log n$, and

$$t=\mu=\binom{|U|}{2}\binom{|W|}{2}\frac{c^4}{n^2}.$$

We omit the easy calculations. □

For more about upper tail estimates, see [14], and for further developments of the above deletion technique, see [13].

We are now ready to plunge into the proof of Lemma 11.

*Proof of Lemma* 11. Let $\sigma_0$, $\lambda$, and $\varrho$ be given. We claim that $\varepsilon = \min\{1/18, \varrho\}$ will do. Let $U' \subseteq U$ be an arbitrary subset of $U$ with

$$|U'| = \varrho|U| \geq \varepsilon|U|.$$

We shall show that $B_U(J_0)$ induces at least $d'\binom{|U'|}{2}$ edges within $U'$ (see (6)), and this will prove Lemma 11.

We first observe that, by the $(p, \lambda; J_0, \varepsilon)$-semi-regularity of $(Z, U)$, at least $(1-\varepsilon)|Z| = (1-\varepsilon)\sigma_0 n$ vertices in $Z$ send at least $\lambda p|U'| = \lambda p \varrho |U|$ edges into $U'$.

By Corollary 22, we may suppose that we may drop a set of edges $E_0$ from $J_0$, with $|E_0| \leq n \log n$, so as to guarantee that the number of 4-cycles left within $Z \cup U'$ is at most

$$\binom{|Z|}{2}\binom{|U'|}{2}\frac{c^4}{n^2} \leq \frac{1}{2}c^4\sigma_0^2\varrho^2u^2, \tag{60}$$

where $u = |U|$. In what follows, we work with the graph $J_0' = J_0 \setminus E_0$. Note that the number of vertices in $Z$ that are incident to at least $n^{1/3}\log n$ edges in $E_0$ is at most $n^{2/3}$. Therefore, for large enough $n$, at least $(1-3\varepsilon/2)|Z| = (1-3\varepsilon/2)\sigma_0 n$ vertices in $Z$ send at least $(3/4)\lambda p\varrho u$ edges into $U'$ in $J_0'$.

Suppose the $\binom{|U'|}{2}$ pairs within $U'$ are $e_1, \ldots, e_{\binom{|U'|}{2}}$. Let $x_i$ be the number of vertices in $Z$ that send edges to both endvertices of $e_i$ in $J_0'$ for all $1 \leq i \leq \binom{|U'|}{2}$. Clearly, the graph $B_U(J_0') \subseteq B_U(J_0)$ induces as many edges within $U'$ as there are indices $i$ for which $x_i > 0$. Adjust the notation so that $x_i > 0$ for $1 \leq i \leq N$ and $x_i = 0$ for $i > N$. It will suffice to show that

$$N \geq d'\binom{|U'|}{2}. \tag{61}$$

For large enough $n$, we have

$$\begin{aligned}\sum_{1\leq i\leq N} x_i &= \sum_{z\in Z}\binom{\deg_{J_0'}(z)}{2} \\ &\geq \frac{1}{2}(1-2\varepsilon)\sigma_0 n\left(\frac{3}{4}\right)^2\lambda^2p^2\varrho^2u^2 \geq \frac{1}{4}\lambda^2c^2\sigma_0\varrho^2u^2.\end{aligned} \tag{62}$$

Let us now observe that the number of $C_4$ in $J_0'[Z, U']$ is

$$\sum_{1\leq i\leq N}\binom{x_i}{2}.$$

By (60), we have that

$$\sum_{1\leq i\leq N}\binom{x_i}{2} \leq \frac{1}{2}c^4\sigma_0^2\varrho^2u^2. \tag{63}$$

On the other hand, by the Cauchy–Schwarz inequality, we have that

$$\sum_{1\leq i\leq N}\binom{x_i}{2} \geq N\binom{N^{-1}\sum_{1\leq i\leq N}x_i}{2}$$
$$= \frac{1}{2}\left(\sum_{1\leq i\leq N}x_i\right)\left(\frac{1}{N}\sum_{1\leq i\leq N}x_i - 1\right). \tag{64}$$

We now analyse two cases. Suppose first that $\sum_{1\leq i\leq N}x_i \leq 2N$. Then, by (62), we have

$$N \geq \frac{1}{2}\sum_{1\leq i\leq N}x_i \geq \frac{1}{8}\lambda^2c^2\sigma_0\varrho^2u^2 \geq \frac{1}{4}\lambda^2c^2\sigma_0\binom{|U'|}{2}. \tag{65}$$

Suppose now that, on the contrary, we have $\sum_{1\leq i\leq N}x_i > 2N$. Then, by (62) and (64), we have

$$\sum_{1\leq i\leq N}\binom{x_i}{2} \geq \frac{1}{4N}\left(\sum_{1\leq i\leq N}x_i\right)^2 \geq \frac{1}{64N}\lambda^4c^4\sigma_0^2\varrho^4u^4. \tag{66}$$

Comparing (63) and (66), we deduce that

$$\frac{1}{64N}\lambda^4c^4\sigma_0^2\varrho^4u^4 \leq \frac{1}{2}c^4\sigma_0^2\varrho^2u^2,$$

whence

$$N \geq \frac{1}{32}\lambda^4\varrho^2u^2 \geq \frac{1}{16}\lambda^4\binom{|U'|}{2}. \tag{67}$$

Lemma 11 follows from (65) and (67). □

4.3.3. *Proof of Lemma* 12. In order to prove Lemma 12, we shall make use of a further auxiliary lemma, Lemma 23 below. Let us introduce some notation to state this lemma.

Let $U^*$ and $W$ be two disjoint sets of vertices. Suppose $F \subseteq [U^*]^2$ is a collection of pairs of vertices in $U^*$. For convenience, we shall often refer to the graph with vertex set $U^*$ and edge set $F$ as $F$ as well. Suppose now that we have constants

$$\sigma_0,\ \lambda,\ \tau,\ \text{and}\ \varepsilon' > 0$$

and integers

$$n\ \text{and}\ T.$$

Suppose further that

(*i*) $|W| = \sigma_0 n$ and $u^* = |U^*| \geq n/\log n$,
(*ii*) $|F| = \tau\binom{|U^*|}{2} = \tau\binom{u^*}{2}$.

We now let

$$\mathcal{B}(U^*, W; F; \sigma_0, \lambda, \tau, \varepsilon'; n, T) \tag{68}$$

be the family of bipartite graphs $H = (U^*, W; E)$ satisfying the following conditions:

(*i*) $|E| = T$ and $\Delta(H) = \max\{\deg_H(x)\colon x \in V(H)\} \le 2c\sqrt{n}$,
(*ii*) $(U^*, W)$ is $(p, \lambda; H, \varepsilon')$-semi-regular, and
(*iii*) we have

$$|F \cap B_{U^*}(H)| < \frac{1}{800}\lambda^2\sigma_0 c^2 \tau |F|. \tag{69}$$

We are finally ready to state our auxiliary lemma, the formulation of which is based on the notation and definitions introduced above.

**Lemma 23.** *Let any positive constants $\sigma_0$, $\lambda$, $\tau$, and $\alpha \le 1$ be given. Then there exist constants $\varepsilon' > 0$ and $n_0$ such that, with the notation as above, for all $n \ge n_0$ and all $T$, we have*

$$|\mathcal{B}(U^*, W; F; \sigma_0, \lambda, \tau, \varepsilon'; n, T)| \le \alpha^T \binom{|U^*||W|}{T}.$$

As observed immediately after its statement, Lemma 12 may be thought of as an assertion concerning the uniform space of certain amalgams. Lemma 23 is a statement concerning the uniform space of the *bipartite graphs* with vertex classes $U^*$ and $W$. Indeed, this lemma asserts that a random element of this space is extremely unlikely to be member of the family in (68).

The proof of Lemma 23 is postponed to Section 4.3.4. We shall now see that Lemma 12 follows from Lemma 23 almost immediately.

*Proof of Lemma* 12. Let $\sigma_0$, $\lambda$, $\tau$, and $\alpha$ as in the statement of Lemma 12 be given. We feed these constants to Lemma 23, and obtain a constant $\varepsilon' > 0$; we claim that this $\varepsilon'$ will do in Lemma 12. We prove this claim as follows.

For any graph

$$H \in \mathcal{B}_1 = \mathcal{B}(Z, U^*, W; \sigma_0, \lambda, \tau, \varepsilon'; n, T, T'), \tag{70}$$

we let $H_0 = H[Z, U^*]$ and $H_1 = H[U^*, W]$. We may count the graphs as in (70) by estimating the number of $H_0$ that we have, and, for each fixed $H_0$, estimating the number of $H_1$ that we may 'put together with $H_0$' to form a graph $H$ in $\mathcal{B}_1$.

We estimate the number of graphs $H_0$ trivially, by saying that there are at most

$$\binom{|Z||U^*|}{T'} \tag{71}$$

such graphs. Let us now fix a graph $H_0 = (Z, U^*; E_0)$ satisfying the conditions $(i)$–$(iv)$ in the definition of the family $\mathcal{B}_1$. Moreover, let

$$F \subseteq B_{U^*}(H_0)$$

with $|F| = \tau\binom{|U^*|}{2}$ be fixed. We now make the following claim.

**Claim 24.** *If $H_1 = (U^*, W; E_1)$ is so that $H = H_0 \cup H_1$ is a member of $\mathcal{B}_1$, then*

$$H_1 \in \mathcal{B}_0 = \mathcal{B}(U^*, W; F; \sigma_0, \lambda, \tau, \varepsilon'; n, T). \tag{72}$$

*Proof.* We have $|W| = \sigma_0 n$ and $|F| = \tau\binom{|U^*|}{2}$. Conditions $(i)$ and $(ii)$ in the definition of $\mathcal{B}_0$ clearly hold for $H_1$. To check condition $(iii)$, we observe that, in view of (9), we have

$$|F \cap B_{U^*}(H_1)| \leq |B_{U^*}(H_0) \cap B_{U^*}(H_1)| < \frac{1}{800}\lambda^2\sigma_0 c^2\tau^2\binom{|U^*|}{2} = \frac{1}{800}\lambda^2\sigma_0 c^2\tau|F|.$$

Therefore (69) holds and so does condition $(iii)$ in the definition of $\mathcal{B}_0$. We conclude that (72) holds. □

It now suffices to put together the bound in (71), Lemma 23, and Claim 24 to deduce that the number of graphs $H$ as in (70) is at most

$$\alpha^T\binom{|Z||U^*|}{T'}\binom{|W||U^*|}{T},$$

as required. □

4.3.4. *Proof of Lemma* 23. In this somewhat long section, we shall prove Lemma 23, the crucial lemma in the proof of Lemma 12. We warn the reader that this will be a fairly technical proof. The basic underlying idea in this proof comes from a lemma in [17, Lemma 11] (see also [12, Lemma 8.30]).

We start by letting arbitrary positive constants $\sigma_0$, $\lambda$, $\tau$, and $\alpha$ as in the statement of Lemma 23 be given. To prove Lemma 23, we need to define an appropriate constant $\varepsilon' > 0$. Put

$$\alpha_0 = \frac{1}{2}\alpha^{48/\lambda\tau\sigma_0}, \tag{73}$$

$$\varepsilon^{**} = \min\left\{\frac{1}{2}\tau\exp\left\{-\frac{1}{18}\lambda^2c^2\sigma_0\right\}, \frac{1}{4}\alpha_0^2e^{-8c^2/\sigma_0}\right\}, \tag{74}$$

and let

$$\varepsilon' = \frac{1}{2}\tau\varepsilon^{**}. \tag{75}$$

We claim that this choice for $\varepsilon'$ will do in Lemma 23. *Our task for the remainder of this section is to justify this claim.*

For later reference, we observe that the following inequalities hold:

$$\varepsilon^{**} = 2\frac{\varepsilon'}{\tau} \leq \frac{1}{2}, \tag{76}$$

$$\varepsilon^{**} \leq \frac{1}{2}\tau\exp\left\{-\frac{1}{18}\lambda^2c^2\sigma_0\right\}, \tag{77}$$

$$2e^{4c^2/\sigma_0}\sqrt{\varepsilon^{**}} \leq \alpha_0, \tag{78}$$

and

$$(2\alpha_0)^{\lambda\tau\sigma_0/48} \leq \alpha. \tag{79}$$

Recall that we wish to estimate the number of bipartite graphs $H = (U^*, W; E)$ in

$$\mathcal{B}_0 = \mathcal{B}(U^*, W; F; \sigma_0, \lambda, \tau, \varepsilon'; n, T).$$

In order to obtain this estimate, let us make the following observation about the graphs $H \in \mathcal{B}_0$.

**Lemma 25.** *Suppose $H \in \mathcal{B}_0$. Then $H$ contains two edge-disjoint subgraphs $H_0 = (U^*, W; E_0)$ and $H_1 = (U^*, W; E_1)$ such that*

$$H = H_0 \cup H_1, \tag{80}$$

*and so that the pair $(U^*, W)$ is $(p, \lambda/3; H_i, \varepsilon')$-semi-regular for both $i = 0$ and $i = 1$. Moreover, we have $|E(H_0)| = |E(H_1)| = T/2$.*

*Proof.* It suffices to use the 'probabilistic method', taking a random partition. The result follows from standard exponential estimates for the tail of the hypergeometric distribution. We omit the details. □

For each graph $H \in \mathcal{B}_0$, we fix a decomposition as in (80). We shall now invoke Lemma 15, which tells us that, roughly speaking, any dense graph contains a large subgraph with large minimum degree. We explicitly state the result of an application of Lemma 15 to $F$ as a fact.

**Fact 26.** *There exists a set $U^{**} \subseteq U^*$ for which we have*

- *(i) $\delta(F[U^{**}]) \geq \tau u^*/2$, that is, the minimum degree of the graph $F[U^{**}]$ induced by $U^{**}$ in $F$ is at least $\tau u^*/2$,*
- *(ii) $u^{**} = |U^{**}| \geq \tau u^*/2 + 1$.*

*Proof.* This follows immediately from Lemma 15 applied to the graph $F$. □

For the remainder of our proof, we fix $U^{**}$ as in Fact 26. The following observation is clear.

**Fact 27.** *Let $H_i^{**} = H_i[U^{**}, W]$ $(i \in \{0,1\})$ be the subgraph of $H_i$ induced by $U^{**} \cup W$. Recall $\varepsilon^{**} = 2\varepsilon'/\tau$. Then $(U^{**}, W)$ is $(p, \lambda/3; H_i^{**}, \varepsilon^{**})$-semi-regular for both $i = 0$ and $i = 1$.*

*Proof.* It suffices to recall that $|U^{**}| \geq \tau u^*/2$ and that the pair $(U^*, W)$ is $(p, \lambda/3; H_i, \varepsilon')$-semi-regular $(i \in \{0,1\})$. □

Recall that if $J$ is a graph and $x$ is a vertex of $J$, then we write

$$N(J; x) = N_J(x)$$

for the neighbourhood of $x$ in $J$.

**Remark 28.** *We observe that $N(H_i^{**}; v) = N(H_i; v)$ for any $v \in U^{**}$ and $i \in \{0,1\}$. Also, clearly, we have $N(H_i^{**}; w) \subseteq N(H_i; w)$ for any $w \in W$ and $i \in \{0,1\}$.*

We now need to fix some notation. We let

$$U^{**} = \{v_i \colon 1 \leq i \leq u^{**} = |U^{**}|\}. \tag{81}$$

Moreover, we let

$$F_i = N(F; v_i) \cap U^{**} \subseteq U^{**} \tag{82}$$

be the neighbourhood of $v_i$ in the graph $F[U^{**}]$. Because of our choice of $U^{**}$ in Fact 26, we have

$$|F_i| \geq \delta(F[U^{**}]) \geq f = \frac{1}{2}\tau u^* \tag{83}$$

for all $1 \leq i \leq u^{**}$.

We now define a concept that will be important in the proof that follows. Let sets $F^* \subseteq U^{**}$ and $W_0 \subseteq W$ be given. We put

$$\widetilde{F}^* = F^* \setminus \bigcup_{w \in W_0} N(H_0; w). \tag{84}$$

We shall say that $W_0$ is *bad with respect to* $F^*$ if

$$|\widetilde{F}^*| \geq |F^*| \exp\left\{-\frac{1}{18}\lambda^2 c^2 \sigma_0\right\}. \tag{85}$$

Roughly speaking, $W_0$ is bad with respect to $F^*$ if the $H_0$-neighbourhood of $W_0$ 'bites out' too little from $F^*$. (We remark in passing that we shall be applying this definition for 'large' sets $W_0$, namely, sets $W_0$ with $|W_0| \geq \lambda p|W|/3$.)

Next, we let $d_i$ $(1 \leq i \leq u^{**})$ be the degree of $v_i$ in the graph $H_1^{**} = H_1[U^{**}, W]$, or, equivalently, in $H_1$. Recalling that $(U^{**}, W)$ is $(p, \lambda/3; H_1^{**}, \varepsilon^{**})$-semi-regular and that

$$\varepsilon^{**} \leq \frac{1}{2}$$

(cf. (76)), we may suppose that

$$\frac{1}{3}\lambda p|W| \leq d_i \leq 2c\sqrt{n} \quad \text{for all} \quad 1 \leq i \leq u^{**}/2 \tag{86}$$

(recall that $\Delta(H) \leq 2c\sqrt{n}$). It will be convenient to set

$$W^{(i)} = N(H_1^{**}; v_i) = N(H_1; v_i) = \{w_1^{(i)}, \ldots, w_{d_i}^{(i)}\} \tag{87}$$

for all $1 \leq i \leq u^{**}/2$. We shall say that $v_i$ $(1 \leq i \leq u^{**}/2)$ is *bad* if $W^{(i)}$ is bad with respect to $F_i = N(F; v_i) \cap U^{**}$.

*Proof strategy, main claims, and the proof.* We now briefly discuss what we are aiming at. Recall that we are considering

$$H \in \mathcal{B}_0 = \mathcal{B}(U^*, W; F; \sigma_0, \lambda, \tau, \varepsilon'; n, T). \tag{88}$$

We shall use the fact that $\mathrm{Base}(H)$ does not contain too many edges from $F$. Explicitly, we shall use that not too many edges $\{a, b\} \in F$ are such that there is a vertex $w \in W$ with both $\{a, w\}$ and $\{w, b\}$ edges in $H$ (see (69)). Note that if $w \in N(H_1^{**}; a)$ and $b \in N(H_0^{**}; w)$, then $\{a, b\}$ belongs to $\mathrm{Base}(H) = \mathrm{Base}(H_0 \cup H_1)$. The neighbourhood $W_0 = N(H_1^{**}; a) = N(H_1; a)$ of $a$ in $H_1$ is bad with respect to $F^*$ if

$$\widetilde{F}^* = F^* \setminus \bigcup_{w \in W_0} N(H_0; w).$$

is large or, equivalently, the neighbourhood

$$F^* \cap \bigcup_{w \in W_0} N(H_0; w)$$

of $W_0$ within $F^* = N(F; a) \cap U^{**}$ in $H_0$ is small. The vertex $a$ is bad if its neighbourhood $W_0$ is bad. Roughly speaking, these definitions make it so that $a$ is bad if too few edges from $F$ incident to $a$ belong to $\mathrm{Base}(H)$.

Since globally $H$ is such that only a small number of edges from $F$ belong to $\mathrm{Base}(H)$, it must be that many vertices $a$ are incident to too few edges from $F \cap \mathrm{Base}(H)$, and hence there are many bad vertices $a$.

We shall estimate from above the number of graphs $H$ as in (88) by proving that, if $H_0$ is fixed, then the great majority of sets $W_0 \subseteq W$ are good, and hence it is very atypical for a vertex $a \in U^{**}$ to be bad. However, since many vertices $a$ must be bad in $H$, this will tell us that $H_1^{**}$ must be quite a peculiar graph. This will imply a strong bound on the number of graphs $H_1^{**}$. Summing over all choices for $H_0$, we shall still have a strong bound on the number of graphs $H$ as in (88).

The following two claims make precise the ideas sketched above. Recall that $H$ as in (88) is fixed, and our notation is so that (86) holds.

**Claim 29.** *For all $1 \le i \le u^{**}/2$, the number of subsets $W_0$ of $W$ with cardinality $d_i = |N(H_1; v_i)|$ that are bad with respect to $F_i$ is at most*

$$\alpha_0^{d_i} \binom{|W|}{d_i}. \tag{89}$$

**Claim 30.** *More than $u^{**}/4$ vertices $v_i$ with $1 \le i \le u^{**}/2$ are bad.*

We complete the proof of Lemma 23 assuming Claims 29 and 30 above.

*Completion of the proof of Lemma* 23. Suppose $H$ as in (88) is given.

- $(a)$ In view of Lemma 25, we may decompose $H$ as a union $H_0 \cup H_1$, where both $H_i$ ($i \in \{0, 1\}$) are such that $(U^*, W)$ is $(p, \lambda/3; H_i, \varepsilon')$-semi-regular, and $|E(H_0)| = |E(H_1)| = T/2$.
- $(b)$ In view of Facts 26 and 27, we know that there is a set $U^{**} \subseteq U^*$ satisfying $(i)$ and $(ii)$ of Fact 26, and such that $(U^{**}, W)$ is $(p, \lambda/3; H_i^{**}, \varepsilon^{**})$-semi-regular for $H_i^{**} = H_i[U^{**}, W]$ for both $i \in \{0, 1\}$.

Using $(a)$ and $(b)$ above, we may estimate the cardinality of

$$\mathcal{B}_0 = \mathcal{B}(U^*, W; F; \sigma_0, \lambda, \tau, \varepsilon'; n, T)$$

as follows. We 'generate' the members $H \in \mathcal{B}_0$ in steps, as outlined below.

(1) We first fix $H_0$, as in $(a)$.
(2) We then define $U^{**}$ using $F$ (see Fact 26).
(3) We generate $H_1$ in two stages:
  (a) We first fix a partition $T/2 = T^{**} + (T/2 - T^{**})$ of the number of edges $T/2$ that we want in our graph $H_1$.

(b) We then generate $H_1^{**} = H_1[U^{**}, W] = (U^{**}, W; E^{**})$, with $|E^{**}| = T^{**}$.
(c) We finally generate $H_1 - E^{**}$.

In order to estimate $|\mathcal{B}_0|$ from above, we estimate the number of ways the steps above may be carried out. (Note that we did not require that $H_0$ and $H_1$ above be disjoint. Since we are after an upper bound for $|\mathcal{B}_0|$, this is justified.)

Let us focus on Step 3b in the procedure above. We start by fixing a degree sequence $\mathbf{d} = \{d(v)\colon v \in U^{**}\}$ for the vertices in $U^{**}$ in the graph $H_1^{**}$. To generate $H^{**}$, it suffices to pick a member of the product

$$\prod \Big\{ [W]^{d(v)} \colon v \in U^{**} \Big\}. \tag{90}$$

We shall use Claims 29 and 30 to estimate how many members from (90) are 'valid' choices (*i.e.*, choices that correspond to graphs $H_1^{**}$ that may be completed to graphs $H_1$ with $H = H_0 \cup H_1 \in \mathcal{B}_0$).

**Fact 31.** *The number of elements of* (90) *that correspond to valid graphs* $H_1^{**}$ *is at most*

$$2^{u^{**}} \alpha_0^D \prod \left\{ \binom{|W|}{d(v)} \colon v \in U^{**} \right\}, \tag{91}$$

*where* $D = (1/12)\lambda p u^{**}|W|$.

*Proof.* The factor $2^{u^{**}}$ takes care of all the possibilities for *which* vertices in $U^{**}$ are bad. We 'gain' the factor $\alpha^D$ in the estimate (91) because Claim 30 tells us that at least $u^{**}/4$ vertices in $U^{**}$ are bad, and each bad vertex $v$ must have a neighbourhood that is chosen from at most

$$\alpha_0^{d(v)} \binom{|W|}{d(v)}$$

many possibilities (cf. (89) in Claim 29). Therefore we 'gain' a factor of

$$\prod \{\alpha_0^{d(v)} \colon v \text{ is a bad vertex}\} \le \alpha_0^D,$$

where $D = (u^{**}/4)(\lambda/3)p|W| = (1/12)\lambda p u^{**}|W|$ (cf. (86)), and Fact 31 is proved. □

To estimate the quantity in (91), we observe that, because $\Delta(H) \le 2c\sqrt{n}$, we have $T \le 2c\sqrt{n}|U^*| = 2cu^*\sqrt{n}$. Moreover, we recall that $u^{**} \ge \tau u^*/2$, $p = c/\sqrt{n}$, and $|W| = \sigma_0 n$, and hence we have

$$\frac{D}{T} \ge \frac{(1/12)\lambda p u^{**}|W|}{2c\sqrt{n}u^*} \ge \frac{1}{48}\lambda\tau\sigma_0.$$

Therefore (91) is at most

$$2^{u^{**}} \alpha_0^{(\lambda\tau\sigma_0/48)T} \prod \left\{ \binom{|W|}{d(v)} \colon v \in U^{**} \right\}. \tag{92}$$

We now sum (92) over all choices for $\mathbf{d} = \{d(v)\colon v \in U^{**}\}$ and over all choices for $U^{**} \subseteq U^*$. Recalling that $\mathbf{d}$ is a partition of $T^{**}$, we see that this sum is at most

$$2^{u^*+u^{**}} \alpha_0^{(\lambda\tau\sigma_0/48)T} \binom{|W||U^{**}|}{T^{**}}. \tag{93}$$

We now sum (93) over all choices for $T^{**}$ and over all choices for $H_1 - E^{**}$ (see Step 3c). This gives us at most

$$2^{u^*+u^{**}} \alpha_0^{(\lambda\tau\sigma_0/48)T} \binom{|W||U^*|}{T} \le 2^{2u^*} \alpha_0^{(\lambda\tau\sigma_0/48)T} \binom{|W||U^*|}{T}. \tag{94}$$

Also (cf. (12) and (13)), we have

$$T \ge \lambda p u^* |W| = \lambda c \sigma_0 u^* \sqrt{n} \gg u^* \log u^*. \tag{95}$$

Inequality (95) gives us that (94) is at most

$$(2\alpha_0)^{(\lambda\tau\sigma_0/48)T} \binom{|W||U^*|}{T} \le \alpha^T \binom{|W||U^*|}{T}, \tag{96}$$

where the inequality above follows from (79). Lemma 23 follows from (96). □

*Proof of the claims.* We are now left with proving Claims 29 and 30.

*Proof of Claim* 29. Fix $1 \le i \le u^{**}/2$. For simplicity, put $F^* = F_i = N(F; v_i) \cap U^{**}$ and let $d = d_i = |N(H_1; v_i)|$. Let $W_0 \subseteq W$ be a $d$-element subset of $W$ and fix and ordering $w_1, \dots, w_d$ of the elements of $W_0$. We define a sequence

$$F^* = F^{(0)} \supseteq \cdots \supseteq F^{(d)} = \widetilde{F}^* = F^* \setminus \bigcup_{w \in W_0} N(H_0; w) \tag{97}$$

of decreasing subsets of $W$, putting

$$F^{(j)} = F^{(j-1)} \setminus N(H_0; w_j) = F^* \setminus \bigcup_{1 \le k \le j} N(H_0; w_k), \tag{98}$$

for all $1 \le j \le d$. Let us call the ratio

$$|F^{(j)}|/|F^{(j-1)}| \tag{99}$$

the *shrinking factor at* $j$ ($1 \le j \le d$). Observe that this factor depends only on the sequence $w_1, \dots, w_j$. Moreover, note that *we like small shrinking factors*; if we have many indices $j$ with small shrinking factors, then $W_0$ is likely to be a good set, because $\widetilde{F}^*$ is likely to be a small set. We make this remark precise in what follows.

Suppose $W_0$ is a bad set with respect to $F^* = F_i = N(F; v_i) \cap U^{**}$.

**Fact 32.** *The number of indices $j$ with shrinking factor less than $1 - \lambda p/3$ at $j$ is smaller than $d/2$.*

*Proof.* Suppose for a contradiction that we have at least $d/2$ indices $j$ with shrinking factor less than $1-\lambda p/3$. Then

$$|\widetilde{F}^*| = |F^{(d)}| < |F^*|\left(1-\frac{1}{3}\lambda p\right)^{d/2} \leq |F^*|\exp\left\{-\frac{1}{6}\lambda p d\right\}. \tag{100}$$

Recall that $p = cn^{-1/2}$ and $d = d_i \geq (\lambda/3)p|W| = (\lambda/3)p\sigma_0 n$ (see (86)). We may deduce from (100) that

$$|\widetilde{F}^*| < |F^*|\exp\left\{-\frac{1}{6}\lambda^2 p^2\sigma_0 n\right\} = |F^*|\exp\left\{-\frac{1}{18}\lambda^2 c^2\sigma_0\right\}, \tag{101}$$

contradicting the assumption that $W_0$ is bad with respect to $F^*$ (cf. (85) and (101)). □

Let us now fix $1 \leq j \leq d$. When is the shrinking factor at $j$ poor? More precisely, when is this shrinking factor at least $1-\lambda p/3$? This is the case if and only if

$$|N(H_0; w_j) \cap F^{(j-1)}| \leq \frac{1}{3}\lambda p|F^{(j-1)}|. \tag{102}$$

We shall now argue that (102) is atypical; that is, if (102) holds, then $w_j$ is an 'atypical vertex'.

Since $W_0$ is a bad set with respect to $F^* = F_i$ and (77) and (83) hold, we have

$$\begin{aligned}|F^{(j-1)}| \geq |\widetilde{F}^*| \geq |F^*|\exp\left\{-\frac{1}{18}\lambda^2 c^2\sigma_0\right\} &\geq f\exp\left\{-\frac{1}{18}\lambda^2 c^2\sigma_0\right\} \\ &\geq \left(\frac{1}{2}\tau\exp\left\{-\frac{1}{18}\lambda^2 c^2\sigma_0\right\}\right)u^* \geq \varepsilon^{**}|U^{**}|. \end{aligned}\tag{103}$$

Recalling that $(U^{**}, W)$ is $(p, \lambda/3; H_0^{**}, \varepsilon^{**})$-semi-regular, we see that (103) implies that the number of vertices $w \in W$ with

$$|N(H_0; w) \cap F^{(j-1)}| < \frac{1}{3}\lambda p|F^{(j-1)}|$$

is smaller than $\varepsilon^{**}|W|$. We have proved the following fact.

**Fact 33.** *For every $w_1, \ldots, w_{j-1} \in W$, there is a set $W^{(j)} \subseteq W$ of cardinality less than $\varepsilon^{**}|W|$ such that if the shrinking factor at $j$ for the sequence*

$$w_1, \ldots, w_{j-1}, w_j$$

*is at least $1-\lambda p/3$, then $w_j \in W^{(j)}$.*

Let $\mathcal{W}^\times \subseteq W^d = W \times \cdots \times W$ be the collection of ordered $d$-tuples $(w_1, \ldots, w_d)$ that have at least $d/2$ indices $j$ with poor shrinking factor, that is, with shrinking factor at least $1-\lambda p/3$. Fact 33 implies that

$$|\mathcal{W}^\times| \leq 2^d(\varepsilon^{**})^{d/2}|W|^d. \tag{104}$$

Now let $\mathcal{W} \subseteq [W]^d$ be the set of $d$-element subsets of $W$ that are bad with respect to $F^*$. We wish to show that $|\mathcal{W}|$ satisfies (89).

Fact 32 implies that, if $W_0 \in \mathcal{W}$, then any ordered $d$-tuple $(w_1, \ldots, w_d)$ with $W_0 = \{w_1, \ldots, w_d\}$ belongs to $\mathcal{W}^\times$ (that is, any ordering of a member of $\mathcal{W}$ is in $\mathcal{W}^\times$). This and (104) give that

$$d!|\mathcal{W}| \le |\mathcal{W}^\times| \le 2^d(\varepsilon^{**})^{d/2}|W|^d, \tag{105}$$

and hence, using that $d \le 2c\sqrt{n}$ (see (86)), for large enough $n$ we have

$$\begin{aligned}|\mathcal{W}| &\le 2^d(\varepsilon^{**})^{d/2}\frac{|W|^d}{d!} \le 2^d(\varepsilon^{**})^{d/2}e^{d^2/|W|}\binom{|W|}{d} \\ &\le 2^d(\varepsilon^{**})^{d/2}e^{4c^2/\sigma_0}\binom{|W|}{d} \le \alpha_0^d\binom{|W|}{d}, \end{aligned} \tag{106}$$

where in the last inequality we used (78). Clearly, Claim 29 follows from (106). □

*Proof of Claim* 30. Suppose for a contradiction that the number of vertices $v_i$ $(1 \le i \le u^{**}/2)$ that are bad is at most $u^{**}/4$. Therefore,

(*) the number of vertices $v_i$ $(1 \le i \le u^{**}/2)$ that are not bad is at least $u^{**}/4$.

Suppose $v_i$ is a vertex that is not bad, where $1 \le i \le u^{**}/2$. This means that $W_0 = N(H_1^{**}; v_i) = N(H_1; v_i)$ is not bad, that is, the neighbourhood

$$F_i \cap \bigcup_{w \in W_0} N(H_0; w)$$

of $W_0$ within $F_i = N(F; v_i) \cap U^{**}$ in $H_0$ has cardinality greater than

$$\left(1 - \exp\left\{-\frac{1}{18}\lambda^2c^2\sigma_0\right\}\right)|F_i| \ge \frac{1}{50}\lambda^2c^2\sigma_0 f, \tag{107}$$

where we used that $|F_i| \ge f$ and that $x = (1/18)\lambda^2c^2\sigma_0 \le 1/2$, and hence $1 - e^{-x} \ge x/2$, say. However, this means that $v_i$ contributes to $F \cap B_{U^*}(H)$ at least as many pairs as in (107) (with some pairs counted twice, when we consider all the vertices $v_i$). Recalling (*) and Fact 26(*ii*), we see that

$$\begin{aligned}|F \cap B_{U^*}(H)| &\ge \frac{1}{4}u^{**} \times \frac{1}{2}\left(\frac{1}{50}\lambda^2c^2\sigma_0 f\right) \\ &\ge \frac{1}{400} \times \frac{1}{2}\tau u^* \times \lambda^2c^2\sigma_0 f = \frac{1}{800}\lambda^2c^2\sigma_0\tau u^* f. \end{aligned} \tag{108}$$

Recall now that $f = \tau u^*/2$; we may then deduce from (108) that

$$\begin{aligned}|F \cap B_{U^*}(H)| &\ge \frac{1}{1600}\lambda^2c^2\sigma_0\tau^2(u^*)^2 \\ &\ge \frac{1}{800}\lambda^2c^2\sigma_0\tau^2\binom{u^*}{2} = \frac{1}{800}\lambda^2c^2\sigma_0\tau|F|. \end{aligned} \tag{109}$$

Note that (109) is in contradiction to the fact that $H \in \mathcal{B}_0$ (see (69) and (88)). We conclude that (*) must fail, and hence that the number

of vertices $v_i$ $(1 \le i \le u^{**}/2)$ that are bad exceeds $u^{**}/4$. Claim 30 is proved. □

## 5. Concluding remarks

We close by mentioning a few remarks and open problems.

*The online game for more colours.* We hope to address the online game when the Painter has $r > 2$ colours in the near future.

*General graphs.* It would be most interesting to generalize our results for general graphs $H$, that is, for avoidance games in which the Painter has to avoid monochromatic copies of a given graph $H$. However, our methods are deeply based on the simple structure of the triangles and hence a completely new approach may be needed.

Our results above suggest that, in the two-round game for arbitrary graph $H$, the interesting range for $N_0$ is around the classical threshold for the corresponding Ramsey property, that is, $N_0 = \Theta(n^{2-1/d_2(H)})$, where $d_2(H) = \max\{(e(F)-1)/(v(F)-2)\colon F \subseteq H\}$ (see [21]).

*Deterministic games.* In the deterministic version of our online game, say for $r = 2$ and $H = K_k$, the edges are generated not randomly, but by an adversary. Hence, this is a two-person game with the payoff to the Painter equal to the number of coloured edges until she is forced to form a monochromatic clique $K_k$.

Define the *online Ramsey number* $\bar{R}(k)$ as the value of this game. The most interesting question is if it is true that $\bar{R}(k) = o(R(k)^2)$ as $k \to \infty$, where $R(k)$ is the classical Ramsey number. Some preliminary results in this direction have been proved by Kurek and Ruciński [18]. In particular, $\bar{R}(3) = 8$ and $\bar{R}(k,l) = o(R(k,l)^2)$ as $k \to \infty$ while $l \ge 3$ is fixed, where $R(k,l)$ and $\bar{R}(k,l)$ are the corresponding off-diagonal numbers.

Finally, we mention that other games inspired by Ramsey's theorem have been investigated in several forms. The reader is referred to, for instance, Beck [2, 3, 4, 5], Beck and Csirmaz [1], Erdős and Selfridge [7], and Seress [22].

## References


1. J. Beck and L. Csirmaz, *Variations on a game*, J. Combin. Theory Ser. A **33** (1982), no. 3, 297–315. MR **84a:**90092
2. József Beck, *van der Waerden and Ramsey type games*, Combinatorica **1** (1981), no. 2, 103–116. MR **84a:**90093
3. ———, *Random graphs and positional games on the complete graph*, Random graphs '83 (Poznań, 1983), North-Holland, Amsterdam, 1985, pp. 7–13. MR **87j:**05130
4. ———, *Deterministic graph games and a probabilistic intuition*, Combin. Probab. Comput. **3** (1994), no. 1, 13–26. MR **95h:**90191
5. ———, *Foundations of positional games*, Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995), vol. 9, 1996, pp. 15–47. MR **99h:**90128

6. Béla Bollobás, *Random graphs*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1985. MR **87f:**05152
7. P. Erdős and J. L. Selfridge, *On a combinatorial game*, J. Combinatorial Theory Ser. A **14** (1973), 298–301. MR 48 #5655
8. P. Frankl and V. Rödl, *Large triangle-free subgraphs in graphs without $K_4$*, Graphs Combin. **2** (1986), no. 2, 135–144. MR **89b:**05124
9. E. Friedgut, V. Rödl, A. Ruciński, and P. Tetali, *A sharp threshold for random graphs with monochromatic triangle in every edge coloring*, in preparation, 2002.
10. Ehud Friedgut, *Sharp thresholds of graph properties, and the k-sat problem*, J. Amer. Math. Soc. **12** (1999), no. 4, 1017–1054, With an appendix by Jean Bourgain. MR **2000a:**05183
11. Svante Janson, Tomasz Łuczak, and Andrzej Ruciński, *An exponential bound for the probability of nonexistence of a specified subgraph in a random graph*, Random graphs '87 (Poznań, 1987), Wiley, Chichester, 1990, pp. 73–87. MR **91m:**05168
12. ———, *Random graphs*, Wiley-Interscience, New York, 2000. MR **2001k:**05180
13. Svante Janson and Andrzej Ruciński, *The deletion method for upper tail estimates*, 30pp, submitted, 2000.
14. ———, *The infamous upper tail*, Random Structures Algorithms **20** (2002), no. 3, 317–342.
15. Y. Kohayakawa, *Szemerédi's regularity lemma for sparse graphs*, Foundations of computational mathematics (Rio de Janeiro, 1997), Springer, Berlin, 1997, pp. 216–230. MR **99g:**05145
16. Y. Kohayakawa and V. Rödl, *Szemerédi's regularity lemma and quasi-randomness*, Recent advances in algorithms and combinatorics, Springer, Berlin, 63pp., to appear, 2002.
17. Yoshiharu Kohayakawa, Tomasz Łuczak, and Vojtěch Rödl, *Arithmetic progressions of length three in subsets of a random set*, Acta Arith. **75** (1996), no. 2, 133–163. MR **97b:**11011
18. Andrzej Kurek and Andrzej Ruciński, *The sparsest Ramsey graphs are complete graphs*, submitted, 2002.
19. Tomasz Łuczak, Andrzej Ruciński, and Bernd Voigt, *Ramsey properties of random graphs*, J. Combin. Theory Ser. B **56** (1992), no. 1, 55–68. MR **94b:**05172
20. Vojtěch Rödl and Andrzej Ruciński, *Random graphs with monochromatic triangles in every edge coloring*, Random Structures Algorithms **5** (1994), no. 2, 253–270. MR **95a:**05102
21. ———, *Threshold functions for Ramsey properties*, J. Amer. Math. Soc. **8** (1995), no. 4, 917–942. MR **96h:**05141
22. Ákos Seress, *On Hajnal's triangle-free game*, Graphs Combin. **8** (1992), no. 1, 75–79. MR **93a:**05095
23. Endre Szemerédi, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), CNRS, Paris, 1978, pp. 399–401. MR **81i:**05095

Einstein Institute of Mathematics, The Hebrew University of Jerusalem, Jerusalem, 91904, Israel
*E-mail address*: ehudf@math.huji.ac.il

Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão 1010, 05508–900 São Paulo, Brazil
*E-mail address*: yoshi@ime.usp.br

Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322, USA
*E-mail address*: rodl@mathcs.emory.edu

Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland, and Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322, USA
*E-mail address*: rucinski@amu.edu.pl

School of Mathematics, Georgia Institute of Technology, Atlanta, GA, 30332, USA
*E-mail address*: tetali@math.gatech.edu