INDUCED RAMSEY NUMBERS

Y. KOHAYAKAWA, H. J. PRÖMEL, AND V. RÖDL

Dedicated to the memory of Professor Paul Erdős

ABSTRACT. We investigate the *induced Ramsey number* $r_{\mathrm{ind}}(G,H)$ of pairs of graphs (G,H). This number is defined to be the smallest possible order of a graph Γ with the property that, whenever its edges are coloured red and blue, either a red induced copy of G arises or else a blue induced copy of H arises. We show that, for any G and H with $k=|V(G)|\le t=|V(H)|$, we have

$$r_{\mathrm{ind}}(G,H)\le t^{Ck\log q},$$

where $q=\chi(H)$ is the chromatic number of H and C is some universal constant. Furthermore, we also investigate $r_{\mathrm{ind}}(G,H)$ imposing some conditions on G. For instance, we prove a bound that is polynomial in both k and t in the case in which G is a tree. Our methods of proof employ certain random graphs based on projective planes.

1. Introduction and Main Results

A fundamental problem in Ramsey theory is to determine or estimate the Ramsey number for complete graphs. Write $R(k)$ for the smallest integer n with the property that, if one colours the edges of the complete graph on n vertices with colours red and blue, one necessarily obtains a complete subgraph on k vertices all of whose edges are coloured red, or else all of whose edges are coloured blue. The number $R(k)$, the existence of which follows from a classical theorem of Ramsey [14], is the *Ramsey number* of the complete graph on k vertices.

The best bounds that are currently known for $R(k)$ imply that

$$2^{1/2}\le\liminf_{k\to\infty}R(k)^{1/k}\le\limsup_{k\to\infty}R(k)^{1/k}\le 4.$$

1991 *Mathematics Subject Classification.* Primary 05C55, 05C80; Secondary: 05C35.

Key words and phrases. Generalised Ramsey theory, induced subgraphs, random graphs.

Research partially supported by PROBRAL project 026/95, a CAPES–DAAD exchange programme and Humboldt Stiftung. The first author was partially supported by FAPESP (Proc. 96/04505–2), by CNPq (Proc. 300334/93–1 and ProTeM-CC-II Project ProComb), and by MCT/FINEP under PRONEX project 107/97. The second author was partially supported DFG project Pr 296/4–1. The third author was partially supported by NSF grant DMS–9401559.

Part of this work was done while the first and the third authors were visiting Humboldt-Universität zu Berlin.

Thus, even the precise rate of the exponential growth of $R(k)$ is unknown. Indeed, deciding whether $\lim_{k\to\infty} R(k)^{1/k}$ exists, and if so determining its value, are outstanding problems in the area. For a pleasant introduction to Ramsey theory in general, the reader is referred to Graham, Rothschild, and Spencer [11]. Nešetřil [13] gives a relatively concise survey on this vast subject. Numerical problems in this area are discussed by Graham and Rödl [10].

In this paper, we are concerned with a counterpart to the theorem of Ramsey for non-complete graphs. Namely, we are interested in a result proved independently by Deuber [4], Erdős, Hajnal, and Pósa [9], and Rödl [15], which states that, for any pair of graphs G and H, there is a graph Γ such that

$$\Gamma \to (G, H), \tag{1}$$

where the arrow notation in (1) above means that whenever we colour the edges of Γ red and blue, either a red induced copy of G arises, or else a blue induced copy of H arises. This variant of Ramsey's theorem immediately raises a numerical problem.

For any given graphs G and H, let $r_{\mathrm{ind}}(G,H)$ denote the smallest integer n for which there exists a graph Γ on n vertices satisfying (1). We refer to $r_{\mathrm{ind}}(G,H)$ as the *induced Ramsey number* of the pair (G,H). Erdős wrote [7, §5] (with a little change in the notation): "Hajnal and I observed that if G_1 and G_2 have at most n vertices, then

$$r_{\mathrm{ind}}(G_1,G_2) \le 2^{2^{n^{1+\varepsilon}}}. \tag{2}$$

We have never published the not entirely trivial proof of (2) since Hajnal and I thought that perhaps

$$\max r_{\mathrm{ind}}(G_1,G_2) = R(n). \tag{3}$$

"Conjecture (3) is perhaps a little too optimistic, but we have no counterexample. Perhaps there is a better chance to prove $r_{\mathrm{ind}}(G_1,G_2) \le 2^{cn}$."

Writing $r_{\mathrm{ind}}(H)$ for $r_{\mathrm{ind}}(H,H)$, this latter problem of Erdős [7], already implicit in [6, §III], is then the following:

Problem 1. *Is there an absolute constant C such that for any graph H on t vertices we have $r_{\mathrm{ind}}(H) \le 2^{Ct}$?*

Techniques in Rödl [15] imply that $r_{\mathrm{ind}}(H,H)$ is indeed exponential in $t = |V(H)|$ if H is bipartite. Note that if the answer to Problem 1 is positive, then the result is best possible up to the value of C, that is, there are graphs H for which $r_{\mathrm{ind}}(H)$ is exponential in $t = |V(H)|$. Indeed, it suffices to take H as the complete graph on t vertices.

Our main result below, Theorem 3, comes close to establishing a positive answer to Problem 1 in general. In particular, our result greatly strengthens the previous doubly exponential bound (2) of Erdős and Hajnal.

Quite curiously, our result also comes close to settling the following conjecture, which deals with the 'asymmetric' induced Ramsey number $r_{\mathrm{ind}}(G,H)$.

Here, we are concerned with the case in which G is some fixed graph and H is very large.

Conjecture 2. *For any graph G, there is a constant $f = f(G)$ that depends only on G such that, for any graph H on t vertices, we have*

$$r_{\text{ind}}(G, H) \le t^f.$$

Our main result is as follows.

Theorem 3. *Let G and H be graphs with $|V(G)| = k$ and $|V(H)| = t$, where $k \le t$, and suppose $q = \chi(H) \ge 2$. Then*

$$r_{\text{ind}}(G, H) \le t^{Ck \log q} \tag{4}$$

for some absolute constant C.

Note that in the 'diagonal case', in which $G = H$, Theorem 3 gives the bound

$$r_{\text{ind}}(H) = r_{\text{ind}}(H, H) \le t^{Ct \log q},$$

which only fails to be a purely exponential bound in $t = |V(H)|$ by a factor of $(\log t)(\log \chi(H)) \le (\log t)^2$ in the exponent. Furthermore, note that (4) in Theorem 3 fails to be polynomial in t only by a factor of $\log \chi(H)$ in the exponent. Thus, Theorem 3 does indeed fall only a little short of settling Problem 1 and Conjecture 2.

Inequality (4) is proved by showing that there is a suitably small graph R for which $R \to (G, H)$ holds. This graph $R = R(\mathcal{P}, H)$ is randomly constructed using a projective plane $\mathcal{P}$ and the graph H. (The graph G plays no rôle in the definition of R.) Roughly speaking, the graph R is obtained from $\mathcal{P}$ and H by randomly embedding 'blown-up copies' of H in each line of $\mathcal{P}$. (A 'blow-up' of H is simply a graph obtained from H by replacing each vertex of H by an independent set and each edge of H by a suitably large complete bipartite graph.)

The random graphs $R = R(\mathcal{P}, H)$ were considered before by Brown and Rödl [2] and Eaton and Rödl [5]. However, in [2, 5], the authors investigate Ramsey properties of these graphs with respect to *vertex* colourings. A similar construction is studied by Rödl and Winkler in [17], where Ramsey properties with respect to orientations of graphs are investigated. More recently, Luczak and Rödl [12] confirmed a conjecture of Trotter by showing that $r_{\text{ind}}(H) = t^{O(1)}$ if the t-vertex graph H has bounded maximum degree. The graph $R = R(\mathcal{P}, H)$ is crucial in [12].

Since (4) fails to prove Conjecture 2 in full generality, we consider graphs G with some special structure. To describe a class of graphs G for which we are able to prove Conjecture 2, we need to introduce a couple of definitions.

If G_1 and G_2 are two graphs, we let the *join* $G_1 \vee G_2$ of G_1 and G_2 be the graph obtained from the disjoint union of G_1 and G_2 by adding all the G_1–G_2 edges. Thus, for instance, $K^{a,b} = E^a \vee E^b$, where $K^{a,b}$ is the complete bipartite graph with vertex classes of cardinalities a and b, and E^a denotes the a-vertex graph with no edges, and similarly for E^b.

Let $\mathcal{S}$ denote the smallest class of finite graphs that contains the 1-vertex complete graph and is closed under taking disjoint unions and joins. Following Erdős and Hajnal [8], we call the graphs in $\mathcal{S}$ *simple.* Theorem 4 below verifies Conjecture 2 in the case in which G is a simple graph.

Theorem 4. *For any simple graph G, there is a constant $f = f(G)$ that depends only on G such that, for any graph H on t vertices, we have*

$$r_{\text{ind}}(G, H) \leq t^f.$$

The proof of Theorem 4 will also be based on the random graph $R = R(\mathcal{P}, H)$. However, quite naturally, the choice of certain parameters in the definition of this graph for this proof will be different from the choice to be made in the proof of Theorem 3.

A basic case that is not covered in Theorem 4 is the one in which G is a tree. As one may expect, this case can be handled quite easily.

Theorem 5. *For any tree T and arbitrary graph H, we have*

$$r_{\text{ind}}(T, H) \leq ck^2t^4 \left(\frac{\log(kt^2)}{\log\log\log(kt^2)}\right)^2, \tag{5}$$

where $k = |V(T)|$, $t = |V(H)|$, and c is some absolute constant.

Thus, Theorem 5 tells us that $r_{\text{ind}}(T, H)$ is polynomial in both $k = |V(T)|$ and $t = |V(H)|$. Our proof of this result is simple, and we have not put much effort in obtaining the best exponents in (5). Let us mention that Beck has investigated the problem of estimating the *induced size-Ramsey number* of a given tree T, which is defined to be the minimal number of edges in a graph $\Gamma = \Gamma(T)$ such that $\Gamma \to (T, T)$. Indeed, in [1], Beck proves that if T is a tree with n edges and n is larger than some absolute constant, then there is a graph Γ with less than

$$n^3(\log n)^4$$

edges such that $\Gamma \to (T, T)$.

This paper is organised as follows. The random graph $R = R(\mathcal{P}, H)$ is defined in Section 2, where auxiliary results concerning projective planes are also given. We also give in Section 2 a couple of properties of $R = R(\mathcal{P}, H)$ that are crucial later on in the proofs of Theorems 3 and 4, given in Sections 3 and 4. The induced Ramsey number for trees versus general graphs is studied in Section 5. Our very short final section is devoted to some concluding remarks.

As the reader will observe, we try to isolate, whenever possible, the probabilistic parts of our arguments in the proofs of Theorems 3 and 4. For instance, Section 3.2 is entirely deterministic in nature. The only probabilistic component in the proof of Theorem 3 is, in fact, an application of Corollary 11 of Section 2.3, made in Section 3.3. Needless to say, this probabilistic component is crucial for our argument.

Below, we write $a \ll b$ to mean that $a/b \to 0$ as $n \to \infty$, and similarly for $\gg$. We also write $a \sim b$ to mean that $a/b \to 1$ as $n \to \infty$. Moreover, for $A > 0$, $B > 0$ and $\delta > 0$, we write $A \sim_\delta B$ if $1/(1+\delta) < A/B < 1+\delta$. In the sequel, the logarithms without specified bases are to the base e.

2. The Construction and Preliminary Lemmas

2.1. **The construction.** Let a projective plane $\mathcal{P} = (V, \mathcal{L})$ with $|V| = |\mathcal{L}| = n$ and a graph H be given. We shall construct a random graph $R = R_n = R_n(\mathcal{P}, H)$ based on $\mathcal{P}$ and H that will have interesting Ramsey properties with respect to 'small' graphs G and the graph H. Note in particular that the definition of R will not depend on the graph G.

To describe our construction, suppose that, for each line $L \in \mathcal{L}$, we have a partition $L = \bigcup_{v \in V(H)} L_v$ of L whose parts L_v are indexed by the vertices $v \in V(H)$ of H. Denote this partition by Π_L, and set $\Pi = (\Pi_L)_{L \in \mathcal{L}}$. We define the graph $R = R_n(\mathcal{P}, H, \Pi)$ as follows. The vertex set of R is V. To define the edge set of R, suppose a, $b \in V$ are two distinct points in $\mathcal{P}$. Then there is a unique line $L \in \mathcal{L}$ that contains both a and b. Consider the partition $\Pi_L \in \Pi$ of L, and suppose $a \in L_u$ and $b \in L_v$. Then

$$ab \in E(R) \text{ if and only if } uv \in E(H).$$

To define the *random* graph $R = R_n(\mathcal{P}, H)$, we pick the family of partitions $(\Pi_L)_{L \in \mathcal{L}}$ randomly. For each $L \in \mathcal{L}$, the partition Π_L is chosen uniformly at random from all the partitions of L indexed by $V(H)$. Equivalently, for each $x \in L$, we choose a vertex $v \in V(H)$ independently and uniformly at random and let $x \in L_v$. The choices of the Π_L are made independently for all $L \in \mathcal{L}$. This defines a random family of partitions $\Pi = (\Pi_L)_{L \in \mathcal{L}}$, and we define our random graph R by letting

$$R = R_n(\mathcal{P}, H) = R_n(\mathcal{P}, H, \Pi).$$

Note that the randomness of $R = R_n(\mathcal{P}, H, \Pi)$ comes solely from the random choice of Π. Once Π is chosen, the graph R is determined.

2.2. **Lemmas on projective planes.** Naturally, the key to a good understanding of the random graph $R = R_n(\mathcal{P}, H)$ is a good understanding of the random family $\Pi = (\Pi_L)_{L \in \mathcal{L}}$. In this section, we shall review a few facts about projective planes that will be needed in Section 2.3, where we analyse the structure of a typical family $\Pi = (\Pi_L)_{L \in \mathcal{L}}$.

We start with a lemma due to Eaton and Rödl [5].

Lemma 6. *Let $(V,\mathcal{L})$ be a projective plane of order p and $n=p^2+p+1$. Let $X\subset V$ and $\mathcal{Y}\subset\mathcal{L}$ be given, and suppose $|X|=\alpha n$ and $|\mathcal{Y}|=\beta n$. Then*

$$\frac{\alpha\beta n^2}{p+\sqrt{p}+1}\left(1-\frac{1+O(n^{-1/4})}{\sqrt{\alpha\beta\sqrt{n}}}\right)\leq\sum_{L\in\mathcal{Y}}|X\cap L|$$
$$\leq\frac{\alpha\beta n^2}{p+\sqrt{p}+1}\left(1+\frac{1+O(n^{-1/4})}{\sqrt{\alpha\beta\sqrt{n}}}\right). \quad\square$$

Corollary 7. *With the notation as above, if $\alpha\beta\geq\omega n^{-1/2}$, then*

$$\sum_{L\in\mathcal{Y}}|X\cap L|=\left(1+O\left(\frac{1}{\sqrt{\omega}}\right)\right)\alpha\beta n^{3/2}. \quad\square$$

In the corollary below, $O_1(a)$ denotes a term b that satisfies $|b|\leq a$.

Corollary 8. *Suppose $X\subset V$ and $\alpha=|X|/n$. Then, if $\beta=\beta(n)$ satisfies $\alpha\beta\geq 2\omega n^{-1/2}$ for some function $\omega=\omega(n)\to\infty$, we have*

$$|X\cap L|=\left(1+O_1\left(\frac{\log\omega}{\sqrt{\omega}}\right)\right)\alpha\sqrt{n} \tag{6}$$

for all but fewer than βn lines $L\in\mathcal{L}$, as long as $n\geq n_0(\omega)$.

Proof. Suppose (6) fails. Then there is a set of lines $\mathcal{Y}\subset\mathcal{L}$ with $|\mathcal{Y}|=\lceil\beta n/2\rceil$ such that either

$$|X\cap L|>\left(1+\frac{\log\omega}{\sqrt{\omega}}\right)\alpha\sqrt{n} \tag{7}$$

for all $L\in\mathcal{Y}$, or else

$$|X\cap L|<\left(1-\frac{\log\omega}{\sqrt{\omega}}\right)\alpha\sqrt{n} \tag{8}$$

for all $L\in\mathcal{Y}$. Suppose (7) happens. Summing (7) over all $L\in\mathcal{Y}$, we have, by Corollary 7,

$$\left(1+\frac{\log\omega}{\sqrt{\omega}}\right)\alpha\sqrt{n}|\mathcal{Y}|<\sum_{L\in\mathcal{Y}}|X\cap L|=\left(1+O\left(\frac{1}{\sqrt{\omega}}\right)\right)\alpha\sqrt{n}|\mathcal{Y}|,$$

which contradicts Corollary 7 if $n\geq n_0(\omega)$, since $\omega=\omega(n)\to\infty$. The case in which (8) holds for all $L\in\mathcal{Y}$ is similar. $\square$

Let us remark that, clearly, the O_1 term in the result above may be chosen to be anything that is sufficiently larger than $\omega^{-1/2}$. Note also that Corollary 8 only concerns sets X with $|X|\gg\sqrt{n}$, since we need $\beta\leq 1$ and hence $\alpha\beta\gg n^{-1/2}$ can only hold if $\alpha\gg n^{-1/2}$.

In view of Corollary 8, we make the following definition. Let $X\subset V$, $L\in\mathcal{L}$, and $\delta>0$ be given. We say that the line L is (X,δ)-*normal* if

$$|X\cap L|\sim_\delta\frac{|X|}{\sqrt{n}}\,.$$

(We recall that we write $A \sim_\delta B$ if $1/(1+\delta) < A/B < 1+\delta$.) Let us set

$$\mathcal{B}(X,\delta) = \{L \in \mathcal{L} \colon L \text{ is } \textit{not } (X,\delta)\text{-normal}\}. \tag{9}$$

Note that Corollary 8 implies that $|\mathcal{B}(X,\delta)| < \beta n$ if $X \subset V$ satisfies $\beta|X| \gg \sqrt{n}$ and n is large enough.

2.3. **A few lemmas on $\Pi = (\Pi_L)_{L\in\mathcal{L}}$.** To study the random graphs $R = R_n(\mathcal{P}, H)$, we need to investigate the random family of partitions $\Pi = (\Pi_L)_{L\in\mathcal{L}}$ that is used to define those random graphs. In fact, in this section, we assume that $\Pi = (\Pi_L)_{L\in\mathcal{L}}$ is a family of random partitions of the $L \in \mathcal{L}$, with each Π_L an independent random partition of L, uniformly chosen from all the partitions of L indexed by $[t] = \{1,\dots,t\}$, for some given integer t. We shall only be concerned with large t and large n. Thus, the inequalities below need only hold for large enough t and n. Before we proceed, let us stress that the graph H will not play any rôle in this section.

We start with a definition that will be very important in the sequel. Let $L \in \mathcal{L}$, $X \subset V$, and $\delta > 0$ be given. Also, let a fixed family of $[t]$-indexed partitions $\Pi = (\Pi_L)_{L\in\mathcal{L}}$ be given. We then say that L is (X,δ,Π)-*normal* if

$$|X \cap L_i| \sim_\delta \frac{|X|}{t\sqrt{n}}$$

for all $1 \le i \le t$, where L_i is the ith block of the partition Π_L of L. A moment's thought shows that if a line L is (X,δ,Π)-normal, then it is also (X,δ)-normal, whereas the converse holds only if the partition Π_L of L splits $X \cap L$ 'judiciously.'

Let us set

$$\mathcal{B}(X,\delta,\Pi) = \{L \in \mathcal{L} \colon L \text{ is } \textit{not } (X,\delta,\Pi)\text{-normal}\}. \tag{10}$$

Our next simple lemma says that the probability that many $(X,\delta/2)$-normal lines should *fail* to be (X,δ,Π)-normal is very small.

Lemma 9. *Fix $\delta > 0$ and $X \subset V$. Then*

$$\mathbb{P}\Big\{\big|\mathcal{B}(X,\delta,\Pi) \setminus \mathcal{B}(X,\delta/2)\big| \ge t^2\sqrt{n}\Big\} < \exp\left(-c_\delta t|X| + \frac{1}{2}t^2\sqrt{n}\log n\right),$$

where c_δ is some positive constant that depends only on δ.

Proof. Let a $(X,\delta/2)$-normal line $L \in \mathcal{L}\setminus\mathcal{B}(X,\delta/2)$ be fixed (cf. (9)). Then, by definition,

$$|X \cap L| \sim_{\delta/2} \frac{|X|}{\sqrt{n}}.$$

Fix $1 \le i \le t$. By Chernoff's inequality, we have

$$\mathbb{P}\left\{|X \cap L_i| \sim_\delta \frac{|X|}{t\sqrt{n}}\right\} \ge 1 - \exp\left(-c_\delta \frac{|X|}{t\sqrt{n}}\right)$$

for some $c_\delta > 0$ that depends only δ. Hence

$$\mathbb{P}\,(L \text{ is not } (X, \delta, \Pi)\text{-normal}) \leq t \exp\left(-c_\delta \frac{|X|}{t\sqrt{n}}\right).$$

Since the partitions Π_L $(L \in \mathcal{L})$ are chosen independently, we have

$$\begin{aligned}\mathbb{P}\Big\{\big|\mathcal{B}(X, \delta, \Pi) \setminus \mathcal{B}(X, \delta/2)\big| \geq t^2\sqrt{n}\Big\}& \\ \leq \binom{n}{\lceil t^2\sqrt{n}\rceil}\left(t \exp\left(-c_\delta \frac{|X|}{t\sqrt{n}}\right)\right)^{\lceil t^2\sqrt{n}\rceil}& \\ \leq \exp\left(-c_\delta t|X| + \frac{1}{2}t^2\sqrt{n}\log n\right),&\end{aligned}$$

as required. □

Corollary 8 and Lemma 9 have the following immediate consequence.

Corollary 10. *Let functions $\alpha = \alpha(n)$ and $\beta = \beta(n)$ such that $\alpha\beta \gg n^{-1/2}$ be given, and suppose that a constant $\delta > 0$ is fixed. Then the probability that there is a set $X \subset V$ with $|X| \geq \alpha n$ such that*

$$|\mathcal{B}(X, \delta, \Pi)| \geq t^2\sqrt{n} + \beta n$$

is at most

$$\sum_{\alpha n \leq r \leq n} \binom{n}{r} \exp\left(-c_\delta tr + \frac{1}{2}t^2\sqrt{n}\log n\right). \tag{11}$$

□

To state the main result of this section concisely, we need to introduce some notation. Fix $\delta > 0$. Suppose $\mathbf{X} = (X_i)_{i=1}^t$ is a family of subsets $X_i \subset V$ of V, and suppose $\Pi = (\Pi_L)_{L\in\mathcal{L}}$ is a fixed family of $[t]$-indexed partitions of the lines of $\mathcal{P} = (V, \mathcal{L})$. Then we let

$$\mathcal{L}(\mathbf{X}, \delta, \Pi) = \{L \in \mathcal{L} \colon L \text{ is } (X_i, \delta, \Pi)\text{-normal for all } i \in [t]\}. \tag{12}$$

Thus, a line $L \in \mathcal{L}$ belongs to $\mathcal{L}(\mathbf{X}, \delta, \Pi)$ if and only if

$$|X_i \cap L_j| \sim_\delta \frac{|X|}{t\sqrt{n}} \tag{13}$$

for all i and $j \in [t]$. Moreover, $\mathcal{L}(\mathbf{X}, \delta, \Pi) = \mathcal{L} \setminus \bigcup_{1\leq i\leq t} \mathcal{B}(X_i, \delta, \Pi)$ (cf. (10)). When studying the subgraphs of the graph $R = R_n(\mathcal{P}, H, \Pi)$, we shall need to look at *families* of t sets of vertices $\mathbf{X} = (X_i)_{i=1}^t$, and we shall be interested in the lines that intersect these t sets in the 'expected way.' The set $\mathcal{L}(\mathbf{X}, \delta, \Pi)$ of lines of $\mathcal{P}$ defined above is the set of such 'interesting' lines.

Given constants $0 < \varepsilon < 1/2$, $\delta > 0$, and $0 < \eta < 1$, and a family $\Pi = (\Pi_L)_{L\in\mathcal{L}}$ of $[t]$-indexed partitions, we let $\mathbf{P}(\varepsilon, \delta, \eta, \Pi)$ denote the following property:

(*) *For all* $\mathbf{X} = (X_i)_{i=1}^t$ *with* $X_i \subset V$ *such that* $|X_i| \geq n^{1/2+\varepsilon}$ *for all* $i \in [t]$, *we have*

$$|\mathcal{L}(\mathbf{X}, \delta, \Pi)| \geq (1-\eta)n.$$

It is probably worth at this point to unfold the definitions to give a more direct formulation of property $\mathbf{P}(\varepsilon, \delta, \eta, \Pi)$: the family Π satisfies $\mathbf{P}(\varepsilon, \delta, \eta, \Pi)$ if and only if, for all $\mathbf{X}$ as in (*), at least $(1-\eta)n$ lines $L \in \mathcal{L}$ are such that (13) holds for all i and $j \in [t]$.

Corollary 10 has the following consequence that will be crucial in our study of the random graph $R = R_n(\mathcal{P}, H)$.

Corollary 11. *Let constants* $0 < \varepsilon \leq 1/3$, $\delta > 0$, *and* $0 < \eta < 1$ *be fixed. Suppose* $t = t(n)$ *is such that* $\log n \ll t \ll n^{\varepsilon/2}$. *Then*

$$\mathbb{P}\big\{\mathbf{P}(\varepsilon, \delta, \eta, \Pi) \text{ holds}\big\} \to 1$$

as $n \to \infty$. *In particular, if* $n \geq n_0(\varepsilon, \delta, \eta, t)$, *there is a family of* $[t]$*-indexed partitions* $\Pi = (\Pi_L)_{L \in \mathcal{L}}$ *for which* $\mathbf{P}(\varepsilon, \delta, \eta, \Pi)$ *holds.*

Proof. Let $\alpha = \alpha(n) = n^{-1/2+\varepsilon}$ and $\beta = \beta(n) = n^{-\varepsilon/2}$. Simple calculations show that

$$t^2\sqrt{n} + \beta n \ll n/t,$$

and that the quantity in (11) tends to 0 as $n \to \infty$. Thus, Corollary 10 says that, with probability tending to 0 as $n \to \infty$, every $X \subset V$ with $|X| \geq n^{1/2+\varepsilon}$ is such that

$$|\mathcal{B}(X, \delta, \Pi)| < t^2\sqrt{n} + \beta n \ll n/t.$$

In particular, given any $\mathbf{X} = (X_i)_{i=1}^t$ as in (*), we have

$$\Big|\bigcup_{1 \leq i \leq t} \mathcal{B}(X_i, \delta, \Pi)\Big| \ll n$$

with probability approaching 1 as $n \to \infty$. Thus Corollary 11 follows. $\square$

2.4. **The pseudorandomness of** $R = R_n(\mathcal{P}, H, \Pi)$**.** Fix a graph H and let $\mathcal{P}$ be a projective plane. Consider the graph $R = R_n(\mathcal{P}, H, \Pi)$ defined in Section 2.1, where $\Pi = (\Pi_L)_{L \in \mathcal{L}}$ is some family of $V(H)$-indexed partitions of the lines of $\mathcal{P}$.

Let γ denote the edge-density $e(H)\binom{t}{2}^{-1}$ of H, where as customary $e(H)$ stands for the number of edges in H. Later we shall see that, in the proofs of Theorems 3 and 4, we may assume that γ is close to $1/2$. Thus, in this section, we assume that

$$3/7 \leq \gamma \leq 4/7. \tag{14}$$

Our aim is to show that if $\mathbf{P}(\varepsilon, \delta, \eta, \Pi)$ holds for suitably small parameters ε, δ, and η, any two disjoint and reasonably large sets A and $B \subset V$ of vertices of the graph $R = R_n(\mathcal{P}, H, \Pi)$ are connected by about $\gamma|A||B|$ edges. In other words, the graph R is 'pseudorandom.'

Lemma 12. *Let $0 < \varepsilon < 1/2$ and $0 < \delta \leq 1$ be fixed. Suppose $\Pi = (\Pi_L)_{L\in\mathcal{L}}$ is a family of $V(H)$-indexed partitions of the lines of $\mathcal{P}$. Furthermore, suppose that the property $\mathbf{P}(\varepsilon, \delta/5, 3\delta/40, \Pi)$ holds and that $t \geq 40/3\delta$. Then the graph $R = R_n(\mathcal{P}, H, \Pi)$ is such that, for every A, $B \subset V(R)$ with $A \cap B = \emptyset$ and $|A|$, $|B| \geq n^{1/2+\varepsilon}$, the number of edges $e(A,B)$ between A and B satisfies*

$$e(A,B) \sim_\delta \gamma|A||B|. \tag{15}$$

Proof. Let $\mathcal{L}_{AB}$ be the set of lines $L \in \mathcal{L}$ that are both $(A, \delta/5, \Pi)$- and $(B, \delta/5, \Pi)$-normal. Since property $\mathbf{P}(\varepsilon, \delta/5, 3\delta/40, \Pi)$ holds, we know that

$$|\mathcal{L}_{AB}| \geq (1 - 3\delta/40)n. \tag{16}$$

By the definition of $\mathcal{L}_{AB}$, for any line $L \in \mathcal{L}_{AB}$ we have

$$|L_v \cap A| \sim_{\delta/5} \frac{|A|}{t\sqrt{n}} \quad \text{and} \quad |L_v \cap B| \sim_{\delta/5} \frac{|B|}{t\sqrt{n}}$$

for every $v \in V(H)$, where $L = \bigcup_{v\in V(H)} L_v$ is the partition Π_L of L. In particular, if $uv \in E(H)$, we have

$$\sim_{3\delta/5} \frac{|A||B|}{t^2 n} \tag{17}$$

edges joining the sets $L_u \cap A$ and $L_v \cap B$, and similary for $L_v \cap A$ and $L_u \cap B$. Furthermore, if $uv \notin E(H)$, the number of non-edges between $L_u \cap A$ and $L_v \cap B$, and similarly between $L_v \cap A$ and $L_u \cap B$, is also given by (17).

Therefore

$$\begin{aligned} e(A,B) \geq \sum_{L\in\mathcal{L}_{AB}} \sum_{uv\in E(H)} &\big\{|L_u \cap A||L_v \cap B| + |L_v \cap A||L_u \cap B|\big\} \\ &\sim_{3\delta/5} |\mathcal{L}_{AB}| 2\frac{|A||B|}{t^2 n}\gamma\binom{t}{2} \geq \frac{1}{1+3\delta/4}\gamma|A||B|, \end{aligned}$$

where in the last inequality we used (16) and the fact that $t \geq 40/3\delta$ and $\delta \leq 1$. Similarly, we infer that the number of non-edges between A and B is at least

$$\frac{1-\gamma}{1+3\delta/4}|A||B|.$$

Using that $\gamma \geq 3/7$ (cf. (14)), we may now deduce from straightforward computations that (15) follows. □

We observe that the proof above only made use of a weak form of property $\mathbf{P}(\varepsilon, \delta/5, 3\delta/40, \Pi)$. Indeed, this property concerns families $\mathbf{X}$ of t elements, but in the proof above we were only concerned about the *pair* of sets $\{A, B\}$.

2.5. **Blow-ups of H in $R = R_n(\mathcal{P}, H, \Pi)$.** In Section 2.4 above, we observed that the graph $R = R_n(\mathcal{P}, H, \Pi)$ tends to have all its edges uniformly spread over $V(R)$. The usual binomial random graphs $G_{n,p}$ also have this property, and hence the random graphs $R = R_n(\mathcal{P}, H)$ and $G_{n,p}$ (with $p = \gamma$) may be expected to behave similarly. However, one sees right away a crucial difference between these graphs: $R = R_n(\mathcal{P}, H)$ is extremely rich in copies of H. Indeed, each line of $\mathcal{P}$ is very likely to induce a 'very fat' blown-up copy of H. This feature will be crucial in our considerations.

In what follows, if $s \geq 1$ is an integer, we shall write H_s for the *s-fold blow-up* of H, namely, the graph on $s|V(H)|$ vertices that we obtain from H by replacing each vertex of H by an independent set of cardinality s and each edge of H by a complete bipartite graph $K^{s,s}$. The following simple lemma will be needed in the sequel.

Lemma 13. *Every red-blue edge colouring of H_s that contains no blue copy of H must contain at least s^2 red edges.*

Proof. Let us say that a copy H' of H in H_s is *transversal* if $|V(H') \cap V_x| = 1$ for all $x \in V(H)$, where V_x is the independent set of vertices of H_s that is naturally associated to $x \in V(H)$. Clearly H_s has s^t transversal copies of H. On the other hand, each edge is contained in s^{t-2} transversal copies of H. Thus, at least $s^t/s^{t-2} = s^2$ red edges are needed to prevent a transversal blue copy of H. This proves Lemma 13. $\square$

3. The General Estimate

Our aim in this section is to give a proof of Theorem 3. Since our proof is fairly long and is based on a few separate ideas, we shall sketch the general approach before we start the proof proper.

Let graphs G and H as in Theorem 3 be given. A simple argument shows that we may assume that the edge-density $e(H)\binom{t}{2}^{-1}$ of H is, say, between $3/7$ and $4/7$. As mentioned in the introduction, the main idea is to make use of the random graphs $R = R_n(\mathcal{P}, H)$, defined in Section 2.1.

The two key properties of R for the arguments in this section are the following: (i) R is a pseudorandom graph, in the sense that its edges are uniformly distributed (with density similar to the density of H), and (ii) any collection of t pairwise disjoint sets $X_1, \ldots, X_t$ of vertices of R induce many 'fat' blow-ups of H, as long as the X_i are all sufficiently large.

Suppose now that the edges of our graph R have been coloured with red and blue. Property (i) may be used to show that any appropriately large set S of vertices of R that is 'uniformly rich' in red edges must induce a red copy of G. Roughly speaking, we shall consider a set S uniformly rich in red edges if S, as well as all large subsets of S, induce sufficiently many red edges. (To be precise, we shall define the term *hereditarily ε-red-rich* in Section 3.2.1 to formalise this notion of uniform richness.)

On the other hand, property (ii) may be used to show that any collection of t suitably large sets of vertices of R is forced to induce a blue copy of H, as long as the density of the red edges across these sets is small. The proof of the fact that $R \to (G, H)$ is then reduced to showing that either R has a set of vertices that is 'uniformly rich' in red edges, or else R has t suitably large sets of vertices that induce, across them, very few red edges.

Let us now start the proof of Theorem 3.

3.1. **The main lemma.** For technical reasons, it will be convenient to prove first the following restricted version of Theorem 3.

Lemma 14. *There exist absolute constants t_0 and C for which the following holds. Let G and H be graphs with $|V(G)| = k \geq 2$ and $|V(H)| = t \geq t_0$, where $t \geq k^2$. Write $\gamma = e(H)\binom{t}{2}^{-1}$ for the edge-density of H, and suppose $3/7 \leq \gamma \leq 4/7$. Suppose furthermore that H has a proper colouring $V(H) = \bigcup_{1\leq i\leq q} V_i$ with all the V_i of cardinality $w = t/q$, where $q = \chi(H) \geq 2$. Then*

$$r_{\mathrm{ind}}(G, H) \leq t^{Ck\log q}.$$

We now sketch an argument that shows that Lemma 14 implies Theorem 3. Thus, let graphs G and H as in Theorem 3 be given. In particular, we have $t \geq k$.

Let us first observe that t may be assumed to be larger than any fixed constant. Indeed, recalling the existence of graphs $\Gamma = \Gamma(G, H)$ for which $\Gamma \to (G, H)$ for any given G and H (see [4, 9, 15]), we may exclude from our consideration any bounded number of pairs (G, H), since we may choose the constant C in (4) sufficiently large to cover this bounded number of cases. Moreover, by adding fewer than $q = \chi(H)$ isolated vertices to H, we may assume that q divides t.

Now fix a proper vertex colouring $V(H) = V_1 \cup \cdots \cup V_q$ of H. Add $t^2/q - |V_i|$ isolated vertices to V_i for all i to obtain a graph H'. Note that now we have t^2 vertices. Now, a little thought shows that we may add edges to H', each of them incident to at least one vertex in $V(H') \setminus V(H)$ and all of them respecting the q-colouring we started with, and, furthermore, in such a way that the graph H'' that we obtain is such that $\gamma = e(H'')\binom{t^2}{2}^{-1}$ lies between $3/7$ and $4/7$ (here we also use that $t \geq t_0$ for some large enough constant t_0).

Let us outline an easy argument to show that we may indeed obtain the graph H'' as required. Note that the only thing to check is whether we can accommodate the condition on the density of H''. If we add to H' all the edges allowed by the other two conditions, we get at least about

$$\left(1 - \frac{1}{q}\right)\binom{t^2}{2} - \binom{t}{2}$$

new edges. Thus, even if $q = 2$, we have at least about $\frac{1}{2}\binom{t^2}{2}$ edges if t is large. On the other hand, if we do not add any edge to H', we have at most $\binom{t}{2}$ edges out of $\binom{t^2}{2}$, which gives a density far smaller than $1/2$ if t is large. Thus the graph H'' as required does exist.

It now suffices to apply Lemma 14 to H'' to deduce that there is a graph Γ such that $\Gamma \to (G, H'')$ and such that $|\Gamma| \le (t+q-1)^{2Ck\log q}$, where C is the constant in Lemma 14. Theorem 3 follows, since clearly $\Gamma \to (G, H)$.

In the rest of Section 3, we investigate $R = R_n(\mathcal{P}, H)$ with the aim of showing that $R \to (G, H)$ with positive probability, even if n is not too large with respect to k and t. More precisely, we shall prove the following statement.

(†) *There exists an absolute constant t_0 such that, for any graphs G and H as in Lemma 14, there is an integer $n = n(k, t, q)$ with*

$$n \le 4t^{1000k\log q},$$

where $q = \chi(H)$, such that

$$R = R_n(\mathcal{P}, H) \to (G, H)$$

holds with positive probability.

We shall not try to optimise our constants. In particular, the constant 1000 in (†) is clearly not optimal. We also remark that our proof will in fact show a stronger Ramsey property of $R = R_n(\mathcal{P}, H)$. Write $\mathcal{G}^k$ for the family of all graphs on k vertices. Our proof of (†) will in fact give that

$$R = R_n(\mathcal{P}, H) \to (\mathcal{G}^k, H),$$

where the arrow notation above means that, when we colour the edges of R with colours red and blue, either we obtain a blue induced copy of H, or else we obtain a red induced copy of *every* graph on k vertices.

3.2. **Monochromatic subgraphs in $R = R_n(\mathcal{P}, H, \Pi)$.** Throughout this section, we assume that graphs G and H as in the statement of Lemma 14 are given.

Let us fix a projective plane $\mathcal{P} = (V, \mathcal{L})$ with $|V| = |\mathcal{L}| = n$. Let us also fix a family $\Pi = (\Pi_L)_{L\in\mathcal{L}}$ of $V(H)$-indexed partitions of the lines of $\mathcal{P}$. We may thus consider the graph $R = R(\mathcal{P}, H, \Pi)$ defined in Section 2.1. Suppose the edges of $R = R(\mathcal{P}, H, \Pi)$ are coloured with colours red and blue.

In Section 3.2.1 below, we describe a situation in which our red-blue coloured graph R may be guaranteed to contain an induced red copy of G. In Section 3.2.2, we study a condition that forces blue induced copies of H in R.

In both sections that follow, we shall assume that Π is fixed. In particular, our results will be deterministic and not probabilistic. For our arguments to work, we shall impose certain hypotheses on Π. Later on in the proof of (†), we shall invoke Corollary 11 to verify that these conditions imposed on Π are met with positive probability if Π is chosen randomly as specified in the definition of the random graph $R = R_n(\mathcal{P}, H)$.

3.2.1. *Red induced copies of G.* Fix $0 < \varepsilon < 1/2$. A set $S \subset V(R)$ is *hereditarily ε-red-rich*, or *ε-HRR* for short, if $|S| \geq n^{1/2+\varepsilon}t^{10k}$ and if, for every disjoint pair of sets A, $B \subset S$ with $|A|$, $|B| \geq |S|/kt^{10k}$, we have

$$e_{\text{red}}(A,B) \geq \frac{1}{10t^2}|A||B|. \tag{18}$$

Here and below, $e_{\text{red}}(A,B)$ denotes the number of red edges between A and B. We shall also write $E_{\text{red}}(A,B)$ for the set of red edges between A and B. Moreover, $N(x) = N_R(x)$ denotes the neighbourhood of the vertex x in R and $N_{\text{red}}(x)$ denotes the set of neighbours of x in R that are joined to x by red edges.

In the beginning of Section 3, we mentioned that a notion of 'uniform richness' in red edges would appear in our arguments. The property of ε-red-richness is the precise formulation of this notion. As inequality (18) may suggest, roughly speaking, t^{-2} will be the 'cut-off' density in our argument: if we find a (large) set S of vertices in our red-blue edge-coloured graph that has red-edge density $\Omega(t^{-2})$ (in fact, not only S but if all its large subsets has this red-edge density), then we search for a red induced copy of G within S. On the other hand, if no such set exists, we shall seek a blue induced copy of H.

Our next lemma, Lemma 15, states that ε-HRR sets S do indeed contain red induced copies of G, as long as some technical conditions are met. These technical conditions are basically to guarantee that our graph $R = R_n(\mathcal{P}, H, \Pi)$ satisfies a suitable pseudorandomness property (cf. Lemma 12). The underlying idea in the proof of Lemma 15 below is a well known one, and the reader is referred to Chvátal, Rödl, Szemerédi, and Trotter [3] and Rödl [16] for former applications of this idea.

Finally, we observe that in later applications of Lemma 15 below, we shall have $\varepsilon = 1/10$.

Lemma 15. *Let $0 < \varepsilon < 1/2$ be fixed, and assume $\Pi = (\Pi_L)_{L\in\mathcal{L}}$ is a family of $V(H)$-indexed partitions of the lines of $\mathcal{P}$ for which $\mathbf{P}(\varepsilon, 1/30, 1/80, \Pi)$ holds. Moreover, suppose $t \geq 80$. Then the graph $R = R_n(\mathcal{P}, H, \Pi)$ is such that every ε-HRR set $S \subset V(R)$ contains an induced red copy of G.*

Proof. Observe first that, since $\mathbf{P}(\varepsilon, 1/30, 1/80, \Pi)$ holds, Lemma 12 implies that

$$e(A,B) \sim_{1/6} \gamma|A||B| \tag{19}$$

holds any pair of disjoint sets A, $B \subset V(R)$ with $|A|$, $|B| \geq n^{1/2+\varepsilon}$.

Let $S_i \subset S$ $(1 \leq i \leq k)$ be k pairwise disjoint subsets of S, each with cardinality $|S_i| = \lfloor |S|/k \rfloor$. Suppose $V(G) = \{v_1, \dots, v_k\}$. We shall find $s_i \in S_i$ $(1 \leq i \leq k)$ so that $\{s_i, s_j\} \in E(R)$ if and only if $\{v_i, v_j\} \in E(G)$ and, moreover, so that all such edges $\{s_i, s_j\}$ are coloured red.

Applying (19) to the pairs of sets (S_1, S_j), where $2 \leq j \leq k$, we may deduce that there are fewer than $(k-1)n^{1/2+\varepsilon}$ many vertices $s \in S_1$ such that $|N(s) \cap S_j| > (2/3)|S_j|$ for some $j = 2, \dots, k$. Furthermore, there are fewer than $(k-1)|S|/kt^{10k}$ vertices $s \in S_1$ with $|N_{\rm red}(s) \cap S_j| < |S_j|/10t^2$ for some $j = 2, \dots, k$. This follows since S is ε-HRR. Thus, since

$$|S_1| > \frac{(k-1)|S|}{kt^{10k}} + (k-1)n^{1/2+\varepsilon},$$

there exist a vertex $s_1 \in S_1$ and sets $S_\ell^{(1)}$ $(1 < \ell \le k)$ with $|S_\ell^{(1)}| = \lceil |S_\ell|/10t^2 \rceil$ so that, for $s \in S_\ell^{(1)}$, if $\{v_1, v_\ell\} \in E(G)$ then $\{s_1, s\} \in E_{\rm red}(R)$ and if $\{v_1, v_\ell\} \notin E(G)$ then $\{s_1, s\} \notin E(R)$. We repeat this procedure $k-1$ times.

On the jth step $(2 \le j \le k)$ we construct sets $S_\ell^{(j)}$ $(j < \ell \le k)$, each of cardinality $\lceil |S_\ell|/(10t^2)^j \rceil$, and select a vertex $s_j \in S_j^{(j-1)}$ such that for $s \in S_\ell^{(j)}$ $(j < \ell \le k)$ we have that if $\{v_j, v_\ell\} \in E(G)$ then $\{s_j, s\} \in E_{\rm red}(R)$ and if $\{v_j, v_\ell\} \notin E(G)$ then $\{s_j, s\} \notin E(R)$. The existence of such a vertex s_j follows similarly as before since

$$|S_j^{(j-1)}| = \left\lceil \frac{|S_j|}{(10t^2)^{j-1}} \right\rceil > \frac{(k-j)|S|}{kt^{10k}} + (k-j)n^{1/2+\varepsilon}.$$

The inductive procedure above constructs a red induced copy of G in R, as required. ☐

Note that the proof above in fact shows that, under the conditions given in the lemma, an ε-HRR set of vertices in R is *red k-universal*, by which we mean that it induces a red copy of *any* graph on k vertices.

3.2.2. *Blue induced copies of H.* We may now turn to the lemma that gives a sufficient condition for the existence of a blue induced copy of H in R. Let us recall that $q = \chi(H)$ and that $w = t/q$. In the sequel, if s is a positive integer, we shall write $\mathrm{Perm}(s)$ for the set of permutations of the set $[s] = \{1, \dots, s\}$.

Perhaps it is worth observing that, in Lemma 16 below, the hypothesis on the number of edges 'across' the sets A_i guarantees that the 'density of red edges' across these sets is at most $1/5t^2$. It is this small density of red edges that will allow us to find induced blue copies of H through an application of Lemma 13.

A comparison between Lemmas 15 and 16 suggests that, roughly speaking, t^{-2} should be a good 'cut-off' density for our arguments (cf. the remark following inequality (18) concerning this 'cut-off' density.)

In the proof of Theorem 3, we shall be concerned with property $\mathbf{P}(\varepsilon, \delta, \eta, \Pi)$ with $\varepsilon = 1/10$, $\delta = 1/30$, and $\eta = 1/80$. However, Lemma 16 below is stated with a weaker (larger δ and η) and more general (arbitrary ε) hypothesis on Π.

Lemma 16. *Let $0 < \varepsilon \le 1/3$ be fixed and suppose $\Pi = (\Pi_L)_{L \in \mathcal{L}}$ is a family of $V(H)$-indexed partitions of the lines of $\mathcal{P}$. If $\mathbf{P}(\varepsilon, 1/3, 1/2, \Pi)$ holds, then $R = R(\mathcal{P}, H, \Pi)$ has the following property. Let $A_1, \dots, A_q$ be pairwise* *disjoint subsets of vertices of* R *with* $|A_1| = \cdots = |A_q| = m$*, where* $m \geq w(n^{1/2+\varepsilon}+1)$*. Assume moreover that*

$$\sum_{1\leq i<j\leq q} e_{\mathrm{red}}(A_i, A_j) \leq \frac{1}{5t^2}\sum_{1\leq i<j\leq q}|A_i||A_j| = \frac{1}{5t^2}\binom{q}{2}m^2. \tag{20}$$

Then there exist $v_i \in \bigcup_{1\leq j\leq q} A_j$ $(1 \leq i \leq t)$ *that induce a blue copy of* H *in* R*.*

Proof. Recall we have a fixed proper vertex colouring $V(H) = \bigcup_{1\leq i\leq q} V_i$ of H with $|V_1| = \cdots = |V_q| = w = t/q$. We write $V_i = \{v_{ij}\colon 1 \leq j \leq w\}$ $(1 \leq i \leq q)$. Let L be a line of our projective plane. The labelling v_{ij} of the vertices of H gives a natural labelling of the members L_v of the partition $L = \bigcup_v L_v$ of L in Π. Namely, we let the partition classes of L be denoted by L_{ij} $(1 \leq i \leq q,\ 1 \leq j \leq w)$ in the natural way. Finally, for each $1 \leq i \leq q$, choose w pairwise disjoint subsets $A_{ij} \subset A_i$ $(1 \leq j \leq w)$ with $|A_{ij}| = \lfloor m/w\rfloor \geq n^{1/2+\varepsilon}$ for each j.

Note that a line L is (A_{ij}, δ, Π)-normal if we have

$$|L_{i'j'} \cap A_{ij}| \sim_\delta \frac{|A_{ij}|}{t\sqrt{n}} = \frac{\lfloor m/w\rfloor}{qw\sqrt{n}} \tag{21}$$

for all $1 \leq i' \leq q$, $1 \leq j' \leq w$. Let $\delta = 1/3$. In this proof, we shall make use of the lines L that are (A_{ij}, δ, Π)-normal for *all* the t sets A_{ij} $(1 \leq i \leq q$, $1 \leq j \leq w)$. Write $\mathcal{L}_\delta$ for the set of all such lines. Then, clearly, we have

$$\mathcal{L}_\delta = \mathcal{L}(\mathbf{A}, \delta, \Pi),$$

where $\mathbf{A} = (A_{ij})_{1\leq i\leq q, 1\leq j\leq w}$ (cf. (12)). Since we are assuming that property $\mathbf{P}(\varepsilon, 1/3, 1/2, \Pi)$ holds, we know that

$$|\mathcal{L}_\delta| \geq n/2. \tag{22}$$

We now consider $(q+1)$-tuples of permutations $\bar\sigma = (\sigma_0; \sigma_1, \ldots, \sigma_q)$, where $\sigma_0 \in \mathrm{Perm}(q)$ and $\sigma_i \in \mathrm{Perm}(w)$ $(1 \leq i \leq q)$. In the sequel, if $(i,j) \in [q]\times[w]$, we write $\bar\sigma(i,j)$ for the pair $(a,b) \in [q]\times[w]$ where $a = \sigma_0(i)$, and $b = \sigma_a(j) = \sigma_{\sigma_0(i)}(j)$.

Now suppose $L \in \mathcal{L}_\delta$. Define $f(L,\bar\sigma)$ to be the number of red edges across the sets $A_{ij} \cap L_{\bar\sigma(i,j)}$, where a red edge is counted only if its endvertices belong to two such sets with distinct i.

We now average over $L \in \mathcal{L}_\delta$ and $\bar\sigma$. We have

$$\sum_{L\in\mathcal{L}_\delta}\sum_{\bar\sigma} f(L,\bar\sigma) \leq \sum_{L\in\mathcal{L}}\sum_{\bar\sigma} f(L,\bar\sigma). \tag{23}$$

The sum on the right-hand side of (23) above is, however,

$$(q-2)!(w-1)!^2 w!^{q-2}$$

times the number of red edges across the A_i (*i.e.*, having endvertices in distinct A_i). To see this, let us fix one such edge e; say, $e = xy$ and $x \in A_{ij}$ and $y \in A_{i'j'}$. Note for later reference that $i \neq i'$. The vertices x and y

determine a unique line $L \in \mathcal{L}$. It now suffices to notice that the values of $i \neq i'$, j, and j' tells us that, out of all the $q!w!^q$ possibilities, exactly $(q-2)!(w-1)!^2 w!^{q-2}$ vectors $\bar{\sigma} = (\sigma_0; \sigma_1, \dots, \sigma_q)$ are such that

$$x \in A_{ij} \cap L_{\bar{\sigma}(i,j)} \quad \text{and} \quad y \in A_{i'j'} \cap L_{\bar{\sigma}(i',j')}.$$

Therefore, we may indeed deduce that the left-hand side of (23) above is, by (20), at most

$$(q-2)!(w-1)!^2 w!^{q-2} \frac{1}{5q^2w^2} \sum_{1 \le i < j \le q} |A_i||A_j| \\ \le (q-2)! \frac{w!^q}{5q^2w^4} \binom{q}{2} m^2 = \frac{q!w!^q}{10q^2w^4} m^2.$$

Since (22) holds and there are $q!w!^q$ many $\bar{\sigma}$, there exists a pair $(L, \bar{\sigma})$ with $L \in \mathcal{L}_\delta$ satisfying

$$f(L, \bar{\sigma}) \le \frac{m^2}{5nq^2w^4}. \tag{24}$$

Fix such a pair $(L, \bar{\sigma})$. Recall (21), and pick a set $X_{ij} \subset A_{ij} \cap L_{\bar{\sigma}(i,j)}$ with $|X_{ij}| = s = \lceil (3/4)m/qw^2\sqrt{n} \rceil$ for each $1 \le i \le q$, $1 \le j \le w$. Note that the t sets X_{ij} induce in R a copy of H_s, the s-fold blow-up of H. We have $m \le (4/3)sqw^2\sqrt{n}$. Hence, we have that the number of red edges in this copy of H_s is at most $f(L, \bar{\sigma}) < s^2$, by (24). It now suffices to apply Lemma 13 to H_s to deduce that there is an induced blue copy of H in R, as required. □

3.3. **The proof of Theorem 3.** As observed in Section 3.1, it suffices to prove Lemma 14 to prove Theorem 3. Lemma 14 is, however, a consequence of assertion (†) in Section 3.1, and hence we are left with proving that assertion.

Proof of (†). We fix $\varepsilon = 1/10$ and pick a prime p such that $n = p^2 + p + 1$ satisfies

$$t^{Ck \log q} \le n \le 4t^{Ck \log q},$$

where $C = 1000$. The existence of such a prime p follows easily from Chebyshev's theorem. Observe now that

$$\log n \ll t \ll n^{\varepsilon/2}. \tag{25}$$

Indeed, the lower bound on t follows since $t \ge k^2$, and the upper bound follows by the choice of C. Let us now observe that, quite crudely,

$$n^{0.4} \ge t^{1+100k(\log_2 q + 1)}. \tag{26}$$

Moreover, note that there is an absolute constant t_0 such that if $t \ge t_0$, then $n \ge n_0(1/10, 1/30, 1/80, t)$, where n_0 is as given in Corollary 11. Clearly, we may assume that $t_0 \ge 80$. We shall show that this choice of t_0 will do in assertion (†).

Suppose graphs G and H as in (†) are given, and suppose that $t \geq t_0$. Let $\mathcal{P} = (V, \mathcal{L})$ be a projective plane of order p. Then $|V| = |\mathcal{L}| = n = p^2 + p + 1$. Recall (25), and invoke Corollary 11 to see that there is a family of $V(H)$-indexed partitions $\Pi = (\Pi_L)_{L \in \mathcal{L}}$ of the lines of $\mathcal{P}$ for which property $\mathbf{P}(\varepsilon, 1/30, 1/80, \Pi) = \mathbf{P}(1/10, 1/30, 1/80, \Pi)$ holds. Fix such a Π, and let $R = R_n(\mathcal{P}, H, \Pi)$. We claim that $R \to (G, H)$. Clearly, this will finish the proof of (†).

To prove our claim, we invoke Lemmas 15 and 16. Since $t \geq t_0 \geq 80$ and property $\mathbf{P}(1/10, 1/30, 1/80, \Pi)$ holds, according to Lemma 15, we may assume that $V(R)$ contains no ε-HRR subset. Indeed, according to that lemma, if there is an ε-HRR set of vertices in R, we must have a red induced copy of G in R. In the remainder of the proof, we show that the existence of no such ε-HRR subset of vertices implies the existence of sets A_i $(1 \leq i \leq q)$ as required in Lemma 16. An application of that lemma will then give us a blue induced copy of H, completing the proof of (†).

To carry out the plan above, we shall need the following technical lemma. Let $\{0,1\}^\ell$ denote the set of all 0–1 strings of length ℓ.

(‡) *Let $\ell = \lceil \log_2 q \rceil$. There are 2^ℓ pairwise disjoint sets $X_I \subset V(R)$ $(I \in \{0,1\}^\ell)$, each of cardinality $m = \lceil n/t^{100\ell k} \rceil$, satisfying*

$$\sum\nolimits_{I \neq J}^{(\ell)} e_{\mathrm{red}}(X_I, X_J) < \left(\frac{1}{1 - kt^{-90k}} \right)^{2(\ell-1)} \frac{1}{10t^2} \sum\nolimits_{I \neq J}^{(\ell)} |X_I||X_J|, \tag{27}$$

where $\sum_{I \neq J}^{(\ell)}$ indicates sum over all I, $J \in \{0,1\}^\ell$ with $I \neq J$.

We postpone the proof of (‡), and carry on with the proof of (†). Assertion (‡) gives us 2^ℓ sets X_I satisfying (27), where $\ell = \lceil \log_2 q \rceil$. In particular, we have $2^\ell \geq q$. A simple averaging argument now gives that some q of the sets X_I, say A_i $(1 \leq i \leq q)$, are such that the number of red edges across these q sets is at most the quantity given in (20), as required in Lemma 16. Recall that property $\mathbf{P}(1/10, 1/30, 1/80, \Pi)$ holds. Moreover, note that the sets A_i have cardinality $m = \lceil n/t^{100\ell k} \rceil \geq w(n^{0.6} + 1)$. It follows from Lemma 16 that there must be a blue induced copy in R. This completes the proof of the claim that $R \to (G, H)$, and the proof of (†) will be complete once we establish (‡). □

Proof of (‡). We prove the existence of the sets X_I $(\ell \in \{0,1\}^\ell)$ by induction on ℓ. This is simple for $\ell = 1$: since $V(R)$ is not ε-HRR, there are sets X_0', $X_1' \subset V(R)$ with $|X_i'| \geq n/kt^{10k}$ $(i \in \{0,1\})$ so that

$$e_{\mathrm{red}}(X_0', X_1') < \frac{1}{10t^2} |X_0'||X_1'|. \tag{28}$$

An easy averaging argument shows that there are in fact $X_0 \subset X_0'$, $X_1 \subset X_1'$ with $|X_0| = |X_1| = \lceil n/t^{100k} \rceil < n/kt^{10k}$ such that (28) remains true if X_i' is replaced X_i $(i \in \{0,1\})$. We now turn to the induction step.

Assume that, for some $1 \leq \ell < \log_2 q = \log_2 \chi(H)$, we have constructed 2^ℓ sets X_J $(J \in \{0,1\}^\ell)$ with $|X_J| = \lceil n/t^{100\ell k} \rceil$ such that inequality (27) holds.

We checked above that we do have this situation for $\ell = 1$. Our aim now is to show that there exist sets X_I ($I \in \{0,1\}^{\ell+1}$) satisfying condition (27) with ℓ replaced by $\ell+1$ and with $|X_I| = \lceil n/t^{100(\ell+1)k} \rceil$ for all I.

Let

$$u = \left\lfloor \frac{1}{2}(t^{100k} - kt^{10k}) \right\rfloor .$$

We now prove the following assertion.

(‡‡) *Each X_J ($J \in \{0,1\}^{\ell}$) contains $u+1$ pairs of disjoint sets $Y_{J,2i}$, $Y_{J,2i+1} \subseteq X_J$ ($0 \le i \le u$) such that*

$$|Y_{J,2i}| = |Y_{J,2i+1}| = \left\lceil \frac{n}{t^{100(\ell+1)k}} \right\rceil \tag{29}$$

and

$$e_{\mathrm{red}}(Y_{J,2i}, Y_{J,2i+1}) < \frac{1}{10t^2} |Y_{J,2i}||Y_{J,2i+1}|. \tag{30}$$

Let us prove (‡‡). We fix $J \in \{0,1\}^{\ell}$ and construct the $Y_{J,j}$ ($0 \le j \le 2u+1$) by induction, using the fact that no subset of X_J is ε-HRR.

Since each X_J has cardinality larger than $n^{1/2+\varepsilon}t^{10k}$ but is not ε-HRR, we can find set $Y_{J,0}$, $Y_{J,1}$ satisfying (29) and such that (30) holds. We now proceed inductively. Suppose the sets $Y_{J,2(i-1)}$, $Y_{J,2(i-1)+1}$ ($1 \le i \le u$) have been found. Put $Z_J = X_J \setminus \bigcup_{0 \le j \le 2u-1} Y_{J,j}$. Observe that

$$|Z_J| = |X_J| - \sum_{0 \le j \le 2u-1} |Y_{J,j}| \ge kt^{10k} \frac{n}{2t^{100k(\ell+1)}},$$

where the last inequality is in fact quite crude. Since $\ell < \log_2 q$ and (26) holds, we have

$$kt^{10k} \frac{n}{2t^{100k(\ell+1)}} \ge kt^{10k} \frac{n}{2t^{100k(\log_2 q+1)}} \ge n^{0.6}t^{10k} = n^{1/2+\varepsilon}t^{10k}.$$

Hence, the fact that Z_J is not ε-HRR implies that there are two sets $Y_{J,2i}$, $Y_{J,2i+1}$ satisfying (29) and (30). The existence of the sets $Y_{J,j}$ ($0 \le j \le 2u+1$) follows by induction. Since J was arbitrary, assertion (‡‡) is proved.

Let us now proceed with the proof of (‡). For all $J \in \{0,1\}^{\ell}$, set

$$Y_J = \bigcup_{i=0}^{2u+1} Y_{J,i}.$$

Then we have that

$$\begin{aligned} |Y_J| &\ge 2(u+1) \left\lceil \frac{n}{t^{100(\ell+1)k}} \right\rceil \\ &\ge (t^{100k} - kt^{10k}) \left\lceil \frac{n}{t^{100(\ell+1)k}} \right\rceil \ge \left(1 - \frac{k}{t^{90k}}\right) |X_J|. \end{aligned}$$

We may therefore infer from (27) that

$$\begin{aligned}\sum\nolimits_{I\neq J}^{(\ell)} e_{\mathrm{red}}(Y_I,Y_J) &< \sum\nolimits_{I\neq J}^{(\ell)} e_{\mathrm{red}}(X_I,X_J)\\ &\leq \left(\frac{1}{1-kt^{-90k}}\right)^{2(\ell-1)}\frac{1}{10t^2}\left(\frac{1}{1-kt^{-90k}}\right)^2\sum\nolimits_{I\neq J}^{(\ell)}|Y_I||Y_J|.\end{aligned}$$

Now note that

$$e_{\mathrm{red}}(Y_I,Y_J)=\sum_{i,j} e_{\mathrm{red}}(Y_{I,2i}\cup Y_{I,2i+1},Y_{J,2j}\cup Y_{J,2j+1}),$$

where the sum is taken over $0\leq i\leq u$ and $0\leq j\leq u$. It follows from an averaging argument that, for each $I\in\{0,1\}^\ell$, there exists r_I $(0\leq r_I\leq u)$ so that

$$\begin{aligned}&\sum\nolimits_{I\neq J}^{(\ell)} e_{\mathrm{red}}(Y_{I,2r_I}\cup Y_{I,2r_I+1},Y_{J,2r_J}\cup Y_{J,2r_J+1})\\ &\quad<\left(\frac{1}{1-kt^{-90k}}\right)^{2\ell}\frac{1}{10t^2}\sum\nolimits_{I\neq J}^{(\ell)}|Y_{I,2r_I}\cup Y_{I,2r_I+1}||Y_{J,2r_J}\cup Y_{J,2r_J+1}|.\end{aligned}$$

For each $I\in\{0,1\}^\ell$, set $X_{I,0}=Y_{I,2r_I}$, $X_{I,1}=Y_{I,2r_I+1}$. In this way we get a collection of $2^{\ell+1}$ sets X_L $(L\in\{0,1\}^{\ell+1})$ with $|X_L|=\lceil n/t^{100(\ell+1)k}\rceil$ for all L and such that

$$\sum\nolimits_{L\neq L'}^{(\ell+1)} e_{\mathrm{red}}(X_L,X_{L'})<\left(\frac{1}{1-kt^{-90k}}\right)^{2\ell}\frac{1}{10t^2}\sum\nolimits_{L\neq L'}^{(\ell+1)}|X_L||X_{L'}|.$$

This finishes the induction step and (‡) is proved. □

4. Erdős–Hajnal Simple Graphs

Our aim in this section is to prove Theorem 4. We shall again use the random graphs $R=R_n(\mathcal{P},H)$ introduced in Section 2.1.

We start by observing that we may assume that the edge-density $\gamma=e(H)\binom{t}{2}^{-1}$ of the graph H in the statement of Theorem 4 satisfies $3/7\leq\gamma\leq 4/7$. We shall make use of this assumption towards the end of this section. Furthermore, we may assume that t is larger than some suitably large constant so that the inequalities below hold.

The proof of Theorem 4 will be by induction on the number of vertices of G. However, to make the induction work, we shall need to boost up the claim in that theorem.

4.1. **A stronger version of Theorem 4.** For graphs Γ, G, and H, and a real number $M>0$, let us write $\Gamma\overset{M}{\to}(G,H)$ if in any red-blue colouring of the edges of Γ, there are either at least M red induced copies of G, or else there is at least one blue induced copy of H.

Suppose a projective plane $\mathcal{P}=(V,\mathcal{L})$ with $|\mathcal{L}|=|V|=n$ and a t-vertex graph H are fixed. Consider the graph $R=R_n(\mathcal{P},H,\Pi)$, where, as usual, $\Pi=(\Pi_L)_{L\in\mathcal{L}}$ is some family of $V(H)$-indexed partitions of the lines L of $\mathcal{P}$.

Let a simple graph G on k vertices be fixed. We shall show below that, if $n \geq t^f$ for some large enough $f = f(k)$ and property $\mathbf{P}(\varepsilon, \delta, \eta, \Pi)$ holds for small enough ε, δ, and $\eta > 0$, then

$$R \stackrel{M}{\to} (G, H),$$

where $M = n^{(1-\varrho)k}$, for any constant $\varrho > 0$ that is fixed in advance. It turns out that the assertion above is still not enough to make the induction work. Indeed, we need a 'local' version of the assertion above. In the result below, as customary, if $X \subset V$, we write $R[X]$ for the subgraph of R induced by X.

Lemma 17. *Let constants ε and $\varrho > 0$ and a k-vertex simple graph G be given. Then there exist constants $f = f(\varepsilon, \varrho, G)$ and $\widetilde{\varepsilon} = \widetilde{\varepsilon}(\varepsilon, \varrho, G) > 0$ for which the following holds. Let a t-vertex graph H be given, where $t \geq 80$, and suppose $\mathcal{P} = (V, \mathcal{L})$ is a projective plane on $n \geq t^f$ points. Consider the graph $R = R_n(\mathcal{P}, H, \Pi)$, where $\Pi = (\Pi_L)_{L \in \mathcal{L}}$ is such that $\mathbf{P}(\widetilde{\varepsilon}, 1/30, 1/80, \Pi)$ holds. Then, for all $X \subset V$ with $|X| \geq n^{1/2+\varepsilon}$, we have*

$$R[X] \stackrel{M}{\to} (G, H), \tag{31}$$

where $M = |X|^{(1-\varrho)k}$.

Proof. The proof will be by induction on $k = |V(G)|$. If $k = 1$, there is nothing to do, as (31) holds trivially in this case for any $\varrho > 0$. Thus, let us assume that $k \geq 2$ and that our lemma holds for smaller values of k. Let ε, $\varrho > 0$, and a k-vertex simple graph G be given. Then $G = G_1 \dot\cup G_2$ or else $G = G_1 \vee G_2$ for a pair of graphs (G_1, G_2) with G_i of order $k_i < k$ ($i = 1$, 2). Let us consider the case in which G is the join of the graphs G_1 and G_2. The other case is similar and a little simpler. We shall make a couple of comments about that case later on.

To find the constants f and $\widetilde{\varepsilon}$ as required in the lemma, we need to make a few numerical considerations. Let $\eta = 1/2 + \varepsilon$.

(††) *There exist constants η', η'', and $\widetilde{\varrho}$ such that $1/2 < \eta'' < \eta' < \eta$, $0 < \widetilde{\varrho} < \varrho$, and*

(*i*) $\eta'' < \eta' + \eta'(1 - \widetilde{\varrho})k_2 - \eta k_2$,

(*ii*) $\eta''(1 - \widetilde{\varrho})k_1 + \eta'(1 - \widetilde{\varrho})k_2 + \eta' - \eta > \eta(1 - \varrho)k$.

We postpone the proof of (††), and carry on with the proof of Lemma 17. Fix η', η'', and $\widetilde{\varrho}$ as in the assertion above. Let $\eta' = 1/2 + \varepsilon'$ and $\eta'' = 1/2 + \varepsilon''$. Clearly, ε', $\varepsilon'' > 0$. Let us apply the induction hypothesis to the triples $(\varepsilon'', \widetilde{\varrho}, G_1)$ and $(\varepsilon', \widetilde{\varrho}, G_2)$. Let $f(\varepsilon'', \widetilde{\varrho}, G_1)$, $f(\varepsilon', \widetilde{\varrho}, G_2)$, $\widetilde{\varepsilon}(\varepsilon'', \widetilde{\varrho}, G_1)$, and $\widetilde{\varepsilon}(\varepsilon', \widetilde{\varrho}, G_2)$ be the constants whose existence is guaranteed by the induction hypothesis. Put

$$f = f(\varepsilon, \varrho, G) = \max\left\{\frac{2}{\varepsilon - \varepsilon'}, f(\varepsilon'', \widetilde{\varrho}, G_1), f(\varepsilon', \widetilde{\varrho}, G_2)\right\} \tag{32}$$

and

$$\widetilde{\varepsilon} = \widetilde{\varepsilon}(\varepsilon, \varrho, G) = \min\{\varepsilon, \widetilde{\varepsilon}(\varepsilon'', \widetilde{\varrho}, G_1), \widetilde{\varepsilon}(\varepsilon', \widetilde{\varrho}, G_2)\}. \tag{33}$$

We claim that the above choice for f and $\widetilde{\varepsilon}$ will do. To prove our claim, let a t-vertex graph H be given, where $t \geq 80$, let $\mathcal{P} = (V, \mathcal{L})$ be a projective plane on $n \geq t^f$ points, and suppose $\Pi = (\Pi_L)_{L \in \mathcal{L}}$ satisfies $\mathbf{P}(\widetilde{\varepsilon}, 1/30, 1/80, \Pi)$. Moreover, fix $X \subset V$ with $|X| \geq n^{\eta} = n^{1/2+\varepsilon}$. We shall now prove that (31) holds for $M = |X|^{(1-\varrho)k}$. A simple averaging argument shows that it suffices to consider the case in which $|X| = n^{\eta} = n^{1/2+\varepsilon}$. (For simplicity, in this proof we ignore ceiling and floor functions.)

Let $R[X]$ have its edges coloured red and blue, and suppose that there is no blue induced of H in $R[X]$. We must find at least M red induced copies of G in $R[X]$. Put $\mathcal{L}(X, 1, \Pi) = \{L \in \mathcal{L} \colon L \text{ is } (X, 1, \Pi)\text{-normal}\}$. Then, since $\mathbf{P}(\varepsilon, 1, 1/2, \Pi)$ clearly holds, we have

$$|\mathcal{L}(X, 1, \Pi)| \geq n/2. \tag{34}$$

Let $s = |X|/2t\sqrt{n}$. Each line L in $\mathcal{L}(X, 1, \Pi)$ induces an s-fold blow-up H_s of s. Thus, by Lemma 13, each such line must induce at least s^2 red edges. Using (34), we see that $R[X]$ must contain at least

$$\frac{1}{2} n s^2 \geq \frac{|X|^2}{8t^2}$$

red edges. Write $d^X_{\mathrm{red}}(x)$ for the number of red edges in $R[X]$ that are incident to the vertex $x \in X$. A simple argument shows that the number of vertices $x \in X$ such that $d^X_{\mathrm{red}}(x) \geq |X|/8t^2$ is at least $|X|/8t^2$. Let us fix one such vertex x, and let us look at the set $X' = N_{\mathrm{red}}(x) \cap X$ of vertices in X that are joined to x by red edges.

Since $f \geq 2/(\varepsilon - \varepsilon')$, we have that $|X'| \geq |X|/8t^2 \geq n^{\eta'}$. Since $f \geq f(\varepsilon', \widetilde{\varrho}, G_2)$ and $\widetilde{\varepsilon} \geq \widetilde{\varepsilon}(\varepsilon', \widetilde{\varrho}, G_2)$, we may deduce that $R[X']$ induces at least $n^{(1-\widetilde{\varrho})\eta' k_2}$ red copies of G_2. Thus, the total number of red induced copies of $K^1 \vee G_2$ within $R[X]$ is at least

$$\frac{|X|}{8t^2} n^{(1-\widetilde{\varrho})\eta' k_2} \geq n^{\eta' + (1-\widetilde{\varrho})\eta' k_2}.$$

Let us say that a red induced copy of G_2 in $R[X]$ is *good* if it belongs to at least $n^{\eta''}$ red induced copies of $K^1 \vee G_2$ in $R[X]$. Again by a simple argument, using inequality (*ii*) above, we may deduce that we have at least

$$\frac{1}{2} n^{(1-\widetilde{\varrho})\eta' k_2 + \eta' - \eta} \tag{35}$$

good red induced copies of G_2 in $R[X]$. Fix one such good copy, and let us look at the set $X'' \subset X$ of vertices x'' that, together with our good copy of G_2, induce a red $K^1 \vee G_2$. We have $|X''| \geq n^{\eta''}$. Since $f \geq f(\varepsilon'', \widetilde{\varrho}, G_1)$ and $\widetilde{\varepsilon} \geq \widetilde{\varepsilon}(\varepsilon'', \widetilde{\varrho}, G_1)$, we may deduce that $R[X'']$ contains at least $n^{(1-\widetilde{\varrho})\eta'' k_1}$ red induced copies of G_1. Thus, our good copy of G_2 in $R[X]$ belongs to at least $n^{(1-\widetilde{\varrho})\eta'' k_1}$ red induced copies of G. Taking into account that the number of such good copies of G_2 in $R[X]$ is at least as large as the quantity in (35), we deduce that the number of red induced copies of G in $R[X]$ is at least

$$\frac{1}{2}n^{(1-\widetilde{\varrho})\eta'' k_1+(1-\widetilde{\varrho})\eta' k_2+\eta'-\eta}.$$

Recalling (i), we see that this quantity is no smaller than $n^{(1-\varrho)k\eta}$. This completes the induction step in the case in which $G = G_1 \vee G_2$.

Let us now briefly look at the case in which $G = G_1 \dot{\cup} G_2$. The argument here is similar, except that we have to make use of the fact that $R[X]$ contains no pair (A, B) of fairly large disjoint sets with $e(A, B) > (2/3)|A||B|$. This follows since $\gamma \leq 4/7$, and inequality (15) holds with $\delta = 1/6$ for any A, $B \subset X$ with $|A|$, $|B| \geq n^{1/2+\varepsilon}$, since $\mathbf{P}(\varepsilon, 1/30, 1/80, \Pi)$ holds. We omit the details. ☐

It now remains to prove (††).

Proof of (††). First we choose η'' such that

$$\eta > \eta'' > 1/2 \tag{36}$$

and $\eta'' k_1 > \eta k_1 - \varrho\eta k$. This second inequality is equivalent to

$$\eta'' k_1 + \eta k_2 > \eta k(1-\varrho). \tag{37}$$

Inequalities (36) and (37) above imply that it is possible to choose an η' close enough to η, with $\eta'' < \eta' < \eta$, ensuring that

$$(\eta' - \eta)k_2 + \eta' > \eta'' \tag{38}$$

and that

$$\eta'' k_1 + \eta' k_2 + \eta' - \eta > \eta k(1-\varrho). \tag{39}$$

Comparing (38) and (39) with (i) and (ii) in (††), we easily see that it is now possible to choose $0 < \widetilde{\varrho} < \varrho$ close enough to 0 to guarantee (i) and (ii). ☐

Write $\mathcal{S}^k$ for the family of simple graphs on k points. In fact, the proof of Lemma 17 given above gives that

$$R = R_n(\mathcal{P}, H, \Pi) \to (\mathcal{S}^k, H),$$

that is, in any red-blue edge colouring of R, we either have a blue induced copy of H, or else we have a red induced copy of *every* simple graph on k vertices.

Corollary 11 and Lemma 17 imply Theorem 4.

5. Small Trees versus General Graphs

Our aim in this section is to prove Theorem 5. Let a graph H of order t be fixed, and let $\mathcal{P} = (V, \mathcal{L})$ be a projective plane with $|V| = |\mathcal{L}| = n$. Let an integer $k \geq 2$ be given. In this section, we consider a random graph $R' = R'_n(\mathcal{P}, H, k)$ that is better suited for the purposes of this section. Indeed, to be able to prove inequality (5), we must certainly consider graphs that contain large induced trees, and hence we are better off by considering sparser versions of the random graph $R = R_n(\mathcal{P}, H)$ that we have considered elsewhere in this paper.

5.1. **The construction of $R' = R'_n(\mathcal{P}, H, k)$.** Let $b = 200$. Put

$$q = bk\frac{\log(kt^2)}{\log\log\log(kt^2)}. \tag{40}$$

(For simplicity, we shall ignore floor and ceiling functions in this section.) Suppose $\mathcal{P} = (V, \mathcal{L})$ is such that $n = |V|$ satisfies

$$k^2t^4\left(\frac{\log(kt^2)}{\log\log\log(kt^2)}\right)^2 \leq n \leq 4k^2t^4\left(\frac{\log(kt^2)}{\log\log\log(kt^2)}\right)^2. \tag{41}$$

Let $L \in \mathcal{L}$ be a line of $\mathcal{P}$. Randomly choose q pairwise disjoint subsets $W_1, \dots, W_q \subset L$ of L, all of cardinality t. Note that this is possible since $tq \leq r + 1$ holds for large enough t. Now, randomly insert q copies of H in L using these W_i as the vertex sets of these copies. Repeat this procedure for all the lines $L \in \mathcal{L}$ of $\mathcal{P}$, making all the random choices independently. The resulting random graph is our $R' = R'_n(\mathcal{P}, H, k)$.

Let us remark that the copies of H in $R' = R'_n(\mathcal{P}, H, k)$ are in fact pairwise edge-disjoint. This will not, however, prevent us from proving an edge-Ramsey result.

Lemma 18. *There exist absolute constants k_0 and t_0 for which the following holds. Let H be a graph of order $t \geq t_0$, and consider $R' = R'_n(\mathcal{P}, H, k)$ for $k \geq k_0$. Then with positive probability we have that*

$$R' = R'_n(\mathcal{P}, H, k) \to (T, H) \tag{42}$$

for any tree T of order k.

Clearly, Lemma 18 above implies Theorem 5. In the rest of Section 5, we give a proof of Lemma 18. Let us mention that, similarly to what happened in previous sections, our proof will in fact give a 'universal' version of (42). Namely, one sees from the proof below that T in (42) may be replaced by $\mathcal{T}^k$, the family of all trees on k points, *i.e.*, if no blue induced copy of H is found in a red-blue edge-coloured R', then *all* trees in $\mathcal{T}^k$ must be present in R' as red induced subgraphs.

5.2. **A technical lemma on $R' = R'_n(\mathcal{P}, H, k)$.** Let a graph H and a projective plane $\mathcal{P}$ as in the previous section be fixed. The integer k is also fixed throughout this section. Consider the random graph $R' = R'_n(\mathcal{P}, H, k)$, and let $x, y \in V(R')$ be two distinct vertices of R'. The common neighbourhood $N(x) \cap N(y)$ of the vertices x and y in R' may be 'unexpectedly large' (with respect to the edge-density of R'). Indeed, this may happen because x and y may belong to one of the copies of H that we embed in R'. However, our technical lemma below asserts that the number of common neighbours of x and y that do *not* belong to the line $L = L(x, y) \in \mathcal{L}$ determined by x and y is fairly small.

Let us say that x and y form a *normal* pair in R' if

$$\big|(N(x) \cap N(y)) \setminus L(x, y)\big| \leq \frac{30\log n}{\log\log\log n}.$$

Lemma 19. *Let $R' = R'_n(\mathcal{P}, H, k)$ be as above. Then, with probability approaching* 1 *as $n \to \infty$, all pairs of vertices of R' are normal.*

Proof. We prove a few assertions that together imply our lemma. Below we shall use the term 'almost surely' to mean 'with probability tending to 1 as $n \to \infty$.' Moreover, we shall tacitly assume that n is large enough whenever needed.

(A1) *Let x and y be two points of $\mathcal{P}$, and fix a line $L \neq L(x, y)$ that contains y. Let $Y = Y(x, y, L) \subset L \setminus \{y\}$ be the set of neighbours of x within $L \setminus \{y\}$. Then the probability that*

$$|Y| < \frac{A \log n}{\log \log n} \tag{43}$$

holds is at least $1 - n^{-6A/7}$, for any fixed constant A.

To prove assertion (A1), fix a vertex $z \in L \setminus \{y\}$. If p is the probability that x and z are adjacent, then, because of (41), we have

$$p = qe(H)\binom{r+1}{2}^{-1} \leq bkt^2 n^{-1} \frac{\log(kt^2)}{\log \log \log(kt^2)} \leq \frac{b}{\sqrt{n}}, \tag{44}$$

where r is such that $n = r^2 + r + 1$, and, as usual, $e(H)$ denotes the number of edges in H. Now, the events $xz \in E(R')$ $(z \in L \setminus \{y\})$ are all independent. Thus, letting

$$\ell = \left\lceil \frac{A \log n}{\log \log n} \right\rceil,$$

the probability that (43) fails is at most

$$\binom{r}{\ell} p^\ell \leq \left(\frac{\mathrm{e} n^{1/2} p}{\ell} \right)^\ell \leq \left(\frac{\mathrm{e} b}{\ell} \right)^\ell \leq n^{-6A/7},$$

as required.

Note that, in particular, the following assertion holds.

(A2) *If $A \geq 3$, then, almost surely, for all triples (x, y, L) as in* (A1)*, the set $Y = Y(x, y, L)$ satisfies* (43).

(A3) *Let two points $x, y \in V$ and a constant B be fixed. Assume that $Y = Y(x, y, L)$ satisfies* (43) *for all lines $L \neq L(x, y)$ and some fixed A. Then the number of lines $L \in \mathcal{L} \setminus \{L(x, y)\}$ that contain a neighbour of both x and y in $R' = R'_n(\mathcal{P}, H, k)$ is less than*

$$\frac{B \log n}{\log \log \log n} \tag{45}$$

with probability at least $1 - n^{-6B/7}$.

To prove assertion (A3), first fix a line $L \neq L(x, y)$. Since (43) holds, in view of (44) the probability $\widetilde{p}$ that L contains a common neighbour of x

and y satisfies

$$\widetilde{p} \leq |Y(x,y,L)|p \leq \frac{Ab\log n}{\sqrt{n}\log\log n}. \tag{46}$$

Now, the events $Y(x,y,L) \cap N(y) \neq \emptyset$ $(L \neq L(x,y))$ are all independent. Hence, the probability that at least

$$\ell = \left\lceil \frac{B\log n}{\log\log\log n} \right\rceil$$

of these events hold is, in view of (46), at most

$$\binom{r}{\ell}\widetilde{p}^{\ell} \leq \left(\frac{\mathrm{e}n^{1/2}\widetilde{p}}{\ell}\right)^{\ell} \leq n^{-6B/7},$$

as required.

(A4) *If $B \geq 3$, then, almost surely, for all pairs x, $y \in V$ of distinct vertices of $R' = R'_n(\mathcal{P}, H, k)$, the number of lines different from $L(x,y)$ that contain a common neighbour of x and y is less than the quantity given in* (45).

Assertion (A4) follows immediately from (A1) and (A3).

(A5) *Let y be a point of $\mathcal{P}$, and let $L \in \mathcal{L}$ contain y. Moreover, suppose that a set $Y \subset L \setminus \{y\}$ is given, and assume that Y satisfies* (43) *for some constant A. Then the probability that $|Y \cap N(y)| \geq 11$ is at most $n^{-11/4}(\log n)^{11}$.*

The probability that $|Y \cap N(y)| \geq \ell$ happens is at most the probability that a randomly chosen t'-element subset Z of $L \setminus \{y\}$ meets Y in at least ℓ elements, where $t' = t - 1$. Writing s for the cardinality of Y, we have that this latter probability is at most

$$\binom{r}{t'}^{-1} \sum_{i \geq \ell} \binom{s}{i}\binom{r-s}{t'-i} \leq \sum_{i \geq \ell} \binom{t'}{i}\left(\frac{s}{r}\right)^{i},$$

which, by (43) and the trivial lower bounds $r \geq \sqrt{n}/2$ and $n \geq t^4$, is seen to be no larger than

$$2\binom{t}{\ell}\left(\frac{2A\log n}{\sqrt{n}\log\log n}\right)^{\ell}.$$

Letting $\ell = 11$, assertion (A5) follows.

A simple application of assertions (A1)–(A5) with $A = B = 3$ proves Lemma 19. □

5.3. **Proof of Lemma 18.** We are now ready to prove Lemma 18, which, as remarked above, implies Theorem 5.

Proof of Lemma 18. Let a projective plane $\mathcal{P} = (V, \mathcal{L})$ on n points be fixed, and suppose a t-vertex graph H and an integer k are given. Suppose further that (40) and (41) hold. Consider the graph $R' = R'_n(\mathcal{P}, H, k)$, and suppose that all pairs of distinct vertices of R' are normal. We shall show below that, under these conditions, we have

$$R' \to (T, H) \tag{47}$$

for any tree T on k points. Invoking Lemma 19, the proof of Lemma 18 will be complete.

To prove (47), let R' be red-blue edge-coloured in such a way that no blue induced copy of H is present. Our aim is to show that, then, any tree T on at most k vertices appears as an induced subgraph of R' with all its edges coloured red. To this end, let us first observe that, since every copy of H in each line of $\mathcal{P}$ must contain a red edge, there are at least qn red edges in R'. In fact, we may choose exactly qn red edges in R' so that, in each line $L \in \mathcal{L}$, we have an independent set of q red edges. Let $\mathcal{R}$ be such a collection of red edges.

A standard argument now shows that there is a subset V' of $V(R')$ such that every vertex $v \in V'$ is incident to at least q red edges in $\mathcal{R} \cap E(R'[V'])$. In other words, the graph spanned by $\mathcal{R} \cap E(R'[V'])$ in $R'[V']$ has minimal degree at least q. In what follows, we restrict our attention to $R'[V']$.

We shall now prove by induction on $|V(T)|$ that any tree T on at most k vertices is present in $R'[V']$ as a red induced subgraph. Thus, $R'[V']$ will be shown to be *red k-tree-universal*, that is, all trees on k vertices appear in $R'[V']$ as red induced subgraphs (cf. the remark following the proof of Lemma 15).

There is nothing to do if $|V(T)| = 1$. So assume that $2 \le |V(T)| \le k$, and that the subtree T' of T obtained from T by removing a leaf of T appears as a red induced subgraph in $R'[V']$. (In the sequel, we identify T' with this copy of T' in $R'[V']$.) Clearly, it suffices to show that an appropriate vertex $x' \in V(T') \subset V'$ has a neighbour $x \in V'$ such that xx' is red and, furthermore, x is not adjacent to any vertex in $F = T' - x'$. Let such a vertex x' be fixed.

Let

$$Z = \{z \in N^{\mathcal{R}}_{\text{red}}(x') \cap V' \colon L(x', z) \cap V(F) = \emptyset\}, \tag{48}$$

where $N^{\mathcal{R}}_{\text{red}}(x')$ denotes the set of vertices connected to x' with edges in $\mathcal{R}$, and $L(x', z)$ is the line of $\mathcal{P}$ determined by x' and z. Since there is at most one red edge in $\mathcal{R}$ that is incident to x' in each line that contains x', and $|V(F)| \le k - 2$, we have that

$$|Z| \ge q - k + 2 \ge \frac{1}{2}q = \frac{1}{2}bk\frac{\log(kt^2)}{\log\log\log(kt^2)}. \tag{49}$$

Using the normality of the pairs (x', y) for $y \in V(F)$, we have that

$$\Big|\bigcup_{y \in V(F)} N(y) \cap Z\Big| \le \frac{30(k-2)\log n}{\log\log\log n}. \tag{50}$$

Comparing (49) and (50) (and recalling that $b = 200$), we see that there is a vertex $x \in Z$ that is *not* adjacent to any vertex in F in R'. Thus, we have found our red induced copy of T in $R'[V']$, as required. This completes the induction step, and hence we have proved that $R'[V']$ is indeed red k-tree-universal.

As remarked earlier, an application of Lemma 19 finishes the proof of Lemma 18. □

6. Concluding Remarks

A refinement of Theorem 3 may be proved in the following fashion. As is well known, the technique applied in the proof of Lemma 15 gives better numerical results if the maximum degree of the graph G has smaller order of magnitude than $k = |V(G)|$. This fact, coupled with slightly more careful calculations, should suffice to improve (4) to

$$r_{\mathrm{ind}}(G,H) \le 2^{c_1 k} t^{c_2 \Delta \log q}, \tag{51}$$

where $\Delta = \Delta(G)$ is the maximum degree of G, and c_1 and c_2 are universal constants.

However, even with (51), Problem 1 and Conjecture 2 remain open.

References

[1] J. Beck, *On size Ramsey number of paths, trees, and circuits II*, Mathematics of Ramsey Theory, Algorithms and Combinatorics, vol. 5, Springer-Verlag, Berlin, 1990, pp. 34–45.

[2] J.I. Brown and V. Rödl, *A Ramsey type problem concerning vertex colourings*, Journal of Combinatorial Theory, Series B **52** (1991), no. 1, 45–52, MR92c:05108.

[3] V. Chvátal, V. Rödl, E. Szemerédi, and W.T. Trotter, *The Ramsey number of a graph with bounded maximum degree*, Journal of Combinatorial Theory, Series B **34** (1983), no. 3, 239–243, MR#85f:05085.

[4] W. Deuber, *A generalization of Ramsey's theorem*, Infinite and Finite Sets (Keszthely, 1973) (A. Hajnal, R. Rado, and V.T. Sós, eds.), Colloquia Mathematica Societatis János Bolyai, vol. 10, North-Holland, 1975, pp. 323–332.

[5] N. Eaton and V. Rödl, *A canonical Ramsey theorem*, Random Structures and Algorithms **3** (1992), no. 4, 427–444.

[6] P. Erdős, *Problems and results on finite and infinite graphs*, Recent advances in graph theory (Proc. Second Czechoslovak Sympos., Prague, 1974), (loose errata) Academia, Prague, 1975, MR52#10500, pp. 183–192.

[7] ———, *On some problems in graph theory, combinatorial analysis and combinatorial number theory*, Graph theory and combinatorics (Cambridge, 1983) (London, New York) (B. Bollobás, ed.), Academic Press, 1984, pp. 1–17.

[8] P. Erdős and A. Hajnal, *Ramsey-type theorems*, Discrete Applied Mathematics **25** (1989), no. 1–2, 37–52, MR90m:05091.

[9] P. Erdős, A. Hajnal, and L. Pósa, *Strong embeddings of graphs into colored graphs*, Infinite and Finite Sets (Keszthely, 1973) (A. Hajnal, R. Rado, and V.T. Sós, eds.), Colloquia Mathematica Societatis János Bolyai, vol. 10, North-Holland, 1975, pp. 585–595.

[10] R.L. Graham and V. Rödl, *Numbers in Ramsey theory*, Surveys in Combinatorics 1987 (C. Whitehead, ed.), London Mathematical Society Lecture Note Series, vol. 123, Cambridge University Press, Cambridge–New York, 1987, pp. 112–153.

[11] R.L. Graham, B. Rothschild, and J.H. Spencer, *Ramsey theory*, second ed., Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, New York, 1990.
[12] T. Łuczak and V. Rödl, *On induced Ramsey numbers for graphs with bounded maximum degree*, Journal of Combinatorial Theory, Series B **66** (1996), no. 2, 324–333.
[13] J. Nešetřil, *Ramsey theory*, Handbook of Combinatorics (R.L. Graham, M. Grötschel, and L. Lovász, eds.), North-Holland, Amsterdam, 1995, pp. 1331–1403.
[14] F.P. Ramsey, *On a problem of formal logic*, Proceedings of the London Mathematical Society **30** (1930), 264–286.
[15] V. Rödl, *The dimension of a graph and generalized Ramsey theorems*, Master's thesis, Charles University, 1973.
[16] ———, *On universality of graphs with uniformly distributed edges*, Discrete Mathematics **59** (1986), no. 1-2, 125–134.
[17] V. Rödl and P. Winkler, *A Ramsey-type theorem for orderings of a graph*, SIAM Journal on Discrete Mathematics **2** (1989), no. 3, 402–406, MR90g:05135.

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

Institut für Informatik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
E-mail address: proemel@informatik.hu-berlin.de

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