A Short Proof of the Random Ramsey Theorem

Rajko Nenadov Angelika Steger
rnenadov@inf.ethz.ch steger@inf.ethz.ch

Institute of Theoretical Computer Science
ETH Zurich, 8092 Zurich, Switzerland

Abstract. In this paper we give a short proof of the Random Ramsey Theorem of Rödl and Ruciński: for any graph F which contains a cycle and $r \geq 2$, there exist constants $c, C > 0$ such that

$$\mathbb{P}[G_{n,p} \to (F)_r^e] = \begin{cases} 1 - o(1), & p \geq Cn^{-1/m_2(F)} \\ o(1), & p \leq cn^{-1/m_2(F)}, \end{cases}$$

where $m_2(F) = \max_{J \subseteq F, v_J \geq 2} \frac{e_J - 1}{v_J - 2}$.
The proof of the 1-statement is based on the recent beautiful hypergraph container theorems by Saxton/Thomason and Balogh/Morris/Samotij. The proof of the 0-statement is elementary.

1. Introduction

For graphs G and F and a constant $r \in \mathbb{N}$, we denote with

$$G \to (F)_r^e$$

the property that every edge-colouring of G with r colours (we call this an *r-colouring*) contains a copy of F with all edges having the same colour. Ramsey's theorem then implies that for all graphs F and r we have $K_n \to (F)_r^e$, for n large enough. At first sight it is not immediately clear whether this follows from the density of K_n or its rich structure. As it turns out, studying Ramsey properties of random graphs shows that the latter is the case, as random graphs give examples of sparse graphs with the desired Ramsey property.

The study of Random Ramsey Theory was initiated by Łuczak, Ruciński, and Voigt [7] who studied the Ramsey property of random graphs in the vertex-colouring case and also established the threshold for the property $G_{n,p} \to (K_3)_2^e$. Thereupon, in a series of papers Rödl and Ruciński [9, 10, 11] determined the threshold of $G_{n,p} \to (F)_r^e$, in full generality. Formally, their result reads as follows.

For every graph G we denote by $V(G)$ and $E(G)$ its vertex and edge sets and by v_G and e_G their sizes. For every graph G on at least 3 vertices we set $d_2(G) = (e_G - 1)/(v_G - 2)$. By $m_2(G)$ we denote for every graph G the so-called *2-density*, defined as

$$m_2(G) = \max_{J \subseteq G, v_J \geq 3} d_2(J).$$

If $m_2(G) = d_2(G)$ then we say that a graph G is *2-balanced*, and if in addition $m_2(G) > d_2(J)$ for every subgraph $J \subset G$ with $v_J \geq 3$, we say that G is *strictly 2-balanced.*

Theorem 1 (Rödl, Ruciński [9, 10, 11])**.** *Let $r \geq 2$ and let F be a fixed graph that is not a forest of stars or, in the case $r = 2$, paths of length 3. Then there exist positive constants $c = c(F, r)$, and $C = C(F, r)$ such that*

$$\lim_{n\to\infty} \mathbb{P}[G_{n,p} \to (F)_r^e] = \begin{cases} 0 & \text{if } p \leq cn^{-1/m_2(F)} \\ 1 & \text{if } p \geq Cn^{-1/m_2(F)}. \end{cases}$$

For the exceptional case of a star with k edges it is easily seen that the threshold is determined by the appearance of a star with $r(k-1)+1$ edges. For the path P_3 of length three the 0-statement only holds for $p \ll n^{-1/m_2(P_3)} = n^{-1}$ since, for example, a C_5 with a pendant edge at every vertex has density one and cannot be 2-coloured without a monochromatic P_3.

Note that $p = n^{-1/m_2(F)}$ is the density where we expect that every edge is contained in roughly a constant number of copies of F. This observation can be used to provide an intuitive understanding of the bounds of Theorem 1. If c is very small, then the number of copies of F is w.h.p. (*with high probability*, i.e., with probability $1-o(1)$ if n tends to infinity) small enough that they are so scattered that a colouring without a monochromatic copy of F can be found. If, on the other hand, C is big then these copies w.h.p. overlap so heavily that every colouring has to induce at least one monochromatic copy of F.

The aim of this paper is to give a short proof of Theorem 1.

2. Proof of the 1-statement

The proof of the 1-statement requires two tools. The first one is a well-known quantitative strengthening of Ramsey's theorem. We include its short proof for convenience of the reader.

Theorem 2 (Folklore)**.** *For every graph F and every constant $r \geq 2$ there exist constants $\alpha > 0$ and n_0 such that for all $n \geq n_0$, every r-colouring of the edges of K_n contains at least αn^{v_F} monochromatic copies of F.*

Proof. From Ramsey's theorem we know that there exists $N := N(F, r)$ such that every r-colouring of the edges of K_N contains a monochromatic copy of F. Thus, in any r-colouring of K_n every N-subset of the vertices contains at least one monochromatic copy of F. As every copy of F is contained in at most $\binom{n-v_F}{N-v_F}$ N-subsets, the theorem follows e.g. with $\alpha = 1/N^{v_F}$. $\square$

We will need in particular the following easy consequence of Theorem 2.

Corollary 3. *For every graph F and every $r \in \mathbb{N}$ there exist constants n_0 and $\delta, \varepsilon > 0$ such that the following is true for all $n \geq n_0$. For any $E_1, \ldots, E_r \subseteq E(K_n)$ such that for all $1 \leq i \leq r$ the set E_i contains at most $\varepsilon n^{v(F)}$ copies of F, we have*

$$|E(K_n) \setminus (E_1 \cup \ldots \cup E_r)| \geq \delta n^2.$$

Proof. Let α and n_0 be as given by Theorem 2 for F and $r+1$, and set $\varepsilon = \alpha/2r$. Further, let $E_{r+1} := E(K_n) \setminus (E_1 \cup \ldots \cup E_r)$ and consider the colouring $\Delta : E(K_n) \to [r+1]$ given by $\Delta(e) = \min\{i \in [r+1] \ : \ e \in E_i\}$. By Theorem 2 there exist at least αn^{v_F} monochromatic copies of F, of which, by assumption on the sets E_i, at least $\frac{1}{2}\alpha \cdot n^{v_F}$ must be contained in E_{r+1}. As every edge is contained in at most $2e_F \cdot n^{v_F - 2}$ copies of F the claim of the corollary follows e.g. for $\delta = \frac{\alpha}{4e_F}$. $\square$

The second tool that we need is a consequence of the beautiful container theorems of Balogh, Morris and Samotij [1] and Saxton and Thomason [12]. The following theorem is from Saxton and Thomason, who obtain it for all graphs F. Balogh, Morris and Samotij obtain a similar statement for all 2-balanced graphs F.

Definition 4. For a given set S and constants $k \in \mathbb{N}$, $s > 0$, let $\mathcal{T}_{k,s}(S)$ be the family of k-tuples of subsets defined as follows,

$$\mathcal{T}_{k,s}(S) = \{(S_1, \ldots, S_k) \mid S_i \subseteq S \text{ for } 1 \leq i \leq k \text{ and } |\bigcup_{i=1}^{k} S_i| \leq s\}.$$

Theorem 5 ([12], Theorem 1.3)**.** *For any graph F and $\varepsilon > 0$, there exist n_0 and $k > 0$ such that the following is true. For every $n \geq n_0$ there exist $t = t(n)$, pairwise distinct tuples $T_1, \ldots, T_t \in \mathcal{T}_{k,kn^{2-1/m_2(F)}}(E(K_n))$ and sets $C_1, \ldots, C_t \subseteq E(K_n)$, such that*

(a) each C_i contains at most εn^{v_F} copies of F,
(b) for every F-free graph G on n vertices there exists $1 \leq i \leq t$ such that $T_i \subseteq E(G) \subseteq C_i$. (Here $T_i \subseteq E(G)$ means that all sets contained in T_i are subsets of $E(G)$.)

Note that the main result in [12] is more general, as it provides a similar structure for independent sets in uniform hypergraphs.

With these two tools in hand the proof of the 1-statement of Theorem 1 is now easily completed.

Proof of Theorem 1 (*1-statement*). Let ε and δ be as in Corollary 3, and let n_0 and k be as in Theorem 5 for F and ε, and assume that $n \geq n_0$. If $G_{n,p} \nrightarrow (F)^e_r$, then there exists a colouring $\Delta : E(G_{n,p}) \to r$ such that for all $1 \leq j \leq r$ the set $E_j := \Delta^{-1}(j)$ does not contain a copy of F. By Theorem 5 we have that for every such E_j there exists $1 \leq i_j \leq t(n)$ such that $T_{i_j} \subseteq E_j \subseteq C_{i_j}$ and C_{i_j} contains at most εn^{v_F} copies of F. The trivial, but nonetheless crucial observation is that $G_{n,p}$ completely avoids $E(K_n) \setminus (C_{i_1} \cup \ldots \cup C_{i_r})$, which by Corollary 3 has size at least δn^2.

Therefore we can bound $\mathbb{P}[G_{n,p} \nrightarrow (F)^e_r]$ by bounding the probability that there exist tuples $T_{i_1}, \ldots, T_{i_r}$ that are contained in $G_{n,p}$ such that $E_0(T_{i_1}, \ldots, T_{i_r}) := E(K_n) \setminus (C_{i_1} \cup \ldots \cup C_{i_r})$ is edge-disjoint from $G_{n,p}$. Thus

$$\mathbb{P}[G_{n,p} \nrightarrow (F)^e_r] \leq \sum_{i_1,\ldots,i_r} \mathbb{P}[T_{i_1}, \ldots, T_{i_r} \subseteq G_{n,p} \ \wedge \ G_{n,p} \cap E_0(T_{i_1}, \ldots, T_{i_r}) = \emptyset],$$

where $i_1, \ldots, i_r$ run over the choices given by Theorem 5. Note that the two events in the above probability are independent and the probability can thus be bounded by

$$p^{|\bigcup_{j=1}^r T^+_{i_j}|} \cdot (1-p)^{\delta n^2},$$

where by $T^+_{i_j}$ we denote the union of the sets of the k-tuple T_{i_j}. The sum can be bounded by first deciding on $s := |\bigcup_{j=1}^r T^+_{i_j}| \leq r \cdot kn^{2-1/m_2(F)}$, then choosing that many edges ($\binom{\binom{n}{2}}{s}$ choices) and finally deciding for every edge in which sets of the k-tuples T_{i_j} it appears ($(2^{rk})^s$ choices). Together, this gives

$$\begin{aligned}\mathbb{P}[G_{n,p} \nrightarrow (F)^e_r] &\leq (1-p)^{\delta n^2} \cdot \sum_{s=0}^{rkn^{2-1/m_2(F)}} \binom{\binom{n}{2}}{s} (2^{rk})^s p^s \\ &\leq e^{-\delta n^2 p} \cdot \sum_{s=0}^{rkn^{2-1/m_2(F)}} \left(\frac{e2^{rk} n^2 p}{2s}\right)^s.\end{aligned}$$

Recall that $p = Cn^{-1/m_2(F)}$. By choosing C sufficiently large (with respect to k) we may assume that

$$\sum_{s=0}^{rkn^{2-1/m_2(F)}} \left(\frac{e2^{rk} n^2 p}{2s}\right)^s \leq n^2 \cdot \left(\frac{e2^{rk} C}{2rk}\right)^{(rk/C)n^2 p} \leq e^{\frac{1}{2}\delta n^2 p}$$

and thus $\mathbb{P}[G_{n,p} \nrightarrow (F)^e_r] = o(1)$, as desired. □

We remark that the same approach, with Theorem 2 and Theorem 5 replaced with the corresponding hypergraph versions, gives an alternative proof of the 1-statement for a Random Ramsey Theorem for hypergraphs obtained by Friedgut, Rödl and Schacht [5] and Conlon and Gowers [3].

3. Proof of the 0-statement

We need to show that with high probability the edges of a random graph $G_{n,p}$ with $p = cn^{-1/m_2(F)}$, for $0 < c = c(F) < 1$ small enough, can be coloured in such a way that we have no monochromatic copy of F. If $m_2(F) = 1$ we have $p \leq cn^{-1}$ with $c < 1$. It is well-known that then every component of $G_{n,p}$ is w.h.p. either a tree or a unicyclic graph (see [4]). One easily checks that we can colour each such component without a monochromatic copy of F whenever F is not a star and not a path of length 3 (or $r \geq 3$ in the latter). In the following we thus assume that $m_2(F) > 1$.

Observe that we may also assume without loss of generality that $r = 2$ and that F is strictly 2-balanced. If not, replace F by a minimal subgraph F' with the same 2-density. Clearly, if we find a 2-colouring of the edges of $G_{n,p}$ without a monochromatic copy of F' this 2-colouring will also contain no monochromatic copy of F.

The expected number of copies of F on any given edge is bounded by

$$2e_F \cdot n^{v_F-2} \cdot p^{e_F-1} \leq 2e_F \cdot c^{e_F-1}.$$

That is, for $0 < c < 1$ small enough we do not expect more than one copy. We now show why this makes the colouring process easier.

Let e be an edge in $G_{n,p}$. Assume that $G_{n,p} - e$ is 2-colourable without a monochromatic copy of F. Consider any such colouring. If this colouring cannot be extended to e then there has to exist both a red and a blue copy of $F - \hat{e}$ (for some $\hat{e} \in E(F)$ which might be different in these two copies) such that e completes both of these copies to a copy of F. Since these two copies of $F - \hat{e}$ are in different colours, and thus edge disjoint, we can conclude that there exist at least two copies of F which intersect in e. In other words: a necessary obstruction for extending a colouring from $G - e$ to G is that e is contained in at least two copies of F that only intersect in e.

To formalize this idea, call an edge e *closed* in G if it is contained in at least two copies of F whose edge sets intersect exactly in e. (Note: we do allow that the vertex sets of these copies intersect in more than two vertices). Otherwise we call the edge *open*. With this notion at hand we can now formulate the following algorithm for obtaining the desired 2-coloring of $G_{n,p}$:

> $\hat{G} := G_{n,p}$;
> **while** there exists an open edge e in $\hat{G}$ **do**
> $\hat{G} \leftarrow \hat{G} - e$;
> colour $\hat{G}$;
> add the edges in reverse order and colour them appropriately.

The critical point, of course, is the statement 'colour $\hat{G}$'. We need to show that this step is indeed possible.

Observe that after termination of the **while**-loop the graph $\hat{G}$ has the following property: every edge of $\hat{G}$ is closed. It is easy to see that $\hat{G}$ is actually the (unique) maximal subgraph of $G_{n,p}$ with the property that every edge is closed (within this subgraph). We call $\hat{G}$ the *F-core* of $G_{n,p}$.

We now further refine $\hat{G}$. Consider an auxiliary graph G_F defined as follows: the set of vertices correspond to the set of copies of F in $\hat{G}$ and two vertices

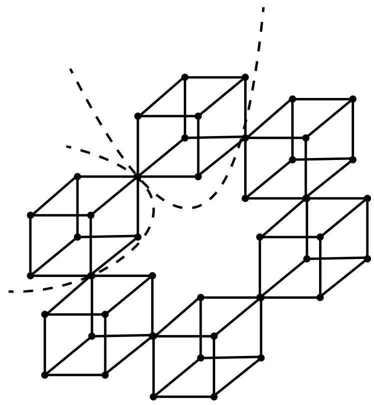

FIGURE 1. Edges of each *cube* represent a C_4-component.

are connected by an edge if and only if the corresponding copies of F have at least one edge in common. Since every edge of $\hat{G}$ belongs to a copy of F, the connected components of G_F naturally partition the edges of $\hat{G}$ into equivalence classes. Observe that, by definition, each equivalence class (an *F-component* for short) can be coloured separately in order to find a valid colouring of the F-core. Note also that within $\hat{G}$ the F-components need not necessarily form components, cf. Figure 1.

The core of our argument is the following lemma, which states that with high probability every F-component in the F-core of $G_{n,p}$ has constant size.

Lemma 6. *Let F be a strictly 2-balanced graph with $e_F \geq 3$. There exist $c = c(F) > 0$ and $L = L(F) > 0$ such that if $p \leq cn^{-1/m_2(F)}$ then w.h.p. every F-component of the F-core of $G_{n,p}$ has size at most L.*

We will prove Lemma 6 in the next subsection. Before doing so we show how it can be used to complete the proof of Theorem 1. For that we also need the following result of Rödl and Ruciński [9], which states that graphs with small enough density do not have the Ramsey property. We include its short proof in the appendix of this paper.

Theorem 7 ([9])**.** *Let G and F be two graphs. If $m(G) \leq m_2(F)$, where $m(G) = \max_{J \subseteq G} \frac{e_J}{v_J}$ is the maximal density of G, and $m_2(F) > 1$ then $G \nrightarrow (F)^e_2$.*

With these results in hand, the proof of the 0-statement of Theorem 1 is straightforward.

Proof of Theorem 1 (*0-statement*). Recall that we may assume w.l.o.g. that F is strictly 2-balanced and that $m_2(F) > 1$. Choose $c = c(F)$ and $L = L(F)$ according to Lemma 6. Then $G_{n,p}$ has w.h.p. the property that every F-component of the F-core of $G_{n,p}$ has size at most L.

Observe that there exist only constantly many different graphs on at most L vertices. Let G be one such graph, and choose $G' \subseteq G$ such that $m(G) = e_{G'}/v_{G'}$. Then the expected number of copies of G' in $G_{n,p}$ is bounded by $n^{v_{G'}}p^{e_{G'}}$. Observe that for $p = cn^{-1/m_2(F)}$ we have $n^{v_{G'}}p^{e_{G'}} = o(1)$ whenever $m(G) = e_{G'}/v_{G'} > m_2(F)$. It thus follows from Markov's inequality that for $p \leq cn^{-1/m_2(F)}$ w.h.p. there is no copy of G', and hence no copy of G in $G_{n,p}$. Therefore, w.h.p. every subgraph G of $G_{n,p}$ of size $|V(G)| \leq L$ satisfies $m(G) \leq m_2(F)$.

Combining both properties we get: with high probability all F-components G of $G_{n,p}$ satisfy $m(G) \leq m_2(F)$ and Theorem 7 thus implies that there exists a

2-colouring of G without a monochromatic copy of F. The union of these edge colourings of all F-components therefore yields the desired colouring of the F-core of $G_{n,p}$. Finally, as explained above this colouring can be extended to a colouring of $G_{n,p}$ without a monochromatic copy of F. □

3.1. **Proof of Lemma 6.** We start by collecting some properties of strictly 2-balanced graphs.

Lemma 8. *If F is strictly 2-balanced, then F is 2-connected.*

Proof. Clearly, F is connected. As then $(e_F - 2)/(v_F - 3) \geq (e_F - 1)/(v_F - 2)$, we deduce that F cannot contain a vertex of degree 1. Assume there exists $v \in V(F)$ that is a cut vertex. Then there exist subgraphs F_1 and F_2 that both contain at least three vertices such that $F_1 \cup F_2 = F$ and $V(F_1) \cap V(F_2) = \{v\}$. As F is strictly 2-balanced we get (using $a/b < x$ and $c/d < x$ implies $(a+c)/(b+d) < x$)

$$e_F - 2 = (e_{F_1} - 1) + (e_{F_2} - 1) < m_2(F) \cdot (v_{F_1} - 2 + v_{F_2} - 2) = m_2(F) \cdot (v_F - 3).$$

Since $m_2(F) = (e_F - 1)/(v_F - 2)$ (as F is balanced), this contradicts the inequality from above. □

Lemma 9. *Let F be strictly 2-balanced and let G be an arbitrary graph. Construct a graph $\hat{G}$ by attaching F to an edge e of G. Then $\hat{G}$ has the property that if $\hat{F}$ is a copy of F in $\hat{G}$ that contains a least one vertex from $F - e$, then $\hat{F} = F$.*

Proof. Assuming the opposite, let $\hat{F}$ be a copy of F which violates the claim and set $F_g = \hat{F}[V(G)]$ and $F_f = \hat{F}[V(F)]$. The fact that $\hat{F}$ violates the claim implies that $\hat{F}$ contains at least one vertex from $F - e$ and at least one from $G - e$. Since every strictly 2-balanced graph is, by Lemma 8, 2-connected, it follows that both vertices of e belong to $V(\hat{F})$, thus $v_F = v_{F_g} + v_{F_f} - 2$ and $v_{F_g}, v_{F_f} \geq 3$.

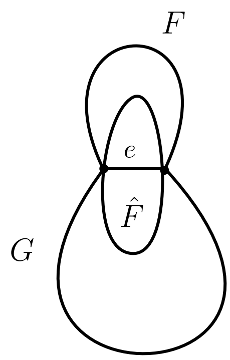

FIGURE 2. Situation of Lemma 9: copy F is attached to G at edge e; copy $\hat{F}$ intersects the interior of both F and G.

If $e \notin \hat{F}$, we add edge e to F_f. Then $e_F = e_{F_g} + e_{F_f} - 1$ regardless of whether $e \in \hat{F}$ or not. As F_g and F_f are strict subgraphs of F, we have

$$\frac{e_{F_g} - 1}{v_{F_g} - 2} < m_2(F) \quad \text{and} \quad \frac{e_{F_f} - 1}{v_{F_f} - 2} < m_2(F)$$

since F is strictly 2-balanced. This, however, yields a contradiction, as

$$m_2(F) = \frac{e_F - 1}{v_F - 2} = \frac{e_{F_g} - 1 + e_{F_f} - 1}{v_{F_g} - 2 + v_{F_f} - 2} < m_2(F).$$

□

In order to prove Lemma 6 we define a process that generates F-closed structures iteratively starting from a single copy of F. Our proof simplifies similar approaches from [6] and [8].

Let G' be an F-component of the F-core of $G_{n,p}$. Then G' can be generated by starting with an arbitrary copy of F in G' and repeatedly attaching copies of F to the graph constructed so far.

Let F_0 be a copy of F in G',
$\ell \leftarrow 0$; $\hat{G} \leftarrow F_0$;
while $\hat{G} \neq G'$ **do**
  $\ell \leftarrow \ell + 1$;
  if $\hat{G}$ contains an open edge **then**
    let $\ell' < \ell$ be the smallest index such that
      $F_{\ell'}$ contains an open edge;
    let e be any open edge in $F_{\ell'}$;
    let F_ℓ be a copy of F in G' that contains e but is
      not contained in $\hat{G}$;
  else
    let F_ℓ be a copy of F in G' that is not contained
      in $\hat{G}$ and intersects $\hat{G}$ in at least one edge;
  $\hat{G} \leftarrow \hat{G} \cup F_\ell$;

In order to finally apply a first moment argument we first collect some properties of this process. Consider a copy F_ℓ for $\ell \geq 1$. We distinguish two cases: a) F_ℓ intersects $\hat{G} := \bigcup_{i<\ell} F_i$ in *exactly* two vertices (that, by definition of the algorithm, have to form an edge), i.e. F_ℓ intersects $\hat{G}$ in exactly one edge (we call this a regular copy) and b) F_ℓ intersects $\hat{G}$ in some subgraph J with $v_J \geq 3$ (we call this a degenerate copy).

In the following we denote the union of the copies $F_0, \ldots, F_\ell$ as the situation at time ℓ. For $0 \leq i \leq \ell$ we say that the copy F_i is *fully-open* at time ℓ if F_i is a regular copy (or $i = 0$) and no vertex of $V(F_i) \setminus (\bigcup_{i'<i} V(F_{i'}))$, is touched by any of the copies $F_{i+1}, \ldots, F_\ell$. Note that F_0 is fully-open only at time 0. Also note that, by Lemma 9, every fully-open copy at time $\ell \geq 1$ contains exactly $e_F - 1$ open edges.

For the analysis of the algorithm it is important to keep track of fully-open copies. For doing so we introduce the following definition. For $\ell \geq 1$ let

$$\kappa(\ell) = |\{0 \leq i < \ell \mid F_i \text{ fully-open at time } \ell - 1 \text{ but not at time } \ell\}|.$$

Clearly, a regular copy can 'destroy' at most one fully-open copy (as it intersects $\hat{G}$ in exactly one edge). Thus $\kappa(\ell) \leq 1$ if F_ℓ is a regular copy. A degenerate copy on the other hand intersects one F_i in an edge and may destroy up to $v_F - 2$ additional fully-open copies. Thus, $\kappa(\ell) \leq v_F - 1$ if F_ℓ is a degenerate copy.

Claim 10. *For any sequence $F_i, \ldots, F_{i+e_F-2}$ of consecutive regular copies such that $\kappa(i) = 1$ we have $\kappa(i+1) = \ldots = \kappa(i + e_F - 2) = 0$.*

Proof. As F_i is a regular copy we know that F_i intersects some copy $F_{i'}$, $i' < i$, in exactly one edge. As $\kappa(i) = 1$ we know that $F_{i'}$ was fully-open at time $i - 1$. Thus at time $i - 1$ the copy $F_{i'}$ had $e_F - 1$ open edges (resp. e_F, if $i' = 0$) and the intersection of F_i with $F_{i'}$ is one of these open edges. At time $i + 1$ the copy $F_{i'}$ thus still has at least $e_F - 2$ open edges and since it was chosen by the process at step i, it will be chosen again in every consecutive step as long as it has an open edge. It easily follows from Lemma 9 that every regular copy closes at most one open edge, thus each of the copies $F_{i+1}, \ldots, F_{i+e_F-2}$ intersects $F_{i'}$ in exactly one open edge, which implies $\kappa(i+1) = \ldots = \kappa(i + e_F - 2) = 0$. □

Next we estimate the number of fully-open copies at time ℓ as a function of the number of regular and degenerate copies. Let us denote with $\mathrm{reg}(\ell)$ and $\deg(\ell)$ the number of copies F_i, $1 \le i \le \ell$, which are regular, resp. degenerate. Furthermore, we denote with $f_o(\ell)$ the number of fully-open copies at time ℓ.

Claim 11. *For every $\ell \ge 1$, assuming the process doesn't stop before adding the ℓ-th copy, we have*

$$f_o(\ell) \ge \mathrm{reg}(\ell)(1 - 1/(e_F - 1)) - \deg(\ell) \cdot v_F.$$

Proof. Let us denote by $\varphi(\ell) := \mathrm{reg}(\ell)(1 - 1/(e_F - 1)) - \deg(\ell) \cdot v_F$ the right hand side. We use induction to prove the following slightly stronger statement

$$f_o(\ell) \ge \begin{cases} \varphi(\ell) & \text{if } F_\ell \text{ is a regular copy} \\ \varphi(\ell) + 1 & \text{if } F_\ell \text{ is a degenerate copy,} \end{cases}$$

for all $\ell \ge 1$. One easily checks that this claim holds for $\ell = 1$: if F_1 is a regular copy then $f_o(1) = 1 > 1 - 1/(e_F - 1)$, otherwise $f_o(1) = 0 > -v_F + 1$. Consider now some $\ell \ge 2$. If F_ℓ is a degenerate copy then $\kappa(\ell) \le v_F - 1$ and so $f_o(\ell) = f_o(\ell-1) - \kappa(\ell) \ge f_o(\ell - 1) - v_F + 1$. The claim thus easily follows from $\mathrm{reg}(\ell) = \mathrm{reg}(\ell - 1)$ and $\deg(\ell) = \deg(\ell - 1) + 1$. Otherwise, assume that F_ℓ is a regular copy and let

$$\ell' := \max\{1 \le \ell' < \ell \mid \kappa(\ell') > 0 \text{ or } F_{\ell'} \text{ is a degenerate copy}\}.$$

Note that ℓ' is well defined, as $\kappa(1) = 1$. Note also that the fact that F_ℓ is regular together with the definition of ℓ' implies that $\varphi(\ell) = \varphi(\ell') + (\ell - \ell')(1 - 1/(e_F - 1))$. In addition, we deduce from $\kappa(i) = 0$ for $\ell' < i < \ell$ that all steps $\ell' < i < \ell$ add a fully-open copy. We thus have

$$f_o(\ell) = f_o(\ell') + (1 - \kappa(\ell)) + (\ell - \ell' - 1) = f_o(\ell') + \ell - \ell' - \kappa(\ell).$$

If $F_{\ell'}$ is a degenerate copy then the induction assumption implies $f_o(\ell') \ge \varphi(\ell') + 1$. As F_ℓ is a regular copy and thus $\kappa(\ell) \le 1$, this implies $f_o(\ell) \ge \varphi(\ell') + \ell - \ell' \ge \varphi(\ell)$, as claimed. Finally, assume that $F_{\ell'}$ is a regular copy. If $\kappa(\ell) = 0$, then the claim follows trivially from the above inequalities. Otherwise we have $\kappa(\ell) = 1$ and Claim 10 thus implies that $\ell \ge \ell' + (e_F - 1)$. Therefore $f_o(\ell) = f_o(\ell') + \ell - \ell' - 1 \ge f_o(\ell') + (\ell - \ell')(1 - 1/(e_F - 1)) \ge \varphi(\ell)$, similarly as before. □

If $f_o(\ell) > 0$ for some $\ell \ge 1$, then F_ℓ cannot be the last copy in the process because there exist edges which are still open. Furthermore, from Claim 11 we have that after adding L copies, out of which at most ξ were degenerate, there are still at least

$$(L - \xi)(1 - 1/(e_F - 1)) - \xi \cdot v_F \tag{1}$$

fully-open copies at time L.

In a first moment calculation we have to multiply the number of choices for F_ℓ with the probability that the chosen copy of F is in $G_{n,p}$. For a regular copy where F_ℓ is attached to an open edge we get that this term is bounded by

$$2e_F^2 \cdot n^{v_F - 2} \cdot p^{e_F - 1} \le 2e_F^2 \cdot c < \frac{1}{2}, \tag{2}$$

for $0 < c < 1/(4e_F^2)$. Here the term $2e_F^2$ bounds the number of choices of the open edge in $F_{\ell'}$ (at most e_F choices) times the number of choices for the edge in F_ℓ that is merged with this open edges (e_F choices) times 2 for the orientation. For a regular copy F_ℓ that is attached to a closed edge we also have to replace the first factor e_F by, say, $\ell \cdot e_F$, as the edge e to which the new copy F_ℓ is attached can be any of the previously added edges. From the above we know that after step L regular copies are always attached to an open edge, as long as the number of degenerate copies is at most ξ. That is, in a first moment argument we may bound the factor attributed to a regular step at time $\leq L$ by $L/2 \leq L$ and after time L by 2^{-1}.

To bound the term for degenerate copies, observe first that for every subgraph $J \subsetneq F$ with $v_J \geq 3$ we have

$$\frac{e_F - 1}{v_F - 2} = m_2(F) > \frac{e_J - 1}{v_J - 2} \quad \text{and thus} \quad \frac{e_F - e_J}{v_F - v_J} = \frac{(e_F - 1) - (e_J - 1)}{(v_F - 2) - (v_J - 2)} > m_2(F).$$

We may thus choose an $\alpha > 0$ so that

$$(v_F - v_J) - \frac{e_F - e_J}{m_2(F)} < -\alpha \qquad \text{for all } J \subsetneq F \text{ with } v_J \geq 3.$$

We can now bound the case that the copy F_ℓ is a degenerate copy by

$$\sum_{J \subsetneq F, v_J \geq 3} (\ell \cdot v_F)^{v_J} \cdot n^{v_F - v_J} \cdot p^{e_F - e_J} < (\ell \cdot v_F \cdot 2^{e_F})^{v_F} \cdot n^{-\alpha}, \tag{3}$$

with room to spare.

We now do a union bound. For that we choose ξ such that $\xi \cdot \alpha > v_F + 1$ and then choose L such that the term in (1) is positive. (Observe that L is a constant that only depends on the graph F.) Finally, choose $\ell_0 = (v_F + 1) \log_2 n + \xi + L$.

Consider first all sequences of length $\ell' \leq \ell_0$ with the property that $F_{\ell'}$ is the ξth degenerate copy. Then the expected number of subgraphs in $G_{n,p}$ that can be built by such a sequence is at most

$$\sum_{\ell' \leq \ell_0} \tbinom{\ell'-1}{\xi-1} n^{v_F} \cdot [(\ell_0 v_F 2^{e_F})^{v_F} \cdot n^{-\alpha}]^{\xi} \cdot L^L \cdot 2^{-(\ell' - L - \xi)} \leq n^{v_F} \cdot o(n) \cdot n^{-\alpha \cdot \xi} = o(1),$$

by choice of ξ. Here the binomial coefficient corresponds to the choices (in time) when the first $\xi - 1$ degenerate steps occurred. The term n^{v_F} bounds the choice of the copy F_0. The next factor bounds the choices of degenerate steps, the last two factors bound those of the regular steps: as explained above, regular copies contribute a term of at most L if they occur before step L and a factor 1/2 if they occur after step L.

So we know that within the first ℓ_0 copies we have less than ξ degenerate ones. Observe that by the choice of L this implies that the sequence that generates G' either has length less than L (which is fine) or length at least ℓ_0. The latter follows from the fact that (1) is increasing in L, and thus $f_o(\ell') > 0$ for every $L < \ell' \leq \ell_0$. It thus suffices to consider all sequences of length ℓ_0. The expected number of subgraphs in $G_{n,p}$ that can be built by such a sequence is at most

$$\sum_{k < \xi} \tbinom{\ell_0}{k} n^{v_F} \cdot [(\ell_0 v_F 2^{e_F})^{v_F} \cdot n^{-\alpha}]^{k} \cdot L^L \cdot 2^{-(\ell_0 - L - k)} \leq n^{v_F} \cdot o(n) \cdot n^{-(v_F + 1)} = o(1),$$

by choice of ℓ_0. This concludes the proof of Lemma 6 and thus also the proof of the 0-statement.

Acknowledgment. The authors would like to thank Luca Gugelmann for his constructive criticism on an earlier version of the paper that allowed us to streamline some arguments.

References

[1] Balogh, J., Morris, R., and Samotij, W. Independent sets in hypergraphs. preprint, arXiv:1204.6530, 2012.

[2] Chen, B., Matsumoto, M., Wang, J. F., Zhang, Z. F., and Zhang, J. X. A short proof of Nash-Williams' theorem for the arboricity of a graph. *Graphs Combin. 10*, 1 (1994), 27–28.

[3] Conlon, D., and Gowers, T. Combinatorial theorems in sparse random sets. preprint, arXiv:1011.4310.

[4] Erdős, P., and Rényi, A. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl. 5* (1960), 17–61.

[5] Friedgut, E., Rödl, V., and Schacht, M. Ramsey properties of random discrete structures. *Random Structures Algorithms 37*, 4 (2010), 407–436.
[6] Gugelmann, L., Person, Y., Steger, A., and Thomas, H. A randomized version of ramsey's theorem. *Random Structures & Algorithms 41*, 4 (2012), 488–505.
[7] Łuczak, T., Ruciński, A., and Voigt, B. Ramsey properties of random graphs. *J. Combin. Theory Ser. B 56*, 1 (1992), 55–68.
[8] Marciniszyn, M., Skokan, J., Spöhel, R., and Steger, A. Asymmetric Ramsey properties of random graphs involving cliques. *Random Structures Algorithms 34*, 4 (2009), 419–453.
[9] Rödl, V., and Ruciński, A. Lower bounds on probability thresholds for Ramsey properties. In *Combinatorics, Paul Erdős is eighty, Vol. 1*, Bolyai Soc. Math. Stud. 1993, pp. 317–346.
[10] Rödl, V., and Ruciński, A. Random graphs with monochromatic triangles in every edge coloring. *Random Structures Algorithms 5*, 2 (1994), 253–270.
[11] Rödl, V., and Ruciński, A. Threshold functions for Ramsey properties. *J. Amer. Math. Soc. 8*, 4 (1995), 917–942.
[12] Saxton, D., and Thomason, A. Hypergraph containers. preprint, arXiv:1204.6595, 2012.

Appendix

For convenience of the reader we provide in this appendix the proof of the deterministic statement (Theorem 7) that was used in the proof of the 0-statement. Our proof essentially follows the approach in [9].

Using standard notation, we denote with $m(G) = \max_{J\subseteq G} \frac{e_J}{v_J}$ the *maximum density* of a graph G, and with

$$ar(G) = \max_{J\subseteq G, v_J\geq 2} \frac{e_J}{v_J - 1}$$

the *arboricity* of the graph G. One easily checks that all graphs G satisfy

$$m(G) \;\leq\; ar(G) \leq m(G) + \tfrac{1}{2}. \tag{4}$$

Let $\delta(G)$ denote the minimum degree in G, i.e., $\delta(G) := \min_{v\in V(G)} \deg(v)$. Furthermore, let $\delta_{max}(G) := \max_{G'\subseteq G} \delta(G')$ be the maximum minimum degree in all subgraphs of G.

The following lemma gives various conditions under which $G \nrightarrow (F)^e_2$.

Lemma 12. *Let G and F be two graphs such that at least one of the following properties is satisfied:*

(i) $ar(G) \leq 2 \cdot \lfloor ar(F) - \varepsilon \rfloor$ *for some* $\varepsilon > 0$,
(ii) $m(G) \leq 2 \cdot \lfloor m(F) - \varepsilon \rfloor$ *for some* $\varepsilon > 0$,
(iii) $\delta_{max}(G) \leq 2(\delta(F) - 1)$,
(iv) $m(G) < \delta_{max}(F)$ *and* $\chi(F) \geq 3$.

Then $G \nrightarrow (F)^e_2$.

Proof. Assume first that *(i)* holds. Nash-Williams' Arboricity Theorem (cf. [2] for a short and self-contained proof) states that for every graph $G = (V, E)$ there exists a partition of the edges into $\lceil ar(G) \rceil$ parts, $E = E_1 \cup \ldots \cup E_{\lceil ar(G)\rceil}$, such that all E_i are forests. If we thus colour the edges in the first $\lfloor ar(F) - \varepsilon \rfloor$ of these sets red and the remaining edges blue, then both the blue as well as the red subgraph have arboricity at most $\lfloor ar(F) - \varepsilon \rfloor < ar(F)$ and thus cannot contain a copy of F.

In case of *(ii)* we proceed similar. We replace Nash-Williams' theorem by the following statement: for every graph $G = (V, E)$ there exists a partition of the edges into $\lceil m(G) \rceil$ parts, $E = E_1 \cup \ldots \cup E_{\lceil m(G)\rceil}$, such that all components in E_i contain at most one cycle. (This follows easily from Hall's theorem applied to a bipartite graph with $\lceil m(G) \rceil$ copies of every vertex in one set and one copy of every edge in the other set.) Similarly as before we now colour the edges in the first $\lfloor m(F) - \varepsilon \rfloor$ of these sets red and the remaining edges blue. Then subgraphs have maximum density at most $\lfloor m(F) - \varepsilon \rfloor < m(F)$ and thus cannot contain a copy of F.

Now assume that *(iii)* holds. Construct a sequence $v_1, v_2, \ldots, v_{v_G}$ of the vertices in G as follows: let v_i be a vertex of minimum degree in $G - \{v_1, \ldots, v_{i-1}\}$. Then every vertex v_i has degree at most $\delta_{max}(G)$ into $G[\{v_{i+1},\ldots,v_{v_G}\}]$, the graph induced by the vertices 'to the right'. Color the vertices of G 'backwards', i.e., starting with v_{v_G} (which is coloured arbitrarily). As every vertex v_i has degree at most $\delta_{max}(G) \leq 2(\delta(F)-1)$ into the part that is already coloured, we can colour $\delta(F)-1$ of these edges blue and the remaining ones red. Clearly, the coloured part can then not contain a monochromatic copy of F that contains v_i. By repeating this procedure for every vertex v_i we obtain a colouring without a monochromatic F.

Finally, assume that *(iv)* holds. Here we proceed in two steps. We first show that we can find a *vertex*-colouring of G without a monochromatic copy of F. Then we show that this implies that there also exists an *edge*-colouring of G without a monochromatic copy of F. Indeed, the latter is easy. Assume there exists a 2-colouring of the vertices, that is, a partition $V = X \cup Y$ such that neither $G[X]$ nor $G[Y]$ contains a copy of F. Then we colour all edges in $G[X]$ and $G[Y]$ with red and all edges in $E(X,Y)$ with blue without inducing a monochromatic copy of F (as F is connected and is not bipartite). So we need to find the vertex-colouring. The following argument is from [7]. Let $F' \subseteq F$ such that $\delta(F') = \delta_{max}(F)$. We need to show that for every graph G with $m(G) < \delta(F')$ we find a vertex colouring of G without a monochromatic F'. Assume this is not true. Then there exists a minimal counterexample G_0. As G_0 is minimal we know that for every vertex $v \in V(G_0)$ the graph $G_0 - v$ does have a vertex colouring without a monochromatic F'. Clearly, if $deg(v) < 2\delta(F')$ then such a colouring can be extended to v. So we know that in G_0 every vertex has degree at least $2\delta(F')$, that is $m(G_0) \geq (\sum_v deg(v))/(2v_{G_0}) \geq \delta(F') = \delta_{max}(F)$. A contradiction. □

Let G' be a subgraph of G with minimum degree $\delta_{max}(G)$. Then $|E(G')| \geq \frac{1}{2}|V(G')| \cdot \delta_{max}(G)$ and we thus see that

$$2m(G) \geq \delta_{max}(G). \tag{5}$$

We use this observation to show that graphs with density at most 2 can be coloured without a monochromatic triangle.

Lemma 13. *Let G be a graph such that $m(G) \leq 2$. Then $G \nrightarrow (K_3)^e_2$.*

Proof. We proceed similarly as in the previous proof. We construct a sequence $v_1, v_2, \ldots,$ by choosing v_i as a vertex of minimum degree in $G - \{v_1,\ldots,v_{i-1}\}$, with the additional condition that the neighborhood of v_i in $G - \{v_1,\ldots,v_{i-1}\}$ is not a K_4. If we do not find a vertex that satisfies this property then we stop. As $\delta_{max}(G) \leq 2m(G) \leq 4$ we will always find a vertex with degree at most four. Also note that if the minimum degree *is* four, then the graph is 4-regular. That is, the above process can only stop if *every* vertex has degree 4 *and* has the property that its neighborhood induces a K_4. In other words, if we cannot find a vertex v_i, then $G' := G - \{v_1,\ldots,v_{i-1}\}$ is a union of vertex-disjoint K_5's. Since there exists a 2-colouring of a K_5 without a monochromatic triangle, we can thus 2-colour G' without a monochromatic triangle. Now we proceed again as in the previous proof and colour the remaining vertices in reverse order. By construction vertex v_i has degree at most 4 into $G[\{v_i,\ldots,v_{v_G}\}]$ and the neighbourhood of v_i in $G[\{v_i,\ldots,v_{v_G}\}]$ is not a K_4. A simple case checking shows that however the neighborhood of v_i is coloured without a monochromatic triangle there is always an extension of the colouring to the edges incident to v_i so that no monochromatic triangle is generated. □

Proof of Theorem 7. Observe that for $v_F = 3$ the only graph with $m_2(F) > 1$ is the triangle, for which the claim of the theorem holds according to Lemma 13. In the remainder of the proof we thus assume $v_F \geq 4$. Observe that we may also assume without loss of generality that F is strictly 2-balanced.

The assumption that F is strictly 2-balanced implies that $m_2(F) = \frac{e_F-1}{v_F-2} > \frac{e_F-\delta(F)-1}{v_F-3}$ from which we deduce $m_2(F) < \delta(F) \leq \delta_{max}(F)$. Thus if $\chi(F) \geq 3$, then F satisfies property (iv) of Lemma 12, which implies $G \nrightarrow (F)^e_2$. Therefore, in the following, we assume that F is bipartite. Then $e_F \leq \frac{1}{4}v_F^2$ implies that

$$m_2(F) \leq m(F) + \tfrac{1}{2} \quad \text{with equality if and only if } e_F = \tfrac{1}{4}v_F^2.$$

If $m_2(F) = k + x$ for some $k \in \mathbb{N}$ and $\frac{1}{2} \leq x < 1$ we thus have $m(F) > k$ whenever $x > \frac{1}{2}$ or $e_F < \frac{1}{4}v_F^2$. In this case we have $m(G) \leq k + 1 \leq 2k = 2\lfloor m(F) - \varepsilon \rfloor$ and F satisfies property (ii) of Lemma 12, which concludes the proof of the theorem in this case. So we may assume that $x = \frac{1}{2}$ and $e_F = \frac{1}{4}v_F^2$. Then, $v_F = 2\ell$ for some $\ell \in \mathbb{N}$, and thus $m_2(F) = (\ell^2 - 1)/(2\ell - 2) = \frac{1}{2}(\ell + 1)$. That is, $k = \frac{1}{2}\ell$ and so $ar(F) = e_F/(v_F - 1) > \frac{1}{4}v_F = k$. By (4) we also have $ar(G) \leq m(G) + \frac{1}{2} \leq m_2(F) + \frac{1}{2} = k + 1$. Thus $ar(G) \leq k + 1 \leq 2k \leq 2\lfloor ar(F) - \varepsilon \rfloor$ and F satisfies property (i) of Lemma 12, which concludes the proof of the theorem for this case.

Finally, if $m_2(F) = k + x$ for some $k \in \mathbb{N}$ and $0 \leq x < 0.5$ then (5) implies that $\delta_{max}(G) \leq 2m(G) \leq 2m_2(F)$ and thus $\delta_{max}(G) \leq 2k$, as $\delta_{max}(G)$ is integral. On the other hand, we have already shown that the assumption that F is strictly 2-balanced implies that $m_2(F) < \delta(F)$. The fact that $\delta(F)$ is integral thus implies $\delta(F) \geq k+1$, and F satisfies property (iii) of Lemma 12, which concludes the proof of the theorem for this case. □