

# Partitioning random graphs into monochromatic components

Deepak Bal

Department of Mathematical Sciences  
Montclair State University  
Montclair, New Jersey; U.S.A.  
deepak.bal@montclair.edu

Louis DeBiasio\*

Department of Mathematics  
Miami University  
Oxford, Ohio; U.S.A.  
debiasld@miamioh.edu

Submitted: Apr 21, 2016; Accepted: Jan 18, 2017; Published: Feb 3, 2017  
Mathematics Subject Classifications: 05E30, 05B25

## Abstract

Erdős, Gyárfás, and Pyber (1991) conjectured that every  $r$ -colored complete graph can be partitioned into at most  $r - 1$  monochromatic components; this is a strengthening of a conjecture of Lovász (1975) and Ryser (1970) in which the components are only required to form a cover. An important partial result of Haxell and Kohayakawa (1995) shows that a partition into  $r$  monochromatic components is possible for sufficiently large  $r$ -colored complete graphs.

We start by extending Haxell and Kohayakawa's result to graphs with large minimum degree, then we provide some partial analogs of their result for random graphs. In particular, we show that if  $p \geq \left(\frac{27 \log n}{n}\right)^{1/3}$ , then a.a.s. in every 2-coloring of  $G(n, p)$  there exists a partition into two monochromatic components, and for  $r \geq 2$  if  $p \ll \left(\frac{r \log n}{n}\right)^{1/r}$ , then a.a.s. there exists an  $r$ -coloring of  $G(n, p)$  such that there does not exist a cover with a bounded number of components. Finally, we consider a random graph version of a classic result of Gyárfás (1977) about large monochromatic components in  $r$ -colored complete graphs. We show that if  $p = \frac{\omega(1)}{n}$ , then a.a.s. in every  $r$ -coloring of  $G(n, p)$  there exists a monochromatic component of order at least  $(1 - o(1))\frac{n}{r-1}$ .

## 1 Introduction

For a graph  $G$  and positive integer  $r$ , the  **$r$ -color tree-partition (tree-cover) number** of  $G$ , denoted by  $\text{tp}_r(G)$  ( $\text{tc}_r(G)$ ), is the minimum  $s$  such that for every  $r$ -edge-coloring of  $G$ , there exists a collection of monochromatic connected subgraphs  $\{H_1, \dots, H_t\}$  with

---

\*Research supported in part by Simons Foundation Collaboration Grant # 283194.

$t \leq s$  such that  $\{V(H_1), \dots, V(H_t)\}$  forms a partition (cover) of  $V(G)$ ; as each subgraph  $H_i$  contains a monochromatic spanning tree we use “connected subgraph”, “tree”, and “component” interchangeably throughout the paper. Similarly define  $\text{pp}_r(G)$ ,  $\text{cp}_r(G)$  to be the  $r$ -color path-partition number and  $r$ -color cycle-partition number of  $G$  respectively.

Gyárfás [18] noted that the following is an equivalent formulation of what is known in the literature as “Ryser’s conjecture” or the “Lovász-Ryser conjecture.”

**Conjecture 1.1** (Ryser 1970 (see [22]), Lovász 1975 [28]). *Let  $r \geq 2$ . For all graphs  $G$ ,  $\text{tc}_r(G) \leq (r-1)\alpha(G)$ .*

If true, this conjecture is best possible when  $r-1$  is a prime power by a well known example using affine planes<sup>1</sup>. For  $r=2$ , this is equivalent to the König-Egerváry theorem. Aharoni [1] proved the  $r=3$  case, and for  $r \geq 4$  it is open. Slightly more is known in the case  $\alpha=1$  (i.e. when  $G=K_n$ ), where it has been proved for  $r \leq 5$  (see [19] and [15] for more details).

In a seminal paper, Erdős, Gyárfás, and Pyber [11] proved that for all  $r \geq 2$ ,

$$\text{tp}_r(K_n) \leq \text{pp}_r(K_n) \leq \text{cp}_r(K_n) = O(r^2 \log r)$$

and made the following conjecture.

**Conjecture 1.2** (Erdős, Gyárfás, Pyber 1991). *For all  $r \geq 2$ ,  $\text{tp}_r(K_n) = r-1$ ,  $\text{pp}_r(K_n) = \text{cp}_r(K_n) = r$ .*

For countably infinite complete graphs, Conjecture 1.2 is known to be true for paths and cycles for all  $r$ , and known to be true for trees when  $r=2, 3$  (with the appropriate notion of paths, cycles, and trees). Rado [33] proved  $\text{pp}_r(K_{\mathbb{N}}) = r$ ; Elekes, D. Soukup, L. Soukup, and Szentmiklóssy [10] proved  $\text{cp}_r(K_{\mathbb{N}}) = r$ ; and Nagy and Szentmiklóssy (see [11]) proved  $\text{tp}_3(K_{\mathbb{N}}) = 2$ . However, for finite complete graphs the story is more complicated.

For trees, an old remark of Erdős and Rado says that a graph or its complement is connected, i.e.  $\text{tp}_2(K_n) = 1$ . Erdős, Gyárfás, and Pyber [11] proved that  $\text{tp}_3(K_n) = 2$ . Later, Haxell and Kohayakawa [21] proved the following.

**Theorem 1.3** (Haxell, Kohayakawa 1995). *Let  $r \geq 2$ . If  $n \geq \frac{3r^4 r! \log r}{(1-1/r)^{3(r-1)}}$ , then*

$$\text{tp}_r(K_n) \leq r.$$

Later, Fujita, Furuya, Gyárfás, and Tóth [14] conjectured that the partition version of Conjecture 1.1 is true and proved it in the case when  $r=2$ .

For paths, Gerencsér and Gyárfás gave a simple proof of  $\text{pp}_2(K_n) = 2$  (see the footnote in [17]). Much later, Pokrovskiy [31] proved  $\text{pp}_3(K_n) = 3$ .

For cycles, Lehel conjectured that  $\text{cp}_2(K_n) = 2$  (in fact, with cycles of different colors). This was proved for large  $n$  by Łuczak, Rödl, and Szemerédi [29], and then for smaller,

<sup>1</sup>In an affine plane of order  $r-1$ , there are  $r$  parallel classes of  $r-1$  lines each. To each of the  $r$  parallel class assign a distinct color.

but still large  $n$  by Allen [2], and finally for all  $n$  by Bessy and Thomassé [5]. For general  $r$ , Gyárfás, Ruszinkó, Sárközy, Szemerédi [20] improved the result from [11] to

$$\text{pp}_r(K_n) \leq \text{cp}_r(K_n) \leq 100r \log r.$$

However for  $r \geq 3$ , Pokrovskiy [31] proved  $\text{cp}_r(K_n) > r$ .

## 1.1 Large minimum degree

Motivated by a new class of Ramsey-Turán type problems raised by Schelp [35], Balogh, Barát, Gerbner, Gyárfás, and Sárközy [4] conjectured (and proved an approximate version of) a significant strengthening of Bessy and Thomassé's result. That is, if  $\delta(G) > 3n/4$ , then  $\text{cp}_2(G) \leq 2$  (with cycles of different colors); they also provided an example which shows that the conjecture would be best possible. DeBiasio and Nelsen [8] proved that this holds for  $G$  with  $\delta(G) > (3/4 + o(1))n$  and then Letzter [26] proved that it holds exactly for sufficiently large  $n$ .

In Observation 3.1, we note that there are graphs with minimum degree  $n - r$  for which  $\text{tp}_r(G) \geq \text{tc}_r(G) \geq r$  and thus it is natural to wonder how small we can make  $\delta(G)$  while maintaining  $\text{tp}_r(G) \leq r$ . In Theorem 4.3, we prove a strengthening of Theorem 1.3 for graphs with large minimum degree. A corollary of our result is the following.

**Corollary 1.4.** *For all  $r \geq 2$  there exists  $n_0$  such that if  $G$  is a graph on  $n \geq n_0$  vertices with  $\delta(G) > (1 - \frac{1}{er!})n$ , then  $\text{tp}_r(G) \leq r$ .*

Furthermore, in Example 3.3 we show that the minimum degree in the above result cannot be improved beyond  $\sim (1 - \frac{1}{r+1})n$ . Additionally, in Theorem 4.6 we show that for *covering* 2-colored graphs with two monochromatic trees, this lower bound on the minimum degree is tight.

Theorem 4.3 actually gives a “robust” tree partition; that is, a collection of trees together with a linear sized set  $L$  such that after deleting any subset of  $L$ , the remaining graph has a tree partition. This is important as we will use it to obtain results on the tree partition number of the random graph  $G(n, p)$ .

Finally, as a consequence of our method of proof, we are able to improve the bound on  $n$  in Theorem 1.3.

**Theorem 1.5.** *For  $r \geq 2$  and  $n \geq 3r^2r! \log r$ ,  $\text{tp}_r(K_n) \leq r$ .*

## 1.2 Random graphs

An active area of current research concerns sparse random analogs of combinatorial theorems (see the survey of Conlon [7]). An early example of such a result is the so called Random Ramsey Theorem. Say  $G \rightarrow_r H$  if every  $r$ -coloring of  $G$  contains a monochromatic copy of  $H$ . For fixed graphs  $H$ , Rödl and Ruciński [34] determined the threshold for which a.a.s.<sup>2</sup>  $G(n, p) \rightarrow_r H$ . For the case of paths, Letzter [25] proved that for  $p = \frac{\omega(1)}{n}$

<sup>2</sup>We say that a sequence of events  $A_n$  happens **a.a.s.** if  $\lim_{n \rightarrow \infty} \mathbb{P}[A_n] \rightarrow 1$ .

a.a.s.,  $G(n, p) \rightarrow_2 P_{(2/3-o(1))n}$ . Random analogs of asymmetric Ramsey problems, hypergraph Ramsey problems, and van der Waerden's Theorem have also been studied (again, see [7]).

In light of these results, it is a natural question to ask whether monochromatic partitioning problems can be extended to the realm of random graphs in an interesting way. Towards this, we prove the following results which provide partial analogs of Conjecture 1.1, and Theorem 1.3 for random graphs.

**Theorem 1.6.** *For all  $r \geq 2$ , there exists  $C \geq r$  such that a.a.s.*

- (i) *if  $p \geq \left(\frac{27 \log n}{n}\right)^{1/3}$  then  $\text{tp}_2(G(n, p)) \leq 2$ , and*
- (ii) *if  $p \geq \left(\frac{C \log n}{n}\right)^{1/(r+1)}$ , then  $\text{tc}_r(G(n, p)) \leq r^2$ , and*
- (iii) *if  $p \geq \left(\frac{C \log n}{n}\right)^{1/r}$ , then there is a collection of  $r$  vertex disjoint monochromatic trees which cover all but at most  $9r \log n/p = O(n^{1/r}(\log n)^{1-1/r})$  vertices.*

**Theorem 1.7.** *For all  $r \geq 2$ ,*

- (i) *if  $p = \left(\frac{r \log n - \omega(1)}{n}\right)^{1/r}$ , then  $\text{tc}_r(G(n, p)) > r$ , and*
- (ii) *if  $p = o\left(\left(\frac{r \log n}{n}\right)^{1/r}\right)$ , then  $\text{tc}_r(G(n, p)) \rightarrow \infty$ .*

It is interesting to compare Conjecture 1.1 to Theorem 1.7 as our results imply that almost every graph  $G \sim G(n, 1/2)$  (the uniform distribution on all graphs with  $n$  vertices) satisfies  $\text{tc}_r(G) \leq r^2$ , which is much smaller than the conjectured upper bound of  $(r-1)\alpha(G)$  since it is known (see e.g. [13]) that a.a.s.  $\alpha(G) \sim 2 \log_2 n$ . So not only are tightness examples rare, examples for which  $\text{tp}_r(G) \geq \text{tc}_r(G) > r^2$  are rare.

### 1.3 Large monochromatic components

We consider one further related line of research. Note that if an  $r$ -colored graph  $G$  can be covered by  $t$  monochromatic components, then  $G$  contains a monochromatic component of order at least  $|V(G)|/t$ . So we may directly ask how large of a monochromatic component we may find in an  $r$ -colored graph<sup>3</sup>. Given a positive integer  $r$  and a graph  $G$ , we let  $\text{tm}_r(G)$  be the maximum integer  $s$  such that the following holds: in every  $r$ -coloring of the edges of  $G$ , there exists a monochromatic component with at least  $s$  vertices. For  $r \geq 2$ , Gyárfás [18] proved  $\text{tm}_r(K_n) \geq \frac{n}{r-1}$  and Füredi [16] proved  $\text{tm}_r(G) \geq \frac{n}{(r-1)\alpha(G)}$  for all graphs  $G$  (see Theorem 5.6 in [19]). Furthermore, this is tight when  $r-1$  is a prime power using the same affine plane example mentioned before. Given the discussion above, note that Füredi's result would be implied by Conjecture 1.1.

Concerning random graphs, two sets of authors [37], [6] independently found the threshold for  $\text{tm}_r(G(n, p)) = \Theta(n)$ . Specifically, they prove that there exists an analytically computable constant  $\psi_r$  such that if  $c < \psi_r$ , then  $\text{tm}_r(G(n, c/n)) = o(n)$  and if  $c > \psi_r$  then  $\text{tm}_r(G(n, c/n)) = \Omega(n)$ .

<sup>3</sup>Historically, the monochromatic partitioning problems mentioned in the first part of the introduction were motivated by their implications for graph Ramsey problems (see [32]).

We prove<sup>4</sup> the following random analog of the fact that  $\text{tm}_r(K_n) \geq \frac{n}{r-1}$ .

**Theorem 1.8.** *For all  $r \geq 2$  and sufficiently small  $\epsilon > 0$ , there exists  $C$  such that for  $p \geq \frac{C}{n}$ , a.a.s. every  $r$ -coloring of  $G(n, p)$  contains either a monochromatic tree of order at least  $(1 - \epsilon)n$  or a monochromatic tree with at least  $(1 - \epsilon)\frac{n}{r-1}$  leaves, which implies  $\text{tm}_r(G(n, p)) \geq (1 - \epsilon)\frac{n}{r-1}$ .*

Again, it is interesting to compare Theorem 1.8 to the corresponding deterministic version, as our result implies that almost every graph  $G$  satisfies  $\text{tm}_r(G) \geq (1 - \epsilon)\frac{n}{r-1}$ , which is much larger than the bound of  $\frac{n}{(r-1)\alpha(G)}$  given by Füredi's result for which there are examples showing tightness.

## 2 Overview and notation

### 2.1 Overview

We consider large minimum degree versions and random versions of some classic results for edge colored complete graphs. In certain cases we will use the large minimum degree results together with the sparse regularity lemma to obtain results for random graphs. In these cases our approach is as follows: First, prove that edge colored graphs of high minimum degree contain (a robust version of) the desired structure. Second, applying the sparse regularity lemma to the random graph gives a reduced graph with high minimum degree and thus we can apply the high minimum degree result. This structure in the reduced graph corresponds to an approximate spanning structure in the original graph. As a simple application of this approach we obtain Theorem 1.8.

A less standard application of this approach is given in the proof of Theorem 1.6(iii) where we are trying to improve the exponent from  $1/(r+1)$  to  $1/r$ . We use the method of multiple exposures to build a tree cover while maintaining a set of vertices which are leaves in each of the monochromatic trees. On each step, the leaf set shrinks by a factor of  $p$  and at the end of the possibly  $r$  steps, we require the leaf set to contain more than  $\log n$  vertices. By using sparse regularity together with the large minimum degree result we are able to begin this process with a tree having  $\Theta(n)$  leaves as opposed to the  $\Theta(pn)$  leaves we would be able to guarantee without sparse regularity.

In Section 3.1, we provide examples of graphs and colorings which give lower bounds on  $\text{tc}_r(G)$  (and hence  $\text{tp}_r(G)$ ). In Section 3.2 we consider the variant where we require that the components in the cover must be of distinct colors. In Section 3.3, we give a simple upper bound on  $\text{tc}_r(G)$  and prove a result about graphs in which every  $r+1$  vertices have a common neighbor.

Section 4 is devoted to proving the large minimum degree versions (including complete versions) of our results. In Section 4.1 we prove Theorems 1.5 and 4.3. The first provides a slight improvement on the bound in Theorem 1.3 and the second extends the theorem to graphs with large minimum degree. In Section 4.2 we prove Theorem 4.6, which provides

---

<sup>4</sup>Essentially the same result was independently discovered by Dudek and Prałat [9].

a tight minimum degree condition on  $G$  such that  $\text{tc}_2(G) \leq 2$ . In Section 4.3, we prove that  $r$ -colored graphs with large minimum degree have large monochromatic components.

In Section 5, we continue with the second step of the method described above by stating the sparse regularity lemma of Kohayakawa [23] and Rödl (see [7]) as well as collecting various lemmas which will be useful for the proof. Lemma 5.7 shows that sparse edge colored random graphs have nearly spanning “robust” tree partitions.

In Section 6 we deduce some properties of  $G(n, p)$  which will be used in Section 7.

In Section 7 we prove Theorem 1.6 and 1.7. Theorem 1.6(ii) and Theorem 1.7(i),(ii) will follow from the results of Section 6. For Theorem 1.6(i), we are able to exploit the fact that there are only two colors to improve the general result of Theorem 1.6(ii). The proof of Theorem 1.6(iii) is discussed in the second paragraph of this section. Finally, in Section 7.3, we prove Theorem 1.8.

In Section 8 we collect some conjectures and open problems.

## 2.2 Notation

We use the following notation throughout the paper. As usual,  $N(v)$  represents the neighborhood of  $v$  and as we deal mainly with colored graphs, if  $c$  is a color then  $N_c(v)$  represents the neighborhood of  $v$  in the subgraph of  $c$  colored edges and  $\deg_c(v) = |N_c(v)|$ . If  $S$  is a set of vertices, then  $N_c(v, S) = N_c(v) \cap S$  and  $\deg_c(v, S) = |N_c(v, S)|$ . We let  $N^\cap(S) = \bigcap_{v \in S} N(v)$  and  $N^\cup(S) = \bigcup_{v \in S} N(v)$ . For two sets of vertices  $X$  and  $Y$ ,  $e(X, Y)$  represents the number of edges with one endpoint in  $X$  and the other in  $Y$ . For two sequences  $a_n, b_n$ , we write  $a_n = o(b_n)$  if  $a_n/b_n \rightarrow 0$  as  $n \rightarrow \infty$  and  $a_n = \omega(b_n)$  if  $a_n/b_n \rightarrow \infty$  as  $n \rightarrow \infty$ . For constants  $a$  and  $b$ , we write  $a \ll b$  to mean that given  $b$ , we can choose  $a$  small enough so that  $a$  satisfies all of necessary conditions throughout the proof. More formally, we say that a statement holds for  $a \ll b$  if there is a function  $f$  such that it holds for every  $b$  and every  $a \leq f(b)$ . In order to simplify the presentation, we will not determine these functions explicitly. We will ignore floors and ceilings when they are not crucial to the calculation. Logarithms are assumed to be base  $e$  unless otherwise noted.

## 3 Examples and observations

### 3.1 Lower bounds on the tree cover number

In this section we provide examples which give lower bounds on  $\text{tc}_r(G)$  in various settings considered throughout the paper. We remind the reader that for all  $r \geq 1$  and all graphs  $G$ ,  $\text{tp}_r(G) \geq \text{tc}_r(G)$ .

**Observation 3.1.** *Let  $r \geq 1$ . For all graphs  $G$ , if  $\alpha(G) \geq r$ , then  $\text{tc}_r(G) \geq r$ . In particular, there exists a graph  $G$  with  $\delta(G) \geq n - r$  with  $\text{tc}_r(G) \geq r$ .*

*Proof.* Choose an independent set  $\{x_1, \dots, x_r\}$  and color every edge incident to  $x_i$  with color  $i$ , then color the remaining edges arbitrarily. None of the vertices  $x_1, \dots, x_r$  are in a tree of the same color.  $\square$

**Observation 3.2.** Let  $G = (V, E)$  be a graph. For all  $1 \leq r \leq s$ , if  $G$  contains an independent set  $X$  of size  $s$  such that every vertex in  $V \setminus X$  has at most  $r - 1$  neighbors in  $X$  and  $X$  is not a dominating set, then  $\text{tc}_r(G) > s$ .

*Proof.* Start by coloring every edge in  $G[V \setminus X]$  with color  $r$ . The only edges not yet colored are those going between  $V \setminus X$  and  $X$ . Since each  $v \in V \setminus X$  is incident with at most  $r - 1$  such edges, we can assign colors so that no vertex in  $V \setminus X$  is incident with more than one edge of color  $i$  for any  $i \in [r - 1]$ .

To see that  $G$  cannot be covered with  $s$  monochromatic trees, note that for any pair  $x, x' \in X$ ,  $x$  and  $x'$  must be in different trees; this follows since  $x$  and  $x'$  are not incident with any edges of color  $r$  and  $x$  and  $x'$  have no neighbors of the same color (by the way colors were assigned to edges from  $V \setminus X$  to  $X$ ), so there are no monochromatic paths from  $x$  to  $x'$ . Furthermore, since  $X$  is not a dominating set, there must exist at least one tree of color  $r$  (since every edge in  $G[V \setminus X]$  has color  $r$ ). This implies that  $\text{tc}_r(G) \geq s + 1$ .  $\square$

**Example 3.3.** For  $r \geq 1$  and  $n \geq 2r + 2$ , there exists a graph  $G$  on  $n$  vertices with  $\delta(G) = \left\lceil \frac{r(n-r-1)+1}{r+1} \right\rceil - 1$  such that  $\text{tc}_r(G) > r$ .

*Proof.* There exist unique integers  $m$  and  $q$  such that  $n = (r+1)m + q$  with  $0 \leq q \leq r$ . Set aside vertices  $u_1, \dots, u_{r+1}$  and then equitably partition the remaining  $n - (r+1)$  vertices into sets  $V_1, \dots, V_{r+1}$ ; that is, partition the remaining vertices into sets  $V_1, \dots, V_{r+1}$  so that  $|V_1| = \dots = |V_{r+1-q}| = m - 1$  and if  $q \geq 1$ ,  $|V_{r+1-q+1}| = \dots = |V_{r+1}| = m$ .

Now add the following colored edges:

- $u_i \sim V_j$  in color  $i$  for  $1 \leq i < j \leq r + 1$
- $u_i \sim V_j$  in color  $i - 1$  for  $1 \leq j < i \leq r + 1$
- $V_i \sim V_j$  in color  $r$  for  $1 \leq i < j \leq r$
- $V_{r+1} \sim V_i$  in color 1 for  $2 \leq i \leq r$
- $V_i \sim V_i$  with arbitrary colors for all  $i$

Note that if  $i \neq j$ , then  $u_i$  and  $u_j$  cannot be in the same monochromatic component. Among  $u_1, \dots, u_{r+1}$ , color  $i$  only appears incident to  $u_i$  and  $u_{i+1}$  but their neighborhoods in color  $i$  are disjoint; these neighborhoods remain disjoint in color  $i$  even when the edges between the  $V_i$ 's are considered.

By construction, it is clear that  $u_{r+1}$  has the smallest degree of the  $u_i$  and vertices in  $V_1$  have the smallest degree of the vertices in the  $V_i$ . Note that  $\deg(u_{r+1}) = |V_1| + \dots + |V_r|$  which is  $r(m - 1)$  if  $q = 0$  and  $r(m - 1) + q - 1$  if  $q = 1, \dots, r$ . Now since  $n = (r+1)m + q$ , we have

$$\left\lceil \frac{r(n - r - 1) + 1}{r + 1} \right\rceil - 1 = r(m - 1) + \left\lceil \frac{rq + 1}{r + 1} \right\rceil - 1.$$

Since  $q \geq \frac{rq+1}{r+1} > q - 1$  for all  $1 \leq q \leq r$  and  $\left\lceil \frac{1}{r+1} \right\rceil = 1$ ,  $u_{r+1}$  satisfies the claimed degree condition for all  $0 \leq q \leq r$ . If  $v_1 \in V_1$ , then  $\deg(v_1) \geq (r - 1)(m - 1) + (m - 2) + r = r(m - 1) + r - 1 \geq \deg(u_{r+1})$ .  $\square$

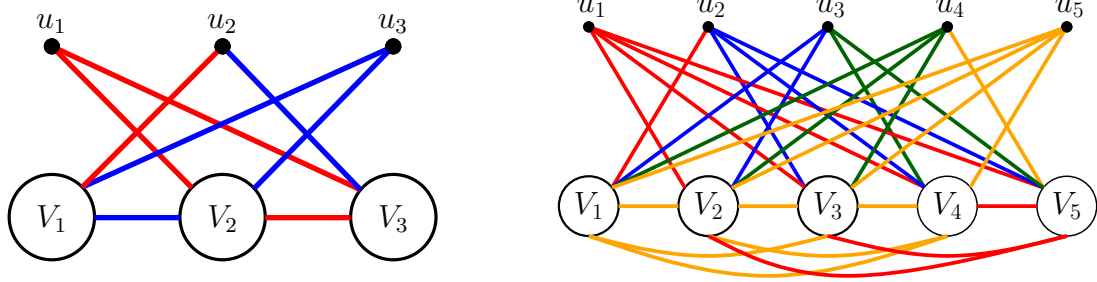


Figure 1: Graphs with  $\text{tc}_2(G) > 2$  and  $\text{tc}_4(G) > 4$  respectively.

### 3.2 Covering with trees of distinct colors

**Definition 3.4.** Let  $G$  be a (multi)graph. Say  $G$  has property  $\mathcal{TP}_r(\mathcal{TC}_r)$  if in every  $r$ -coloring of the edges of  $G$  there is a partition (cover) of  $V(G)$  with at most  $r$  trees of distinct colors.

Next we provide examples of graphs which cannot be covered by  $r$  trees of distinct colors. We note however, that these graphs can be partitioned into just two components of the same color.

**Example 3.5.** For all  $r \geq 1$  and  $n \geq 2^r$ , there exists a graph  $G$  on  $n$  vertices with  $\delta(G) = \lfloor (1 - \frac{1}{2^r})n \rfloor - 1$  such that  $G$  does not have property  $\mathcal{TC}_r$ .

*Proof.* We construct a graph  $G$  with  $n = m \cdot 2^r$  vertices. The vertices are partitioned into  $2^r$  sets of size  $m$ . We index these sets by binary strings of length  $r$ . So  $V = \bigcup_{\mathbf{b} \in \{0,1\}^r} V_{\mathbf{b}}$ . Every vertex within such a set will also be referred to by the index of its set. For a binary string  $\mathbf{b}$ , let  $\neg \mathbf{b}$  represent the string with all bits flipped.

Include every edge between vertices which agree on at least one index. So  $V_{\mathbf{b}} \sim V_{\mathbf{b}'}$  iff  $\mathbf{b} \neq \neg \mathbf{b}'$  (all edges within the  $V_{\mathbf{b}}$  are present as well). This graph has  $\delta(G) = n - 1 - m = (1 - \frac{1}{2^r})n - 1$ . Now color each edge with the smallest coordinate on which the endpoints indices agree. For example, two vertices with indices  $(0, 1, 0, 0)$  and  $(1, 0, 0, 1)$ , would be connected by an edge of color 3.

We claim that  $G$  cannot be covered by  $r$  monochromatic components of distinct colors. Any connected subgraph of color  $i$  can only contain vertices which agree on the  $i$ th coordinate. Suppose the component of color  $i$  only covers vertices with  $b_i$  in the  $i$ th component. Then the vertices with index  $\neg(b_1, \dots, b_r)$  are not covered by any of the components. When  $n = m \cdot 2^r + q$  with  $q < 2^r$ , we proceed in the same way, but partition the vertices into  $q$  sets of size  $m + 1$  and  $2^r - q$  sets of size  $m$ .  $\square$

### 3.3 Simple upper bounds on the tree cover number

**Observation 3.6.** Let  $r \geq 1$ . For all (multi)graphs  $G$ ,  $\text{tc}_r(G) \leq r\alpha(G)$ .

*Proof.* Let  $a := \alpha(G)$  and let  $X = \{x_1, \dots, x_a\}$  be a maximum independent set. So every vertex in  $V(G) \setminus X$  has a neighbor in  $X$ . Taking the stars centered at  $x_1, \dots, x_a$  gives a collection of at most  $r\alpha(G)$  monochromatic components which cover  $V(G)$ .  $\square$



**Proposition 3.7.** *Let  $r \geq 2$  and let  $G$  be a graph having the property that every set of  $r + 1$  vertices have a common neighbor, then  $\text{tc}_r(G) \leq r^2$ .*

*Proof.* Consider any  $r$ -coloring of  $G$  and let  $H$  be an  $r$ -colored auxiliary (multi)graph on  $V(G)$  where  $uv \in E(H)$  of color  $i$  if and only if there is a path in  $G$  of color  $i$  from  $u$  to  $v$ . Since every set of  $r + 1$  vertices of  $G$  have a common neighbor and there are at most  $r$  colors, this implies  $\alpha(H) \leq r$ . Thus by Observation 3.6, we have  $\text{tc}_r(H) \leq r\alpha(H) \leq r^2$ . Note that a monochromatic component in  $H$  corresponds to a monochromatic component in  $G$  giving the result.  $\square$

## 4 Monochromatic trees in graphs with large minimum degree

### 4.1 Partitions

We start by proving a lemma which we will use in the proofs of Theorems 1.5, 4.3, and 1.6(iii).

**Lemma 4.1.** *Let  $k \geq 2$ . If  $G$  is an  $Y, Z$ -bipartite graph such that for all  $v \in Z$ ,  $\deg(v, Y) > k \log |Z|$ , then in every  $k$ -coloring of the edges of  $G$ , there exists a partition  $\{Y_1, \dots, Y_k\}$  of  $Y$  such that for all  $v \in Z$ , there exists  $i \in [k]$  such that  $N_i(v) \cap Y_i \neq \emptyset$ .*

*Proof.* Randomly color the vertices of  $Y$  with colors from  $[k]$ , giving us a partition  $\{Y_1, \dots, Y_k\}$  of  $Y$  (with possibly empty parts). The probability that some vertex  $v \in Z$  has  $N_i(v) \cap Y_i = \emptyset$  for all  $i \in [k]$  is

$$\left(1 - \frac{1}{k}\right)^{\deg(v, Y)} < \left(1 - \frac{1}{k}\right)^{k \log |Z|} \leq e^{-k \log |Z|/k} = \frac{1}{|Z|}.$$

So by the union bound the probability of at least one failure is less than 1, and thus there exists a partition of  $Y$  with the desired property.  $\square$

We now prove Theorem 1.5 which says that  $\text{tp}_r(K_n) \leq r$  holds provided  $n \geq 3r^2 r! \log r$  (improving the lower bound of  $n \geq \frac{3r^4 r! \log r}{(1-1/r)^{3(r-1)}}$  from Theorem 1.3) and illustrates the idea for both Theorem 4.3 and Theorem 1.6(iii). We note that our proof follows a similar procedure as the proof in [21], except that at the end of the process we use Lemma 4.1 instead of a greedy algorithm.

*Proof of Theorem 1.5. Step 1:* Let  $x_1 \in V(G)$  and let  $Y_1$  be the largest monochromatic neighborhood of  $x_1$ , say the color is 1. Note that  $|Y_1| \geq (n-1)/r$ . If every vertex in  $V \setminus Y_1$  has a neighbor of color 1 in  $Y_1$ , then stop as we would already have the desired tree partition. So some vertex  $x_2$  has at least  $\frac{1}{r-1}|Y_1|$  neighbors of say color 2 in  $Y_1$ . Set  $Y_2 := Y_1 \cap N_2(x_2)$ .

For  $2 \leq i \leq r-1$ , assuming  $Y_i$  has already been defined, we do the following: if for all  $v \in V \setminus Y_i$ ,  $|(\bigcup_{j=1}^i N_j(v)) \cap Y_i| > i \log n$ , then set  $k := i$ ,  $Y := Y_k$ , and  $Z = V \setminus (\{x_1, \dots, x_k\} \cup Y_k)$  then proceed to Step 2. Otherwise some vertex  $x_{i+1} \in V \setminus Y_i$  has at most  $i \log n$  neighbors having colors from  $[i]$  in  $Y_i$  and thus  $x_{i+1}$  has at least

$\frac{1}{r-i}(|Y_i| - i \log n)$  neighbors of color say  $i + 1$ , in  $Y_i$ . Set  $Y_{i+1} := Y_i \cap N_{i+1}(x_{i+1})$ . Continue in this manner until we go to Step 2 or until  $Y_r$  has been defined. After we complete the  $i = (r - 1)$ -th step, we have

$$|Y_r| \geq \frac{n-1}{r!} - \log n \sum_{j=1}^{r-2} \frac{r-j}{j!} \geq r \log n$$

where the last inequality holds provided  $\frac{n}{\log n} \geq r! \sum_{j=0}^{r-2} \frac{r-j}{j!}$ . Recall that  $n \geq 3r^2 r! \log r$  and note that  $\sum_{j=0}^{r-2} \frac{r-j}{j!} \leq er$ . For  $r \geq 6$ , we have  $\frac{n}{\log n} \geq err! \geq r! \sum_{j=0}^{r-2} \frac{r-j}{j!}$ , and for  $2 \leq r \leq 5$  we directly verify  $\frac{n}{\log n} \geq r! \sum_{j=0}^{r-2} \frac{r-j}{j!}$ . Now set  $k := r$ ,  $Y := Y_k$ , and  $Z = V \setminus (\{x_1, \dots, x_r\} \cup Y_r)$  then proceed to Step 2.

**Step 2:** Note that for all  $v \in Z$ ,  $|\bigcup_{j=1}^k N_j(v) \cap Y_k| \geq k \log n > k \log |Z|$ . Thus we may apply Lemma 4.1 to get a partition  $\{Y_1, \dots, Y_k\}$  of  $Y$  such that for all  $v \in Z$ , there exists  $i \in [k]$  such that  $N_i(v) \cap Y_i \neq \emptyset$ . Let each  $v \in Z$  choose an arbitrary such  $i$  and an arbitrary neighbor in  $N_i(v) \cap Y_i$ . Then  $x_i$  along with  $Y_i$  and all the  $v \in Z$  which chose neighbors in  $Y_i$  form a tree of color  $i$  and radius at most 2. Thus we have a partition into  $k \leq r$  monochromatic trees.  $\square$

A key element in the preceding proof was to first build a monochromatic tree cover in which the common intersection of all of the trees was a large enough set of leaves. We now explicitly define this structure.

**Definition 4.2.** A  **$(k, l, n)$ -absorbing tree partition** is a collection of trees  $T_1, \dots, T_k$  together with a **common leaf set**  $L$  of size  $l$  such that

- (i)  $|\bigcup_{i \in [k]} V(T_i)| = n$ ,
- (ii) the edges of  $T_i$  have color  $i$  for all  $i \in [k]$ ,
- (iii) every vertex in  $L$  is a leaf of  $T_i$  for all  $i \in [k]$ , and
- (iv) for all  $i \neq j$ ,  $V(T_i) \cap V(T_j) = L$ .

Note that if every  $r$ -coloring of a graph  $G$  on  $n$  vertices contains a  $(k, l, n)$ -absorbing tree partition for some  $1 \leq k \leq r$  and  $l \geq 0$ , then by arbitrarily assigning the leaves to the trees we have  $\text{tp}_r(G) \leq r$  (with trees of distinct colors).

We will consider absorbing tree partitions in two different settings: first, in Theorem 4.3 we wish to optimize the bound on the minimum degree so that  $\text{tp}_r(G) \leq r$ , and second, we will apply Theorem 4.3 in a setting where the graph is nearly complete, in which case we do not need so much control over the minimum degree as we need control over the size of the common leaf set. So for the purposes of streamlining, we combine everything we want into the following statement, which has a parameter  $\epsilon$  related to the minimum degree and a parameter  $\alpha$  which is related to the size of the leaf set and the lower bound on  $n$ . The method of proof will be similar to that of Theorem 1.5; however, the calculations are different as here we are attempting to optimize the minimum degree instead of the lower bound on  $n$ .

**Theorem 4.3.** Let  $r \geq 2$ ,  $0 < \epsilon < \frac{1}{er!}$ ,  $\alpha = \frac{1}{er!} - \epsilon$ , and  $n_0 = \max\{\frac{12}{\alpha^2} \log(\frac{6}{\alpha^2}), \frac{4r}{\alpha} \log(\frac{2r}{\alpha})\}$ . If  $G$  is a graph on  $n \geq n_0$  vertices with  $\delta(G) \geq (1 - \epsilon)n$ , then in every  $r$ -coloring of  $G$  there either exists a  $(1, l, n)$ -absorbing tree partition with  $l \geq n - \frac{2}{\alpha} \log n$  (i.e. a monochromatic spanning tree with at least  $n - \frac{2}{\alpha} \log n$  leaves) or a  $(k, l, n)$ -absorbing tree partition with  $2 \leq k \leq r$  and  $l \geq \alpha n/2$ .

We will need the following two statements in the proof of Theorem 4.3.

**Observation 4.4.** Let  $x \in \mathbb{R}$  with  $x \geq 2$ . If  $n \geq 2x \log x$ , then  $\frac{\log n}{n} < \frac{1}{x}$ .

*Proof.* We first note that  $\frac{\log n}{n}$  is strictly decreasing since  $n \geq 2x \log x > e$ . Now since  $2x \log x < x^2$ , we have  $\frac{\log(2x \log x)}{2x \log x} = \frac{\log(2x \log x)}{x \log x^2} < \frac{1}{x}$ .  $\square$

**Lemma 4.5.** Let  $0 < \alpha \leq 1$  and  $n_0 = \frac{12}{\alpha^2} \log(\frac{6}{\alpha^2})$  and let  $G$  be a graph on  $n \geq n_0$  vertices. If there exists  $x \in V(G)$  such that for all  $v \in V(G)$ ,  $\deg(v, N(x)) \geq \alpha n$ , then  $G$  has a spanning tree with at least  $n - \frac{2}{\alpha} \log n$  leaves.

*Proof.* We will show that  $x$  along with at most  $\frac{2}{\alpha} \log n$  of its neighbors form a dominating set. Set  $Y_1 := N(x)$  and  $Z_1 = V \setminus (\{x\} \cup Y_1)$ . For  $1 \leq i \leq \lceil \frac{3}{2\alpha} \log n \rceil - 1$ , do the following: If  $Z_i \neq \emptyset$ , let  $y_i$  be the vertex in  $Y_i$  with the largest degree to  $Z_i$ . Set  $Y_{i+1} = Y_i \setminus \{y_i\}$  and  $Z_{i+1} = Z_i \setminus N(y_i)$ . Since  $\deg(y_i, Z_i) \geq \frac{\alpha n - i}{|Y_1| - i} |Z_i| > \frac{2\alpha}{3} |Z_i|$  (where the second inequality holds by Observation 4.4 and the bound on  $n$ ) and thus

$$|Z_{i+1}| < (1 - 2\alpha/3) |Z_i| \leq (1 - 2\alpha/3)^i |Z_1| \leq (1 - 2\alpha/3)^{i+1} n \leq 1$$

when  $i + 1 \geq \frac{3}{2\alpha} \log n$ . Thus when the process stops, we have a spanning tree with at most  $\lceil \frac{3}{2\alpha} \log n \rceil \leq \frac{2}{\alpha} \log n$  non-leaves.  $\square$

*Proof of Theorem 4.3. Step 1:* Let  $x_1 \in V(G)$  and let  $Y_1$  be the largest monochromatic neighborhood of  $x_1$ , say the color is 1. Note that  $|Y_1| \geq \frac{1}{r}(1 - \epsilon)n$ . If for all  $v \in V \setminus Y_1$ ,  $\deg_1(v, Y_1) \geq \alpha n$ , then since  $n \geq \frac{12}{\alpha^2} \log(\frac{6}{\alpha^2})$ , we may apply Lemma 4.5 to get a monochromatic spanning tree (in color 1) with at least  $n - \frac{2}{\alpha} \log n$  leaves and we are done. Otherwise some vertex  $x_2$  has at least  $\frac{1}{r-1}(|Y_1| - \epsilon n - \alpha n)$  neighbors of say color 2 in  $Y_1$ . Set  $Y_2 := Y_1 \cap N_2(x_2)$ .

For  $i \geq 2$ , do the following: if for all  $v \in V \setminus Y_i$ ,  $|(\bigcup_{j=1}^i N_j(v)) \cap Y_i| \geq \alpha n$ , then set  $k := i$ ,  $Y := Y_k$ , and  $Z = V \setminus (\{x_1, \dots, x_k\} \cup Y_k)$  then proceed to Step 2. Otherwise some vertex  $x_{i+1} \in V \setminus Y_i$  has at least  $\frac{1}{r-i}(|Y_i| - \epsilon n - \alpha n)$  neighbors of color say  $i + 1$ , in  $Y_i$ . Set  $Y_{i+1} := Y_i \cap N_{i+1}(x_{i+1})$ . Continue in this manner, until we go to Step 2 or until  $i = r$ . If  $i = r$ , then

$$|Y_r| \geq \left( \frac{1}{r!} - \epsilon \sum_{j=1}^r \frac{1}{j!} - \alpha \sum_{j=1}^{r-1} \frac{1}{j!} \right) n \geq \left( \frac{1}{r!} - \epsilon(e-1) - \alpha(e-1) \right) n$$

and thus every vertex in  $V \setminus Y_r$  has at least

$$|Y_r| - \epsilon n \geq \left( \frac{1}{r!} - \epsilon(e-1) - \alpha(e-1) \right) n - \epsilon n = \alpha n$$

neighbors in  $Y_r$ . Now set  $k := r$ ,  $Y := Y_k$ , and  $Z = V \setminus (\{x_1, \dots, x_k\} \cup Y_k)$  then proceed to Step 2.

**Step 2:** First set aside  $\alpha n/2$  vertices from  $Y$  to be the common leaf set of the absorbing tree partition and let  $Y'$  be the remaining vertices in  $Y$ . Every vertex in  $Z$  still has at least  $\alpha n/2$  neighbors in  $Y'$ . Since  $n \geq \frac{4r}{\alpha} \log(\frac{2r}{\alpha})$ , Observation 4.4 implies that  $\alpha n/2 > r \log n$  and thus we can apply Lemma 4.1 to get a partition  $\{Y'_1, \dots, Y'_k\}$  of  $Y'$  such that for all  $v \in Z$ , there exists  $i \in [k]$  such that  $N_i(v) \cap Y'_i \neq \emptyset$ . By arbitrarily choosing such a  $Y'_i$  for each  $v \in Z$ , we have a  $(k, l, n)$ -absorbing tree partition with  $2 \leq k \leq r$  and  $l \geq \alpha n/2$ .  $\square$

While we are not able to prove that the bound on the minimum degree in Theorem 4.3 is optimal, Observation 3.1 shows that there are graphs with  $\delta(G) \geq n - r$  for which  $\text{tp}_r(G) \geq \text{tc}_r(G) \geq r$ , and thus the goal in the minimum degree version of the problem (optimizing  $\delta(G)$  while maintaining  $\text{tp}_r(G) \leq r$ ) is different from the goal in the case of complete graphs (proving  $\text{tp}_r(K_n) \leq r - 1$ ).

In Theorem 4.3 we actually prove that the trees have distinct colors, so it is natural to ask the question of how the minimum degree threshold for partitioning (covering) into  $r$  trees compares to the minimum degree threshold for partitioning (covering) into  $r$  trees of distinct colors (see Section 3.2).

## 4.2 Covers

For the cover version of the  $r = 2$  case, we can actually prove a tight bound on the minimum degree (see Example 3.3).

**Theorem 4.6.** *Let  $n \geq 1$ . For all graphs  $G$  on  $n$  vertices, if  $\delta(G) \geq \frac{2n-5}{3}$ , then  $\text{tc}_2(G) \leq 2$ .*

*Proof.* Suppose that  $n = 3m + q$  where  $q \in \{0, 1, 2\}$ . Then  $\delta(G) \geq \frac{2n-5}{3}$  translates to  $\delta(G) \geq 2m - 1 + \lfloor \frac{q}{2} \rfloor$ .

Suppose  $G$  is 2-colored and let  $\mathcal{T} = \{R_1, \dots, R_k, B_1, \dots, B_l\}$  be a monochromatic component cover of  $G$  with the fewest number of components, where each component is maximal; and with respect to this, choose  $\mathcal{T}$  so that as many different colors are represented as possible. Without loss of generality suppose  $k \geq l$ . We are done unless  $|\mathcal{T}| \geq 3$ . It is clear from minimality of the number of components, that for each component  $T \in \mathcal{T}$ , there is a non-empty subset of vertices  $\phi(T)$  such that every vertex in  $\phi(T)$  is not contained in any other component  $S \in \mathcal{T} \setminus \{T\}$ .

**Case 1** (There is at least one component of each color).

Since  $k \geq l$  and  $k + l \geq 3$ , we have  $k \geq 2$ . Suppose first that there exist vertices  $u_i \in \phi(R_i)$ ,  $u_j \in \phi(R_j)$ , and  $v_h \in \phi(B_h)$  such that  $u_i u_j \notin E(G)$ . By the maximality of the components,  $u_i v_h, u_j v_h \notin E(G)$ . So

$$|N(u_i) \cap N(u_j) \cap N(v_h)| \geq n - 3 - 3(n - 3 - \delta(G)) = 3\delta(G) - 2n + 6 \geq 1$$

So let  $w \in N(u_i) \cap N(u_j) \cap N(v_h)$ . If  $w$  is in a blue component  $B' \in \mathcal{T}$ , then  $w$  cannot be adjacent to  $u_i$  or  $u_j$  via a blue edge (as this would imply that  $u_i$  or  $u_j$  is contained in  $B'$ ). So  $w$  is adjacent to  $u_i$  and  $u_j$  via red edges, but this contradicts the fact that  $u_i$

and  $u_j$  are in different red components. So suppose that for all blue components  $B' \in \mathcal{T}$ ,  $w \notin V(B')$ . This implies that  $wv_h$  must be a red edge and that  $w$  is in a red component of  $\mathcal{T}$ , but again this contradicts the fact that  $v_h$  is not contained in any red component of  $\mathcal{T}$ . In either case, we get a contradiction.

So we may assume that the vertices of  $\phi(R_1), \dots, \phi(R_k)$  induce a complete  $k$ -partite blue graph  $B'$ . However, this implies that  $\{B_1, \dots, B_l, B'\}$  is a cover with fewer components.

**Case 2** (All of the components have the same color).

Since  $k \geq l$ ,  $\mathcal{T} = \{R_1, \dots, R_k\}$ . Without loss of generality suppose  $|R_1| \leq |R_2| \leq \dots \leq |R_k|$ . Note that since  $k \geq 3$  and all components are red, we have  $2 \leq |R_1| \leq m$ . Since  $R_1$  is maximal, every edge leaving  $R_1$  is blue and thus for all  $v \in R_1$  we have

$$|N_B(v) \cap (V(G) \setminus R_1)| \geq \delta(G) - (|R_1| - 1) \geq 2m + \left\lfloor \frac{q}{2} \right\rfloor - |R_1| \geq m + \left\lfloor \frac{q}{2} \right\rfloor. \quad (1)$$

From (1), and the fact that  $|V(G) \setminus R_1| \leq 3m + q - 2$ , we see that for any set of three vertices  $\{x, y, z\}$  in  $R_1$ , some pair of  $\{x, y, z\}$  must have a common blue neighbor in  $V(G) \setminus R_1$ . This implies that there are either one or two blue components which cover the vertices of  $R_1$ . If there were only one, we would be in Case 1. So assume that there are two blue components  $B_1$  and  $B_2$  which cover every vertex in  $R_1$ . Now using (1), we get that

$$|B_1| + |B_2| \geq |R_1| + 2(2m + \left\lfloor \frac{q}{2} \right\rfloor - |R_1|) = 4m + 2 \left\lfloor \frac{q}{2} \right\rfloor - |R_1| \geq 3m + 2 \left\lfloor \frac{q}{2} \right\rfloor \geq 3m + q - 1.$$

So either  $B_1$  and  $B_2$  form a cover with two components, or  $B_1$ ,  $B_2$ , and the red component containing the leftover vertex form a cover with two blue components and a red component and thus we are in Case 1.  $\square$

### 4.3 Large monochromatic components

We use the following lemma of Liu, Morris, and Prince [27] (an essentially equivalent version of this lemma was independently proved by Mubayi [30]). A **double-star** is a tree having at most two vertices which are not leaves.

**Lemma 4.7** (Lemma 9 in [27]). *Let  $c \geq 0$  and let  $G$  be an  $X, Y$ -bipartite graph on  $n$  vertices. If  $e(G) \geq c|X||Y|$ , then  $G$  has a double-star of order at least  $cn$ .*

**Theorem 4.8.** *Let  $r \geq 2$ , and let  $0 < \epsilon \leq \frac{1}{2}$ . If  $G$  is a graph on  $n$  vertices with  $\delta(G) \geq (1 - \epsilon)n$ , then every  $r$ -coloring of  $G$  contains either a monochromatic tree of order at least  $(1 - \epsilon)n$  or a monochromatic double-star of order at least  $(1 - 2\epsilon)\frac{n}{r-1}$ .*

*Proof.* Let  $H$  be the largest monochromatic tree in  $G$ , say of color  $r$  and suppose that  $|V(H)| < (1 - \epsilon)n$ . Set  $X = V(H)$  and  $Y = V(G) \setminus X$  and let  $B = G[X, Y]$  be the bipartite graph induced by the bipartition  $\{X, Y\}$ . Without loss of generality suppose  $|X| \geq |Y|$  and note that  $e(B) \geq |Y|(|X| - \epsilon n) \geq (1 - 2\epsilon)|X||Y|$ . Note that by the maximality of  $H$ , there are no edges of color  $r$  in  $B$ , so at least  $(1 - 2\epsilon)\frac{1}{r-1}|X||Y|$  of the edges of  $B$  are say color 1. So by Lemma 4.7, we have a monochromatic double-star with at least  $(1 - 2\epsilon)\frac{n}{r-1}$  vertices.  $\square$

## 5 Sparse Regularity and Basic Applications

We will use of a variant of Szemerédi's regularity lemma [38] for sparse graphs, which was proved independently by Kohayakawa [23] and Rödl (see [7]) and later generalized by Scott [36].

Say that a  $U, V$ -bipartite graph is **weakly- $(\epsilon, q)$ -regular** if for all  $U' \subseteq U$  and  $V' \subseteq V$  with  $|U'| \geq \epsilon|U|$  and  $|V'| \geq \epsilon|V|$ ,  $e(U', V') \geq q$ . Given disjoint sets  $X, Y$  and  $0 < p \leq 1$ , define the  **$p$ -density** of  $(X, Y)$ , denoted  $d_p(X, Y)$ , by

$$d_p(X, Y) = \frac{e(X, Y)}{p|X||Y|}.$$

We say the bipartite graph  $G[U, V]$  induced by disjoint sets  $U$  and  $V$  is  **$(\epsilon, p)$ -regular** or that  $(U, V)$  is an  **$(\epsilon, p)$ -regular pair** if for all subsets  $U' \subseteq U$ ,  $V' \subseteq V$  with  $|U'| \geq \epsilon|U|$ ,  $|V'| \geq \epsilon|V|$  we have  $|d_p(U', V') - d_p(U, V)| \leq \epsilon$ . Given a graph  $G = (V, E)$ , a partition  $\{V_1, \dots, V_k\}$  of  $V$  is said to be an  **$(\epsilon, p)$ -regular partition** if it is an equitable partition (i.e.  $||V_i| - |V_j|| \leq 1$  for all  $i, j \in [k]$ ) and all but at most  $\epsilon \binom{k}{2}$  pairs  $(V_i, V_j)$  induce  $(\epsilon, p)$ -regular pairs.

We will use the following  $r$ -colored version of the sparse regularity lemma due to Scott [36, Theorem 4.1].

**Lemma 5.1.** *For every  $\epsilon > 0$  and  $m, r \geq 1$ , there exists  $M$  such that if  $G_1, \dots, G_r$  are edge-disjoint graphs on vertex set  $V$  with  $|V| \geq m$  and where  $p_i = \frac{e(G_i)}{\binom{|V|}{2}} > 0$  is the density of  $G_i$  for all  $i \in [r]$ , there exists a partition  $\{V_1, \dots, V_k\}$  of  $V$  with  $m \leq k \leq M$  such that for all  $i \in [r]$ ,  $\{V_1, \dots, V_k\}$  is an  $(\epsilon, p_i)$ -regular partition of  $G_i$ .*

Let  $G$  be an  $r$ -colored graph on  $n$  vertices with  $p = \frac{e(G)}{\binom{n}{2}}$  and  $r$ -coloring  $\{G_1, \dots, G_r\}$ , where for all  $i \in [r]$ ,  $p_i = \frac{e(G_i)}{\binom{n}{2}}$ ; note that  $p = \sum_{i=1}^r p_i$ . We define the  **$(\epsilon, p, \delta)$ -reduced graph**  $\Gamma$  as follows: Let  $\{V_1, \dots, V_k\}$  be the partition of  $G$  obtained from an application of Lemma 5.1. Let  $V(\Gamma) = \{V_1, \dots, V_k\}$  and say that  $\{V_i, V_j\}$  is a  $c$ -colored edge of  $\Gamma$  if the  $p_c$ -density of  $(V_i, V_j)$  in  $G_c$  is at least  $\delta$ . Note that  $\Gamma$  is a (possibly) multicolored graph.

The following simple lemma is typically applied to the reduced graph obtained after an application of Lemma 5.1.

**Lemma 5.2.** *Let  $\alpha > 0$ . If  $H$  is a graph on  $k \geq \frac{2}{\sqrt{\alpha}}$  vertices with  $e(H) \geq (1 - \alpha) \binom{k}{2}$ , then there exists a subgraph  $H' \subseteq H$  such that  $|V(H')| \geq (1 - \sqrt{\alpha})k$  and  $\delta(H') \geq (1 - 2\sqrt{\alpha})k$ .*

*Proof.* As there are at most  $\alpha \binom{k}{2} \leq \alpha \frac{k^2}{2}$  non-edges, there are at most  $\sqrt{\alpha}k$  vertices  $V'$  which have at least  $\frac{\sqrt{\alpha}}{2}k$  non-neighbors. Let  $H' = H[V']$ . We have  $|V(H')| \geq k - \sqrt{\alpha}k = (1 - \sqrt{\alpha})k$  and  $\delta(H') \geq k - 1 - \frac{\sqrt{\alpha}}{2}k - \sqrt{\alpha}k \geq (1 - 2\sqrt{\alpha})k$ .  $\square$

We will apply Lemma 5.1 to an  $r$ -colored graph  $G \sim G(n, p)$  with  $p = \frac{\omega(1)}{n}$  to get a partition  $\{V_1, \dots, V_k\}$ . We will only require the very weak condition that each pair which

is  $(\epsilon, p_i)$ -regular for all  $i \in [r]$  is weakly- $(\epsilon, 1)$ -regular in some color  $c \in [r]$ . This will follow as a simple consequence of the union bound and the Chernoff bound, which we state first for convenience (see e.g. Corollary 21.7 in [13]).

**Theorem 5.3** (Chernoff). *Let  $X$  be a binomially distributed random variable. Then for  $\alpha > 0$ ,*

$$\mathbb{P}[X \leq (1 - \alpha)\mathbb{E}[X]] \leq \exp(-\alpha^2\mathbb{E}[X]/2).$$

**Lemma 5.4.** *For all  $\eta > 0$ , there exists  $C$  such that if  $p \geq \frac{C}{n}$ , then a.a.s.  $G \sim G(n, p)$  has the property that for all  $X, Y \subseteq V(G)$  with  $X \cap Y = \emptyset$  and  $|X|, |Y| \geq \eta n$ ,*

$$3p|X||Y|/2 \geq e(X, Y) \geq p|X||Y|/2.$$

**Lemma 5.5.** *Suppose we have applied Lemma 5.1 with  $0 < \epsilon < \frac{1}{2r}$  to an  $r$ -colored graph  $G \sim G(n, p)$  with  $p = \frac{\omega(1)}{n}$  to get a partition  $\{V_1, \dots, V_k\}$ . Then every pair  $(V_i, V_j)$  which is  $(\epsilon, p_l)$ -regular for all  $l \in [r]$  is weakly- $(\epsilon, 1)$ -regular in some color  $c \in [r]$ .*

*Proof.* Suppose the pair  $(V_i, V_j)$  is  $(\epsilon, p_l)$ -regular for all  $l \in [r]$ . Lemma 5.4 implies that  $e(V_i, V_j) \geq p|V_i||V_j|/2$  and thus for some color  $c \in [r]$ , we have  $e_c(V_i, V_j) \geq p|V_i||V_j|/2r$ . So if  $V'_i \subseteq V_i$  and  $V'_j \subseteq V_j$  with  $|V'_i|, |V'_j| \geq \epsilon \frac{n}{k}$ , then by  $(\epsilon, p_c)$ -regularity, we have

$$\frac{e_c(V'_i, V'_j)}{p_c|V'_i||V'_j|} = d_{p_c}(V'_i, V'_j) \geq d_{p_c}(V_i, V_j) - \epsilon = \frac{e_c(V_i, V_j)}{p_c|V_i||V_j|} - \epsilon \geq \frac{p}{2rp_c} - \epsilon$$

and so

$$e_c(V'_i, V'_j) \geq \left(\frac{p}{2r} - \epsilon p_c\right)|V'_i||V'_j| > 0,$$

and thus  $(V_i, V_j)$  is weakly- $(\epsilon, 1)$ -regular in color  $c$ . □

## 5.1 Nearly spanning tree partitions

**Lemma 5.6.** *Let  $0 < \epsilon < 1/3$  and let  $G$  be a  $U, V$ -bipartite graph. If  $G$  is weakly  $(\epsilon, 1)$ -regular, then  $G$  contains a tree  $T$  with leaf set  $L$  such that*

- (i)  $|V(T) \cap U| \geq (1 - \epsilon)|U|$ ,  $|V(T) \cap V| \geq (1 - \epsilon)|V|$ , and
- (ii)  $|L \cap U| \geq (1 - 4\epsilon)|U|$ ,  $|L \cap V| \geq (1 - 4\epsilon)|V|$ .

*Proof.* Let  $U'$  be a subset of  $U$  of size exactly  $\lfloor 3\epsilon|U| \rfloor$  and let  $V'$  be a subset of  $V$  of size exactly  $\lfloor 3\epsilon|V| \rfloor$ . Let  $G'$  be the graph induced by  $U', V'$  and suppose first that no component of  $G'$  intersects  $U'$  in at least  $\epsilon|U|$  vertices or intersects  $V'$  in at least  $\epsilon|V|$  vertices. However, now we may partition  $U'$  into sets  $U_1$  and  $U_2$  and  $V'$  into sets  $V_1$  and  $V_2$  such that  $\epsilon|U| \leq |U_1| \leq 2\epsilon|U|$ ,  $\epsilon|V| \leq |V_1| \leq 2\epsilon|V|$ , and there are no edges from  $U_1 \cup V_1$  to  $U_2 \cup V_2$ ; however, this contradicts the fact that  $G$  is weakly  $(\epsilon, 1)$ -regular. So suppose that some component  $H'$  of  $G'$  intersects say  $U'$  in at least  $\epsilon|U|$  vertices. Then since  $G$  is weakly- $(\epsilon, 1)$ -regular, we see that all but at most  $\epsilon|V|$  vertices of  $V'$  are also in  $H'$  and thus there exists a connected subgraph  $H'$  of  $G$  which intersects each of  $U$  and  $V$  in at most  $3\epsilon|U|$  and  $3\epsilon|V|$  vertices respectively.

Finally, using the fact that  $G$  is weakly  $(\epsilon, 1)$ -regular, we see that all but at most  $\epsilon|U|$  vertices of  $U$  and  $\epsilon|V|$  vertices of  $V$  have a neighbor in  $H'$ , and thus  $G$  contains a tree in which at least  $(1 - 4\epsilon)|U|$  vertices of  $U$  and at least  $(1 - 4\epsilon)|V|$  vertices of  $V$  are leaves.  $\square$

Say that a graph  $G$  on  $n$  vertices has property  $T(r, l, \lambda)$  if every  $r$ -coloring of  $G$  either has a monochromatic tree with at least  $(1 - \lambda)n$  leaves or  $G$  has a  $(s, l, n')$ -absorbing tree partition for some  $2 \leq s \leq r$  and  $n' \geq (1 - \lambda)n$ .

**Lemma 5.7.** *Let  $r \geq 1$  and  $0 < \epsilon < \frac{1}{10r}$ . If  $p = \frac{\omega(1)}{n}$ , then a.a.s.  $G(n, p)$  has property  $T(r, \frac{n}{8er!}, 6\epsilon)$ .*

*Proof.* Let  $G \sim G(n, p)$  be  $r$ -colored. Let  $m$  be large enough so that  $\frac{\log m}{m} < \frac{\epsilon}{4r}(\frac{1}{er!} - \epsilon)$  and let  $\epsilon' := \frac{\epsilon^2}{4r}$ . Apply Lemma 5.1 to  $G$  with  $\epsilon'$ ,  $m$ , and  $r$ . Let  $\Gamma$  be the  $(\epsilon', p, 1/2r)$ -reduced graph obtained.

By Lemma 5.2, we can pass to a subgraph  $\Gamma' \subseteq \Gamma$  with  $k' := |\Gamma'| \geq (1 - \sqrt{r\epsilon'})k = (1 - \epsilon/2)k$  and  $\delta(\Gamma') \geq (1 - 2\sqrt{r\epsilon'})k \geq (1 - \epsilon)k'$ . We color the edges of  $\Gamma'$  by an arbitrary color  $c$  guaranteed by Lemma 5.5 and recall that this says the  $c$ -colored edges of  $\Gamma'$  represent  $c$ -colored weakly- $(\epsilon, 1)$ -regular pairs in  $G$ .

Now apply Theorem 4.3 with  $\epsilon$  to  $\Gamma'$  to get an  $(s, l, k')$ -absorbing tree partition  $\mathcal{T}$  with  $l \geq k' - \frac{2}{\alpha} \log k' \geq (1 - \epsilon)k'$  (where the second inequality holds by the choice of  $m$  and since  $k \geq m$ ) if  $s = 1$  and  $l \geq k'/(4er!)$  if  $2 \leq s \leq r$ . Since the edges of  $\Gamma'$  represent weakly- $(\epsilon, 1)$ -regular pairs, we can apply Lemma 5.6 to each edge of  $\mathcal{T}$  to see that  $G$  a.a.s. has property  $T(r, \frac{n}{8er!}, 6\epsilon)$ ; that is, to get the absorbing tree partition  $\mathcal{T}'$  of  $G$ . Note that if there are  $\tau k'$  leaves in  $\mathcal{T}$ , each of which is a leaf in  $s$  different trees, we will get (by Lemma 5.6(ii)) at least  $\tau k' \cdot (1 - 4s\epsilon) \frac{n}{k} \geq (1 - 5s\epsilon)\tau n$  leaves in the original graph which are leaves in all  $s$  trees. So the total number of leaves in  $\mathcal{T}'$  is at least

$$(1 - 5s\epsilon)\tau n \geq \begin{cases} (1 - 5s\epsilon)(1 - \epsilon)n \geq (1 - 6\epsilon)n & \text{if } s = 1 \\ (1 - 5s\epsilon)\frac{1}{4er!}n \geq \frac{n}{8er!} & \text{if } 2 \leq s \leq r \end{cases}$$

Also note that since  $\mathcal{T}$  covers the  $k'$  vertices of  $\Gamma'$ ,  $\mathcal{T}'$  will cover (by Lemma 5.6(i)) at least  $k' \cdot (1 - \epsilon) \frac{n}{k} \geq (1 - \epsilon/2)k \cdot (1 - \epsilon) \frac{n}{k} \geq (1 - 2\epsilon)n$  vertices of  $G$  as desired.  $\square$

## 6 Lemmas for Random Graphs

The following lemma follows from standard applications of the Chernoff and union bounds.

**Lemma 6.1.** *Let  $r \geq 1$ . If  $p \geq \left(\frac{9r \log n}{n}\right)^{1/r}$ , then a.a.s., every set  $R$  of  $r$  vertices in  $G(n, p)$  satisfies  $|N^\cap(R)| \geq np^r/2$ . Furthermore, if  $p = \omega\left(\left(\frac{\log n}{n}\right)^{1/r}\right)$ , then for any  $\epsilon > 0$ , a.a.s., every set  $R$  of  $r$  vertices satisfies*

$$(1 - \epsilon)np^r \leq |N^\cap(R)| \leq (1 + \epsilon)np^r.$$



Erdős, Palmer, and Robinson [12] determined the exact threshold for when the neighborhood of every vertex (of degree at least 2) in  $G(n, p)$  induces a connected subgraph. We need the following lemma which gives us a bound on the value of  $p$  for which the common neighborhood of every set of  $r$  vertices in  $G(n, p)$  is non-empty and induces a connected graph.

**Lemma 6.2.** *Let  $r \geq 1$ . If  $p \geq \left(\frac{C \log n}{n}\right)^{1/(r+1)}$  with  $C$  sufficiently large, then a.a.s. in  $G \sim G(n, p)$ ,  $G[N^\cap(R)]$  is connected for every set  $R$  of  $r$  vertices.*

*Proof.* Let  $0 < \epsilon \leq 1/2$ , let  $3(r+1)^2 < C' < \frac{C^{1/(r+1)}}{2}$ , and suppose  $(1-\epsilon)np^r \leq m \leq (1+\epsilon)np^r$ . Then

$$C' \frac{\log m}{m} \leq C' \frac{\log n}{(1-\epsilon)np^r} \leq \frac{C'}{(1-\epsilon)} \left(\frac{\log n}{n}\right)^{1/(r+1)} < \left(\frac{C \log n}{n}\right)^{1/(r+1)} \leq p. \quad (2)$$

Using (2), the probability that the subgraph of  $G$  induced by a set of size  $m$  is disconnected is at most

$$\begin{aligned} \sum_{k=1}^{m/2} \binom{m}{k} (1-p)^{k(m-k)} &\leq \sum_{k=1}^{m/2} \exp \left( k \log \left( \frac{me}{k} \right) - k(m-k) \frac{C' \log m}{m} \right) \\ &\leq \sum_{k=1}^{m/2} \exp \left( k \left( \log m - \frac{C'}{2} \log m \right) \right) = o \left( \frac{1}{n^{r+1}} \right). \end{aligned}$$

There are  $O(n^r)$   $r$ -sets  $R$  for which we must consider  $N^\cap(R)$  (which by Lemma 6.1 satisfy  $(1-\epsilon)np^r \leq |N^\cap(R)| \leq (1+\epsilon)np^r$ ), and thus the expected number of  $N(R)$  which induce a disconnected subgraph tends to 0.  $\square$

**Lemma 6.3.** *Consider  $G \sim G(n, p)$  with vertex set  $V$ . Then a.a.s. for any set  $L \subset V$  with  $\frac{80 \log n}{p} \leq |L| \leq n$ , all but at most  $\frac{9 \log n}{p}$  of  $v \in V \setminus L$  satisfy  $|N(v, L)| \geq |L|p/2$ .*

*Proof.* For a fixed set  $L$ , we have that  $|N(v, L)| \sim \text{Bin}(|L|, p)$ , so the Chernoff Bound implies that

$$\mathbb{P}[|N(v, L)| < |L|p/2] \leq e^{-\frac{|L|p}{8}}.$$

Call  $v$  *bad* for  $L$  if  $|N(v, L)| < |L|p/2$ . Since  $N(v, L)$  and  $N(u, L)$  are independent for different vertices  $v$  and  $u$ , we have that the probability that there exists an  $L$  with at least  $9 \log n/p$  many bad vertices is at most

$$\begin{aligned} \sum_{\ell=80 \log n/p}^n \binom{n}{\ell} \binom{n}{9 \log n/p} \cdot \exp \left( -\frac{\ell p 9 \log n}{8 p} \right) &\leq \sum_{\ell=80 \log n/p}^n \exp \left( -\frac{\ell \log n}{8} + \frac{9(\log n)^2}{p} \right) \\ &\leq \exp \left( \log n - \frac{(\log n)^2}{p} \right) = o(1) \quad \square \end{aligned}$$

The following Lemma will be used to prove Theorem 1.7. We prove it here as it may be of independent interest.

**Lemma 6.4.** *Let  $r \geq 1$ .*

- (i) *If  $p = \left(\frac{r \log n - \omega(1)}{n}\right)^{1/r}$  and  $G \sim G(n, p)$ , then a.a.s. there exists a set  $S \subseteq V(G)$  such that  $|S| = r$ ,  $S$  is independent,  $S$  is not a dominating set, and for all  $v \in V(G)$ ,  $\deg(v, S) \leq r - 1$ .*
- (ii) *For all  $s > r$  with  $s \log s = o(\log n)$  and  $0 < c < \frac{1}{2\binom{s}{r}}$ , if  $p \leq \left(\frac{c \log n}{n}\right)^{1/r}$  and  $G \sim G(n, p)$ , then a.a.s. there exists a set  $S \subseteq V(G)$  such that  $|S| = s$ ,  $S$  is independent,  $S$  is not a dominating set, and for all  $v \in V(G)$ ,  $\deg(v, S) \leq r - 1$ .*

*Proof.* We begin with the proof of (ii). Choose  $s$ ,  $c$ , and  $p$  as in the statement. We first show that we can find an independent set  $S$  of  $s$  vertices such that no  $r$  vertices in  $S$  have a common neighbor (or equivalently, no  $v \in V \setminus S$  has more than  $r - 1$  neighbors in  $S$ ). For  $S \in \binom{[n]}{s}$  let  $X_S$  be the indicator random variable for the event

$$A_S := \{S \text{ is independent and no } r \text{ vertices in } S \text{ have a common neighbor}\}$$

and let

$$X = \sum_{S \in \binom{[n]}{s}} X_S.$$

For a fixed set  $S$ , let  $q$  represent the probability that a vertex  $v \notin S$  has at least  $r$  neighbors in  $S$ . Then

$$q := \sum_{k=r}^s \binom{s}{k} p^k (1-p)^{s-k} = \binom{s}{r} p^r \left( (1-p)^{s-r} + O\left(\sum_{t=1}^{s-r} (sp)^t\right) \right) = \binom{s}{r} p^r (1 + O(sp)).$$

To see the last two equalities, note that the conditions on  $s$  imply that  $sp = o(1)$  and for all  $t > 0$ ,

$$\frac{\binom{s}{r+t} p^{r+t}}{\binom{s}{r} p^r} \leq (sp)^t.$$

Now we have that

$$\mathbb{P}[X_S = 1] = (1-q)^{n-s} (1-p)^{\binom{s}{2}}$$

and

$$\mathbb{E}[X] = \binom{n}{s} (1-q)^{n-s} (1-p)^{\binom{s}{2}} = \exp\left(s \log\left(\frac{ne}{s}\right) - \binom{s}{r} c \log n (1 + o(1)) + O(1)\right) \rightarrow \infty$$

since  $\binom{s}{r} c < 1 < s$  and  $s \log s = o(\log n)$ .

An application of the second moment method will show that such a set exists a.a.s. It suffices to show that  $\mathbb{E}[X^2] / \mathbb{E}[X]^2 \leq 1 + o(1)$  (see e.g. Corollary 4.3.2 in [3]).

$$\begin{aligned}
\mathbb{E}[X^2] &= \mathbb{E} \left[ \left( \sum_{S \in \binom{[n]}{s}} X_S \right)^2 \right] \\
&= \mathbb{E}[X] + \sum_{k=0}^{s-1} \sum_{|S_1 \cap S_2|=k} \mathbb{P}[A_{S_1} \wedge A_{S_2}] \\
&\leq \mathbb{E}[X] + \binom{n}{s}^2 (1-q)^{2n-4s} (1-p)^{2\binom{s}{2}} + \sum_{k=1}^{s-1} O(n^{2s-k}).
\end{aligned} \tag{3}$$

The second sum in (3) is over ordered pairs of sets. Since  $\mathbb{E}[X]^2 = \binom{n}{s}^2 (1-q)^{2n-2s} (1-p)^{2\binom{s}{2}} = \Omega\left(\left(\frac{n}{s}\right)^{2s} e^{-2\binom{s}{r} c \log n}\right)$ , we have

$$\mathbb{E}[X^2] / \mathbb{E}[X]^2 \leq \frac{1}{\mathbb{E}[X]} + (1-q)^{-2s} + O\left(\frac{s^{2s+1} e^{2\binom{s}{r} c \log n}}{n}\right) \leq 1 + o(1)$$

since  $s \log s = o(\log n)$  and  $2\binom{s}{r} c < 1$ .

Now note that a.a.s.,  $|N^\cup(S)| = O(snp) = o(n)$  and so  $S$  is not a dominating set and (ii) is proved.

The proof of (i) is similar, but we are more careful in the calculation of  $\mathbb{E}[X^2]$ . Set  $\omega := \omega(1)$  and suppose  $p = \left(\frac{r \log n - \omega}{n}\right)^{1/r}$ . We want to prove the existence of a set  $S$  of size  $r$ . In this case we simply have  $q = p^r$ , and thus

$$\mathbb{P}[X_S = 1] = (1 - p^r)^{n-r} (1 - p)^{\binom{r}{2}}$$

and

$$\begin{aligned}
\mathbb{E}[X] &= \binom{n}{r} (1 - p^r)^{n-r} (1 - p)^{\binom{r}{2}} \\
&= \exp \left( r \log n - n \cdot \frac{r \log n - \omega}{n} + O(1) \right) = \exp(\omega(1 + o(1))) \rightarrow \infty.
\end{aligned}$$

For  $r$ -sets  $R_1, R_2$  with  $|R_1 \cap R_2| = k$ , we have

$$\begin{aligned}
\mathbb{P}[A_{R_1} \wedge A_{R_2}] &\leq (p^k (1 - p^{r-k})^2 + (1 - p^k))^{n-2r} (1 - p)^{\binom{r}{2} + k(r-k) + \binom{r-k}{2}} \\
&\leq (p^k (1 - p^{r-k})^2 + (1 - p^k))^{n-2r}
\end{aligned}$$

Thus

$$\begin{aligned}
\mathbb{E}[X^2] &= \mathbb{E}[X] + \sum_{k=0}^{r-1} \sum_{|R_1 \cap R_2|=k} \mathbb{P}[A_{R_1} \wedge A_{R_2}] \\
&\leq \mathbb{E}[X] + \sum_{k=0}^{r-1} \binom{n}{r} \binom{r}{k} \binom{n-r}{r-k} (p^k(1-p^{r-k})^2 + (1-p^k))^{n-2r} \\
&\leq \mathbb{E}[X] + \binom{n}{r}^2 (1-p^r)^{2n-4r} + \sum_{k=1}^{r-1} \exp(2\omega - k \log n + O(1)).
\end{aligned}$$

Since  $\mathbb{E}[X]^2 = \binom{n}{r}^2 (1-p^r)^{2n-2r} (1-p)^{2\binom{r}{2}} = \exp(2\omega(1+o(1)))$ , we again have that

$$\mathbb{E}[X^2] / \mathbb{E}[X]^2 \leq 1 + o(1).$$

□

## 7 Monochromatic trees in random graphs

### 7.1 Upper bounds on the tree cover/partition number

**Theorem 7.1.** *For all  $r \geq 2$ , there exists  $C \geq r$  such that a.a.s.*

- (i) *if  $p \geq \left(\frac{27 \log n}{n}\right)^{1/3}$  then  $\text{tp}_2(G(n, p)) \leq 2$ , and*
- (ii) *if  $p \geq \left(\frac{C \log n}{n}\right)^{1/(r+1)}$ , then  $\text{tc}_r(G(n, p)) \leq r^2$ , and*
- (iii) *if  $p \geq \left(\frac{C \log n}{n}\right)^{1/r}$ , then there is a collection of  $r$  vertex disjoint monochromatic trees which cover all but at most  $9r \log n/p$  vertices.*

*Proof.* Part (ii) follows directly from Proposition 3.7 and Lemma 6.1.

For part (i), suppose the edges have been colored with colors 1 and 2 and consider two vertices  $u$  and  $v$  with no monochromatic path between them; if there were no such pair, we would have a spanning monochromatic component by the remark of Erdős and Rado. By Lemma 6.1,  $|N^\cap(\{u, v\})| \geq np^2/2$ . Let  $N_{x,c}$  with  $x \in \{u, v\}$  and  $c \in \{1, 2\}$  represent the color  $c$  neighbors of vertex  $x$  in  $N^\cap(\{u, v\})$ . Then we must have that  $N_{u,1} \cap N_{v,1} = N_{u,2} \cap N_{v,2} = \emptyset$ . Thus  $N_{u,1} = N_{v,2} =: A$  and  $N_{u,2} = N_{v,1} =: B$ . Now by Lemma 6.2,  $N^\cap(\{u, v\})$  induces a connected subgraph. If both  $A$  and  $B$  are non-empty then there is an edge between them, but this edge would give a monochromatic path between  $u$  and  $v$ . Thus wlog,  $N^\cap(\{u, v\}) = A$ . By Lemma 6.1, every vertex in  $Z := V - N^\cap(\{u, v\}) - \{u, v\}$  has at least  $np^3/2 \geq 27 \log n/2$  many neighbors in  $N^\cap(\{u, v\})$ . So applying Lemma 4.1, with  $k = 2$ ,  $Y = N^\cap(\{u, v\})$ , and  $Z$  as above, we have obtained the desired partition into two monochromatic trees.

In order to prove part (iii), we will use the method of multiple exposures. Let  $p \geq \left(\frac{C \log n}{n}\right)^{1/r}$  with  $C > 2000er!r(4r^2)^r$  and let  $\hat{p}$  be such that  $(1 - \hat{p})^{r+1} = (1 - p)$ . Note that in this case,  $p \geq \hat{p} \geq \frac{p}{r+1}$ . Then we may view  $G(n, p)$  as  $G_0 \cup \dots \cup G_r$  where each  $G_i \sim G(n, \hat{p})$  for  $0 \leq i \leq r$ . We will expose the  $G_i$  one at a time along with the colors

assigned to their edges. Note that if an edge belongs to more than one  $G_i$ , then when it is revealed a second time, we already know its color. This does not affect our argument.

Let  $\alpha = \frac{1}{8er!}$ . First we expose  $G_0$  and apply Lemma 5.7 with  $\epsilon$  as any small constant. This provides us with an  $(s, \alpha n, n')$ -absorbing tree partition  $\mathcal{T}$  and common leaf set  $L_0$ , such that  $1 \leq s \leq r$  and  $n' \geq (1 - 6\epsilon)n$ . If  $n' = n$ , then we are done. Our goal is now to “attach” as many of the vertices of  $V_0 := V(G) \setminus V(\mathcal{T})$  to  $L_0$  as possible. Expose  $G_1$  (and all the colors assigned to its edges). Apply Lemma 6.3 to  $G_1$  with  $L_0$  as  $L$ . This is possible since  $|L_0|\hat{p} \geq \alpha n \hat{p} \geq 80 \log n$ . Let  $V'_0 = \{v \in V_0 : |N(v, L_0)| \geq |L_0|\hat{p}/2\}$  and let  $V''_0 = V_0 \setminus V'_0$ . Then by the lemma,  $|V''_0| \leq 9 \log n/p$ . Now if every vertex in  $V'_0$  has at least  $r \log n$  neighbors in  $L_0$  with colors from  $[s]$  (note that if  $s = r$ , then this must be the case), then we may apply Lemma 4.1 with  $L_0$  as  $Y$  and  $V'_0$  as  $Z$  to get a partition  $\{Y_1, \dots, Y_s\}$  of  $L_0$  such that for all  $v \in V'_0$  there exists  $\ell \in [s]$  such that  $N_\ell(v) \cap Y_\ell \neq \emptyset$ . By arbitrarily choosing such a  $Y_\ell$  for each  $v \in V'_0$ , we have the desired tree partition of  $V \setminus V''_0$ . Otherwise there is a vertex  $x_1$  in  $V'_0$  satisfying

$$N_j(x_1, L_0) \geq \left(\frac{|L_0|\hat{p}}{2} - r \log n\right)/(r - s) \geq \frac{\alpha n \hat{p}}{2r}$$

for some  $j \in [r] \setminus [s]$ . Without loss of generality,  $j = s + 1$ . Set  $L_1 := N_{s+1}(x_1, L_0)$  and  $V_1 = V'_0 - \{x_1\}$ .

Now suppose that for some  $1 \leq i \leq r - s$ , we have found vertices  $\{x_1, \dots, x_i\} \subset V_0$  and sets  $L_j$  for  $1 \leq j \leq i$  such that  $|L_i| \geq \alpha n \left(\frac{\hat{p}}{2r}\right)^i$ ,  $L_i \subseteq L_{i-1}$  and  $L_i \subset N_j(x_j)$  for  $1 \leq j \leq i$ . Expose  $G_{i+1}$  and apply Lemma 6.3 to  $G_{i+1}$  with  $L_i$  as  $L$ . We may apply the lemma since

$$|L_i|\hat{p} \geq \alpha n \left(\frac{\hat{p}}{2r}\right)^i \hat{p} \geq \alpha n \left(\frac{p}{2r(r+1)}\right)^r \geq 80r \log n.$$

Let  $V'_i = \{v \in V_i : |N(v, L_i)| \geq |L_i|\hat{p}/2\}$  and let  $V''_i = V_i \setminus V'_i$ . If every vertex in  $V'_i$  has at least  $r \log n$  neighbors in  $L_i$  with colors from  $[s + i]$  (if  $s + i = r$ , then this is the case by the calculation above), then we may apply Lemma 4.1 and we are done as in the base case. Otherwise, there is a vertex  $x_{i+1} \in V'_i$  satisfying

$$N_j(x_{i+1}, L_i) \geq \left(\frac{|L_i|\hat{p}}{2} - r \log n\right)/(r - (s + i)) \geq \alpha n \left(\frac{\hat{p}}{2r}\right)^{i+1}$$

for some  $j \in [r] \setminus [s + i]$ . Without loss of generality,  $j = s + i + 1$ . Set  $L_{i+1} := N_{s+i+1}(x_{i+1}, L_i)$  and  $V_{i+1} = V'_i - \{x_{i+1}\}$ . Thus we will find the desired partition after at most  $r - s$  iterations of the above procedure. At each stage we lose at most  $9 \log n/p$  many vertices and thus we lose at most  $9r \log n/p$  in total.  $\square$

## 7.2 Lower bounds on the tree cover number

**Theorem 7.2.** *For all  $r \geq 2$ ,*

- (i) *if  $p = \left(\frac{r \log n - \omega(1)}{n}\right)^{1/r}$ , then  $\text{tc}_r(G(n, p)) > r$ , and*

(ii) if  $p = o\left(\left(\frac{r \log n}{n}\right)^{1/r}\right)$ , then  $\text{tc}_r(G(n, p)) \rightarrow \infty$ .

*Proof.* (i) We apply Lemma 6.4 (i) to get an independent set of size  $r$  which is not dominating and with no common neighbor. But then Observation 3.2 applied with  $s = r$  finishes the proof.

(ii) Choose  $s$  as a function of  $n$  so that  $s \rightarrow \infty$ , but  $s \log s = o(\log n)$ . Similarly to the previous part, the proof follows by applying Lemma 6.4 (ii) and Observation 3.2.  $\square$

### 7.3 Large monochromatic components

Finally, we prove that for all  $r \geq 2$  and  $0 < \epsilon \ll 1/r$ , there exists  $C > 0$  such that if  $p \geq \frac{C}{n}$ , then a.a.s.  $\text{tm}_r(G(n, p)) \geq (1 - \epsilon) \frac{n}{r-1}$ .

*Proof of Theorem 1.8.* Let  $r \geq 2$ ,  $0 < \epsilon < \min\{1/2r, 1/9\}$ ,  $m \geq 1/\epsilon^2$ , and let  $M$  be given by Lemma 5.1. Choose  $C$  sufficiently large for an application of Lemma 5.4, let  $p \geq \frac{C}{n}$ , and let  $G_1, \dots, G_r$  be an edge coloring of  $G \sim G(n, p)$ .

Apply Lemma 5.1 with  $\epsilon^4/r^2$ ,  $m$ , and  $r$  and then Lemma 5.2 to get a “cleaned-up” reduced graph  $\Gamma'$  on  $k' \geq (1 - \epsilon^2)k$  vertices with  $\delta(\Gamma') \geq (1 - 2\epsilon^2)k$ . Apply Theorem 4.8 to  $\Gamma'$  to get a monochromatic tree  $\mathcal{T}$  in  $\Gamma'$  with either  $|\mathcal{T}| \geq (1 - 2\epsilon^2)k'$  or  $\mathcal{T}$  having at least  $(1 - 4\epsilon^2)k'/(r - 1)$  leaves. Apply Lemma 5.6 to each edge of  $\mathcal{T}$  to get the desired tree  $T$ . In the second case, note that the number of leaves of  $T$  is at least

$$\frac{(1 - 4\epsilon^2)k'}{r - 1} \cdot (1 - 4\epsilon^2) \frac{n}{k} \geq (1 - 4\epsilon^2)^2 (1 - \epsilon^2) \frac{n}{r - 1} \geq (1 - \epsilon) \frac{n}{r - 1},$$

where the last inequality holds since  $\epsilon \leq 1/9$ .  $\square$

## 8 Open Problems

The main open problem is to improve Theorem 1.6.(iii) so as to avoid the need for the leftover vertices. We make the following conjecture and note that for  $r \geq 3$  it would be interesting to get any non-trivial bound on  $p$  such that  $\text{tp}_r(G(n, p)) \leq r$ .

**Conjecture 8.1.** For all  $\epsilon > 0$ ,  $r \geq 1$ , if  $p \geq \left(\frac{(1+\epsilon)r \log n}{n}\right)^{1/r}$ , then a.a.s.  $\text{tp}_r(G(n, p)) \leq r$ .

Note: While this paper was under review, Kohayakawa, Mota, and Schacht [24] proved the  $r = 2$  case of Conjecture 8.1.

It would be interesting to extend Theorem 4.6 to a partition version.

**Conjecture 8.2.** For all graphs  $G$  on  $n$  vertices, if  $\delta(G) \geq \frac{2n-5}{3}$ , then  $\text{tp}_2(G) \leq 2$ .

For the following conjecture, Theorem 4.6 provides the  $r = 2$  case, and for the  $r = 1$  case, note that if  $\delta(G) \geq \frac{n-1}{2}$ , then  $G$  is connected, i.e.  $\text{tp}_1(G) = \text{tc}_1(G) = 1$ . Furthermore, Example 3.3 shows that this is best possible if true.

**Conjecture 8.3.** For all  $r \geq 1$ , if  $G$  is a graph on  $n$  vertices with  $\delta(G) \geq \frac{r(n-r-1)+1}{r+1}$ , then  $\text{tc}_r(G) \leq r$ .

Regarding the distinct colors variant of these problems, we make the following conjecture which is true for  $r = 1$  as above, and for  $r = 2$  by Letzter's result [26]. Furthermore, Example 3.5 shows that this is best possible if true.

**Conjecture 8.4.** Let  $r \geq 1$ . If  $\delta(G) \geq (1 - \frac{1}{2r})n$ , then  $G$  has property  $\mathcal{TC}_r$  ( $\mathcal{TP}_r$ ).

Finally, in Theorem 4.3 we prove that if an  $r$ -colored graph  $G$  has sufficiently large minimum degree, then  $G$  can be partitioned into  $r$  monochromatic trees, each of which implicitly has many leaves. What about partitioning into trees with few leaves?

**Problem 8.5.** For all  $r \geq 2$ , sufficiently small  $\epsilon > 0$ , and sufficiently large  $n_0$ , if  $G$  is a graph on  $n \geq n_0$  vertices with  $\delta(G) \geq (1 - \epsilon)n$ , then in every  $r$ -coloring of  $G$  there exists a partition of  $G$  into  $O(r)$  monochromatic trees so that each tree has  $O(1)$  leaves.

## Acknowledgements

We thank Rajko Nenadov, Frank Mousset, Nemanja Škorić and independently Hiệp Hán for drawing our attention to an error in an earlier version of this paper related to Theorem 1.6. We also thank the referees for making many useful comments which helped us improve the organization of the paper.

## References

- [1] R. Aharoni. Ryser's conjecture for tripartite 3-graphs. *Combinatorica*, 21(1):1–4, 2001.
- [2] P. Allen. Covering two-edge-coloured complete graphs with two disjoint monochromatic cycles. *Combinatorics, Probability and Computing*, 17(04):471–486, 2008.
- [3] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, third edition, 2008. With an appendix on the life and work of Paul Erdős.
- [4] J. Balogh, J. Barát, D. Gerbner, A. Gyárfás, and G. N. Sárközy. Partitioning 2-edge-colored graphs by monochromatic paths and cycles. *Combinatorica*, 34(5):507–526, 2014.
- [5] S. Bessy and S. Thomassé. Partitioning a graph into a cycle and an anticyle, a proof of lehel's conjecture. *Journal of Combinatorial Theory, Series B*, 100(2):176–180, 2010.
- [6] T. Bohman, A. Frieze, M. Krivelevich, P. Loh, and B. Sudakov. Ramsey games with giants. *Random Structures & Algorithms*, 38(1-2):1–32, 2011.
- [7] D. Conlon. Combinatorial theorems relative to a random set. [arXiv:1404.3324](https://arxiv.org/abs/1404.3324), 2014.

- [8] L. DeBiasio and L. Nelsen. Monochromatic cycle partitions of graphs with large minimum degree. *Journal of Combinatorial Theory, Series B*, 122:634–667, 2017.
- [9] A. Dudek and P. Prałat. On some multicolour ramsey properties of random graphs. [arXiv:1601.02564](#), 2016.
- [10] M. Elekes, D. T Soukup, L. Soukup, and Z. Szentmiklóssy. Decompositions of edge-colored infinite complete graphs into monochromatic paths. [arXiv:1502.04955](#), 2015.
- [11] P. Erdős, A. Gyárfás, and L. Pyber. Vertex coverings by monochromatic cycles and trees. *Journal of Combinatorial Theory, Series B*, 51(1):90–95, 1991.
- [12] P. Erdős, E. M. Palmer, and R. W. Robinson. Local connectivity of a random graph. *Journal of Graph Theory*, 7(4):411–417, 1983.
- [13] A. Frieze, M. Karonski. *Introduction to Random Graphs*. Cambridge University Press, 2016
- [14] S. Fujita, M. Furuya, A. Gyárfás, and Á. Tóth. Partition of graphs and hypergraphs into monochromatic connected parts. *Electronic Journal of Combinatorics*, 19(3):#P27, 2012.
- [15] S. Fujita, H. Liu, and C. Magnant. Monochromatic structures in edge-coloured graphs and hypergraphs-a survey, 2015.
- [16] Z. Füredi. Maximum degree and fractional matchings in uniform hypergraphs. *Combinatorica*, 1(2):155–162, 1981.
- [17] L. Gerencsér and A. Gyárfás. On Ramsey-type problems. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math*, 10:167–170, 1967.
- [18] A. Gyárfás. Partition covers and blocking sets in hypergraphs. *MTA SZTAKI Tanulmányok*, 71, 1977.
- [19] A. Gyárfás. Large monochromatic components in edge colorings of graphs: a survey. In *Ramsey Theory*, pages 77–96. Springer, 2011.
- [20] A. Gyárfás, M. Ruszinkó, G. N. Sárközy, and E. Szemerédi. An improved bound for the monochromatic cycle partition number. *Journal of Combinatorial Theory, Series B*, 96(6):855–873, 2006.
- [21] P. E. Haxell and Y. Kohayakawa. Partitioning by monochromatic trees. *J. Combin. Theory Ser. B*, 68(2):218–222, 1996.
- [22] J. R. Henderson. *Permutation Decompositions of  $(0, 1)$ -Matrices and Decomposition Transversals*. PhD thesis, California Institute of Technology, 1970.
- [23] Y. Kohayakawa. Szemerédi’s regularity lemma for sparse graphs. In *Foundations of computational mathematics*, pages 216–230. Springer, 1997.
- [24] Y. Kohayakawa, G. O. Mota, and M. Schacht. Monochromatic trees in random graphs. [arXiv:1611.10299](#), 2016.
- [25] S. Letzter. Path Ramsey number for random graphs. *Combinatorics, Probability and Computing*, 25(4):612–622, 2016.



- [26] S. Letzter. Monochromatic cycle partitions of 2-coloured graphs with minimum degree  $3n/4$ . [arXiv:1502.07736](https://arxiv.org/abs/1502.07736), 2015.
- [27] H. Liu, R. Morris, and N. Prince. Highly connected monochromatic subgraphs of multicolored graphs. *Journal of Graph Theory*, 61(1):22–44, 2009.
- [28] L. Lovász. On minimax theorems of combinatorics. *Math. Lapok*, 26:209–264, 1975.
- [29] T. Łuczak, V. Rödl, and E. Szemerédi. Partitioning two-coloured complete graphs into two monochromatic cycles. *Combinatorics, Probability and Computing*, 7(04):423–436, 1998.
- [30] D. Mubayi. Generalizing the Ramsey problem through diameter. *Electronic Journal of Combinatorics*, 9(1):#R41, 2002.
- [31] A. Pokrovskiy. Partitioning edge-coloured complete graphs into monochromatic cycles and paths. *Journal of Combinatorial Theory, Series B*, 106:70–97, 2014.
- [32] A. Pokrovskiy. Calculating Ramsey numbers by partitioning colored graphs. *Journal of Graph Theory*, 2016.
- [33] R. Rado. Monochromatic paths in graphs. *Annals of Discrete Mathematics*, 3:191–194, 1978.
- [34] V. Rödl and A. Ruciński. Threshold functions for Ramsey properties. *Journal of the American Mathematical Society*, 8(4):917–942, 1995.
- [35] R. H. Schelp. Some Ramsey–Turán type problems and related questions. *Discrete Mathematics*, 312(14):2158–2161, 2012.
- [36] A. Scott. Szemerédi’s regularity lemma for matrices and sparse graphs. *Combinatorics, Probability and Computing*, 20(03):455–466, 2011.
- [37] R. Spöhel, A. Steger, and H. Thomas. Coloring the edges of a random graph without a monochromatic giant component. *Electronic Journal of Combinatorics*, 17(1):#R133, 2010.
- [38] E. Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.