MONOCHROMATIC COMPONENTS WITH MANY EDGES IN RANDOM GRAPHS

HANNAH FOX, SAMMY LUO

ABSTRACT. In an r-coloring of edges of the complete graph on n vertices, how many edges are there in the largest monochromatic connected component? A construction of Gyárfás shows that for infinitely many values of r, there exist colorings where all monochromatic components have at most $\left(\frac{1}{r^2-r} + o(1)\right)\binom{n}{2}$ edges. Conlon, Luo, and Tyomkyn conjectured that components with at least this many edges are attainable for all $r \geq 3$. Conlon, Luo, and Tyomkyn proved this conjecture for $r = 3$ and Luo proved it for $r = 4$, along with a lower bound of $\frac{1}{r^2-r+\frac{5}{4}}\binom{n}{2}$ for all $r \geq 2$ and n.

In this paper, we look at extensions of this problem where the graph being r-colored is a sparse random graph or a graph of high minimum degree. By extending several intermediate technical results from previous work in the complete graph setting, we prove analogues in both the sparse random setting and the high minimum degree setting of the bounds for $r = 3$ and general r.

1. INTRODUCTION

Given a graph G with vertex set V and edge set E, an r-coloring of the edges of G (or an r-edge-coloring of G) is a map $\chi : E \to [r]$, i.e. a way to assign one of r given colors to each edge of G. Given such a coloring, we can consider the *color classes* $G_i = \chi^{-1}(i)$, i.e. the subgraphs consisting of edges of a given color i. This paper studies the sizes of the connected components (i.e. maximal connected subgraphs) within the color classes in such an r-edge-coloring of a graph G.

A classical observation of Erdős and Rado is that in any 2-coloring of the edges of the complete graph K_n, one of the two color classes consists of a single connected component on all n vertices. Gyárfás [4] extended this result, proving that for $r \geq 2$, in any r-coloring of the edges of K_n, there is a monochromatic connected component with at least $\frac{n}{r-1}$ vertices. He also exhibited a construction showing that this bound is tight for arbitrarily large n and infinitely many values of r.

Since then, Gyárfás's result has been further generalized in many different ways. One notable example is due to Bal and DeBiasio [1], who showed that a similar result holds when K_n is replaced by an Erdős-Rényi random graph $G(n, p)$, which is a random graph

---

on n vertices where each pair of vertices is adjacent with an independent probability of p.

Theorem 1.1 ([1]). *For all $r \geq 2$ and sufficiently small $\varepsilon > 0$, there exists C such that for $p \geq \frac{C}{n}$, with high probability every r-coloring of $G(n,p)$ contains a monochromatic connected component with at least $(1 - \varepsilon)\frac{n}{r-1}$ vertices.*

We can also ask analogous questions about the largest number of *edges* in a monochromatic component.

Question 1.2. *If the edges of K_n are colored with r colors, what is the largest number of edges that some monochromatic component is guaranteed to have?*

This question was first studied by Conlon and Tyomkyn [2], who showed that in the case $r = 2$, every such coloring contains a monochromatic component with at least $\frac{2}{9}n^2 + O(n^{3/2})$ edges. Later work by Luo [5] and by Conlon, Luo, and Tyomkyn [3] resolved the cases $r = 3$ and $r = 4$ respectively, showing that when the edges of K_n are colored with 3 (resp. 4) colors, there always exists a monochromatic connected component with at least $\frac{1}{6}\binom{n}{2}$ (resp. $\frac{1}{12}\binom{n}{2}$) edges, and that these bounds are tight.

The lower bounds in the $r = 3$ and $r = 4$ cases match the upper bounds arising from a close examination of Gyárfás's construction, which shows that for infinitely many values of r (in particular, whenever $r - 1$ is a prime power), there is an r-coloring of K_n where every monochromatic component has at most $(\frac{1}{r(r-1)} + o(1))\binom{n}{2}$ edges. Conlon, Luo, and Tyomkyn conjectured that a matching lower bound holds for every value of r.

Conjecture 1.3 ([3]). *For any positive integers n, r with $r \geq 3$, in every r-coloring of the edges of K_n, there is some monochromatic connected component with at least $\frac{1}{r(r-1)}\binom{n}{2}$ edges.*

In [5], Luo proves a general lower bound that approximates the conjectured lower bound within a factor of $1 - O(\frac{1}{r^2})$.

Theorem 1.4 ([5]). *In any r-coloring of the edges of K_n for $r \geq 2$, there is a monochromatic connected component with at least $\frac{1}{r^2 - r + \frac{5}{4}}\binom{n}{2}$ edges.*

In analogy to the generalization of Gyárfás's original result to Theorem 1.1, in this paper we study a generalization of Question 1.2 for random graphs: In an r-coloring of the edges of the random graph $G(n,p)$, what lower bounds can we obtain (with high probability) for the number of edges in the largest monochromatic component?

The answer to this question naturally depends on our choice of the density p. We may ask if there is a *threshold* for values of p above which the answer to the above question about $G(n,p)$ starts to behave in a certain way with high probability, e.g. when our bounds on the number of edges in the largest monochromatic component start to behave similarly to the analogous bounds in the complete graph setting.

Our main result is an analogue of Theorem 1.4 in the setting of $G(n,p)$ for any $p = \omega(\frac{\log n}{n})$.

Theorem 1.5. *For any fixed integer $r \geq 2$ and any $p = \omega(\frac{\log n}{n})$, $G := G(n, p)$ satisfies the following property with high probability: In every r-coloring of the edges of G, there is a monochromatic connected component with at least $\frac{1}{r^2 - r + \frac{5}{4}}|E(G)|$ edges.*

We prove this result by extending the argument used in proving Theorem 1.4, including several technical lemmas, to the sparse random graph setting. Applying similar extensions yields a sparse random analogue of the lower bound for $r = 3$ as well.

The paper proceeds as follows. We begin in Section 2 by providing some background on previous work on Question 1.2 for complete graphs, including some technical tools that we will adapt for the random graph setting. We also introduce some relevant results on random graphs that will be used later in the proofs. Then, in Section 3, we prove our main results on $G(n, p)$. We follow in Section 4 by investigating analogues of our questions in the setting of graphs of high minimum degree and discussing how this setting relates to the sparse random setting. We conclude by discussing some potential future directions.

2. Preliminaries

All graphs studied in this paper are finite and simple. For a graph G, let $E(G)$ and $V(G)$ denote the sets of edges and vertices of G, respectively, and write $e(G) = |E(G)|$ and $v(G) = |V(G)|$. Given subsets of vertices $S, T \subseteq V(G)$, let $G[S]$ denote the induced subgraph of G on the vertex set S (i.e. $G[S]$ is the graph on S such that two vertices $v, w \in S$ are adjacent in $G[S]$ if and only if they are adjacent in G). Let $E_G(S) = E(G[S])$, $e_G(S) = |E_G(S)|$, and let $e_G(S, T)$ denote the number of ordered pairs $(s, t) \in S \times T$ such that s, t are adjacent in S. We write $e(S, T) = e_G(S, T)$, $E(S) = E_G(S)$, and $e(S) = e_G(S)$ when the choice of G is clear from context. Given two graphs G and M on the same set of vertices V, we write $G \cap M$ for the graph on vertex set V whose edge set is $E(G) \cap E(M)$.

For real-valued functions f and g, we write $f = o(g)$ or $f \ll g$ if $\lim_{n \to \infty} \frac{f(n)}{g(n)} = 0$, and we write $f = \omega(g)$ or $f \gg g$ if $\lim_{n \to \infty} \frac{f(n)}{g(n)} = \infty$. Unless otherwise noted, all such asymptotic notation will be in terms of the variable n going to infinity.

To extend results about edge colorings of complete graphs to the random graph setting, we will frequently apply the following form of the Chernoff bound:

Proposition 2.1 (Chernoff bound)**.** *Let $X = X_1 + \cdots + X_n$, where the X_i are i.i.d. Bernoulli random variables. Let $\mu = \mathbb{E}[X]$. Then for any $\delta > 0$,*

$$\mathbb{P}(X \geq (1 + \delta)\mu) \leq e^{-\delta^2 \mu/(2+\delta)},$$

while for $\delta \in (0, 1)$,

$$\mathbb{P}(X \leq (1 - \delta)\mu) \leq e^{-\delta^2 \mu/2}.$$

In particular, for $\delta \in (0, 1)$,

$$\mathbb{P}(|X - \mu| \geq \delta\mu) \leq 2e^{-\delta^2\mu/3}.$$

2.1. **Upper bounds and Gyárfás's construction.** Gyárfás's construction [4] gives, for certain values of r, an upper bound corresponding to the conjectured lower bound in Conjecture 1.3, showing that the conjecture, if true, is tight for infinitely many values of r. We summarize the relevant details of the construction below.

The construction depends on the existence of a certain family of objects called affine planes. An affine plane of order t consists of a set of t^2 points and $t^2 + t$ lines, where every pair of points shares exactly one line, and the set of lines can be partitioned into $t + 1$ classes (or "directions") of parallel lines, such that each pair of lines intersects in one point if they are not parallel (and no points if they are parallel). Affine planes of order t are known to exist whenever t is a prime power; it is open whether affine planes of any other order exist.

When an affine plane of order $r - 1$ exists and $n = (r - 1)^2$, we can construct an r-coloring of K_n as follows: The vertices are the points of an affine plane of order $r - 1$, and the edge between any pair of vertices is colored based on the direction of the line between the corresponding points. Examples for $r - 1 = 2$ and $r - 1 = 3$ are shown below.
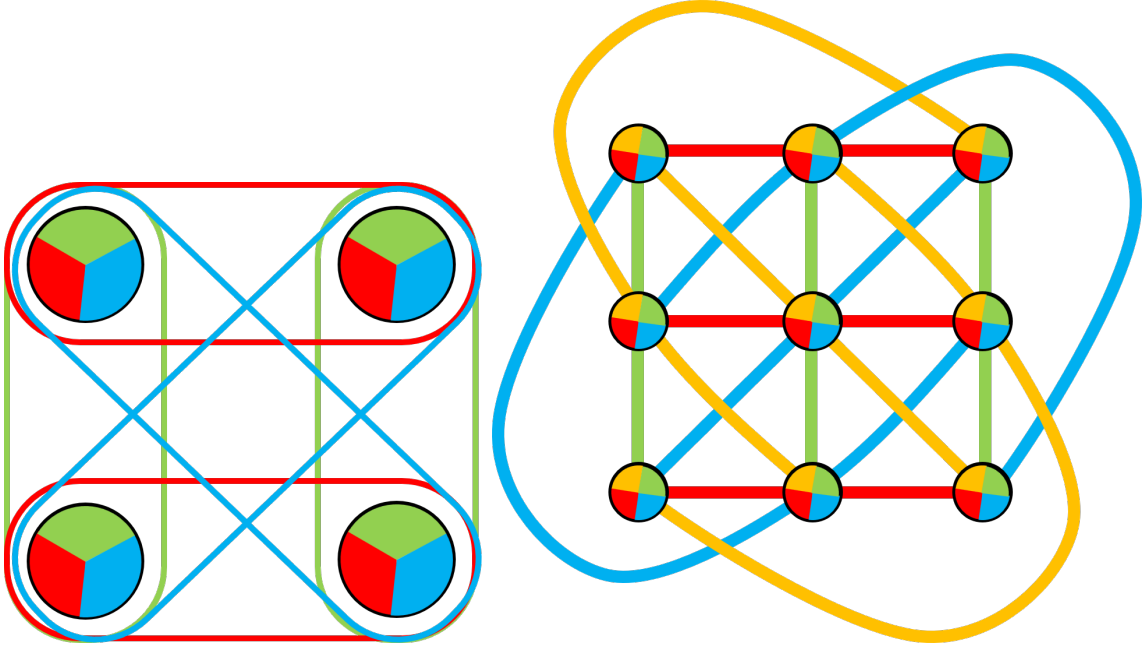
To extend this coloring to larger values of n, we partition the vertices of K_n into $(r - 1)^2$ clusters, each corresponding to a point in the affine plane of order $r - 1$. For each edge between two distinct clusters, we color it based on the direction of the line between the two corresponding points in the affine plane. Within each cluster, we color the edges such that each cluster has a roughly equal number of edges of each color. Each

connected component in a given color k then consists of the clusters corresponding to the points on one line in the affine plane with the direction corresponding to k. The number of components then equals the number of lines in the affine plane of order $r - 1$, which is $r(r - 1)$. By symmetry, each component contains roughly the same number of edges, so each component contains roughly $\frac{1}{r^2-r}\binom{n}{2}$ edges.

We can use the same construction to show an analogous upper bound in the random graph setting, by identifying $G = G(n, p)$ as a subgraph of K_n and taking the coloring induced by the r-edge-coloring of K_n given by Gyárfás's construction. For fixed r and large n, if p is not too small (e.g. $p = \omega(1/n)$ suffices), each monochromatic connected component in the r-edge-coloring of K_n will have, with high probability, close to a p fraction of its edges remain in G, meaning that for any constant $\varepsilon > 0$, with high probability, $G = G(n, p)$ will have the property that for some r-coloring of its edges, every monochromatic connected component has at most $(1 + \varepsilon)p\frac{1}{r^2-r}\binom{n}{2}$ edges.

2.2. **Lower bounds in the complete graph setting.** An important tool in the proof of Theorem 1.4 in [5] is the following inequality, which gives a density version of the relevant coloring results about sizes of connected components in subgraphs. Recall that a *complete k-partite graph* is a graph on a vertex set $V = V_1 \cup \cdots \cup V_k$, where the V_i are pairwise disjoint, and a pair of vertices $v \in V_i$, $v' \in V_j$ is connected by an edge if and only if $i \neq j$. A graph M is called a *complete multipartite graph* if it is a complete k-partite graph for some $k \geq 2$.

Theorem 2.2 ([5]). *In a subgraph H of a complete multipartite graph M, there must be a connected component with at least $\frac{e(H)^2}{e(M)}$ edges.*

This result, in turn, follows from the following even simpler looking (yet still nontrivial) inequality, which is Corollary 6 in [5].

Proposition 2.3 ([5]). *Let M be a complete multipartite graph. For any $S, T \subseteq V(M)$, we have*

$$e(S, T)^2 \geq 4e_M(S)e_M(T).$$

In order to prove Theorem 1.5, we will need to adapt the proofs of Proposition 2.3 and Theorem 2.2, in order to extend them to the random graph setting.

3. PROOF OF RESULTS ON RANDOM GRAPHS

Our first aim is to prove an analogue of Proposition 2.3 for random graphs. We begin with a series of technical results that are essentially direct applications of the Chernoff bound.

Proposition 3.1. *For any choice of $p = p_n$ such that $p = \omega(\frac{\log n}{n})$, there exists a sequence of values $\varepsilon = \varepsilon_n = o(1)$ such that the following holds: In $G = G(n, p)$, with high*

probability, for every graph H on $V(G)$ satisfying $e(H) \geq \varepsilon\binom{n}{2}$ that is either a clique on a subset of vertices or a complete multipartite graph on $V(G)$,

$$(1 - \varepsilon)\mu \leq e(H \cap G) \leq (1 + \varepsilon)\mu,$$

where $\mu = \mathbb{E}[e(H \cap G)] = pe(H)$.

Proof. For any $H \subseteq E(K_n)$ and $\delta \in (0, 1)$, the Chernoff bound yields

$$\mathbb{P}(|e(H \cap G) - \mu| \geq \delta\mu) \leq e^{-\delta^2\mu/3}.$$

When $|H| \geq \varepsilon\binom{n}{2}$ and $\delta = \varepsilon$, this yields

$$\mathbb{P}(|e(H \cap G) - \mu| \geq \varepsilon\mu) \leq \exp(-\varepsilon^2 p\varepsilon\binom{n}{2}/3) \leq \exp(-(1 + o(1))\varepsilon^3 pn^2/6).$$

Since there are at most $2^n + n^n \leq n^{2n}$ possible choices of H, the overall probability that there exists some valid choice of H with $|e(H) - \mu| \geq \varepsilon\mu$ is then bounded above by

$$\exp(-(1 + o(1))\varepsilon^3 pn^2/6 + 2n \log n).$$

For this probability to go to 0, it suffices to have

$$\varepsilon^3 pn^2/(12n \log n) > 1.$$

Thus it suffices to take $\varepsilon = 3(\frac{\log n}{pn})^{1/3}$. Since $p = \omega(\frac{\log n}{n})$, this yields $\varepsilon = o(1)$ as desired. $\qquad\square$

Proposition 3.2. *For any $c_0 \in (0, 1)$ and any choice of $p = p_n$ such that $p = \omega(\frac{\log n}{n})$, there exists a sequence of values $\varepsilon = \varepsilon_n = o(1)$ such that the following holds. In $\hat{G} := G(n, p)$, with probability $1 - o(\frac{1}{n})$, for all $c \geq c_0$, every subgraph H with average degree $\bar{d}(H) \geq c_0 p(n - 1)$ has $v(H) > \bar{d}(H)(1 - \varepsilon)\frac{n}{p(n-1)}$.*

Proof. It suffices to consider the induced subgraphs $H := G[S]$ on vertex sets $S \subseteq V(G)$. We fix some such S and consider the probability that

$$\bar{d}(H) \geq \max(c_0 p(n - 1), p|S|(1 - \varepsilon)^{-1}\tfrac{n-1}{n}).$$

Note that a subgraph with average degree at least $c_0 p(n - 1)$ must have at least $c_0 p(n - 1) + 1$ vertices, so we can assume $s := |S| \geq c_0 p(n - 1) + 1 \geq c_0 pn$. By the Chernoff bound, we obtain that

$$\mathbb{P}(\bar{d}(H) \geq \max(c_0 p(n - 1), ps(1 - \varepsilon)^{-1}\tfrac{n-1}{n})) \leq \mathbb{P}(e(H) \geq \frac{s^2}{2}\frac{p(n - 1)}{(1 - \varepsilon)n})$$

$$\leq \mathbb{P}(e(H) \geq \frac{1}{1 - \varepsilon}\mathbb{E}[e(H)])$$

$$\leq e^{-\frac{\varepsilon^2}{(1-\varepsilon)(2-\varepsilon)}p\binom{s}{2}} \leq e^{-\frac{\varepsilon^2}{4}(n-1)c_0 sp}.$$

For any fixed value of s, a union bound over all S of size s shows that the probability that G has a subgraph on s vertices not satisfying the relevant event is at most

$$\binom{n}{s} e^{-\frac{\varepsilon^2}{4}(n-1)c_0 sp} \leq e^{s(1+\log n - \log s) - \frac{\varepsilon^2}{4}(n-1)c_0 sp} \leq e^{s(1+\log n - \frac{\varepsilon^2}{4}(n-1)c_0 p)}.$$

Since $p = \omega(\frac{\log n}{n})$, for n large enough such that $p > \frac{8}{\varepsilon^2 c_0} \frac{\log n}{n-1}$ and $\log n > 1$, this last quantity is maximized when s is as small as possible, at a value of $e^{c_0 pn(1+\log n - \frac{\varepsilon^2}{4}(n-1)c_0 p)}$. Taking a union bound over all possible values of $s \leq n$, the probability of any such subgraph existing is at most $ne^{c_0 pn(1+\log n - \frac{\varepsilon^2}{4}(n-1)c_0 p)} = e^{\log n + c_0 pn(1+\log n - \frac{\varepsilon^2}{4}(n-1)c_0 p)}$. Since $p = \omega(\frac{\log n}{n})$, we can choose $\varepsilon = o(1)$ such that this upper bound is $e^{-\omega((\log n)^2)} = o(\frac{1}{n})$, as claimed. $\square$

Next, we study the intersection of $G = G(n,p)$ with an arbitrary complete multipartite graph M on $V(G)$. In the proof of Theorem 1.5, we will not have much control over the specific choice of M after fixing an instance of $G(n,p)$, so it is important that our intermediate claims hold for all choices of M.

Proposition 3.3. *For any choice of $p = p_n$ such that $p = \omega(\frac{\log n}{n})$ and any constant $c \in (0,1)$, with high probability, the following holds for $G := G(n,p)$. For any complete multipartite graph M on $V(G)$ with minimum degree $\delta(M) \geq cn$, the minimum degree of $G \cap M$ is at least $\frac{p}{2}\delta(M)$.*

Proof. For any vertex $v \in V(G)$, we have $\mathbb{E}[\deg_{G \cap M}(v)] = p \deg_M(v) \geq p\delta(M)$, so the Chernoff bound gives

$$\Pr[\deg_{G \cap M}(v) < \frac{p}{2}\delta(M)] \leq e^{-p\delta(M)/8} \leq e^{-cpn/8}.$$

Recalling that $p = \omega(\log n/n)$, a union bound gives

$$\Pr[\min_{v \in V(G)} \deg_{G \cap M}(v) < \frac{p}{2}\delta(M)] \leq ne^{-cpn/8} = o(1),$$

as claimed. $\square$

We are now ready to prove the following random graph analogue of Proposition 2.3.

Theorem 3.4. *For any choice of $p = p_n$ such that $p = \omega(\frac{\log n}{n})$ and any constant $c \in (0,1)$, there exists a sequence of values $\varepsilon = \varepsilon_n = o(1)$ such that with $1 - o(\frac{1}{n})$ probability, in $G := G(n,p)$, the following holds for every complete multipartite graph M on the vertices of G:*

For every pair of subsets $S, T \subseteq V(G)$ with $e_M(S,T) \geq cn^2$, in the graph $M' = G \cap M$, we have

$$e_{M'}(S,T)^2 \geq 4e_{M'}(S)e_{M'}(T)(1-\varepsilon).$$

We will prove this result by arguing that, without the factor of $1 - \varepsilon$, the inequality is true in expectation, and that with high probability, the actual values of both sides of the equation are close to their expected values.

Proof. There exists a constant d such that the number of ways to choose M, S, and T is at most $e^{dn \log n}$. Fix such a choice of M, S, and T. By Proposition 2.3, $e_M(S,T)^2 \geq 4e_M(S)e_M(T)$. Therefore,

$$\mathbb{E}[e_{M'}(S,T)]^2 = p^2 e_M(S,T)^2 \geq p^2 4e_M(S)e_M(T)$$
$$= 4\mathbb{E}[e_{M'}(S)]\mathbb{E}[e_{M'}(T)].$$

As in the proof of Proposition 3.1, by the Chernoff bound, for any $\delta \in (0,1)$ we have

$$\mathbb{P}(e_{M'}(S,T) < (1-\delta)\mathbb{E}[e_{M'}(S,T)]) \leq e^{-\frac{\delta^2 \mathbb{E}[e_{M'}(S,T)]}{2}} \leq e^{-\frac{\delta^2 pcn^2}{2}}.$$

Taking $\delta = a\sqrt{\frac{\log n}{np}} = o(1)$ for a sufficiently large constant a yields $e^{-\frac{\delta^2 pcn^2}{2}} \cdot e^{dn \log n} = e^{n \log n(d - \frac{a^2 c}{2})} = o(\frac{1}{n})$, so that with $1 - o(\frac{1}{n})$ probability,

$$e_{M'}(S,T) \geq (1-\delta)\mathbb{E}[e_{M'}(S,T)]$$

for all choices of M, S, and T. We also have

$$\mathbb{P}(e_{M'}(S) > (1+\delta)\mathbb{E}[e_{M'}(S)]) \leq e^{-\frac{\delta^2 \mathbb{E}[e_{M'}(S)]}{2}} \leq e^{-\frac{\delta^2 pcn^2}{2}}.$$

Therefore, with $1 - o(\frac{1}{n})$ probability,

$$e_{M'}(S,T)^2 \geq \frac{1-\delta}{1+\delta}4e_{M'}(S)e_{M'}(T),$$

for all choices of M, S, and T. Choosing $\varepsilon = 2\delta$ gives the desired conclusion. $\square$

Next, we prove the following random graph analogue of Theorem 2.2.

Theorem 3.5. *For any choice of $p = p_n$ such that $p = \omega(\frac{\log n}{n})$ and any constants $c_1, c_2 \in (0,1)$, there exists a sequence of values $\varepsilon = \varepsilon_n = o(1)$ such that with $1 - o(\frac{1}{n})$ probability, in $G := G(n,p)$, the following holds for every complete multipartite graph M on $V(G)$ with minimum degree at least $c_1 n$:*

Let $G' := G \cap M$. Then every subgraph H of G' with $\frac{e(H)}{e(G')} \geq c_2$ has a connected component H' with $e(H') \geq \frac{e(H)^2}{e(G')}(1-\varepsilon)$.

Proof. Choose $\varepsilon = \varepsilon_n$ small enough to satisfy the conclusions of Proposition 3.2 and Theorem 3.4, with parameters $c_0 = \frac{c_1 c_2}{2}$ and $c = \frac{c_1^2 c_2}{2}$ respectively, and condition on the high probability events in the conclusions of those results. Let $H_1, \ldots, H_k$ be the connected components of H. For any set $S \subset V := V(G)$, let $f(S) = \frac{1}{2}e_{G'}(S,V)$. In particular, we have $f(H_1) + \cdots + f(H_k) = f(V) = \frac{1}{2}e_{G'}(V,V) = e(G')$. Then, by the generalized mediant inequality, for some $H' = H_\ell$, we have

$$\frac{e(H')}{f(V(H'))} \geq \frac{e(H_1) + \cdots + e(H_k)}{f(V(H_1)) + \cdots + f(V(H_k))} = \frac{e(H)}{e(G')} \geq c_2.$$

By Proposition 3.3, since M has minimum degree $\delta(M) \geq c_1 n$, G' has minimum degree at least $\frac{p}{2} c_1 n$, so

$$f(V(H')) = \frac{1}{2}e(V(H'), V) \geq \frac{1}{2}v(H')\frac{p}{2}c_1 n.$$

This means that the average degree of H' is

$$\frac{2e(H')}{v(H')} \geq \frac{2f(H')}{v(H')}c_2 \geq \frac{c_1 c_2}{2}pn.$$

The assumption of Proposition 3.2 then yields $v(H') \geq (1-\varepsilon)\frac{c_1 c_2}{2}n$. Then

$$e_M(V(H'), V) \geq v(H')\delta(M) \geq (1-\varepsilon)\frac{c_1^2 c_2}{2}n^2,$$

so the assumption of Theorem 3.4 with $S = V(H')$ and $T = V$ yields

$$f(V(H'))^2 \geq e(H')e(G')(1-\varepsilon),$$

for some $\varepsilon = o(1)$. Therefore, we conclude that

$$e(H') \geq \frac{e(H')}{f(V(H'))}f(V(H')) \geq \frac{e(H)}{e(G')}\sqrt{e(H')e(G')(1-\varepsilon)},$$

which rearranges to give

$$e(H') \geq \frac{e(H)^2}{e(G')}(1-\varepsilon),$$

in the high probability event we have conditioned on, as desired. $\square$

We finally have all the tools we need to prove our main result, Theorem 1.5.

Proof of Theorem 1.5. Let $z_0 = \frac{1}{r^2-r+\frac{5}{4}}$, $c_1 = \frac{1}{r+1}$, $c_2 = \frac{1}{r-1}$, and $c_0 = \sqrt{z_0\frac{1+\varepsilon_0}{1-\varepsilon_0}}$, where $\varepsilon_0 \leq c_1(1-c_1)$ is some constant depending on r to be chosen later. For a given choice of $p = p_n = \omega(\frac{\log n}{n})$, take a sequence of values $\varepsilon = \varepsilon_n = o(1)$ satisfying the conclusions of Proposition 3.1, Proposition 3.2, and Theorem 3.5 with parameters c_0, c_1, c_2 as chosen above. Assume n is large enough so that $\varepsilon_n < \varepsilon_0$, and condition on the high probability events in the conclusions of Proposition 3.1 and Theorem 3.5, as well as the high probability event in the conclusion of Proposition 3.2 with parameters $\varepsilon = \varepsilon_0$ and c_0. In particular, we can assume that $(1-\varepsilon)p\binom{n}{2} \leq e(G) \leq (1+\varepsilon)p\binom{n}{2}$.

Fix an r-edge-coloring of G. Without loss of generality, let red be the color with the most edges. Let R be the subgraph of G consisting of all its red edges, so $e(R) \geq \frac{1}{r}e(G) \geq (1-\varepsilon)p\frac{1}{r}\binom{n}{2}$. Let $x = \frac{e(R)}{e(G)} \geq \frac{1}{r}$. Let the red connected components be $R_1, \ldots, R_k$, with $v(R_1) \geq \cdots \geq v(R_k)$. Let z be the fraction of the edges of G that are in the monochromatic component with the most edges. It suffices to show that $z \geq z_0$, conditioned on the high probability events specified earlier.

First, suppose that $v(R_1) \geq (1 - \frac{1}{r+1})n$. Then by Proposition 3.1, we have

$$e(R_2) + \cdots + e(R_k) \leq e_G(V(G) \setminus V(R_1)) \leq (1+\varepsilon)p\binom{\frac{1}{r+1}n}{2} \leq (1+\varepsilon)p\frac{1}{(r+1)^2}\binom{n}{2},$$

which means

$$e(R_1) \geq (1-\varepsilon)p\frac{1}{r}\binom{n}{2} - (1+\varepsilon)p\frac{1}{(r+1)^2}\binom{n}{2} \geq \frac{1+\varepsilon}{r^2 - r + \frac{5}{4}}p\binom{n}{2},$$

for $r \geq 2$ and sufficiently small ε, so that $z \geq z_0$ as desired. Thus, we can assume that $v(R_1) < (1 - \frac{1}{r+1})n = (1 - c_1)n$.

For $0 \leq j \leq k-1$, let $G_j = G[V(R_{j+1}) \cup \cdots \cup V(R_k)]$ and $H_j = G_j \cap R$. Let $c = \sqrt{z\frac{1+\varepsilon_0}{1-\varepsilon_0}} \geq c_0$. If any component R_i has average degree at least $pc(n-1) \geq pc_0(n-1)$, then the assumption of Proposition 3.2 gives $v(R_j) > (1 - \varepsilon_0)cn$, so that $e(R_j) > (1-\varepsilon_0)c^2p\binom{n}{2} = z(1+\varepsilon_0)p\binom{n}{2} \geq ze(G)$, contradicting our choice of z. Thus we can assume every R_i, and hence every H_j, has average degree less than $pc(n-1) \leq (1 - \varepsilon_0)^{-1}p\sqrt{z}(n-1)$. In particular, $j = 0$ gives

$$xe(G) = e(R) < (1-\varepsilon_0)^{-1}p\sqrt{z}\binom{n}{2} \leq \frac{\sqrt{z}}{(1-\varepsilon_0)^2}e(G),$$

so that $z \geq (1-\varepsilon_0)^4 x^2 \geq (1-2\varepsilon_0)^2 r^{-2}$. Let $\delta = r - \frac{1-2\varepsilon_0}{\sqrt{z}}$, so $z = \frac{(1-2\varepsilon_0)^2}{(r-\delta)^2}$. If $\delta \geq 1$, then $z \geq z_0$ as long as we pick ε_0 small enough in terms of r, so we can assume $\delta < 1$. More generally, since $e(R_i) \leq ze(G)$ for $1 \leq i \leq k$, we have for any j that

$$e(H_j) = \sum_{i=j+1}^{k} e(R_i) \geq (x - jz)e(G),$$

meaning

$$(1-\varepsilon_0)^{-1}p\sqrt{z}(n-1) > \frac{2e(H_j)}{v(H_j)} \geq \frac{2(x-jz)e(G)}{n - \sum_{i=1}^{j} v(R_i)},$$

which rearranges to give

$$\sum_{i=1}^{j} v(R_i) < n - \frac{2(x-jz)e(G)}{(1-\varepsilon_0)^{-1}p\sqrt{z}(n-1)}$$

$$\leq n(1 - \frac{(1-\varepsilon_0)^2(x-jz)}{\sqrt{z}}) \leq (1+2\varepsilon_0)(1 - x/\sqrt{z} + j\sqrt{z})n.$$

The same smoothing argument as in the proof of [5, Lemma 3.2] then yields

$$\sum_{i=1}^{j} \binom{v(R_i)}{2} \leq \frac{n^2((1 - x/\sqrt{z} + \sqrt{z})^2 + (x/z - 1)z) - n}{2}(1 + 5\varepsilon_0).$$

Let M be the complete k-partite graph on $V(G)$ with $V(R_1), \ldots, V(R_k)$ as the independent sets. Because $G' := M \cap G$ has no red edges, one of the other colors (say, blue) has at least $\frac{1}{r-1}e(G')$ edges in G'. Let H be the subgraph of G' consisting of all its blue edges, so $\frac{e(H)}{e(G')} \geq \frac{1}{r-1} = c_2$. Since $v(R_1) < (1 - c_1)n$, the minimum degree of M is at least $c_1 n$. Then, by Theorem 3.5, there is a connected component H' of H with $e(H') \geq (1 - \varepsilon)\frac{e(H)^2}{e(G')} \geq (1 - \varepsilon)\frac{1}{(r-1)^2}e(G')$. Since $e(M) \geq c_1 n(1 - c_1)n \geq \varepsilon_0 \binom{n}{2}$, Proposition 3.1 yields $e(G') \geq (1 - \varepsilon)pe(M)$, so

$$
\begin{aligned}
z &\geq (1 - \varepsilon)\frac{1}{(r-1)^2}\frac{e(G')}{e(G)} \geq (1 - \varepsilon)^2\frac{1}{(r-1)^2}p\frac{e(M)}{e(G)} \\
&\geq (1 - \varepsilon)^2\frac{1}{(r-1)^2}(1 - ((1 - x/\sqrt{z} + \sqrt{z})^2 + (x/z - 1)z)(1 + 5\varepsilon_0)) \\
&\geq (1 - 7\varepsilon_0)\frac{1}{(r-1)^2}(1 - (1 - x/\sqrt{z} + \sqrt{z})^2 - (x/z - 1)z).
\end{aligned}
$$

This rearranges to yield

$$
\begin{aligned}
-\frac{(r-1)^2 z}{1 - 7\varepsilon_0} &\leq (1 - x/\sqrt{z} + \sqrt{z})^2 + (x/z - 1)z - 1 \\
&= \frac{1}{z}x^2 - (1 + 2/\sqrt{z})x + 2\sqrt{z}.
\end{aligned}
$$

For ε_0 small enough in terms of r, we have $\frac{z}{2} + \sqrt{z} \geq x \geq \frac{1}{r}$, so the right hand side is at most

$$
\frac{1}{zr^2} - (1 + 2/\sqrt{z})\frac{1}{r} + 2\sqrt{z}.
$$

Recall that $z = \frac{(1 - 2\varepsilon_0)^2}{(r-\delta)^2}$ for some $\delta \in (0, 1)$. Substituting this in yields

$$
-\frac{(1 - 2\varepsilon_0)^2(r-1)^2}{(1 - 7\varepsilon_0)(r-\delta)^2} \leq \frac{(r-\delta)^2}{(1 - 2\varepsilon_0)^2 r^2} - \frac{(1 - 2\varepsilon_0) + 2(r-\delta)}{(1 - 2\varepsilon_0)r} + \frac{2(1 - 2\varepsilon_0)}{r-\delta},
$$

which can be rearranged to get

$$
\begin{aligned}
0 &\leq \frac{(1 - 2\varepsilon_0)^4}{1 - 7\varepsilon_0}r^2(r-1)^2 \\
&+ (r-\delta)^4 + 2(1 - 2\varepsilon_0)^3(r-\delta)r^2 - (1 - 2\varepsilon_0)^2 r(r-\delta)^2 - 2(1 - 2\varepsilon_0)r(r-\delta)^3 \\
&\leq (r-\delta^2)^2 - r(r^2 - \delta^2)(1 - 2\delta) + 4\varepsilon_0 r(r-\delta)^2(r-\delta+1),
\end{aligned}
$$

and thus, taking $\varepsilon_0 \leq (32(r+1))^{-4}$,

$$
1 - 2\delta \leq \frac{(r-\delta^2)^2}{r(r^2 - \delta^2)} + \frac{1}{8r^3} < \frac{1}{r} + \frac{1}{8r^3},
$$

so that $\delta > \frac{r-1}{2r} - \frac{1}{16r^3}$, and thus

$$z = \frac{(1-2\varepsilon_0)^2}{(r-\delta)^2} \geq \frac{1 - 32^{-3}r^{-4}}{r^2 - r + \frac{1}{4} + 1 - \frac{1}{4r} + \frac{1}{4r^2} + 2\frac{r}{16r^3}}.$$

For $r \geq 2$, it can be verified that this expression is at least $\frac{1}{r^2-r+\frac{5}{4}} = z_0$, as desired. □

4. High minimum degree argument

We now turn to the closely related setting of graphs of high minimum degree. Unlike in the problem of finding monochromatic components with many vertices, for the edge version of the problem it suffices to assume a global condition of an edge density close to 1, rather than a local condition about degrees. We obtain a result that approximates the bound in Theorem 1.4.

Theorem 4.1. *Let $r \geq 4$ be an integer and let $\beta \in (0, \frac{1}{r-1})$. Then for any graph G on n vertices with at least $(1-\beta)\binom{n}{2}$ edges, every r-coloring of the edges of G contains a monochromatic component with at least $\frac{(1-\beta)^2}{r^2-r+\frac{5}{4}}\binom{n}{2}$ edges.*

Proof. We adapt the proof of Theorem 1.4 in [5], needing to make only a few adjustments. Instead of considering an edge coloring of a subgraph of K_n, we essentially consider an edge coloring of K_n that leaves up to a β fraction of edges uncolored. Fix such a partial r-edge-coloring, let z be the fraction of the edges of K_n that are in the monochromatic component with the most edges, and let red be the color with the most edges. Let R be the subgraph of K_n consisting of all the edges colored red, and let $x = \frac{e(R)}{e(K_n)} \geq \frac{1-\beta}{r}$. By Theorem 2.2 applied trivially with $M = K_n$, we have $z \geq x^2 \geq \frac{(1-\beta)^2}{r^2}$.

Let $R_1, \ldots, R_k$ be the red connected components, with $v(R_1) \geq \cdots \geq v(R_k)$. For $j \in [0, k-1]$, let G_j be the complete graph on $\bigcup_{i=j+1}^{k} V(R_i)$, with its (partial) coloring induced by the coloring on K_n. Let $H_j = G_j \cap R$. Applying Theorem 2.2 with $M = G_j$ and $H = H_j$ yields a red component R_i with $ze(K_n) \geq e(R_i) \geq \frac{e(H_j)^2}{e(G_j)}$. We have

$$e(H_j) = \sum_{i=j+1}^{k} e(R_i) \geq (x - jz)e(K_n),$$

while $e(G_j) = \binom{n-\sum_{i=1}^{j} v(R_i)}{2}$. This is the exact same setup as in the proof of Theorem 1.4 given in [5]. Letting $v_i = \frac{V(R_i)}{n}$ and solving the optimization via [5, Lemma 3.2] yields

$$\sum_{i=1}^{m} v_i^2 \leq (1 - x/\sqrt{z} + \sqrt{z})^2 + (x/z - 1)z,$$

so that

$$\sum_{i=1}^{m} \binom{v(R_i)}{2} \leq \frac{n^2((1 - x/\sqrt{z} + \sqrt{z})^2 + (x/z - 1)z) - n}{2}.$$

Let M be the complete multipartite graph obtained by removing all edges within each $V(R_i)$ from K_n, with its (partial) coloring induced by the coloring on K_n. Then no edges of M are colored red, and at most $\beta e(K_n)$ edges of M are uncolored, so by the pigeonhole principle, some color class H in M has $e(H) \geq \frac{1}{r-1}(e(M) - \beta e(K_n)) \geq \frac{1}{r-1}e(K_n)(1 - \beta - ((1 - x/\sqrt{z} + \sqrt{z})^2 + (x/z - 1)z))$. By Theorem 2.2, some component H' of H has $e(H') \geq \frac{e(H)^2}{e(M)} \geq \frac{1}{r-1}e(H)$. Therefore,

$$z \geq \frac{e(H')}{e(K_n)} \geq \frac{1}{(r-1)^2}(1 - \beta - ((1 - x/\sqrt{z} + \sqrt{z})^2 + (x/z - 1)z)).$$

Applying the same sequence of rearrangements and simplifications as in the proof of Theorem 1.4 (which are very similar to those in the proof of Theorem 1.5), with the distinction being that we only have $x \geq \frac{1-\beta}{r}$ in this case, then yields

$$-(r-1)^2 z \leq (1-\beta)^2 \frac{1}{zr^2} - (1 + 2/\sqrt{z})\frac{1-\beta}{r} + 2\sqrt{z} + \beta.$$

Letting $z = \frac{(1-\beta)^2}{(r-\delta)^2}$ and further expanding and simplifying the above then gives

$$0 \leq (r-\delta)^4 + (r-1)^2 r^2 (1-\beta)^2 - (1-\beta)r(r-\delta)^2 - 2(r-\delta)^3 r + 2(1-\beta)(r-\delta)r^2 + \beta(r-\delta)^2 r^2$$
$$= (r-\delta^2)^2 - r(r^2-\delta^2)(1-2\delta) + \beta((r-\delta)^2 r^2 + r(r-\delta)^2 - 2(r-\delta)r^2 - (2-\beta)(r-1)^2 r^2)$$
$$\leq (r-\delta^2)^2 - r(r^2-\delta^2)(1-2\delta) + \beta r^2((r-\delta)^2 - 2(r-1)^2),$$

where the last step uses the fact that $\beta \leq \frac{1}{r-1}$. Since $r \geq 4$, the last term $\beta r^2((r-\delta)^2 - 2(r-1)^2)$ is negative, so we obtain $\delta > \frac{r-1}{2r}$ just like in the (fully colored) complete graph setting. Then we have $z \geq \frac{(1-\beta)^2}{(r-\frac{1}{2}+\frac{1}{2r})^2} \geq \frac{(1-\beta)^2}{r^2-r+\frac{5}{4}}$, as desired.

$\square$

Theorem 4.2. *Let G be a graph on n vertices where every vertex has degree at least $(1-\beta)n - 1$, where $\beta \leq \frac{1}{25}$. Then in every 3-coloring of the edges of G, there is a monochromatic connected component with at least $\frac{1}{6}e(G)$ edges.*

Proof. Call the three colors red, green, and blue. Assume for the sake of contradiction that every monochromatic component has less than $\frac{1}{6}e(G)$ edges.

First suppose that one of the colors (say, red) consists of exactly one component. Then the number of red edges is at most $\frac{1}{6}e(G)$, so there are at least $\frac{5}{6}e(G)$ total green and blue edges. Without loss of generality, there are at least $\frac{5}{12}e(G)$ green edges. By Theorem 2.2, there is a green component H' with

$$e(H') \geq \frac{(\frac{5}{12}e(G))^2}{\binom{n}{2}} \geq \frac{25}{144}(1-\beta)e(G) \geq \frac{1}{6}e(G),$$

a contradiction. So, we can assume there are multiple components of each color. Let red be the color with the most edges, so there are at least $\frac{1}{3}e(G)$ red edges. Let H' be the

red component with the largest average degree, so

$$v(H') \geq \bar{d}(H') + 1 \geq \frac{1}{3}\bar{d}(G) + 1 \geq \frac{1}{3}(1 - \beta)n.$$

If $v(H') > \frac{n}{2}$, then $e(H') = \frac{1}{2}v(H')\bar{d}(H') \geq \frac{1}{6}e(G)$, so we can assume that $v(H') \leq \frac{n}{2}$.
Let $V_1 = V(H')$ and $V_2 = V(G) \setminus V_1$, and let G' be the bipartite subgraph of G consisting of all its edges between V_1 and V_2.

Fix a green component C_1 of G'. If $|C_1 \cap V_1| \geq |V_1| - 3\beta n - 1$, then the edges of G' between $V_1 \cap C_1$ and $V_2 \setminus C_1$ are all blue. So, all the edges between $V_1 \cap C_1$ and V_2 are split into two monochromatic components. One of those components has a number of edges that is at least

$$\frac{1}{2}e_G(V_1 \cap C_1, V_2) \geq \frac{1}{2}|V_1 \cap C_1||V_2| - \beta(n-1)|V_1| \geq \frac{2}{9}(1 - \beta)\binom{n}{2} - \beta\binom{n}{2} \geq \frac{1}{6}e(G),$$

a contradiction. So, no green (or blue) component can cover all but at most $3\beta n$ vertices of V_1 (or, by symmetry, of V_2).

Let $A_1 = C_1 \cap V_1$, $A_2 = C_1 \cap V_2$, $B_1 = V_1 \setminus C_1$, and $B_2 = V_2 \setminus C_1$. By assumption, B_1 and B_2 each have at least $3\beta n + 1$ vertices. Since every vertex in A_1 is connected in blue to all but at most βn vertices of B_2, by pigeonhole every pair of vertices in A_1 shares a neighbor in B_2. Thus, A_1 is contained in a single blue component. The same is true of A_2 by symmetry. Let M_1 (resp. M_2) be the subset of B_1 (resp. B_2) contained in the same blue component as A_2 (resp. A_1). Then M_1 and M_2 each have at least $2\beta n + 1$ vertices, so the bipartite graph between M_1 and M_2 is connected. If any edge between M_1 and M_2 was blue, then there would be a blue component with more than $|V_1| - 3\beta n - 1$ vertices in V_1, a contradiction. So, $M_1 \cup M_2$ is contained in a green component. Let K be the subgraph of G induced on $A_1 \cup A_2 \cup M_1 \cup M_2$, with its coloring induced from the coloring on G. Every edge between A_1 and M_1, or between A_2 and M_2, is red; every edge between A_1 and A_2, or between M_1 and M_2, is green; and every edge between A_1 and M_2, or between A_2 and M_1, is blue.

Let $R = V(G) \setminus V(K)$, so $|R| \leq 2\beta n$. Suppose for the sake of contradiction that there is another green component contained in R. Let v be a vertex of R. Then, all of the edges between v and $V(K)$ are red or blue. Without loss of generality, assume there is a red edge from v to a vertex of A_1. Then, because A_1 is in a different red component from A_2 and M_2, the edges from v to A_2 and M_2 are blue. However, A_2 and M_2 must be in different blue components, a contradiction. By similar logic, R cannot contain any extra red or blue components. So, in G, there are exactly two components of each color, and thus one of the 6 components has at least $\frac{1}{6}e(G)$ edges, as desired. $\square$

5. Acknowledgements

References

[1] D. Bal and L. DeBiasio, Partitioning random graphs into monochromatic components, *Electron. J. Combin.* **24** (1) (2017), Paper No.1.18, 25.

[2] D. Conlon and M. Tyomkyn, Ramsey numbers of trails and circuits, *J. Graph Theory* **102** (1) (2023), 194-196.

[3] D. Conlon, S. Luo, and M. Tyomkyn, Monochromatic Components with Many Edges, *J. Comb.* **15.1** (2024), 59-75.

[4] A. Gyárfás, Partition coverings and blocking sets in hypergraphs, *Communications of the Computer and Automation Institute of the Hungarian Academy of Sciences* **71** (1977), 62.

[5] S. Luo, On Connected Components with Many Edges, *SIAM J. Discr. Math.* **37** (2) (2023), 573-585.