

Monochromatic cycle covers in random graphs

Dániel Korándi ^{*} Frank Mousset [†] Rajko Nenadov [‡] Nemanja Škorić [†]
Benny Sudakov [§]

July 16, 2018

Abstract

A classic result of Erdős, Gyárfás and Pyber states that for every coloring of the edges of K_n with r colors, there is a cover of its vertex set by at most $f(r) = O(r^2 \log r)$ vertex-disjoint monochromatic cycles. In particular, the minimum number of such covering cycles does not depend on the size of K_n but only on the number of colors. We initiate the study of this phenomenon in the case where K_n is replaced by the random graph $\mathcal{G}(n, p)$. Given a fixed integer r and $p = p(n) \geq n^{-1/r+\varepsilon}$, we show that with high probability the random graph $G \sim \mathcal{G}(n, p)$ has the property that for every r -coloring of the edges of G , there is a collection of $f'(r) = O(r^8 \log r)$ monochromatic cycles covering all the vertices of G . Our bound on p is close to optimal in the following sense: if $p \ll (\log n/n)^{1/r}$, then with high probability there are colorings of $G \sim \mathcal{G}(n, p)$ such that the number of monochromatic cycles needed to cover all vertices of G grows with n .

1 Introduction

In this paper, we consider a question of the following type: For a certain class of graphs \mathcal{G} , is it true that the vertex set of every r -edge-colored graph $G \in \mathcal{G}$ can be covered with a number of monochromatic paths or cycles¹ that only depends on the number of colors r ?

The study of such questions dates back to the 1960s, when Gerencsér and Gyárfás [7] observed that every 2-coloring of the edges of the complete graph K_n contains two vertex-disjoint monochromatic paths that together cover all vertices of the graph. Later Gyárfás [9] conjectured that the analogous statement for r colors should also be true, namely, that every r -edge-colored K_n can be covered with r vertex-disjoint monochromatic paths. He made a step towards this conjecture by showing that there is always a cover that uses $O(r^4)$ (not necessarily disjoint) monochromatic paths. The case $r = 3$ was only recently resolved by Pokrovskiy [19] and for $r \geq 4$ the conjecture remains open.

Strengthening Gyárfás' result, Erdős, Gyárfás and Pyber [6] showed that the vertices of every r -colored K_n can be covered by $O(r^2 \log r)$ vertex-disjoint monochromatic *cycles*. It is worth noting that their proof is one of the first applications of the *absorbing method*,

^{*}Institute of Mathematics, EPFL, Lausanne, Switzerland. Email: daniel.korandi@epfl.ch. Research supported in part by SNSF grants 200020-162884 and 200021-175977.

[†]Department of Computer Science, ETH, 8092 Zürich, Switzerland.
Email: {frank.mousset|nskoric}@inf.ethz.ch.

[‡]School of Mathematical Sciences, Monash University, VIC 3800, Australia.
Email: rajko.nenadov@monash.edu.

[§]Department of Mathematics, ETH, 8092 Zürich, Switzerland. Email: benjamin.sudakov@math.ethz.ch.
Research supported in part by SNSF grant 200021-175573.

¹Single edges and vertices count as “degenerate” cycles for the purposes of this paper.

a technique that has turned out to be extremely useful for embedding-type problems. This was subsequently improved to $O(r \log r)$ cycles by Gyárfás, Ruszinkó, Sárközy and Szemerédi [12]. For $r = 2$, Lehel conjectured that, just like for paths, the vertices can be covered by two vertex-disjoint monochromatic cycles of different colors. This was eventually established by Bessy and Thomassé [3]. Some generalizations of these results concerning more complicated graphs other than paths or cycles were obtained in [8, 21]. Similar properties of host graphs other than complete graphs were also studied: complete bipartite graphs are considered in [9, 13, 17], complete graphs with only few edges missing in [11], graphs with large minimum degree in [2, 5, 18] and graphs with small independence number in [20]. For further results and research directions we refer the reader to the recent survey by Gyárfás [10].

In this paper, we consider the same problem in the setting of the binomial random graph model $\mathcal{G}(n, p)$. The study of covering $\mathcal{G}(n, p)$ by monochromatic pieces was initiated by Bal and DeBiasio [1], who showed that if $p = \tilde{\Omega}(n^{-1/3})$ then with high probability (w.h.p), $G \sim \mathcal{G}(n, p)$ has the property that every 2-coloring of the edges of G contains two vertex-disjoint monochromatic trees that cover its vertex set. They proposed a conjecture that already $p \gg (\log n/n)^{1/2}$ suffices, which was recently verified by Kohayakawa, Mota and Schacht [15]. Here we continue this line of research by studying random analogs of the theorems of Gyárfás [9] and Erdős, Gyárfás and Pyber [6]. Our main result reads as follows:

Theorem 1.1. *Fix $\varepsilon > 0$ and an integer $r \geq 2$. If $p = p(n) > n^{-1/r+\varepsilon}$, then $G \sim \mathcal{G}(n, p)$ w.h.p has the property that for every r -coloring of the edges of G , there is a collection of at most $(100r)^8 \log r$ monochromatic cycles covering all the vertices of G .*

Although we believe that it should be possible to choose the cycles so that they are vertex-disjoint, our result does not give this. We remark that the bound on p in the theorem is almost best possible. Indeed, a result of Bal and DeBiasio [1, Theorem 1.7] shows that for $p \ll (\log n/n)^{1/r}$ w.h.p there exists an r -coloring of $G \sim \mathcal{G}(n, p)$ which requires an unbounded number of monochromatic components (and in particular, cycles) to cover all vertices. Their coloring is based on the fact that such a G contains an independent set X of unbounded size with the property that every vertex has at most $r - 1$ neighbors in X . Now one can color all edges outside of X with the color r , and for every $v \in V(G) \setminus X$ color the edges from v to X using each of the colors from $[r - 1]$ at most once. It is not difficult to verify that every monochromatic component can cover at most one vertex of X , and so at least $|X|$ such components are needed to cover the whole graph. With this in mind, it seems likely that $(\log n/n)^{1/r}$ is the correct order of magnitude of the threshold for the property of always having a cover by a bounded number of monochromatic cycles.

We remark that Theorem 1.1 is an example of the more general phenomenon that sufficiently dense – but still very sparse – random graphs $\mathcal{G}(n, p)$ often have global properties that are remarkably similar to those of the (much denser) complete graph K_n . Transferring classic results about complete graphs to the random graph setting is an active line of research with some of the milestones achieved only recently (see the survey by Conlon [4]).

Structure of the paper The paper is organized as follows. In the next section we give the proof of Theorem 1.1, assuming two key lemmas: the first one shows that we can cover all but $O(1/p)$ vertices, while the other one takes care of the remaining vertices. In Section 3 we collect some tools and properties of random graphs that are used in the proof of these lemmas. The two lemmas are then proved in Sections 4 and 5, respectively. In the last section we discuss some open problems and future research directions.

Notation We use the common notation $[r] = \{1, \dots, r\}$ for the first r positive integers. Instead of saying that a set has size r , we sometimes say that it is a r -set. We occasionally

write $A = B_1 \cup \dots \cup B_t$ to mean that B_1, \dots, B_t form a partition of A . For $a, b > 0$, we write $a \pm b$ to denote the interval $[a - b, a + b]$.

If G is a graph, we let $N_G(v)$ denote the neighborhood of v , that is, $N_G(v) = \{w : \{v, w\} \in E(G)\}$. Similarly, if $A \subseteq V(G)$, then $N_G(A) = \bigcup_{v \in A} N_G(v)$ is the neighborhood of A . On the other hand, we denote by $N_G^*(A) = \bigcap_{v \in A} N_G(v)$ the *common* neighborhood of A . For subsets $A, B \subseteq V(G)$ and a vertex $v \in V(G)$, we write $N_G(v, B)$, $N_G(A, B)$ and $N_G^*(A, B)$ for the sets $N_G(v) \cap B$, $N_G(A) \cap B$ and $N_G^*(A) \cap B$, respectively. If $A, B \subseteq V(G)$ are disjoint, we write $e_G(A, B)$ for number of edges with one endpoint in A and another in B . If it is clear which graph G we are talking about, we omit the subscript G in the above notations.

Given a coloring of the edges of G with colors $\{1, \dots, r\}$, we define G_i to be the spanning subgraph of G containing the edges of color i . In this setting, we abbreviate $N_{G_i}(v)$, $N_{G_i}(A, B)$, $N_{G_i}^*(A)$, etc. by $N_i(v)$, $N_i(A, B)$, $N_i^*(A)$, and so on. So for example $N_i(v, B)$ is the neighborhood of v in the set B via edges of color i .

The vertex set of the random graph $\mathcal{G}(n, p)$ is understood to be $[n]$. We say that $G \sim \mathcal{G}(n, p)$ satisfies some property with high probability (w.h.p) if the property holds for G with probability tending to 1 as n tends to infinity.

We routinely omit floor and ceiling signs if they are not essential.

2 Proof of Theorem 1.1

Theorem 1.1 states that if the edges of $G \sim \mathcal{G}(n, p)$ (where $p > n^{-1/r+\varepsilon}$) are colored with r colors, then there are $O(r^8 \log r)$ monochromatic cycles covering all its vertices. We use the following strategy to find such cycles.

Our first step is to partition the vertex set of G into $s = (101r)^4$ disjoint sets whose sizes differ by at most 1:

$$V(G) = W_1 \cup \dots \cup W_s,$$

where each W_i has size at most $n/(100r)^4$. Next, we consider one particular set W_i and try to find $O(r^4 \log r)$ monochromatic cycles in G that cover all the vertices in W_i . Importantly for our proof method, these cycles can (and will) use vertices *outside* of W_i . Since there are $s = O(r^4)$ sets, finding such cycles for each W_i independently results in a cover of all vertices by $O(r^8 \log r)$ cycles (although many vertices might be covered multiple times).

We cover the vertices in the set W_i in two steps: first, we cover all but $O(1/p)$ vertices, and then we cover the remaining ones. Here the quantity $1/p$ comes into play as the “threshold” size of a set X to expand to all other vertices (note that each vertex has roughly np neighbors). Indeed, the proof of the first step relies on the fact that every vertex set of size $\Omega(1/p)$ is adjacent to $\Omega(n)$ other vertices. On the other hand, it is key to our second step that the individual neighborhoods of $O(1/p)$ vertices are almost disjoint. In any case, our arguments for the two steps are entirely different, so it is natural to split the proof accordingly. More precisely, we establish the following two lemmas.

Lemma 2.1. *For every integer $r \geq 2$, there is a constant $K = K(r) > 0$ such that the following holds. Let $W \subseteq [n]$ be a fixed set of at most $n/(100r)^4$ vertices, and let $G \sim \mathcal{G}(n, p)$ where $p = p(n) \gg (\log n/n)^{1/2}$. Then w.h.p for every r -coloring of the edges of G , there is a collection of $3r^2$ vertex-disjoint monochromatic cycles that cover all but at most K/p vertices of W .*

Lemma 2.2. *Let $r \geq 2$ be a fixed integer and let $\varepsilon, K > 0$ be some constants. Let $W \subseteq [n]$ be a fixed set of at most $n/2$ vertices, and let $G \sim \mathcal{G}(n, p)$ where $p = p(n) \geq n^{-1/r+\varepsilon}$. Then w.h.p in every r -coloring of the edges of G , every subset $Q \subseteq W$ of size at most K/p can be covered by $400r^4 \log(4r^2)$ monochromatic cycles.*

Observe that while Lemma 2.1 provides an approximate cover of one fixed set W , Lemma 2.2 applies to all small subsets $Q \subseteq W$ simultaneously. We note that Lemma 2.1

applies to a somewhat larger probability range than we actually need it to be; also, the lemma gives vertex-disjoint cycles, but we do not really use this fact.

Combining the two lemmas, we immediately get that w.h.p each fixed set W_i can be covered by $3r^2 + 400r^2 \log(4r^2)$ monochromatic cycles. By the union bound, this is true for each of the constantly many sets W_1, \dots, W_s , and so we can cover all of $V(G)$ using

$$s(3r^2 + 400r^4 \log(4r^2)) \leq (101r)^4 \cdot 1000r^4 \log r \leq (100r)^8 \log r$$

monochromatic cycles. This completes the proof of Theorem 1.1, although of course we still have to prove Lemmas 2.1 and 2.2.

3 Tools and preliminaries

In the proof of Lemma 2.2 we use the following generalization of the result of Erdős, Gyárfás and Pyber [6] to graphs with a given independence number:

Theorem 3.1 (Sárközy [20]). *If the edges of a graph G with independence number α are colored with r colors then G contains a collection of at most $25(\alpha r)^2 \log(\alpha r)$ vertex-disjoint monochromatic cycles covering the vertex set of G .*

Of course, we cannot apply this directly to $G \sim \mathcal{G}(n, p)$ because $\alpha(G)$ is w.h.p unbounded (unless p is very close to 1). Instead, we will use it to find cycle covers in a certain auxiliary graph. To turn the cycles from the auxiliary graph into real cycles in G , we use a generalization of Hall's criterion, due to Haxell [13], for the existence of saturating matchings in hypergraphs (see Section 5). Given a family \mathcal{E} of subsets of some ground set V , the *vertex cover number* $\tau(\mathcal{E})$ is the smallest size of a set $X \subseteq V$ intersecting every set in \mathcal{E} .

Theorem 3.2 (Haxell [13]). *Let $\{H_i = (V, \mathcal{E}_i)\}_{i \in \mathcal{I}}$ be a family of r -uniform hypergraphs on the same vertex set, for some positive integer r . If $\tau(\bigcup_{i \in \mathcal{I}'} \mathcal{E}_i) > (2r - 1)(|\mathcal{I}'| - 1)$ for every $\mathcal{I}' \subseteq \mathcal{I}$, then there is a family of hyperedges $\{e_i\}_{i \in \mathcal{I}}$ such that $e_i \subseteq \mathcal{E}_i$ and $e_i \cap e_j = \emptyset$ for every $i \neq j \in \mathcal{I}$.*

3.1 A Ramsey-type lemma

Next, let K_k^m denote the complete k -partite graph with parts of size m . The proof of Lemma 2.1 relies on the following auxiliary result:

Lemma 3.3. *Let $m \geq 1$ and $k \geq 2$ be integers, and assume that G is a graph whose complement is K_k^m -free. Then G contains a collection of at most $k - 1$ vertex-disjoint cycles covering all but at most $2k^2m + k^3$ vertices.*

The case $k = 2$ of Lemma 3.3 states that if a graph does not have a large bipartite hole then it contains a large cycle. This was proved with a slightly smaller leftover by Krivelevich and Sudakov [16]. We remark that the number of cycles given by the lemma is best possible if one wants a leftover that can be bounded by a function of m and k . This can be seen by considering the disjoint union of $k - 1$ cliques of size $n/(k - 1)$: the complement of such a graph does not contain a K_k^m for any $m \geq 1$ and yet any collection of $k - 2$ cycles must necessarily leave $n/(k - 1)$ vertices uncovered. Although it does not matter for the present paper, it would be interesting to see how much the size of the leftover in Lemma 3.3 can be reduced.

We start the proof of Lemma 3.3 with the following claim, which is essentially already contained in [16], included here for completeness:

Claim 3.4. *Let G be a graph with at least $m \geq 1$ vertices. Then there is a (possibly empty) path P in G and subsets $D, U \subseteq V(G)$ such that $|D| = m$, $e(D, U) = 0$, and*

$$V(G) = D \cup V(P) \cup U.$$

Proof. Let $G = (V, E)$ where $|V| \geq m$. To prove the claim, we analyze the depth-first-search (DFS) algorithm when run on G . The state at the i -th step of this algorithm can be described by disjoint sets $D_i, U_i \subseteq V$ and a possibly empty path P_i in G . The set D_i contains the *discarded* vertices from which the DFS has already back-tracked. The set U_i represents the set of *unexplored* vertices, that is, of those vertices that have not yet been visited by the DFS algorithm. The path P_i contains the vertices that have been visited, but are not discarded yet. The initial state of the algorithm is $(D_0, P_0, U_0) = (\emptyset, (), V)$: all vertices are unexplored. Given the state (D_i, P_i, U_i) at the i -th step, the next state $(D_{i+1}, P_{i+1}, U_{i+1})$ is obtained using the following rules:

1. (TERMINATE) If $P_i = ()$ and $U_i = \emptyset$:
The algorithm terminates (there is no next state).
2. (RESTART) If $P_i = ()$ and $U_i \neq \emptyset$:
Choose an arbitrary vertex $w \in U_i$ and set $(D_{i+1}, P_{i+1}, U_{i+1}) = (D_i, (w), U_i \setminus \{w\})$.
3. (EXPLORE) If $P_i = (v_1, \dots, v_\ell)$ is non-empty and $U_i \cap N(v_\ell) \neq \emptyset$:
Choose an arbitrary vertex $w \in U_i \cap N(v_\ell)$;
the next state is $(D_i, (v_1, \dots, v_\ell, w), U_i \setminus \{w\})$.
4. (BACK-TRACK) If $P_i = (v_1, \dots, v_\ell)$ is non-empty but $U_i \cap N(v_\ell) = \emptyset$:
The next state is $(D_i \cup \{v_\ell\}, (v_1, \dots, v_{\ell-1}), U_i)$.

Note that at every step of the algorithm, the sets $D_i, V(P_i), U_i$ form a partition of V . It is also easy to see that the algorithm eventually terminates – indeed, the value $|D_i| - |U_i|$ increases by exactly one in each step and is capped at $|V|$. As the initial state is $(D_0, P_0, U_0) = (\emptyset, (), V)$ and the terminal state is $(V, (), \emptyset)$, and because in every round, $|D_i|$ can increase only by at most 1, we see that there is a state (D_i, P_i, U_i) where $|D_i| = m$ ($\leq |V|$).

The required path is then $P = P_i$ and the required sets are $D = D_i$ and $U = U_i$. Note that since vertices are moved to D_i only if they no longer have any neighbors in U_i , we have $e(D_i, U_i) = 0$ for all $i \geq 0$ and in particular $e(D, U) = 0$. \square

It will be convenient to first prove Lemma 3.3 under the additional assumption that the given graph G contains a Hamiltonian path.

Claim 3.5. *Let $m \geq 1$ and $k \geq 2$ be integers, and let G be a graph whose complement is K_k^m -free. If G contains a Hamiltonian path, then G contains a collection of at most $k - 1$ vertex-disjoint cycles covering all but at most km vertices of G .*

Proof. We prove the claim by induction on k . Let $n = |V(G)|$. We may assume $n > km$, as otherwise the statement is trivially satisfied. Let $P = (v_1, \dots, v_n)$ be a Hamiltonian path in G .

In the base case $k = 2$, let $S_1 = \{v_1, \dots, v_m\}$ and $S_2 = \{v_{n-m+1}, \dots, v_n\}$ be the sets containing the first and the last m elements of the path P , respectively. As $n > 2m$, the sets S_1 and S_2 are disjoint. Since the complement of G does not contain a complete bipartite graph K_2^m , there must be an edge between S_1 and S_2 . Together with P , such an edge forms a cycle containing all but at most $2m$ vertices.

Let us now assume that the claim holds for all $k' < k$, for some $k \geq 3$. Consider the set $S_1 = \{v_1, \dots, v_m\}$ consisting of the first m vertices of the path P . Let $m < i \leq n$ be the largest index i such that v_i has a neighbor in S_1 , and let $S = \{v_1, \dots, v_i\}$. Then $G[S]$ clearly contains a cycle that covers all but at most m vertices from S . Moreover, there is no edge between S_1 and $S_2 = V(P) \setminus S = \{v_{i+1}, \dots, v_n\}$, and so, as the complement of G is K_k^m -free, the complement of $G[S_2]$ is K_{k-1}^m -free. By induction $G[S_2]$ contains $k - 2$ vertex-disjoint cycles that cover all but at most $(k - 1)m$ vertices from S_2 . In total, there are $k - 1$ vertex-disjoint cycles that cover all but at most km vertices from $S \cup S_2 = V(G)$. \square

Next, we have the following simple observation:

Claim 3.6. *Let $m \geq 1$ and $k \geq 2$ be integers, and let G be a graph whose complement is K_k^m -free. Let $S_1, \dots, S_k \subseteq V(G)$ be disjoint sets of size at least $m + k - 1$. Then for some $i \neq j$ there are two disjoint edges between S_i and S_j .*

Proof. The fact that the complement of G does not contain a K_k^m means that for every choice of disjoint subsets S'_1, \dots, S'_k of size at least m each, we can find an edge going from S'_i to S'_j for some $i \neq j$. So let us start with S_1, \dots, S_k of size at least $m + k - 1$, find such an edge, and remove its endpoints from the sets S_i and S_j . As long as the remaining sets (with the endpoints removed) still have size at least m , we can repeat this procedure. Eventually some S_i will only have $m - 1$ vertices left. This means that we have removed k disjoint edges touching S_i , each going to some S_j with $i \neq j$. Then by the pigeonhole principle, two of these edges must go to the same S_j . \square

Proof of Lemma 3.3. Fix a graph $G = (V, E)$ on $n = |V|$ vertices and suppose that the complement of G does not contain a copy of K_k^m . We start by proving a slightly weaker statement:

$$G \text{ contains } \binom{k}{2} \text{ vertex-disjoint cycles covering all but at most } k^2m \text{ vertices.} \quad (1)$$

The proof of (1) is by induction on k . We first prove the base case $k = 2$. By Claim 3.4, we can find a path P and sets $D, U \subseteq V$ such that $|D| = m$, $e(D, U) = 0$, and $V = D \cup V(P) \cup U$. Since the complement of G is K_2^m -free, it follows that $|U| < m$, and so $|V(P)| \geq n - |D| - |U| > n - 2m$. Now $G[V(P)]$ is a graph that contains a Hamiltonian path, and whose complement is K_2^m -free. Thus, by Claim 3.5, it contains a cycle covering all but at most $2m$ vertices of $G[V(P)]$, i.e., all but at most $4m$ vertices of G . This completes the proof for $k = 2$.

For the induction step, assume that $k \geq 3$ and that (1) holds for $k - 1$. Using Claim 3.4 we find a path P and sets D, U such that $|D| = m$, $e(D, U) = 0$, and $V = D \cup V(P) \cup U$. As the complement of G is K_k^m -free, and since there are no edges going between D and U , the complement of $G[U]$ is K_{k-1}^m -free. By induction, there is a collection of $\binom{k-1}{2}$ vertex-disjoint cycles in $G[U]$ that covers all but at most $(k-1)^2m$ vertices from U . Moreover, as $G[V(P)]$ is a graph with a Hamiltonian path whose complement is K_k^m -free, we can use Claim 3.5 to obtain $k - 1$ vertex-disjoint cycles in $G[V(P)]$ covering all but at most km vertices of $G[V(P)]$. In total, we have a collection of $k - 1 + \binom{k-1}{2} = \binom{k}{2}$ vertex-disjoint cycles in G covering all but at most $(k-1)^2m + km + m \leq k^2m$ vertices, establishing (1).

All in all, we have collection of $\binom{k}{2}$ vertex-disjoint cycles covering all but at most k^2m vertices. We reduce the size of this collection as follows. If the collection contains a cycle shorter than $m + k - 1$, we remove this cycle from the collection, increasing the number of uncovered vertices by at most $m + k - 1$. Otherwise, if there are at least k cycles C_1, \dots, C_k of length at least $m + k - 1$ left, then we apply Claim 3.6 to a set S_i of $m + k - 1$ consecutive vertices from each cycle C_i . The claim provides two edges that merge two of these cycles into a single one, while increasing the size of the leftover by at most $2(m + k - 1)$. If we repeat this until the collection contains at most $k - 1$ cycles, then we end up with $k - 1$ vertex-disjoint cycles covering all except at most $k^2m + \binom{k}{2}2(m + k - 1) \leq 2k^2m + k^3$ vertices, as required. \square

3.2 Properties of random graphs

In this section we collect some properties of random graphs that are used throughout the proof. The following Chernoff-type bounds on the tails of the binomial distribution are used throughout.

Lemma 3.7 ([14, Theorem 2.1]). *Let $X \sim \text{Bin}(n, p)$ be a binomial random variable. Then*

- $\Pr[X < (1 - a)np] < e^{-a^2 np/2}$ for every $a > 0$, and
- $\Pr[X > (1 + a)np] < e^{-a^2 np/3}$ for every $0 < a < 3/2$.

The next result says that w.h.p a random graph contains approximately the expected number of edges between any two sufficiently large vertex sets.

Lemma 3.8. *Fix $0 < \alpha, \beta < 1$ and let $C = 6/(\alpha^2 \beta)$ and $D = 9/\alpha^2$. Then for every $p = p(n) \in (0, 1)$, the random graph $G \sim \mathcal{G}(n, p)$ satisfies the following property w.h.p: For any two disjoint subsets $X, Y \subseteq V(G)$ satisfying either of*

1. $|X|, |Y| \geq D(\log n)/p$, or
2. $|X| \geq C/p$ and $|Y| \geq \beta n$,

we have

$$e(X, Y) \in (1 \pm \alpha)|X||Y|p.$$

Proof. For fixed sets X and Y , the quantity $e(X, Y)$ follows the binomial distribution $\text{Bin}(|X||Y|, p)$, so we can apply Lemma 3.7 to get

$$\Pr[e(X, Y) \notin (1 \pm \alpha)|X||Y|p] \leq 2e^{-\alpha^2 |X||Y|p/3}.$$

Hence the probability that there exist sets $X, Y \subseteq V(G)$ such that $|X| \geq |Y| \geq D(\log n)/p$ and $e(X, Y) \notin (1 \pm \alpha)|X||Y|p$ is at most

$$\begin{aligned} \sum_{\frac{D \log n}{p} \leq y \leq x \leq n} \binom{n}{x} \binom{n}{y} 2e^{-\alpha^2 xy p/3} &\leq \sum_{\frac{D \log n}{p} \leq y \leq x \leq n} e^{2x \log n} e^{-x \cdot \alpha^2 D(\log n)/3} \\ &\leq \sum_{\frac{D \log n}{p} \leq y \leq x \leq n} e^{-x \log n} \leq n^2 e^{-\log^2 n} = o(1), \end{aligned}$$

using $D = 9/\alpha^2 > 1$. Similarly, the probability that there are subsets $X, Y \subseteq V(G)$ such that $|X| \geq C/p$, $|Y| \geq \beta n$, and $e(X, Y) \notin (1 \pm \alpha)|X||Y|p$ is at most

$$\sum_{x=C/p}^n \sum_{y=\beta n}^n \binom{n}{x} \binom{n}{y} 2e^{-\alpha^2 xy p/3} \leq \sum_{x=C/p}^n \sum_{y=\beta n}^n 2^{2n} e^{-\alpha^2 C \beta n/3} = o(1),$$

for $C \geq 6/\alpha^2 \beta$. □

The following lemma studies how the common neighborhoods of given vertex pairs intersect an arbitrary set.

Lemma 3.9. *For every $p = p(n) \in (0, 1)$, the random graph $G \sim \mathcal{G}(n, p)$ satisfies the following property w.h.p: For every family \mathcal{L} of ℓ disjoint pairs of vertices, and for every set Y of 3ℓ vertices that is disjoint from each pair in \mathcal{L} , we have*

$$\sum_{\{v,w\} \in \mathcal{L}} |N(v, Y) \cap N(w, Y)| \leq \begin{cases} 72\ell \log n & \text{if } \ell \leq 6 \log n/p^2, \\ 2\ell |Y| p^2 & \text{otherwise.} \end{cases}$$

Proof. Fix sets \mathcal{L} and Y as in the statement of the lemma. Since the pairs in \mathcal{L} are pairwise disjoint and disjoint from Y , the sum

$$Q(\mathcal{L}, Y) = \sum_{\{v,w\} \in \mathcal{L}} |N(v, Y) \cap N(w, Y)|$$

has the same distribution as a sum of $\ell|Y| = 3\ell^2$ independent Bernoulli random variables with probability p^2 . In other words, $Q(\mathcal{L}, Y) \sim \text{Bin}(3\ell^2, p^2)$. Therefore, for $\ell \leq 6 \log n/p^2$ we have

$$\begin{aligned} \Pr[Q(\mathcal{L}, Y) > 72\ell \log n] &\leq \binom{3\ell^2}{72\ell \log n} (p^2)^{72\ell \log n} \leq \left(\frac{e \cdot 3\ell^2 p^2}{72\ell \log n}\right)^{72\ell \log n} \\ &\leq (e/4)^{72\ell \log n} \leq e^{-6\ell \log n}, \end{aligned}$$

where the last inequality follows from $(e/4)^{12} \leq 1/e$. Otherwise, if $\ell > 6 \log n/p^2$ then we can apply the Chernoff bound (Lemma 3.7) to get

$$\Pr[Q(\mathcal{L}, Y) > 2\ell|Y|p^2] \leq e^{-\ell|Y|p^2/3} = e^{-\ell^2 p^2} \leq e^{-6\ell \log n}.$$

Taking a union-bound over all choices of \mathcal{L} and Y , we obtain that the probability that for some \mathcal{L} and Y the desired upper bound does not hold is at most

$$\sum_{\ell=1}^n \binom{n^2}{\ell} \binom{n}{3\ell} e^{-6\ell \log n} \leq \sum_{\ell=1}^n n^{2\ell+3\ell} e^{-6\ell \log n} = \sum_{\ell=1}^n n^{-\ell} \rightarrow 0,$$

completing the proof. \square

Our last tool shows that the common neighborhoods of not too many distinct small sets are close to disjoint.

Lemma 3.10. *Let $r \geq 2$ be a fixed integer and let $\tilde{\varepsilon} = \tilde{\varepsilon}(n) \in (0, 1)$. Let $p = p(n) \in (0, 1)$ and let $L \subseteq [n]$ be a fixed set of at least $50r \log n / (\tilde{\varepsilon} p^r)$ vertices. Then $G \sim \mathcal{G}(n, p)$ w.h.p satisfies the following property: For every family of at most $\tilde{\varepsilon}/p$ different sets $X_1, \dots, X_t \subseteq [n] \setminus L$ of size r , we have*

$$\left| \bigcup_{i=1}^t N^*(X_i, L) \right| \in (1 \pm \sqrt{\tilde{\varepsilon}}) t |L| p^r.$$

Proof. Let X_1, \dots, X_t be a fixed family of $t \leq \tilde{\varepsilon}/p$ different r -sets contained in $[n] \setminus L$. We consider the random set $W = \bigcup_{i=1}^t N^*(X_i, L)$. Note that W contains each vertex of L independently with the same probability, so $|W| \sim \text{Bin}(|L|, q)$ for some probability q . We will use the inclusion-exclusion principle (Bonferroni's inequality) to estimate the expectation of $|W|$, and then apply Chernoff bounds to get concentration.

Let $A_i = N^*(X_i, L)$. Then, according to Bonferroni's inequality, we have

$$\sum_{i=1}^t |A_i| - \sum_{1 \leq i < j \leq t} |A_i \cap A_j| \leq \left| \bigcup_{i=1}^t A_i \right| = |W| \leq \sum_{i=1}^t |A_i|.$$

Here $\mathbb{E}[|A_i|] = \mathbb{E}[|N^*(X_i, L)|] = |L| p^r$ and $\mathbb{E}[|A_i \cap A_j|] = \mathbb{E}[|N^*(X_i \cup X_j, L)|] = |L| p^{|X_i \cup X_j|}$. In particular, $\mathbb{E}[|A_i \cap A_j|] \leq |L| p^{r+1}$. Thus, on the one hand we have

$$\mathbb{E}[|W|] \leq \sum_{i=1}^t \mathbb{E}[|A_i|] = t |L| p^r,$$

and on the other hand,

$$\mathbb{E}[|W|] \geq \sum_{i=1}^t \mathbb{E}[|A_i|] - \sum_{1 \leq i < j \leq t} \mathbb{E}[|A_i \cap A_j|] \geq t |L| p^r - \binom{t}{2} |L| p^{r+1} \geq t |L| p^r - t \tilde{\varepsilon} |L| p^r / 2,$$

using $t \leq \tilde{\varepsilon}/p$ in the last inequality. Hence

$$(1 - \tilde{\varepsilon}/2)t|L|p^r \leq \mathbb{E}[|W|] \leq t|L|p^r,$$

so in particular, $q \geq (1 - \tilde{\varepsilon}/2)tp^r \geq tp^r/2$. Then by a Chernoff bound, i.e., Lemma 3.7 with $a = \sqrt{\tilde{\varepsilon}}/2$, we get

$$\Pr[|W| \notin (1 \pm \sqrt{\tilde{\varepsilon}})t|L|p^r] \leq 2e^{-\tilde{\varepsilon}t|L|p^r/24} \leq e^{-2tr \log n},$$

using $|L| \geq 50r \log n/(\tilde{\varepsilon}p^r)$ in the last inequality. Finally, a union-bound over all choices of X_1, \dots, X_t shows that the property in the lemma fails with probability bounded by

$$\sum_{t=1}^{\tilde{\varepsilon}/p} \binom{n}{r}^t e^{-2tr \log n} \leq \sum_{t=1}^{\infty} e^{-tr \log n} \leq \sum_{t=1}^{\infty} n^{-t} \rightarrow 0. \quad \square$$

4 Approximate covering – proof of Lemma 2.1

Let W be a fixed set of at most $n/(100r)^4$ vertices and let $G \sim \mathcal{G}(n, p)$, where $p = p(n) \gg (\log n/n)^{1/2}$. Consider some coloring of the edges of G with r colors. Our goal is to find $3r^2$ vertex-disjoint monochromatic cycles that cover all but at most K/p vertices of W , for some sufficiently large constant $K = K(r)$.

Let $U = V(G) \setminus W$. It is convenient to work with different colors separately, so we partition W into r sets

$$W = W_1 \cup \dots \cup W_r$$

such that $v \in W_i$ if the most commonly used color on edges between v and U is i :

$$v \in W_i \implies |N_i(v, U)| = \max_{j \in [r]} |N_j(v, U)|.$$

Recall here that $N_i(v, U)$ is the set of vertices in U joined to v by an edge of color i .

Next, for every $i \in [r]$, we define an auxiliary graph H_i on W_i where two vertices $v, w \in W_i$ are connected by an edge if and only if

$$|N_i(v, U) \cap N_i(w, U)| \geq \frac{np^2}{(50r)^4}.$$

The main idea is to apply Lemma 3.3 on H_i to find a small collection of “auxiliary cycles” covering most of W_i . Then we will use Hall’s condition to turn each such auxiliary cycle into a cycle in G_i (the subgraph of G of edges in color i) that covers the same vertices in W_i . We thus have two claims:

Claim 4.1. *Let $c = 384r$. Then w.h.p each auxiliary graph H_i contains $3r - 1$ vertex-disjoint cycles covering all but at most $18r^2c/p + (3r)^3$ vertices of H_i .*

Claim 4.2. *The following holds w.h.p: For every choice of $k_i \leq 3r - 1$ vertex-disjoint cycles $C_1^i, \dots, C_{k_i}^i$ in each auxiliary graph H_i , the graph G contains $k_1 + \dots + k_r$ vertex-disjoint monochromatic cycles covering all the sets $V(C_j^i)$.*

From these two claims, Lemma 2.1 follows immediately with $K = r \cdot 18r^2 \cdot 384r + 27r^4 \leq (20r)^4$. It remains to prove the two claims.

Proof of Claim 4.1. In light of Lemma 3.3, it is enough to show that the complement of each H_i is $K_{3r}^{c/p}$ -free. There is an easy intuition as to why this is the case: in G , each vertex of W is adjacent to about $p|U|$ vertices in U , so we expect a subset of size c/p to expand to almost the whole U . As a vertex from W_i has at least $p|U|/r$ neighbors in U in color i , this suggests that a subset of c/p vertices of W_i should have at least $|U|/r$

neighbors in this color i . But then if we take a bit more than r sets of this size, then their i -colored neighborhoods overlap significantly, so there should be two sets with linearly many common neighbors in color i . This readily implies that some two vertices in H_i are joined by an edge. We will now make this argument precise.

First, w.h.p G satisfies the properties of Lemma 3.8 with $\alpha = 1/4$ and $\beta = 1/(4r)$. Let $c = C_{3.8}(\frac{1}{4}, \frac{1}{4r}) = 384r$ be the corresponding constant given by Lemma 3.8. Then Lemma 3.8 states that w.h.p any two disjoint subsets $X, Y \subseteq V(G)$ such that $|X| \geq c/p$ and $|Y| \geq n/(4r)$ satisfy

$$e(X, Y) \leq \frac{5}{4}|X||Y|p. \quad (2)$$

Next, every vertex $v \in W$ has $e(v, U) \sim \text{Bin}(|U|, p)$ neighbors in U , thus Lemma 3.7 shows that the probability of $e(v, U) \notin (1 \pm r^{-2})p|U|$ is at most $2e^{-|U|p/(3r^4)} = e^{-\Omega(np)}$, using $|U| \geq n/2$. As $p > 1/\sqrt{n}$, a union bound over all vertices of W gives that w.h.p $e(v, U) \in (1 \pm r^{-2})p|U|$ for every $v \in W$. By the definition of W_i , it follows that w.h.p we have

$$e_i(v, U) \geq e(v, U)/r \geq p|U|/(r+1) \quad \text{for every } v \in W_i. \quad (3)$$

In the following, fix some $i \in [r]$. We will show that properties (2) and (3) already imply that the complement of H_i is $K_{3r}^{c/p}$ -free. In other words, we show that if $X_1, \dots, X_{3r} \subseteq W_i$ are disjoint sets of size c/p , then there exist $j \neq j'$ such that $e_{H_i}(X_j, X_{j'}) > 0$.

Fix any such sets, and let $Y_j = N_i(X_j, U)$ denote the set of vertices in U that have a neighbor in X_j in color i . We first show that

$$|Y_j| > |U|/(2r) \quad \text{for every } j \in [3r]. \quad (4)$$

First, by (3) we have

$$e_i(X_j, Y_j) = e_i(X_j, U) \geq |X_j||U|p/(r+1).$$

Suppose that $|Y_j| \leq |U|/(2r)$ and choose an arbitrary set $Y_j \subseteq Y_{j'} \subseteq U$ of size $|U|/(2r) \geq n/(4r)$. Then by (2), we have

$$e_i(X_j, Y_j) \leq e(X_j, Y_{j'}) \leq \frac{5}{4}|X_j||Y_{j'}|p = \frac{5}{8r}|X_j||U|p < |X_j||U|p/(r+1),$$

which is a contradiction. This establishes (4).

Now we can bound how much these colored neighborhoods intersect using Bonferroni's inequality:

$$\sum_{1 \leq j < j' \leq 3r} |Y_j \cap Y_{j'}| \geq \left(\sum_{j=1}^{3r} |Y_j| \right) - |U| \geq |U|/2,$$

using (4) for the last inequality. In particular, we must have $|Y_j \cap Y_{j'}| \geq |U|/(3r)^2$ for some $j \neq j'$. Note that $Y_j \cap Y_{j'}$ can also be written as the union of the common neighborhoods $N_i(v, U) \cap N_i(w, U)$ over all pairs of vertices $v \in X_j$ and $w \in X_{j'}$. But then for some choice of v and w , the size of this is at least the average:

$$|N_i(v, U) \cap N_i(w, U)| \geq \frac{|U|}{(3r)^2} \cdot \frac{1}{|X_j||X_{j'}|} = \frac{|U|p^2}{(3r)^2 c^2} \geq \frac{np^2}{(50r)^4}$$

using $|U| \geq n/2$ and $c = 384r$. Thus v and w are connected by an edge in H_i and so $e_{H_i}(X_j, X_{j'}) > 0$, which is what we set out to prove. \square

Proof of Claim 4.2. First note that w.h.p G satisfies the conclusion of Lemma 3.9, which says that for every family \mathcal{L} of ℓ disjoint pairs of vertices, and for every set Y of 3ℓ vertices that is disjoint from each pair in \mathcal{L} , we have

$$\sum_{\{v,w\} \in \mathcal{L}} |N(v, Y) \cap N(w, Y)| \leq \begin{cases} 72\ell \log n, & \text{if } \ell \leq 6 \log n/p^2 \\ 2\ell|Y|p^2, & \text{otherwise.} \end{cases} \quad (5)$$

This property will be enough to deduce the claim.

Let $\mathcal{E}_i = \bigcup_{j \in [k_i]} E(C_j^i)$ be the edge set of the given cycles in the auxiliary graph H_i , and let $\mathcal{E} = \bigcup_{i \in [r]} \mathcal{E}_i$. We define an auxiliary bipartite graph B with parts U and \mathcal{E} where an edge $vw \in \mathcal{E}_i$ is joined to a vertex $u \in U$ if and only if $u \in N_i(v, U) \cap N_i(w, U)$. We will use Hall's condition to show B contains a matching covering \mathcal{E} . In other words, we will show that there exists an injection $f: \mathcal{E} \rightarrow U$ such that for every $vw \in \mathcal{E}_i$, the vertex $f(vw) \in U$ is a common neighbor of both v and w in color i . This will immediately imply the statement of Claim 4.2, because we can then convert every cycle C_j^i in H_i into a monochromatic cycle in G by replacing each edge vw of C_j^i by an i -colored path $(v, f(vw), w)$. The injectivity of f ensures that the cycles we get are vertex-disjoint.

To verify Hall's condition, we need to show that for every subset $\mathcal{L} \subseteq \mathcal{E}$ we have $|N_B(\mathcal{L})| \geq |\mathcal{L}|$. We instead prove the somewhat different statement that for every subset $\mathcal{L} \subseteq \mathcal{E}$ consisting of *pairwise disjoint* edges in \mathcal{E} we have $|N_B(\mathcal{L})| \geq 3|\mathcal{L}|$. This second statement actually implies the first: as \mathcal{E} is a disjoint union of cycles, every set of edges $\mathcal{L}' \subseteq \mathcal{E}$ contains a subset $\mathcal{L} \subseteq \mathcal{L}'$ of size at least $|\mathcal{L}'|/3$ such that the edges in \mathcal{L} are vertex-disjoint, and so

$$|N_B(\mathcal{L}')| \geq |N_B(\mathcal{L})| \geq 3|\mathcal{L}| \geq |\mathcal{L}'|.$$

So take any set of pairwise disjoint edges $\mathcal{L} \subseteq \mathcal{E}$ and suppose for contradiction that $N_B(\mathcal{L})$ is properly contained in some set $Y \subseteq U$ of size $3|\mathcal{L}|$. Recall that every vertex pair $vw \in \mathcal{L}$ is also an auxiliary edge in some H_i , so v and w have at least $np^2/(50r)^4$ common neighbors in U in color i . By definition, these common neighbors are also neighbors of the pair vw in B , so they are contained in Y . Hence

$$\frac{np^2|\mathcal{L}|}{(50r)^4} \leq \sum_{vw \in \mathcal{L}} |N(v, Y) \cap N(w, Y)|.$$

Note that since $|Y| = 3|\mathcal{L}|$, we can apply (5) to the sum on the right-hand side. There are now two cases, depending on the size of \mathcal{L} .

If $|\mathcal{L}| \leq 6 \log n/p^2$, then by (5)

$$\frac{np^2|\mathcal{L}|}{(50r)^4} \leq \sum_{vw \in \mathcal{L}} |N(v, Y) \cap N(w, Y)| \leq 72|\mathcal{L}| \log n,$$

contradicting our assumption that $np^2 \gg \log n$.

If $|\mathcal{L}| > 6 \log n/p^2$, then by (5)

$$\frac{np^2|\mathcal{L}|}{(50r)^4} \leq \sum_{vw \in \mathcal{L}} |N(v, Y) \cap N(w, Y)| \leq 2|\mathcal{L}||Y|p^2,$$

and thus $|Y| \geq n/(2 \cdot (50r)^4) > 3n/(100r)^4 \geq 3|W| \geq 3|\mathcal{L}|$, which contradicts the assumption that $|Y| = 3|\mathcal{L}|$. This concludes the proof of the claim. \square

5 Covering a set of size $O(1/p)$ – proof of Lemma 2.2

Let $r, \varepsilon, K > 0$ be constants, where $r \geq 2$ is an integer. Consider a subset $W \subseteq [n]$ of size at most $n/2$ and let $G \sim \mathcal{G}(n, p)$ for $p = p(n) > n^{-1/r+\varepsilon}$. We need to show that w.h.p every subset $Q \subseteq W$ of size at most K/p can be covered using at most $400r^4 \log(4r^2)$ monochromatic cycles.

5.1 Proof overview

The main idea of the proof is the following: We define an edge-colored auxiliary graph H on the vertex set W , where two vertices $v, w \in W$ are joined by an edge of color i if they are “robustly connected” by monochromatic paths of color i , whose interior vertices belong to $U = V(G) \setminus W$. This auxiliary graph should have two properties. First, we want the independence number of H to be bounded by a function of r , as this will imply, via the result of Sárközy (Theorem 3.1), that every induced subgraph $H[Q]$ can be covered by a number of monochromatic cycles in $H[Q]$ that depends only on r . Second, we want the notion of “robustly connected” to be sufficiently strong to allow us to convert such a cover of $H[Q]$ by auxiliary monochromatic cycles into a cover of Q by monochromatic cycles in G , at least if $|Q| \leq K/p$.

It is instructive to note that this task would be significantly easier for $p \gg n^{-1/(r+1)}$. In this case we could define H by saying that $v, w \in W$ are joined by an edge of color i if there are many (i.e., $\Omega(np^{r+1})$) vertices in the i -colored common neighborhood $N_i(v, U) \cap N_i(w, U)$. It is not hard to see that if $p \gg (\log n/n)^{1/(r+1)}$, then this graph has independence number at most r . Indeed, the high density implies that every set $X \subseteq W$ of $r+1$ vertices has a large common neighborhood $N^*(X, U)$ in U (of size $\approx np^{r+1}$). Of course, for every vertex in $N^*(X, U)$, at least two of the edges coming from X must have the same color (by the pigeonhole principle). This in turn implies that some two vertices in X will have a large common neighborhood in the same color, so $H[X]$ contains an edge. We could then apply Theorem 3.1 to cover $H[Q]$ with few disjoint monochromatic cycles, and, as in the proof of Lemma 2.1, use Hall’s condition to turn each of these auxiliary cycles into a monochromatic cycle in G by replacing each auxiliary edge with a path of length 2 in G in the same color.

Unfortunately, when p is smaller than $n^{-1/(r+1)}$ a typical set of $r+1$ vertices does not have any common neighbor, and the graph H as defined in the previous paragraph might have unbounded independence number. We overcome this issue by using slightly longer paths to create monochromatic cycles in G . So here the edges of H will correspond to short monochromatic paths whose lengths are possibly greater than 2. We now describe informally how we are going to do this, assuming for simplicity that $r = 2$, i.e., that there are only two colors, called red and blue.

To define H , consider an arbitrary set $\hat{X} \subseteq W$ consisting of 3 vertices u, v, w (see Figure 1). Since we assume $p > n^{-1/r+\varepsilon} = n^{-1/2+\varepsilon}$, any two vertices in \hat{X} will have $\Theta(np^2)$ common neighbors in U . Let Z_1 be the set of common neighbors of u and v in U (and keep the vertex w for later). If there are $\Omega(np^2)$ vertices $x \in Z_1$ that have an edge to both u and v in the same color, then we add an edge in that color in H between u and v . However, it could be that most vertices in Z_1 are connected to u and v by edges of different colors (i.e., there is a red edge to u and a blue edge to v , or vice-versa). In this case, we can find many vertices in Z_1 which are of the “majority color profile”, that is, we can find a set $S_1 \subseteq Z_1$ of $\Omega(np^2)$ common neighbors of u and v such that either every vertex $x \in S_1$ has a red edge to u and a blue edge to v , or every vertex $x \in S_1$ has a blue edge to u and a red edge to v . In any case, we can relabel $\hat{X} = \{v_{\text{free}}, v_{\text{red}}, v_{\text{blue}}\}$ such that every vertex in S_1 is connected by a red edge to v_{red} and by a blue edge to v_{blue} (and v_{free} will just be the vertex w).

We now define the set Z_2 of all vertices in U that have an edge to both v_{free} and a vertex in S_1 . The properties of $\mathcal{G}(n, p)$ will make sure that this set is about as large as expected: $|Z_2| = \Omega(n^2 p^4)$. If there are $\Omega(n^2 p^4)$ vertices in Z_2 that have an edge to both v_{free} and a vertex of S_1 in the same color – say, both are blue – then we add a blue edge to H between v_{free} and v_{blue} . And of course, if both edges are red, then we add a red edge between v_{free} and v_{red} . This way an edge of color i corresponds to many i -colored paths of length 3 between the two involved vertices. However, as before, it could be that most vertices in Z_2 are connected to v_{free} and S_1 in both colors. Then, again, there must be a

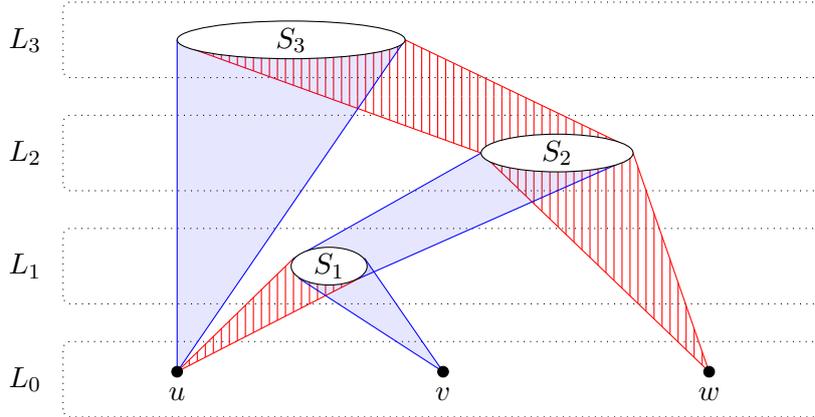


Figure 1: Building towers when $r = 2$.

majority color profile of such vertices, and we can find a subset $S_2 \subseteq Z_2$ of size $\Omega(n^2 p^4)$ such that every vertex in S_2 has either a blue edge to v_{free} and a red edge to a vertex of S_1 , or the other way around. The important observation is that it is again possible to relabel the vertices $\hat{X} = \{v_{\text{free}}, v_{\text{red}}, v_{\text{blue}}\}$ such that every vertex in S_2 is now connected by a blue path to v_{blue} and by a red path to v_{red} (for example, if every vertex in S_2 has a blue edge to S_1 and a red edge to v_{free} , we exchange the identities of v_{red} and v_{free}).

If we continue like this, the following pattern emerges: starting from a set \hat{X} of three vertices, either we are able to place an edge in H between two vertices in \hat{X} , or we get a sequence S_1, S_2, S_3, \dots of larger and larger sets ($|S_i|$ will be around $(np^2)^i$) such that every vertex in S_i is connected by a red path to some $v_{\text{red}} \in \hat{X}$ and by a blue path to some $v_{\text{blue}} \in \hat{X}$. This statement is formalized in Claim 5.5 below.

Now take any set X of 6 vertices, and split it into two sets \hat{X} and \hat{X}' of three vertices each. If H contains an edge inside \hat{X} or \hat{X}' , then X is not independent. Otherwise, since $p > n^{-1/2+\varepsilon}$, we see that after about $m = 1/\varepsilon$ iterations of the above procedure, we reach a set S_m from \hat{X} and a set S'_m from \hat{X}' , both of size much larger than $(\log n)/p$, such that every vertex in the set S_m (S'_m) is connected in both colors to the set \hat{X} (\hat{X}'). In $\mathcal{G}(n, p)$ there are many edges between any two sets of size larger than $(\log n)/p$ (see Lemma 3.8), in particular, we can find many (say) red edges between S_m and S'_m . But then there is a vertex in \hat{X} that is connected to a vertex in \hat{X}' by many red paths, and we can add a red edge connecting these vertices to H . This gives us an H that has independence number at most 5 such that an i -colored edge in H corresponds to many i -colored paths between the two vertices in G .

The same general approach works if there are more than two colors. In the rest of this section, we explain the details of the above outline to get a real proof, and show how to turn the auxiliary cycles in $H[Q]$ into monochromatic cycles in G .

5.2 Towers, cascades, and the auxiliary graph

Without loss of generality, we may assume that $p = n^{-1/r+\varepsilon}$ where $\varepsilon = 1/qr$ for some large integer q such that r does not divide $q - 1$. (So we have $p = n^{-(q-1)/qr} = n^{-(q-1)\varepsilon}$.)

Let us fix a partition of $U = V(G) \setminus W$ into $1/\varepsilon$ levels of the same size:

$$U = L_1 \cup \dots \cup L_{1/\varepsilon},$$

where $|L_k| = \varepsilon|U|$. For notational convenience, we also define $L_0 = W$. As this partition

does not depend on the choice of $Q \subseteq W$, we may fix it before exposing G . Set

$$\mu = \frac{|L_k|p^r}{2r^r} = \Theta(n^{r\varepsilon}), \quad (6)$$

and let $m \in \mathbb{N}$ be such that $m - 1 < \frac{q-1}{r} < m$. Note that this implies $m \leq 1/\varepsilon$, and

$$\mu^{m-1} = O(n^{-\varepsilon}/p) \quad \text{and} \quad \mu^m = \Omega(n^\varepsilon/p). \quad (7)$$

Indeed, $\mu^{m-1} \leq \mu^{(q-2)/r} = O(n^{(q-1)\varepsilon-\varepsilon})$ and $\mu^m \geq \mu^{q/r} = \Omega(n^{(q-1)\varepsilon+\varepsilon})$.

Next, we describe the structures needed for the proof. The point of the argument we sketched in Section 5.1 was to find monochromatic paths from a vertex v through a sequence of sets (S_s, \dots, S_f) (see Figure 1). We will call such a monochromatic piece a *tower* on v . Due to technical reasons, the formal definition below needs to include some auxiliary sets, as well.

Definition 5.1 (Towers). Let $1 \leq s \leq f \leq m$, let $i \in [r]$ be a color and $v \in L_0$. We call a sequence of sets $(S_{s-1}, S_s, \dots, S_f)$ an *i -tower on v* if there is a sequence of sets $T_s, \dots, T_f \subseteq L_0$ of size $r-1$ such that the following properties hold:

- (T1) $S_k \subseteq L_k$ and $|S_k| = \mu^k$ for all $k \in \{s-1, \dots, f\}$,
- (T2) $S_s \subseteq N_i(v) \cap N(S_{s-1}) \cap N^*(T_s)$,
- (T3) $S_k \subseteq N_i(S_{k-1}) \cap N^*(T_k)$ for all $k \in \{s+1, \dots, f\}$,
- (T4) $v \in T_s$ if $s > 1$, and $S_0 = \{v\}$ with $v \notin T_1$ if $s = 1$.

We say that (T_s, \dots, T_f) is a *witness sequence* for the tower (S_{s-1}, \dots, S_f) .

Note that in the case $r = 2$ that we discussed in Section 5.1, the v_{free} of step i will be the witness T_{i+1} that ‘‘generates’’ S_{i+1} from S_i . For example, in Figure 1, (u, S_1) is a red tower on u with witness sequence (v) , (v, S_1, S_2) is a blue tower on v with witness sequence (u, w) , (S_1, S_2, S_3) is a red tower on w with witness sequence (w, u) , and (S_2, S_3) is a blue tower on u with witness sequence (u) .

In general, if $(S_{s-1}, S_s, \dots, S_f)$ is an i -tower on v , then it follows from (T2) and (T3) that every vertex in S_f is reachable from v by a path in color i passing through each S_s, \dots, S_{f-1} exactly once. Property (T4), the set S_{s-1} and the witness sequence T_s, \dots, T_f are needed to establish some pseudorandom properties of the sets S_s, \dots, S_f , such as expansion.

To define our auxiliary graph, we need one additional structure, a *cascade* that we define on a pair of vertices in L_0 . As we will see later, it guarantees that the two vertices are connected by monochromatic paths in a very robust way (Claim 5.9).

Definition 5.2 (Cascades). Let $i \in [r]$ be a color. We say that two vertices $v, w \in L_0$ are *connected by an i -cascade* if there is an i -tower $(S_{s_v-1}^v, S_{s_v}^v, \dots, S_f^v)$ on v and an i -tower $(S_{s_w-1}^w, S_{s_w}^w, \dots, S_f^w)$ on w for some $1 \leq s_v, s_w \leq f \leq m$, such that either

- (C1) $S_f^v = S_f^w$, or
- (C2) $f = m$ and $e_i(R_v, R_w) \geq e(R_v, R_w)/r$, where $R_v = S_m^v \setminus S_m^w$ and $R_w = S_m^w \setminus S_m^v$.

Note that there is no need to impose any conditions on the disjointness of the sets in (C2). This is because if R_v and R_w are small then S_m^v and S_m^w have a significant overlap, so the situation is similar to (C1).

We now define the auxiliary graph H on the vertex set L_0 by adding an edge of color i between two vertices $v, w \in L_0$ if v and w are connected by an i -cascade. There are two central claims:

Claim 5.3. *W.h.p, H has independence number at most $4r$.*

Claim 5.4. *W.h.p, for every monochromatic cycle C in H of length at most K/p , there is a monochromatic cycle C^* (of the same color) in G such that $V(C) \subseteq V(C^*)$.*

From these two claims, Lemma 2.2 follows easily:

Proof of Lemma 2.2. Suppose that $G \sim \mathcal{G}(n, p)$ is such that the two properties in Claims 5.3 and 5.4 hold (this happens w.h.p), so H has independence number at most $4r$. Let $Q \subseteq L_0$ be an arbitrary set of size at most K/p . As the independence number of $H[Q]$ is also bounded by $4r$, we can apply Theorem 3.1 to find a collection of at most $25(\alpha r)^2 \log(\alpha r) \leq 400r^4 \log(4r^2)$ vertex-disjoint monochromatic cycles in $H[Q]$ covering Q .

We know that every cycle in the collection has length at most $|Q| \leq K/p$, so we can replace each such cycle C by a monochromatic cycle C^* in G that covers the vertex set of C (using the property in Claim 5.4). This gives a collection of at most $400r^4 \log(4r^2)$ monochromatic cycles in G covering the whole set Q . \square

5.3 Proof of Claim 5.3

We first prove an auxiliary claim:

Claim 5.5. *W.h.p, for every set $\hat{X} = \{x_1, \dots, x_{2r-1}\}$ of $2r - 1$ distinct vertices in L_0 , one of the following two statements holds:*

- (i) *there are two vertices $v, v' \in \hat{X}$ and a color $i \in [r]$ such that v and v' are connected by an i -cascade, or*
- (ii) *there exists a subset $X = \{v_1, \dots, v_r\} \subseteq \hat{X}$ and for each $i \in [r]$ an i -tower $(S_{s_{i-1}}^i, S_{s_i}^i, \dots, S_m^i)$ on v_i , such that $S_m^1 = \dots = S_m^r$.*

Proof. First, note that w.h.p the property in Lemma 3.10 holds with $\tilde{\varepsilon} = 1/4$ simultaneously for every set $L = L_k$ where $k \in [m]$. This is because $|L_k| = \Theta(n) \gg \log n / (\tilde{\varepsilon} p^r)$ and m is a constant. Thus, we may assume the following property: for every $k \in [m]$ and every list of $t \leq \tilde{\varepsilon}/p$ distinct r -sets $X_1, \dots, X_t \subseteq [n] \setminus L_k$, we have

$$\left| \bigcup_{i=1}^t N^*(X_i, L_k) \right| \geq t|L_k|p^r/2. \quad (8)$$

This property will be enough to imply Claim 5.5.

To prove the claim, assume $\hat{X} = \{x_1, \dots, x_{2r-1}\}$ does not satisfy (i), that is, no two vertices in \hat{X} are connected by an i -cascade for some $i \in [r]$. We show that for every $k \in [m]$

$$\text{there is a set } X_k = \{v_1, \dots, v_r\} \subseteq \hat{X} \text{ and an } i\text{-tower } (S_{s_{i-1}}^i, S_{s_i}^i, \dots, S_k^i) \text{ on each } v_i \text{ such that } S_k^1 = \dots = S_k^r. \quad (9)$$

Note that $X = X_m$ then satisfies (ii).

We prove (9) by induction on k . For the base case $k = 1$, let $X_1 = \{x_1, \dots, x_r\} \subseteq \hat{X}$. Let $Z_1 = N^*(X_1, L_1)$ be the common neighborhood of X_1 in L_1 . By (8) applied to the single r -set X_1 , we have

$$|Z_1| = |N^*(X_1, L_1)| \geq |L_1|p^r/2 = \mu r^r.$$

Let us say that a vertex $z \in Z_1$ has the *color pattern* (i_1, \dots, i_r) if for each $a \in [r]$, the edge zx_a has color i_a . There are r^r different color patterns, so we can find a subset $S_1 \subseteq Z_1$ of size $|S_1| = \mu$ such that all vertices in S_1 have the same color pattern (i_1, \dots, i_r) .

Let $S_0^a = \{x_a\}$ for every $a \in [r]$, and $S_1^a = S_1$. We now claim that (S_0^a, S_1) is an i_a -tower on x_a , for every $a \in [r]$. For this, choose $T_1^a = X_1 \setminus \{x_a\}$, and let us check the

conditions (T1)–(T4) separately. First, (T1) requires that $|S_0^a| = 1$, $|S_1^a| = \mu$, $S_0^a \subseteq L_0$ and $S_1 \subseteq L_1$, all of which are true. For (T2) note that $S_1^a \subseteq N_{i_a}(x_a) \cap N(S_0^a) \cap N^*(T_1^a)$ holds because every vertex in S_1^a lies in the common neighborhood of $X_1 = T_1 \cup S_0^a$ and has an edge of color i_a to x_a . The condition (T3) is vacuous in this case. Finally, for (T4), note that $S_0^a = \{x_a\}$ and $x_a \notin T_1^a$ hold by definition.

If $i_a = i_b$ for some distinct $a, b \in [r]$, then x_a and x_b are connected by an i_a -cascade (condition (C1)), contrary to our assumption. So all r different colors appear in the color pattern (i_1, \dots, i_r) , and by relabeling the vertices in $X_1 = \{x_1, \dots, x_r\}$ as $X_1 = \{v_1, \dots, v_r\}$ so that $v_{i_a} = x_a$, we get the required i -tower on each v_i , proving (9) for $k = 1$.

Now suppose that (9) holds for some $k - 1 \geq 1$, and let us prove that it holds for k , as well. By the induction hypothesis, there is a set of vertices $X_{k-1} = \{v_1, \dots, v_r\} \subseteq \hat{X}$ and an i -tower $(S_{s_{i-1}}^i, S_{s_i}^i, \dots, S_{s_{k-1}}^i)$ on each v_i such that $S_{k-1}^1 = \dots = S_{k-1}^r =: S_{k-1}$, where by (T1) we have $|S_{k-1}^i| = \mu^{k-1}$. Let $X' = \hat{X} \setminus X_{k-1} = \{w_1, \dots, w_{r-1}\}$ be the remaining $r - 1$ vertices (recall that \hat{X} has size $2r - 1$).

Next, we define $Z_k \subseteq L_k$ as the set of all common neighbors of the sets $X' \cup \{v\}$ for $v \in S_{k-1}$, i.e.,

$$Z_k = \bigcup_{v \in S_{k-1}} N^*(X' \cup \{v\}, L_k).$$

Here we have $\mu^{k-1} \leq \mu^{m-1} \leq \varepsilon/p$ (see (7)) distinct r -sets, so (8) gives

$$|Z_k| = \left| \bigcup_{v \in S_{k-1}} N^*(X' \cup \{v\}, L_k) \right| \geq \mu^{k-1} |L_k| p^r / 2 = \mu^k r^r.$$

Each vertex $z \in Z_k$ is connected to w_j for every $j \in [r - 1]$ and has a neighbor in S_{k-1} . Fix one such neighbor $w'_z \in S_{k-1}$ (chosen arbitrarily) and define the color pattern of z to be (i_1, \dots, i_r) if the edge zw_j has color i_j for every $j \in [r - 1]$ and zw'_z has color i_r . Once again, there are r^r such patterns, so there is a subset $S_k \subseteq Z_k$ of $|S_k| = \mu^k$ vertices that all have the same pattern (i_1, \dots, i_r) .

We first note that for every $j \in [r - 1]$, the sequence (S_{k-1}, S_k) is an i_j -tower on w_j with the witness sequence (T_k) , where $T_k = X'$. Verifying the conditions (T1)–(T3) is the same as in the $k = 1$ case, and (T4) again holds by definition: $w_j \in T_k$. Similarly, $(S_{s_{i_r-1}}^{i_r}, S_{s_{i_r}}^{i_r}, \dots, S_{s_{k-1}}^{i_r}, S_k)$ is an i_r -tower on the vertex $v_{i_r} \in X_{k-1}$. For this, we take a witness sequence $T_{s_{i_r-1}}^{i_r}, \dots, T_{s_{k-1}}^{i_r}$ of the tower $(S_{s_{i_r-1}}^{i_r}, S_{s_{i_r}}^{i_r}, \dots, S_{s_{k-1}}^{i_r})$ and extend it by setting $T_k^{i_r} = X'$; the conditions (T1)–(T4) are then easy to check using the fact that $(S_{s_{i_r-1}}^{i_r}, S_{s_{i_r}}^{i_r}, \dots, S_{s_{k-1}}^{i_r})$ is already an i_r -tower on v_{i_r} : for (T1), we already have $|S_k| = \mu^k$, whereas for (T3), we know that every $z \in S_k$ is in the common neighborhood of X' , and is also connected to S_{k-1} by an i_r -colored edge, so $S_k \subseteq N_{i_r}(S_{k-1}^{i_r}) \cap N^*(T_k^{i_r})$. Every other requirement holds by induction.

If $i_a = i_b$ for some distinct $a, b \in [r]$, then two vertices in $X_k = \{w_1, \dots, w_{r-1}, v_{i_r}\} \subseteq \hat{X}$ are connected by an i_a -cascade (again by condition (C1)), which we had ruled out. So all r colors appear in (i_1, \dots, i_r) exactly once, and thus we can relabel the vertices in X_k as v_1, \dots, v_r to get the desired i -towers on the (new) v_i 's. \square

Proof of Claim 5.3. Recall that H is the graph on vertex set L_0 where we add an edge vw in color i whenever v and w are connected by an i -cascade in G .

Assume that G satisfies the property in Claim 5.5, as it does w.h.p, and let $X \subseteq L_0$ be a set of $4r - 2$ vertices. We show that X contains two distinct vertices connected by an i -cascade, for some $i \in [r]$. For this, split X into two disjoint sets \hat{X}, \hat{X}' of $2r - 1$ vertices each. If either set contains two vertices connected by an i -cascade for some $i \in [r]$, then we are done. Otherwise, by Claim 5.5, there are subsets $\{v_1, \dots, v_r\} \subseteq \hat{X}$ and $\{v'_1, \dots, v'_r\} \subseteq \hat{X}'$ and an i -tower $(S_{s_{i-1}}^i, S_{s_i}^i, \dots, S_m^i)$ on v_i and another i -tower $(S_{s'_i-1}^i, S_{s'_i}^i, \dots, S_m^i)$ on v'_i for every color $i \in [r]$, such that $S_m^1 = \dots = S_m^r =: S_m$ and $S_m^1 = \dots = S_m^r =: S_m'$. Let

$i \in [r]$ be a color such that $e_i(R, R') \geq e(R, R')/r$, where $R = S_m \setminus S'_m$ and $R' = S'_m \setminus S_m$. Then $v = v_i$ and $v' = v'_i$ are connected by an i -cascade, and we are again done. \square

5.4 Proof of Claim 5.4

Let us recall the statement: given, say, a red cycle C of size $O(1/p)$ in the auxiliary graph H , we want to find a red cycle C^* in G that covers all the vertices of C . We use the following strategy.

For each edge e in C , we have a cascade on top of it, so we know that there are many short (of length at most $2m + 1$) red paths connecting the endpoints of e , all with internal vertices in U . Let us denote the set of internal vertices of these paths by $\mathcal{H}_e = \{P_e^1, \dots, P_e^\ell\}$. In order to create a cycle C^* , it suffices to choose exactly one $P_e \in \mathcal{H}_e$ for each e so that they are all pairwise disjoint. We use Theorem 3.2 to achieve this. Taking the \mathcal{H}_e as the $(2m)$ -uniform hypergraphs, it is enough to check that for any subset $E' \subset E(C)$ and any $Y \subseteq U$ of size $|Y| \leq 4m|E'|$ there is a path $P_e \in \mathcal{H}_e$, for some $e \in E'$ that completely avoids Y .

The proof relies on two ingredients. First, we will show that most cascades (or, rather, the associated towers) corresponding to the edges in E' are disjoint on almost all levels (Claim 5.8), thus it is impossible for Y to significantly intersect each of them. Second, if we remove a small fraction of the vertices from each level of a cascade connecting v and w , then it still contains a path from v to w (Claim 5.9).

The following notion of *independent towers* will be crucial for our applications of Lemma 3.10.

Definition 5.6 (Independent towers). Suppose that $1 \leq s \leq f \leq m$ and that we are given a collection $\{(S_{s-1}^j, S_s^j, \dots, S_f^j)\}_{j \in [t]}$ of t towers². We say that this collection is *independent* if there exists a witness sequence (T_s^j, \dots, T_f^j) for every tower $(S_{s-1}^j, \dots, S_f^j)$ such that all sets of the form $T_s^j \cup \{v\}$, where $j \in [t]$ and $v \in S_{s-1}^j$, are *distinct*.

Our next claim says that among towers on distinct vertices, there is always a large subset of independent towers.

Claim 5.7. *Let $1 \leq s \leq f \leq m$ and let $t \geq 0$. Let v_1, \dots, v_t be distinct vertices in L_0 and for each $j \in [t]$, let $\mathcal{T}_j = (S_{s-1}^j, \dots, S_f^j)$ be a tower on v_j . Then there is a set $\mathcal{I} \subseteq [t]$ of at least t/r indices such that the towers $\{\mathcal{T}_j\}_{j \in \mathcal{I}}$ are independent.*

Proof. We define a graph $G_{\mathcal{T}}$ on the vertex set $\{\mathcal{T}_1, \dots, \mathcal{T}_t\}$ where \mathcal{T}_j and $\mathcal{T}_{j'}$ for $j \neq j'$ are connected by an edge if and only if $T_s^j \cup \{v\} = T_s^{j'} \cup \{v'\}$ for some $v \in S_{s-1}^j$ and $v' \in S_{s-1}^{j'}$. We will show that the maximum degree of $G_{\mathcal{T}}$ is at most $r - 1$. This clearly implies that $G_{\mathcal{T}}$ contains an independent set of size at least t/r (e.g. by choosing its vertices greedily), which is exactly what we need.

We first show that if \mathcal{T}_j and $\mathcal{T}_{j'}$ are adjacent in $G_{\mathcal{T}}$ then $v_{j'} \in T_s^j$. Indeed, if $s = 1$, then for each $j \in [t]$ we have $S_0^j = \{v_j\}$ by (T4). Hence \mathcal{T}_j and $\mathcal{T}_{j'}$ can only be adjacent if $T_1^j \cup \{v_{j'}\} = T_1^{j'} \cup \{v_j\}$, but then $v_j \neq v_{j'}$ implies $v_{j'} \in T_1^j$. On the other hand, if $s > 1$ then $T_s^j \subseteq L_0$ and $S_{s-1}^j \subseteq L_{s-1}$ where $L_0 \cap L_{s-1} = \emptyset$, so \mathcal{T}_j and $\mathcal{T}_{j'}$ can only be adjacent if $T_s^j = T_s^{j'}$. Once again, (T4) then implies that $v_{j'} \in T_s^{j'} = T_s^j$. But T_s^j has $r - 1$ elements, so there are at most $r - 1$ different choices for this $v_{j'}$. Thus \mathcal{T}_j has at most $r - 1$ neighbors in $G_{\mathcal{T}}$. \square

The next claim shows that in every small collection of independent towers, a significant fraction of them are almost mutually disjoint on any given level k .

²More precisely, each $(S_{s-1}^j, S_s^j, \dots, S_f^j)$ is an i_j -tower on some vertex $v_j \in L_0$, but the precise values for i_j and v_j do not matter here. We use this kind of sloppiness throughout the proof.

Claim 5.8. *The following holds with high probability. For every family $\{(S_{s-1}^j, \dots, S_f^j)\}_{j \in [t]}$ of t independent i -towers, where $1 \leq s \leq f \leq m$ and $t \leq n^\varepsilon / (\mu^k p)$ for some $k \in \{s, \dots, f\}$, there is a set $\mathcal{I}_k \subseteq [t]$ of size $|\mathcal{I}_k| \geq t/2^k$ such that*

$$\left| S_k^j \cap \bigcup_{j' \in \mathcal{I}_k \setminus \{j\}} S_k^{j'} \right| \leq (80r^r)^k \cdot n^{-\varepsilon/4} \mu^k$$

for every $j \in \mathcal{I}_k$.

Proof. Note that with high probability Lemma 3.10 applies with $\tilde{\varepsilon} = n^{-\varepsilon/2}$ and $L = L_k$ simultaneously for every $k \in [m]$. Thus, we may assume that for every $k \in [m]$ and every family of $T \leq \tilde{\varepsilon}/p$ different r -sets $X_1, \dots, X_T \subseteq [n] \setminus L_k$, we have

$$\left| \bigcup_{\ell=1}^T N^*(X_\ell, L_k) \right| \in (1 \pm \sqrt{\tilde{\varepsilon}}) T |L_k| p^r = (1 \pm \sqrt{\tilde{\varepsilon}}) T \cdot 2r^r \mu. \quad (10)$$

We derive Claim 5.8 from this property. For convenience, let $\xi_k = (80r^r)^k n^{-\varepsilon/4}$ for $k \in [m]$.

Let $\{(S_{s-1}^j, S_s^j, \dots, S_f^j)\}_{j \in [t]}$ be a collection of t independent i -towers as in the statement of the claim. For each $j \in [t]$, let (T_s^j, \dots, T_f^j) be a corresponding witness sequence such that all sets of the form $T_s^j \cup \{v\}$, where $j \in [t]$ and $v \in S_{s-1}^j$, are distinct (this is possible because of the independence). We prove the claim by induction on k .

Suppose $k = s$. For each $j \in [t]$ define $Z_k^j = N(S_{k-1}^j) \cap N^*(T_k^j, L_k)$, and set

$$Z_k = \bigcup_{j \in [t]} Z_k^j \quad \text{and} \quad \hat{Z}_k = \bigcup_{j \neq j' \in [t]} Z_k^j \cap Z_k^{j'}.$$

Note that the definition Z_k^j is almost the same as the definition of S_k^j , however without restricting the neighborhood of S_{k-1}^j to a specific color. Therefore, we have $S_k^j \subseteq Z_k^j$. Alternatively, we can define Z_k^j as

$$Z_k^j = \bigcup_{v \in S_{k-1}^j} N^*(T_k^j \cup \{v\}, L_k).$$

By (10), each Z_k^j is of size $|Z_k^j| \in (1 \pm \sqrt{\tilde{\varepsilon}}) 2r^r \mu^k$ (recall that $|S_{k-1}^j| = \mu^{k-1}$). We prove a somewhat stronger statement than needed, namely that the set $\mathcal{I}_k \subseteq [t]$ consisting of all indices $j \in [t]$ such that

$$\left| Z_k^j \cap \hat{Z}_k \right| = \left| Z_k^j \cap \bigcup_{j' \in [t] \setminus \{j\}} Z_k^{j'} \right| \leq 10\sqrt{\tilde{\varepsilon}} \cdot 2r^r \mu^k < \xi_k \mu^k,$$

is of size $|\mathcal{I}_k| \geq t/2$.

To this end, we first estimate the size of Z_k and \hat{Z}_k . As the towers are independent, there are exactly $\sum_{j \in [t]} |S_{k-1}^j| = t\mu^{k-1} < 1/(n^{\varepsilon/2} p) = \tilde{\varepsilon}/p$ sets of the form $T_k^j \cup \{v\}$ where $j \in [t]$ and $v \in S_{k-1}^j$ (recall $\mu = \Theta(n^{r\varepsilon})$). Thus applying (10) again we obtain $|Z_k| \in (1 \pm \sqrt{\tilde{\varepsilon}}) t \cdot 2r^r \mu^k$, which in turn gives the following estimate on the size of \hat{Z}_k .

$$|\hat{Z}_k| \leq \left(\sum_{j \in [t]} |Z_k^j| \right) - |Z_k| \leq 2\sqrt{\tilde{\varepsilon}} t \cdot 2r^r \mu^k.$$

Indeed, here $|\hat{Z}_k|$ is the number of elements in L_k that are counted at least twice by the sum $\sum_{j \in [t]} |Z_k^j|$, whereas $|Z_k|$ is the number of elements counted at least once.

Putting everything together, we have the following:

$$\begin{aligned}
(1 - \sqrt{\varepsilon})t \cdot 2r^r \mu^k &\leq |Z_k| = \sum_{j \in [t]} |Z_k^j \setminus \hat{Z}_k| + |\hat{Z}_k| \\
&\leq \sum_{j \in [t]} |Z_k^j| - \sum_{j \in [t] \setminus \mathcal{I}_k} |Z_k^j \cap \hat{Z}_k| + |\hat{Z}_k| \\
&\leq (1 + \sqrt{\varepsilon})t \cdot 2r^r \mu^k - (t - |\mathcal{I}_k|)10\sqrt{\varepsilon} \cdot 2r^r \mu^k + 2\sqrt{\varepsilon}t \cdot 2r^r \mu^k \\
&\leq (1 + 3\sqrt{\varepsilon} - 5(t - |\mathcal{I}_k|)\frac{2\sqrt{\varepsilon}}{t})t \cdot 2r^r \mu^k.
\end{aligned}$$

This implies $|\mathcal{I}_k| > t/2$, as required.

Next, suppose that $k > s$ and the claim holds for $k-1$. As $t \leq n^\varepsilon / (\mu^k p) \leq n^\varepsilon / (\mu^{k-1} p)$, we can apply the induction hypothesis to obtain a family $\mathcal{I}_{k-1} \subseteq [t]$ of $t' = |\mathcal{I}_{k-1}| \geq t/2^{k-1}$ almost disjoint towers on the $(k-1)$ 'st level. For each $j \in \mathcal{I}_{k-1}$ let

$$\hat{S}_{k-1}^j = S_{k-1}^j \setminus \bigcup_{j' \in \mathcal{I}_{k-1} \setminus \{j\}} S_{k-1}^{j'}.$$

Then these sets are all disjoint and have size

$$(1 - \xi_{k-1})\mu^{k-1} \leq |\hat{S}_{k-1}^j| \leq \mu^{k-1}. \quad (11)$$

Note that $\hat{S}_{k-1}^j \subseteq L_{k-1}$ with $k > 1$, so \hat{S}_{k-1}^j is also disjoint from $T_k^j \subseteq L_0$. More importantly, these facts imply that all the sets of the form $T_k^j \cup \{v\}$ for $j \in \mathcal{I}_{k-1}$ and $v \in \hat{S}_{k-1}^j$ are distinct. Now we can argue as in the base case.

For each $j \in \mathcal{I}_{k-1}$ define

$$Z_k^j = N(\hat{S}_{k-1}^j) \cap N^*(T_k^j, L_k) = \bigcup_{v \in \hat{S}_{k-1}^j} N^*(T_k^j \cup \{v\}, L_k)$$

and let

$$Z_k = \bigcup_{j \in \mathcal{I}_{k-1}} Z_k^j \quad \text{and} \quad \hat{Z}_k = \bigcup_{j \neq j' \in \mathcal{I}_{k-1}} Z_k^j \cap Z_k^{j'}.$$

Observe that, unlike in the base case, we do not have $S_k^j \subseteq Z_k^j$: some vertices in S_k^j might only have neighbors in $S_{k-1}^j \setminus \hat{S}_{k-1}^j$. However, $S_{k-1}^j \setminus \hat{S}_{k-1}^j$ is small, so $S_k^j \setminus Z_k^j$ will turn out to be negligible. We will deal with these vertices at the end.

From $|\hat{S}_{k-1}^j| \leq \mu^{k-1} \leq \varepsilon/p$ and (10) we have

$$|Z_k^j| \leq (1 + \sqrt{\varepsilon}) \cdot 2r^r \mu^k \leq (1 + \xi_{k-1}) \cdot 2r^r \mu^k.$$

As already mentioned, sets of the form $T_k^j \cup \{v\}$ for $j \in \mathcal{I}_{k-1}$ and $v \in \hat{S}_{k-1}^j$ are all distinct, thus by (11), there are

$$\sum_{j \in \mathcal{I}_{k-1}} |\hat{S}_{k-1}^j| \geq (1 - \xi_{k-1})\mu^{k-1}t'$$

many of them. Therefore, applying (10) again we get $|Z_k| \geq (1 - 2\xi_{k-1})t' \cdot 2r^r \mu^k$ which, in turn, gives the following bound on the size of \hat{Z}_k .

$$|\hat{Z}_k| \leq \left(\sum_{j \in \mathcal{I}_{k-1}} |Z_k^j| \right) - |Z_k| \leq 3\xi_{k-1}t' \cdot 2r^r \mu^k.$$

We now define $\mathcal{I}_k \subseteq \mathcal{I}_{k-1}$ as the set of all indices $j \in \mathcal{I}_{k-1}$ such that

$$|Z_k^j \cap \hat{Z}_k| < 20\xi_{k-1} \cdot 2r^r \mu^k = \frac{\xi_k}{2} \mu^k.$$

Using the above estimates on the size of Z_k and \hat{Z}_k , we get the following:

$$\begin{aligned}
(1 - 2\xi_{k-1})t' \cdot 2r^r \mu^k &\leq |Z_k| = \sum_{j \in \mathcal{I}_{k-1}} |Z_k^j \setminus \hat{Z}_k| + |\hat{Z}_k| \\
&\leq \sum_{j \in \mathcal{I}_{k-1}} |Z_k^j| - \sum_{j \in \mathcal{I}_{k-1} \setminus \mathcal{I}_k} |Z_k^j \cap \hat{Z}_k| + |\hat{Z}_k| \\
&\leq (1 + \xi_{k-1})t' \cdot 2r^r \mu^k - (t' - |\mathcal{I}_k|)20\xi_{k-1} \cdot 2r^r \mu^k + 3\xi_{k-1}t' \cdot 2r^r \mu^k \\
&\leq (1 + 4\xi_{k-1} - (t' - |\mathcal{I}_k|)20\frac{\xi_{k-1}}{t'})t' \cdot 2r^r \mu^k,
\end{aligned}$$

which implies $|\mathcal{I}_k| > t'/2 \geq t/2^k$.

So far we have shown that \mathcal{I}_k has the desired size, and that for each $j \in \mathcal{I}_k$, the intersection of S_k^j with other sets $S_k^{j'}$ inside Z_k^j contains at most $\xi_k \mu^k / 2$ vertices. Therefore, it suffices to prove $|S_k^j \setminus Z_k^j| \leq \frac{\xi_k}{2} \mu^k$ to finish the proof. Note that

$$S_k^j \setminus Z_k^j \subseteq N(S_{k-1}^j \setminus \hat{S}_{k-1}^j) \cap N^*(T_k^j, L_k) = \bigcup_{v \in S_{k-1}^j \setminus \hat{S}_{k-1}^j} N^*(T_k^j \cup \{v\}, L_k).$$

As $|S_{k-1}^j \setminus \hat{S}_{k-1}^j| \leq \xi_{k-1} \mu^{k-1}$, from (10) we get

$$|S_k^j \setminus Z_k^j| \leq (1 + \sqrt{\varepsilon})\xi_{k-1} \cdot 2r^r \mu^k < \xi_k \mu^k / 2,$$

as required. This concludes the argument. \square

Next, we show that cascades are resilient to small changes.

Claim 5.9. *There is a constant $c > 0$ such that the following holds with high probability. Suppose $v, w \in L_0$ are connected by an i -cascade with underlying i -towers $\mathcal{T}_v = (S_{s_v-1}^v, \dots, S_f^v)$ and $\mathcal{T}_w = (S_{s_w-1}^w, \dots, S_f^w)$, for some $1 \leq s_v, s_w \leq f \leq m$. Then for any subset $Y \subseteq U$ such that $|S_k^u \cap Y| \leq c\mu^k$ for every $u \in \{v, w\}$ and $k \in \{s_u, \dots, f\}$, the graph $G[(\{v, w\} \cup U) \setminus Y]$ contains an i -colored v - w path of length at most $2f + 1$.*

Proof. With high probability, Lemma 3.10 applies with $\tilde{\varepsilon} = 1/4$ and $L = L_k$ simultaneously for every $k \in [m]$. Thus, we may assume that for every $k \in [m]$ and every family of $t \leq \tilde{\varepsilon}/p$ different r -sets $X_1, \dots, X_t \subseteq [n] \setminus L_k$ we have

$$\left| \bigcup_{j \in [t]} N^*(X_j, L_k) \right| \leq 2t|L_k|p^r = 4r^r t \mu. \quad (12)$$

Furthermore, by Lemma 3.8, we can assume that for every disjoint $X, X' \subseteq V(G)$ of size $|X|, |X'| \gg \log n/p$ we have

$$e(X, X') \in (1 \pm \alpha)|X||X'|p, \quad (13)$$

for $\alpha = 1/(8r + 8)$. We show that these properties suffice to derive the claim.

Consider some $u \in \{v, w\}$. We first show that there is a subset $B^u \subseteq S_f^u$ of size at most $\alpha|S_f^u| = \alpha\mu^f$, such that for every vertex $u' \in S_f^u \setminus B^u$, there is an i -colored u - u' path avoiding Y of length at most f .

To this end, we define the sets $B_{s_u}^u, \dots, B_f^u$ level by level as follows. Let $B_{s_u}^u = Y \cap S_{s_u}^u$ and then iteratively set $B_k = N(B_{k-1}, S_k^u) \cup (Y \cap S_k^u)$ for $k \in \{s_u + 1, \dots, f\}$. It is easy to see by induction on $k \in \{s_u, \dots, f\}$ (and using the definition of an i -tower) that for every vertex $u' \in S_k^u \setminus B_k^u$, the graph $G[(\{v\} \cup U) \setminus Y]$ contains an i -colored u - u' path of length at most k . We will prove, using induction on k , that

$$|B_k^u| \leq (8r^r)^k c\mu^k \quad \text{for every } k \in \{s_u, \dots, f\}. \quad (14)$$

Choosing a $c < \alpha/(8r^r)^f$ in the assumptions of the claim then ensures that $B^u = B_f^u$ is of size at most $\alpha\mu^f$.

The case $k = s_u$ follows from the assumptions on Y and the definition of $B_{s_u}^u$, so let $k > s_u$ and assume that (14) holds for $k - 1$. By the definition of S_k^u , we have

$$N(B_{k-1}^u, S_k^u) \subseteq N(B_{k-1}^u) \cap N^*(T_k^u, L_k) \subseteq \bigcup_{v \in B_{k-1}^u} N^*(T_k^u \cup \{v\}, L_k).$$

As B_{k-1}^u is asymptotically smaller than $1/p$ (see (7)), we can use (12) to get

$$|N(B_{k-1}^u, S_k^u)| \leq 4r^r |B_{k-1}^u| \mu \leq 4r^r (8r^r)^{k-1} c \mu^k.$$

The assumption of the claim states $|S_k^u \cap Y| \leq c\mu^k$, which implies the desired bound on B_k^u (with room to spare).

We now use these sets to find a desired path from v to w . The towers \mathcal{T}_v and \mathcal{T}_w form an i -cascade connecting v and w , thus by definition, we either have $S_f^v = S_f^w$, or $f = m$ and $e_i(R_v, R_w) \geq e(R_v, R_w)/r$ where $R_v = S_m^v \setminus S_m^w$ and $R_w = S_m^w \setminus S_m^v$. In the former case we are immediately done: since $|B^v \cup B^w| < 2\alpha|S_f^v|$ there is a vertex $z \in S_f^v \setminus (B^v \cup B^w)$ and hence an i -colored v - z - w walk, containing an i -colored path, of length at most $2f \leq 2m + 1$, disjoint from Y . Similarly, if we are in the latter case and $|S_m^v \cap S_m^w| > 2\alpha\mu^m$, then we are done for the same reason.

Let us therefore assume that we are in the latter case and $|R_v|, |R_w| \geq (1 - 2\alpha)\mu^m$. It is enough to show that there is an edge $z_v z_w \in G$ of color i such that $z_v \in R_v \setminus B^v$ and $z_w \in R_w \setminus B^w$. Indeed, such an edge would connect an i -colored v - z_v path and an i -colored w - z_w path that both avoid Y , thus providing a desired path from v to w of length at most $2m + 1$. To show that such an edge exists, note that $|R_v|, |R_w| \gg (\log n)/p$ (see (7)), so (13) gives

$$e_i(R_v, R_w) \geq e(R_v, R_w)/r \geq |R_v||R_w|p/2r \geq (1 - 4\alpha)|S_m^v||S_m^w|p/r.$$

On the other hand, as B^v is contained in some set B of size $\alpha|S_m^v|$, it touches at most

$$e(B^v, R_w) \leq e(B, R_w) \leq 2|B||R_w|p \leq 2\alpha|S_m^v||S_m^w|p.$$

of these edges (again, using (13)). We similarly get that B^w touches at most $2\alpha|S_m^v||S_m^w|p$ such edges, which means that

$$e_i(R_v \setminus B^v, R_w \setminus B^w) \geq (1 - (4 + 4r)\alpha)|S_m^v||S_m^w|p/r > 0$$

with $\alpha = 1/(8r + 8)$, so at least one edge avoids both B^v and B^w . \square

We finally have all the necessary tools to prove Claim 5.4.

Proof of Claim 5.4. Suppose that G satisfies the statement in Claim 5.8 and Claim 5.9, as it does w.h.p, and let $c > 0$ be a constant given by Claim 5.9. We show that these two properties are enough to prove Claim 5.4.

Let C be an i -colored cycle in H for some color $i \in [r]$. For every edge $e = vv'$ of C , we define a $2m$ -uniform hypergraph $\mathcal{H}_e = (U, \mathcal{E}_e)$ on the vertex set U , where a set $A \subseteq U$ of size $2m$ belongs to \mathcal{E}_e if $G[A \cup \{v, v'\}]$ contains an i -colored v - v' path. Note that if for every edge $e \in E(C)$ we can find a hyperedge $f_e \in \mathcal{E}_e$ such that all these hyperedges are pairwise disjoint, then the corresponding paths form an i -colored cycle C^* in G such that $V(C^*) \cap L_0 = V(C) \cap L_0$. It is therefore enough to show that such a family of hyperedges exists for every monochromatic cycle C in H of length at most K/p .

This will follow from Theorem 3.2, provided that for every $E' \subseteq E(C)$, we have

$$\tau\left(\bigcup_{e \in E'} \mathcal{E}_e\right) > (4m - 1)(|E'| - 1), \quad (15)$$

where $\tau(\mathcal{E})$ is the smallest size of a set $X \subseteq U$ intersecting every hyperedge in \mathcal{E} .

Consider some $E' \subseteq E(C)$, and label the endpoints of each edge $e \in E'$ with v_e and w_e . We first pass to a large subset $F \subseteq E'$ that has some convenient properties. This is done in three steps.

First, let $E'' \subseteq E'$ be a subset of size $|E'|/3$ such that the edges in E'' are pairwise disjoint (i.e., no two edges in E'' share an endpoint). This is possible because all the edges of E' lie on a cycle. Next, recall from the definition of H that each edge $e \in E''$ represents an i -cascade between v_e and w_e formed by an i -tower $\mathcal{T}_{v_e} = (S_{s_{v_e}-1}^{v_e}, \dots, S_{f_e}^{v_e})$ on v_e and another i -tower $\mathcal{T}_{w_e} = (S_{s_{w_e}-1}^{w_e}, \dots, S_{f_e}^{w_e})$ on w_e (see Definition 5.2). Note that f_e is the same for both vertices v_e and w_e . As there are m possible values for each of s_v, s_w and f_e , we can find a subset $E''' \subseteq E''$ of size at least $|E''|/m^3 \geq |E'|/(3m^3)$ and levels $s_v, s_w, f \in [m]$ such that $s_{v_e} = s_v, s_{w_e} = s_w$ and $f_e = f$ for every $e \in E'''$. Finally, applying Claim 5.7 twice (once for the the collection $\{\mathcal{T}_{v_e}\}_{e \in E'''}$ and once for the collection $\{\mathcal{T}_{w_e}\}_{e \in E'''}$), we find a subset $F \subseteq E'''$ of size $|E'|/(3m^3 r^2)$ such that $\{\mathcal{T}_{v_e}\}_{e \in F}$ and $\{\mathcal{T}_{w_e}\}_{e \in F}$ are both independent collections of towers (but their union might not be).

Let $t = |F| \leq K/p$ and let e_1, \dots, e_t be the edges in F . Rephrasing (15), we need to prove that no set $Y \subseteq U$ of size $|Y| = 4m|E'| \leq 12m^4 r^2 t$ covers all the hyperedges in $\bigcup_{e \in F} \mathcal{E}_e$. That is, there is an edge $e = v_e w_e$ in F and an i -colored path P of length at most $2m + 1$ connecting v_e and w_e such that $V(P) \subseteq (\{v_e, w_e\} \cup U) \setminus Y$. Claim 5.9 suggests that it suffices to show that there is an edge $e \in F$ whose cascade mostly evades Y .

Let $Y_k^{v_e} = S_k^{v_e} \cap Y$ for every $e \in F$ and $k \in \{s_v, \dots, f\}$. We will show that for each $k \in \{s_v, \dots, f\}$ most sets $Y_k^{v_e}$ are quite small. More precisely, the set $\mathcal{B}_k \subseteq F$ of all edges $e \in F$ such that $|Y_k^{v_e}| \geq c\mu^k$ is of size

$$|\mathcal{B}_k| < t/(2m). \quad (16)$$

Consider some $k \in \{s_v, \dots, f\}$. First, we show that $|\mathcal{B}_k| \leq n^\varepsilon/(\mu^k p)$. Indeed, if this is not the case then choose an arbitrary subset $\mathcal{J} \subseteq \mathcal{B}_k$ of size $n^\varepsilon/(\mu^k p)$ and let $\mathcal{I}_k \subseteq \mathcal{J}$ be the subset provided by Claim 5.8 when applied to the towers $\{\mathcal{T}_{v_e}\}_{e \in \mathcal{J}}$. For each $e \in \mathcal{I}_k$, the set $S_k^{v_e}$ intersects $\bigcup_{e' \in \mathcal{I}_k} S_k^{v_{e'}}$ on $o(\mu^k)$ vertices, thus it contains at least $c\mu^k/2$ unique elements from Y (i.e. elements which do not appear in any other $S_k^{v_{e'}}$ for $e' \in \mathcal{I}_k$). This implies

$$|Y| \geq \frac{c}{2}\mu^k |\mathcal{I}_k| \geq \frac{c}{2}\mu^k \cdot n^\varepsilon/(2^k \mu^k p) \gg t,$$

which is a contradiction. Therefore $|\mathcal{B}_k| \leq n^\varepsilon/(\mu^k p)$, thus we can apply Claim 5.8 to all the towers $\{\mathcal{T}_{v_e}\}_{e \in \mathcal{B}_k}$. Let \mathcal{I}_k be the obtained set of indices and, again, note that each $S_k^{v_e}$ contains at least $c\mu^k/2$ unique elements from Y . Assuming $|\mathcal{B}_k| \geq t/(2m)$, the contradiction follows similarly as in the previous case:

$$|Y| \geq \frac{c}{2}\mu^k |\mathcal{I}_k| \geq \frac{c}{2}\mu^k \cdot t/(2^k \cdot 2m) \gg t.$$

Finally, it follows from (16) that there is a subset $\mathcal{I}_v \subseteq F$ of size $|\mathcal{I}_v| > t/2$ such that for all $e \in \mathcal{I}_v$ and all $k \in \{s_v, \dots, f\}$, we have $|Y_k^{v_e}| \leq c\mu^k$. By the same argument, a subset $\mathcal{I}_w \subseteq F$ of size $|\mathcal{I}_w| > t/2$ exists for the other endpoints of the edges, as well, such that $|Y_k^{w_e}| \leq c\mu^k$ for every $e \in \mathcal{I}_w$ and all $k \in \{s_w, \dots, f\}$. In particular, there is an edge $e \in \mathcal{I}_v \cap \mathcal{I}_w$ such that Y intersects at most a c -fraction of each level of the cascade on e . The existence of a desired path now follows from Claim 5.9. \square

6 Concluding remarks

In this paper we made a step towards the random analog of the theorem of Erdős, Gyárfás and Pyber [6] on monochromatic cycle covers. Our result leaves a few interesting open problems:

- The most interesting open problem is to show that there is a *partition* of the vertices of $\mathcal{G}(n, p)$ into constantly many monochromatic cycles (or paths), even for some larger values of p .
- It would be nice to give a more precise estimate on the threshold for the property that every r -coloring of the edges of $\mathcal{G}(n, p)$ admits a vertex cover by a number of monochromatic cycles depending only on r . In view of the construction of Bal and DeBiasio that we mentioned in the introduction, it seems natural to guess that the threshold should be of the order $(\log n/n)^{1/r}$. Note that in the proof of Lemma 2.2 we heavily rely on the fact that there are only constantly many levels in a cascade, which requires $p \geq n^{-1/r+\varepsilon}$ for some constant $\varepsilon > 0$.
- We did not put much effort into optimizing the number of cycles we cover with in Theorem 1.1, thus it could most likely be improved. It would be interesting to see if one could obtain similar bounds to the case $G = K_n$, e.g., do $O(r \log r)$ cycles suffice?

References

- [1] D. Bal and L. DeBiasio. Partitioning random graphs into monochromatic components. *The Electronic Journal of Combinatorics*, 24(1):P1–18, 2017.
- [2] 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.
- [3] S. Bessy and S. Thomassé. Partitioning a graph into a cycle and an anticycle, a proof of Lehel’s conjecture. *Journal of Combinatorial Theory, Series B*, 100(2):176–180, 2010.
- [4] D. Conlon. Combinatorial theorems relative to a random set. In *Proceedings of the International Congress of Mathematicians 2014*, volume 4, pages 303–328. 2014.
- [5] L. DeBiasio and L. L. Nelsen. Monochromatic cycle partitions of graphs with large minimum degree. *Journal of Combinatorial Theory, Series B*, 122:634–667, 2017.
- [6] 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.
- [7] L. Gerencsér and A. Gyárfás. On Ramsey-type problems. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.*, 10:167–170, 1967.
- [8] A. Grinshpun and G. N. Sárközy. Monochromatic bounded degree subgraph partitions. *Discrete Mathematics*, 339(1):46–53, 2016.
- [9] A. Gyárfás. Covering complete graphs by monochromatic paths. In *Irregularities of partitions*, pages 89–91. Springer, 1989.
- [10] A. Gyárfás. Vertex covers by monochromatic pieces—a survey of results and problems. *Discrete Mathematics*, 339(7):1970–1977, 2016.
- [11] A. Gyárfás, A. Jagota, and R. H. Schelp. Monochromatic path covers in nearly complete graphs. *J. Combin. Math. Combin. Comput.*, 25:129–144, 1997.
- [12] 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.
- [13] P. E. Haxell. A condition for matchability in hypergraphs. *Graphs Combin.*, 11(3):245–248, 1995.
- [14] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. John Wiley & Sons, 2000.

- [15] Y. Kohayakawa, G. O. Mota, and M. Schacht. Monochromatic trees in random graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, to appear.
- [16] M. Krivelevich and B. Sudakov. The phase transition in random graphs: A simple proof. *Random Structures and Algorithms*, 43(2):131–138, 2012.
- [17] R. Lang and M. Stein. Local colourings and monochromatic partitions in complete bipartite graphs. *European Journal of Combinatorics*, 60:42–54, 2017.
- [18] S. Letzter. Monochromatic cycle partitions of 2-coloured graphs with minimum degree $3n/4$. *arXiv preprint arXiv:1502.07736*, 2015.
- [19] A. Pokrovskiy. Partitioning edge-coloured complete graphs into monochromatic cycles and paths. *Journal of Combinatorial Theory, Series B*, 106:70–97, 2014.
- [20] G. N. Sárközy. Monochromatic cycle partitions of edge-colored graphs. *Journal of graph theory*, 66(1):57–64, 2011.
- [21] G. N. Sárközy, S. M. Selkow, and F. Song. An improved bound for vertex partitions by connected monochromatic k -regular graphs. *Journal of Graph Theory*, 73(2):127–145, 2013.