Anti-Ramsey Type Problems

Sean Elliott
Mentor: Dr. Asaf Ferber

Abstract

A classical theorem due to Ramsey says the following: Given a finite number of colors and a positive integer p, any edge-coloring of the complete graph K_n will contain a monochromatic copy of K_p as long as n is sufficiently large. A related problem is to consider colorings of K_n for which every copy of K_4 uses at least 3 distinct colors, and ask for the minimum number of colors that can be used to produce such a coloring. Here we present an alternate proof of the best known upper bound, which is $2^{O(\sqrt{\log n})}$.

We also consider the problem of covering a regular graph with regular bipartite subgraphs. The motivation for this problem comes from the example of covering K_n with complete bipartite subgraphs, which can be done with $\log_2(n)$ many subgraphs. Here we show that with high probability, a random d-regular graph with an even number of vertices can be covered with $c \log d$ many regular bipartite subgraphs for an absolute constant c.

1 Introduction

Ramsey theory is an area intersecting various fields of mathematics. This area began with a theorem by Ramsey which said that for any positive integers p and q, any edge-coloring of K_n with the colors red and blue contains a red copy of K_p or a blue copy of K_q, as long as n is sufficiently large. Let $R(p, q)$ denote the minimum possible such n. Early work in this area produced a lower bound for $R(p, p)$ of essentially $2^{\frac{p}{2}}$ and an upper bound of essentially 4^p. In the many years since then, these bounds have only seen marginal improvements [1].

Ramsey's theorem can be extended to more colors and also to different subgraphs and to hypergraphs. For example, one can define the multicolor Ramsey number $r_k(p)$ to be the minimum n such that every edge-coloring of K_n with k colors will produce a monochromatic copy of K_p [2]. Again, improving the existing bounds on such a function is difficult, even in the case $p = 3$ [3]. Note that finding the minimum n such that every edge-coloring of K_n will produce a monochromatic copy of K_p is equivalent to finding the maximum n for which there exists an edge-coloring of K_n where every copy of K_p uses at least 2 distinct colors. This motivates the following more general definition, as given by Erdős and Gyárfás in [3]:

Definition 1.1. *For positive integers p and q with $p \geq 3$ and $2 \leq q \leq \binom{p}{2}$, a (p,q)-*coloring *is an edge-coloring of K_n (from an arbitrary set of colors) where every copy of*

K_p has at least q distinct colors. We let $f(n,p,q)$ be the minimal possible number of colors of a (p,q)-coloring of K_n.

To see that $f(n,p,q)$ must exist, note that simply giving every edge a distinct color will produce a (p,q)-coloring. From our remark following the definition of the multicolor Ramsey number, we see that determining $f(n,p,2)$ is equivalent to determining $r_k(p)$. As stated above, finding asymptotics for these functions is difficult. However, for $q \geq 3$ there have been many results found about the nature of $f(n,p,q)$, for example in [3]. In this paper, we present an alternate proof of the best known upper bound on $f(n,4,3)$.

Another way of viewing graph coloring problems is that we are *covering* the graph with subgraphs that have certain properties. In the problem above, this means choosing subgraphs such that no $q-1$ subgraphs together contain a K_p, and so that every edge appears in exactly one subgraph. More generally, one can consider problems in which we are covering a graph with subgraphs but we allow each edge to be covered more than once. Here we consider the problem of covering a graph with regular bipartite subgraphs.

Fishburn and Hammer proved in [4] that one can find a collection of $\lceil \log_2 n \rceil$ complete bipartite subgraphs of K_n that together cover every edge at least once. This means finding a collection of pairs (A_i, B_i), where i ranges from 1 to $\lceil \log_2 n \rceil$, such that $A_i, B_i \subset V(K_n), A_i \cap B_i = \emptyset$, and for all $xy \in E$ there exists an i with $x \in A_i, y \in B_i$ or $y \in A_i, x \in B_i$. Indeed, one can define A_i and B_i as follows: Label each vertex with a distinct binary vector of dimension $\lceil \log_2 n \rceil$, and let A_i be the set of vertices with a 1 in the ith coordinate and $B_i = V(K_n) \backslash A_i$. Since any two vertices differ in at least one coordinate, every edge appears in at least one complete subgraph.

More generally, one can consider graphs that are d-regular for some d; the above is the special case $d = n-1$. In this paper, we find a covering of d-regular graphs that have an even number of vertices by regular bipartite subgraphs, where the number of subgraphs needed grows logarithmically with the regularity d. This covering works asymptotically almost surely, in the sense that if a graph is randomly chosen from all d-regular graphs, the probability that the covering works approaches 1 as n approaches infinity.

In Section 2 we discuss previous results on $f(n,p,q)$ for various values of p and q. In Section 3 we introduce our $(4,3)$-coloring and prove its correctness. In Section 4 we prove our result on covering with bipartite graphs. In Section 5 we discuss conclusions and future work.

2 Previous Results on $f(n,p,q)$

The problem of finding upper and lower bounds on $f(n,p,q)$ has a wide range of difficulty, depending on the values p and q chosen. For example, the case $p = q = 3$ is easy because $(3,3)$-colorings are equivalent to proper edge-colorings (those in which no two adjacent edges have the same color). Thus $f(n,3,3)$ equals the chromatic index $\chi(K_n)$, which is $n-1$ for n even and n for n odd, as shown in [3].

The next case, $(p,q) = (4,3)$, is substantially more difficult. Mubayi constructed a coloring in [5] which uses $2^{O(\sqrt{\log n})}$ colors; this is the best upper bound known. The lower bound was recently improved to $\Omega(\log n)$ by Fox and Sudakov in [6].

Mubayi resolved the case $(p, q) = (4, 4)$ by extending their coloring for $(4, 3)$ to a $(4, 4)$ coloring which uses $n^{1/2+o(1)}$ colors [7]. Because $f(n, 4, 4) \geq n^{1/2} - 1$ was already known [3], this is optimal.

For more general formulas, Erdős and Gyárfás proved a variety of results about $f(n, p, q)$ in [3]. They proved the general bound $f(n, p, q) \leq c_{p,q} n^{\frac{p-2}{\binom{p}{2}-q+1}}$. However, their proof is nonconstructive; they considered random colorings and showed that at least one worked. They also showed that $f(n, p, p) \geq n^{\frac{1}{p-2}} - 1$, and raised the question of whether $f(n, p, p-1)$ would be polynomial or subpolynomial in n. This was resolved in [2], in which Conlon et al. showed that $f(n, p, p-1)$ is subpolynomial. Their proof used a generalization of Mubayi's (4,3)-coloring.

The problem of $(4, 3)$-colorings is the one we have chosen to focus on. Below we present another coloring that achieves the same upper bound asymptotically as Mubayi's. However, this coloring shows room for improvement, because in one of the steps we choose the simplest possible coloring that works, while it is possible that a more clever choice would produce a better bound.

3 A (4,3)-Coloring

It will be helpful to first discuss the ideas behind Mubayi's coloring in [5]. The first step is to eliminate monochromatic triangles. One can check that, up to isomorphism, the only K_4's that have at most 2 distinct colors but no monochromatic triangles are the two shown in Figure 1.

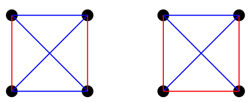

Figure 1: K_4's with 2 distinct colors but no monochromatic triangles

Mubayi eliminated monochromatic triangles by labelling each vertex with a subset of $[m]$ with exactly t elements, for some t and m depending on n, and defining a coloring as follows: An edge between A and B is colored by the smallest element of $A \Delta B$, the symmetric difference of A and B. This removes monochromatic triangles: If AB, BC, CA have the same color, without loss of generality let the smallest element of $A \Delta B$ be in A, and call this element i. Since $A \Delta C$ also has color i, then $i \notin C$ and $i \notin B$, so $B \Delta C$ cannot have the color i. One can think of this as partitioning the edges into edge-disjoint bipartite graphs, each of which is made monochromatic. But in fact, one can show that this coloring prevents the existence of any copies of the rightmost K_4 in Figure 1. And based on how m and t are chosen, this uses only $O(\log n)$ edges. So the main difficulty is removing the leftmost K_4.

Now we discuss our (4, 3)-coloring. Let $k = \lceil 2\sqrt{\log_2 n} \rceil$ and $t = \lceil 2^{\sqrt{\log_2 n}} \rceil$. For each i from 1 to k, let $S_1^{(i)}, \ldots, S_t^{(i)}$ be a partition of $[n]$ into sets whose cardinalities are in $\{\lfloor \frac{n}{t} \rfloor, \lceil \frac{n}{t} \rceil\}$, chosen uniformly at random among all such partitions. We say an edge $e = \{x, y\}$ is *crossing* if there exists an $i \in [k]$ and $a, b \in [t], a \neq b$, such that $x \in S_a^{(i)}, y \in S_b^{(i)}$. For each edge e let $X_e = \begin{cases} 1 & \text{if e is noncrossing} \\ 0 & \text{if e is crossing} \end{cases}$.

Note that

$$\mathbb{E}[X_e] = \Pr(e \text{ is noncrossing}) \leq (\frac{\lceil \frac{n}{t} \rceil - 1}{n-1})^k \leq \frac{1}{t^k}.$$

Then the expected number of noncrossing edges is

$$\mathbb{E}[\sum_e X_e] = \sum_e \mathbb{E}[X_e] \leq \frac{\binom{n}{2}}{t^k} \leq \frac{\binom{n}{2}}{(2^{\sqrt{\log_2 n}})^{2\sqrt{\log_2 n}}} < 1.$$

So there exists a collection of partitions in which every edge is crossing.

Thus, we can assume that $(S_1^{(i)}, \ldots, S_t^{(i)})_{i=1}^k$ has only crossing edges. Now, suppose we have some coloring c on K_t. We color every edge e in our K_n with a triple of colors $(c_1(e), c_2(e), c_3(e))$, where the 1st coordinate $c_1(e)$ is the smallest i such that e crosses between two sets $S_a^{(i)}$ and $S_b^{(i)}$ in the ith partition, the 2nd coordinate $c_2(e)$ is the color of the edge $\{a, b\}$ in the K_t, and the 3rd coordinate $c_3(e)$ is a binary string of length k, where the jth entry is 1 iff e is crossing in the jth partition. Depending on which c we chose, this coloring will satisfy certain properties we want. To actually get a (p, q)-coloring, we choose c to be the simplest possible coloring - one in which every edge in K_t has a unique color. This is where we see a place to improve our coloring.

Claim 3.1. *If c contains no monochromatic triangles, then this coloring contains no monochromatic triangles. If c produces no K_4 with at most 2 distinct colors, then this coloring does not contain a copy of the rightmost K_4 in Figure 1. If c assigns to any two edges in K_t different colors, then this coloring also does not contain a copy of the leftmost K_4 in Figure 1, and so it is a (4,3)-coloring that uses $2^{O(\sqrt{\log n})}$ colors.*

Proof. If xyz is a monochromatic triangle in K_n, then $x \in S_a^{(i)}, y \in S_b^{(i)}, z \in S_c^{(i)}$ for some i and a, b, c. We also have $c_2(xy) = c_2(yz) = c_2(xz)$, so abc form a monochromatic triangle in the K_t colored with c. So if c contains no monochromatic triangles, this coloring also contains no monochromatic triangles.

Now suppose that c produces no K_4 with at most 2 distinct colors. The rightmost bad K_4 can be written as $xyzw$ with xy, yz, zw the same color and xz, xw, yw the same color. If $c_1(xy) = c_1(xz)$, then x, y, z, w are in different sets $S_a^i, S_b^i, S_c^i, S_d^i$. But since $c_2(xy) = c_2(yz) = c_2(zw)$ and $c_2(xz) = c_2(xw) = c_2(yw)$, $abcd$ is a K_4 with at most 2 distinct colors, which is impossible. So the 1st coordinates are different. Let the smaller one be i, and without loss of generality let this be the 1st coordinate for xy, yz, zw. Then in the ith partition, x and w are in the same set, as are x and z and y and z (as otherwise $c_1(xw)$ would be $\leq i$). But this is impossible since x and y are in different sets. So this configuration is impossible.

Now suppose c assigns different colors to any two edges in K_t. The other bad K_4 is when xy, yz, zw, wx are the same color and xz and yw are the same color. Let $c_1(xy) = i$ and $c_1(xz) = j$. Then $j \geq i$, since x, y, z, w were in the same set until the ith partition. And $j \neq i$, since otherwise in the ith partition x, y, z, w would be in four different sets $S_a^i, S_b^i, S_c^i, S_d^i$, and so $abcd$ would be a monochromatic K_4 in the K_t, which is impossible. So $j > i$. Thus in the ith partition x and z must be in the same set, and y and w must be in the same set (otherwise the edge between them would have 1st coordinate $\leq i$). Since $c_2(xz) = c_2(yw)$, xz and yw must cross between the same pair of sets in the jth partition (because our coloring c assigns any two edges in K_t different colors). So in the jth partition we can also guarantee that x and y (without loss of generality; otherwise swap y and w) are in the same set, as are z and w. But then $c_3(xy)$ contains a 0 in the jth entry, whereas $c_3(xw)$ contains a 1 in the jth entry. So this configuration is impossible.

Finally, to see that in this case we achieve the upper bound, note that we use $t^2 2^k$ colors in total (as c_1 is redundant with c_3). This is asymptotically equal to $(2^{\sqrt{\log_2 n}})^2 2^{2\sqrt{\log_2 n}} = 2^{O(\sqrt{\log n})}$, as desired.

□

4 Covering with Bipartite Graphs

To begin, we introduce a few concepts from spectral graph theory.

Definition 4.1. *Let $G = (V, E)$ be a graph whose vertices are labelled $v_1, v_2, \ldots, v_{|V|}$. The* adjacency matrix *of G, denoted $A(G)$, is a $|V| \times |V|$ matrix that has a 1 in entry ij if $v_i v_j \in E$ and a 0 otherwise.*

Suppose that G is d-regular. Then d is an eigenvalue of $A(G)$ for the eigenvector $(1, 1, \ldots, 1)^T$. Furthermore, d is the largest eigenvalue of $A(G)$, in absolute value, see [8]. If $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the eigenvalues of $A(G)$, let $\lambda(G) = \max\{|\lambda_2|, |\lambda_n|\}$. Now we introduce the following definition, as done in [9].

Definition 4.2. *We say that G is an (n, d, λ)*-graph if $|V| = n$, G is d-regular, and $\lambda(G) \leq \lambda$.

Bounding the size of $\lambda(G)$ gives us control over the regularity of G, through the following well-known lemma, known as the Expander Mixing Lemma [8]. First, we introduce the following notation: For subsets S, T of the vertex set of a graph $G = (V, E)$, we let $e(S, T) = |\{(x, y) \in S \times T \mid xy \in E\}|$.

Lemma 4.3 (Expander Mixing Lemma)**.** *Let $G = (V, E)$ be an (n, d, λ) graph, and let $S, T \subset V$. Then*

$$\left| e(S, T) - \frac{d|S||T|}{n} \right| \leq \lambda \sqrt{|S||T|}.$$

An *r-factor* is a spanning r-regular subgraph. We will also need the following lemma, which follows from a generalization of the Gale-Ryser theorem due to Mirsky [10], to find r-factors.

Lemma 4.4. *Let $G = (A \cup B, E)$ be a balanced bipartite graph with $|A| = |B| = m$, and let r be a positive integer. Then G contains an r-factor if and only if for every $X \subset A$ and $Y \subset B$,*

$$e(X, Y) \geq r(|X| + |Y| - m). \tag{1}$$

To find our regular bipartite subgraphs, we first need a collection of almost-regular bipartitions that together cover every edge of G at least once. We find this collection using the following lemma, which is a variant on [9, Lemma 3.1] by Ferber and Jain.

Lemma 4.5. *Let $G = (V, E)$ be a d-regular graph on n vertices, where $d = \omega(1)$ and n is an even and sufficiently large positive integer. Let $t = \lceil 20 \log d \rceil$. Then there exists a collection $(A_i, B_i)_{i=1}^t$ of balanced bipartitions such that:*

Let G_i be the subgraph of G induced by $E_G(A_i, B_i)$. For all $1 \leq i \leq t$ we have $\frac{d}{2} - d^{2/3} \leq \delta(G_i) \leq \Delta(G_i) \leq \frac{d}{2} + d^{2/3}$.

For all $e \in E(G)$, there exists an $i \in [t]$ such that $e \in E(G_i)$.

Proof. The proof is the same as in [9], except that there is a slight difference in the inequalities due to the change in the size of t and the relaxation of the second condition. We will choose our bipartitions (A_i, B_i) according to a random process. Let $D_{i,v}$ be an indicator variable for the event $d_{G_i}(v) \notin \left[\frac{d}{2} - d^{2/3}, \frac{d}{2} + d^{2/3}\right]$, and A_e be an indicator variable for the event $|\{i \in [t] \mid e \in E(G_i)\}| \leq \frac{t}{100}$. We wish to find a collection of bipartitions for which none of these events occur, which for n and d sufficiently large will imply that both of our desired properties hold.

First consider the case where $d > \frac{\sqrt{n}}{2}$. Here we randomly choose t subsets $A_1, A_2, \ldots, A_t$ independently from the uniform distribution on all subsets of V of size $\frac{n}{2}$, and let $B_i = V \backslash A_i$. By Chernoff's inequality [11],

$$\Pr[A_e] \leq \exp\left(-\frac{(\frac{49}{50})^2 t}{4}\right) \leq \frac{1}{e^{4.802 \log d}} = \frac{1}{d^{4.802}}.$$

For n sufficiently large, this is at most $\frac{1}{n^{2.4}}$. Taking the union bound over all edges e, we obtain $\Pr[\bigcup_{e \in E} A_e] \leq \frac{1}{n^{0.4}}$. Again using Chernoff's inequality, we obtain $\Pr[D_{i,v}] \leq \exp\left(-\frac{d^{1/3}}{10}\right)$, which we note is at most $\frac{1}{n^3}$ for n sufficiently large. Taking the union bound over all i and v, we obtain $\Pr[\bigcup_{i \in [t], v \in V} D_{i,v}] \leq \frac{1}{n}$. So by the union bound, the entire collection satisfies both of the desired properties with probability at least $1 - \frac{1}{n^{0.4}} - \frac{1}{n}$. Hence we can find such a collection.

Now consider the case where $d \le \dfrac{\sqrt{n}}{2}$. Define the graph $G' = (V, E')$ by the following rule: For all $x, y \in V$, we have $xy \in E'$ if and only if $xy \notin E$ and there is no vertex $v \in V$ with $xv, yv \in E$. For any $x \in V$ there are at most d^2 many $y \in V$ that do not have this property, so $\delta(G') \ge n - d^2 \ge \frac{n}{2}$. Then we claim that G' must have a perfect matching. Suppose for the sake of contradiction that the maximum size of a matching is $k < \frac{n}{2}$, and let $\{x_1y_1, \dots, x_ky_k\}$ be any such matching. Define $S = \{x_1, x_2, \dots, x_k\} \cup \{y_1, y_2, \dots, y_k\}$, and let $S' = V \backslash S$. Note that for any $u, v \in S'$ we have $uv \notin E'$, as otherwise the matching would not be maximal. Thus for any $u, v \in S'$, we know that u and v are each adjacent to at least $\frac{n}{2}$ elements in S, so there exists an $i \in [k]$ such that $x_iu, y_iv \in E'$ (without loss of generality; otherwise swap u and v). But then $\{x_jy_j \mid j \in [k], j \ne i\} \cup \{x_iu, y_iv\}$ is a larger matching, contradicting maximality.

So, let $s = \frac{n}{2}$ and take a perfect matching $M = \{x_1y_1, \dots, x_sy_s\}$ of G'. For each $i \in [t]$ let f_i be a random function chosen independently and uniformly from the set of functions mapping $\{x_1, \dots, x_s\}$ to $\{\pm 1\}$. Then we define the following partition: $A_i = \{x_j \mid f_i(x_j) = -1\} \cup \{y_j \mid f_i(x_j) = +1\}$ and $B_i = [n] \backslash A_i$. For each $i \in [t]$, let $g_i : V(G) \to \{A_i, B_i\}$ indicate whether a vertex is mapped to A_i or B_i. Then for each $v \in V(G)$, the choices $g_i(w)$ for $w \in N_G(v)$ (the neighborhood of v in G) are mutually independent.

We wish to apply the Symmetric Local Lemma [11] to the collection of events consisting of all the $D_{i,v}$'s and all the A_e's. Note that $D_{i,v}$ and $D_{j,u}$ are independent unless $i = j$ and either $\mathrm{dist}_G(u, v) \le 2$ or $uv \in M$. Also, $D_{i,v}$ and A_e are independent unless an endpoint of e is within distance 1 of v in either G or M. Therefore, each $D_{i,v}$ depends on at most $2d^2$ events in the collection. Furthermore, A_e and $A_{e'}$ are independent unless e and e' share an endpoint in G or if any of the endpoints of e are matched to any of the endpoints in M. So the maximum degree of the dependency graph is $2d^2$. Using Chernoff's inequality as before we see that $2ed^2 \Pr[A_e]$ and $2ed^2 \Pr[D_{i,v}]$ are bounded above by 1, so by the Local Lemma there exists a collection of bipartitions satisfying both properties.

□

Next we require the following lemma, which is based on [12, Lemma 16] by Kühn, Osthus, and Treglownin, but is adapted for our purposes.

Lemma 4.6. *Let G be a bipartite graph with parts U and V satisfying $|U| = |V| = \dfrac{n}{2}$. Let the edges of G be partitioned into sets E_1 and E_2 where the degree of any vertex is at most m in the subgraph induced by E_1. If for some positive integer $t \ge m$ we have*

$$e_{E_2}(A, B) \ge \min\left(|A|t - \left(\frac{n}{2} - |B|\right)(t - m), |B|t - \left(\frac{n}{2} - |A|\right)(t - m)\right) \qquad (2)$$

for all $A \subset U$ and $B \subset V$, then there exists a subset $E' \subset E_2$ such that the subgraph induced by $E_1 \cup E'$ is t-regular.

Proof. For each i from 1 to $\frac{n}{2}$, let x_i be the degree of the ith vertex u_i in U and y_i be the degree of the ith vertex v_i in V, both in the subgraph induced by E_1 (implying that these satisfy $\sum x_i = \sum y_i$). Then we wish to find a subset $E' \subset E_2$ that induces a graph in which u_i has degree $c_i = t - x_i$ and v_i has degree $d_i = t - y_i$.

Like in [12], we create a network by adding vertices s and t to the graph induced by E_2, along with edges su_i and v_it for each i, and directing every edge in E_2 from U to V. We give each edge su_i capacity c_i, each edge v_it capacity d_i, and every other edge capacity 1. Suppose we have some valid flow on this network. We define a subset $E_2' \subset E_2$ consisting of all edges u_iv_j that have flow 1. The degree of u_i in the graph induced by E_2' is then the flow coming out of u_i, which equals the flow of edge su_i. Similarly, the degree of v_i in the graph induced by E_2' is the flow coming into v_i, which equals the flow of edge v_it. Therefore, if we can find a flow such that every edge su_i has flow c_i and every edge v_it has flow d_i, E_2' will be our desired set E'. Since c_i and d_i are the capacities of these edges, any flow is bounded above by $\sum c_i = \sum d_i$, so it suffices to show that there exists a flow with value $\sum c_i$. By the max-flow min-cut theorem, it suffices to show that any cut has capacity at least $\sum c_i$.

For a given cut, let A be the subset of U contained in the source set and B be the subset of V contained in the sink set. Let $\overline{A} = U \backslash A$ and $\overline{B} = V \backslash B$. Then the capacity of the cut is $\sum_{u_i \in \overline{A}} c_i + \sum_{v_i \in \overline{B}} d_i + e_{E_2}(A, B)$, so we just need to show that $e_{E_2}(A, B) \geq \sum_{u_i \in A} c_i - \sum_{v_i \in \overline{B}} d_i$. Since $x_i, y_i \in [0, m]$, we have $c_i, d_i \in [t - m, t]$, so

$$\sum_{u_i \in A} c_i - \sum_{v_i \in \overline{B}} d_i \leq |A|t - \left(\frac{n}{2} - |B|\right)(t - m).$$

Similarly,

$$\sum_{v_i \in B} d_i - \sum_{u_i \in \overline{A}} c_i \leq |B|t - \left(\frac{n}{2} - |A|\right)(t - m).$$

Note that $\sum_{u_i \in A} c_i - \sum_{v_i \in \overline{B}} d_i = \sum_{v_i \in B} d_i - \sum_{v_i \in \overline{A}} c_i$. So by (2) and the inequalities above, we obtain $e_{E_2}(A, B) \geq \sum_{u_i \in A} c_i - \sum_{v_i \in \overline{B}} d_i$ as desired.

□

Using the above lemmas, we can now prove our main theorem of this section.

Theorem 4.7. *There exist $n_0, d_0 \in \mathbb{N}$ such that if G is a random d-regular graph on n vertices with $n \geq n_0, d \geq d_0$, and n even, then with high probability there exists a collection of at most $c \log d$ regular bipartite subgraphs that together cover every edge of G at least once, where c is an absolute constant.*

Proof. By [13], for d sufficiently large G is an (n, d, λ) graph with $\lambda = O(\sqrt{d})$ asymptotically almost surely. If we can find a covering collection of regular bipartite subgraphs for each such graph, then it follows that we can find such a collection for a random graph with high probability. So in what follows, we assume that $\lambda(G)$ is negligible compared to $d^{0.5+\epsilon}$ for constants $\epsilon > 0$ by taking n and d sufficiently large.

Consider the collection of bipartitions found in Lemma 4.5. Let $r = \left\lfloor \frac{d}{2} - d^{0.9} \right\rfloor$. For each bipartition (A_i, B_i), we first need an r-factor of the subgraph induced by $E_G(A_i, B_i)$.

We show that an r-factor exists using Lemma 4.4. To see that (1) holds, take any $X \subset A$ and $Y \subset B$, where without loss of generality $|X| \leq |Y|$. If $|X| < |\overline{Y}|$ then the right hand side of Equation 1 is negative, so there is nothing to show. So we assume that $|X| \geq |\overline{Y}|$. First consider the case $|X| \leq \frac{n}{d^{0.3}}$. By Lemma 4.3 we then have

$$e(X, \overline{Y}) \leq \frac{d|X||\overline{Y}|}{n} + \lambda\sqrt{|X||\overline{Y}|} \leq d^{0.7}|X| + \lambda X.$$

From the regularity of our bipartition we know that $e(X, B_i) \geq \left(\frac{d}{2} - d^{2/3}\right)|X|$. Combining this with the inequality above, we obtain

$$e(X, Y) = e(X, B_i) - e(X, \overline{Y}) \geq \left(\frac{d}{2} - d^{2/3} - d^{0.7} - \lambda\right)|X|.$$

The desired lower bound is $\left(\frac{d}{2} - d^{0.9}\right)(|X| - |\overline{Y}|)$, so we can rewrite the desired inequality as

$$(d^{0.9} - d^{2/3} - d^{0.7} - \lambda)|X| + \left(\frac{d}{2} - d^{0.9}\right)|\overline{Y}| \geq 0,$$

which holds for d sufficiently large.

Now consider the case $|Y| \geq |X| > \frac{n}{d^{0.3}}$. We assume that $\frac{d^{0.9}\sqrt{|X||Y|}}{n} > \lambda$ because this holds for d sufficiently large. Now using Lemma 4.3, we have

$$e(X, Y) \geq \frac{d|X||Y|}{n} - \lambda\sqrt{|X||Y|} \geq \frac{|X||Y|}{n}(d - d^{0.9}).$$

The desired lower bound is $\left(\frac{d}{2} - d^{0.9}\right)(|X| + |Y| - \frac{n}{2})$, so we can rewrite the desired inequality as

$$\frac{|X||Y|}{n}(d - d^{0.9}) - \left(\frac{d}{2} - d^{0.9}\right)(|X| + |Y| - \frac{n}{2}) \geq 0.$$

For fixed $|X|$, consider the left hand side as a function of $|Y|$. At $|Y| = |X|$ the left hand side is a quadratic in $|X|$ with discriminant $(d - 2d^{0.9})^2 - (d - d^{0.9})(d - 2d^{0.9}) < 0$ and positive constant term, so the inequality holds. At $|Y| = \frac{n}{2}$, the inequality holds trivially. Since it holds at both endpoints of $\left[|X|, \frac{n}{2}\right]$, and the left hand side is linear in $|Y|$ for fixed $|X|$, it follows that it holds for the whole interval.

In either case, we have that Equation 1 holds, so by Lemma 4.4 we can find an r-factor H_i of each bipartition. Now, we wish to apply Lemma 4.6 to each bipartition, where $t = \lfloor \frac{d}{10} \rfloor, m = 2\lfloor d^{0.9} \rfloor + 2$, $E_2 = E_{H_i}$ and $E_1 = E_G(A_i, B_i) \backslash E_{H_i}$. To see why Equation 2 holds, consider any two subsets $X \subset A_i$ and $Y \subset B_i$, where without loss of generality $|X| \leq |Y|$. If $\frac{n}{2} - |Y| \geq |X|(1 + 20d^{-1/3})$ then the right hand side of our desired inequality is negative, and hence is less than $e(X,Y)$. Otherwise suppose that $\frac{n}{2} - |Y| < |X|(1+20d^{-1/3})$. For d sufficiently large the right hand side is at most $2|X|$, so we may assume that $\left|\overline{Y}\right| < 2|X|$.

First suppose that $|X| < \frac{n}{10}$. By Lemma 4.3 we then have

$$e_{H_i}(X,\overline{Y}) \le e_{(A_i,B_i)}(X,\overline{Y}) \le \frac{d|X|\left|\overline{Y}\right|}{n} + \lambda\sqrt{|X|\left|\overline{Y}\right|} \le \frac{2d|X|^2}{n} + \lambda\sqrt{2}|X| < d|X|\left(\frac{1}{5} + \frac{\lambda\sqrt{2}}{d}\right).$$

From the regularity of H_i we know that $e_{H_i}(X,B_i) = r|X| = d|X|\left(\frac{1}{2} + O(d^{-0.1})\right)$. Combining this with the inequality above, we obtain

$$e_{H_i}(X,Y) = e_{H_i}(X,B_i) - e_{H_i}(X,\overline{Y}) \ge d|X|\left(\frac{3}{10} + O(d^{-0.1})\right),$$

using our assumption that λ is negligible compared to d. The right hand side above is certainly at least $\frac{d|X|}{10}$ for d sufficiently large. Hence we satisfy the desired lower bound from Equation 2, which is $\frac{d|X|}{10} - \left(\frac{n}{2} - |Y|\right)\left(\frac{d}{10} - m\right)$.

Now consider the case when $|X| \ge \frac{n}{10}$. For d sufficiently large we have $\frac{d}{20} - m \ge \lambda$, which combined with our bound on $|X|$ and the fact that $|X| \le |Y|$ gives us $\frac{d\sqrt{|X||Y|}}{2n} - m\sqrt{\frac{|X|}{|Y|}} \ge \lambda$. Thus by Lemma 4.3 we have

$$e_{H_i}(X,Y) \ge e_{(A_i,B_i)}(X,Y) - m|X| \ge \frac{d|X||Y|}{n} - \lambda\sqrt{|X||Y|} - m|X| \ge \frac{d|X||Y|}{2n}.$$

The desired lower bound from Equation 2 is $\left(|X| + |Y| - \frac{n}{2}\right)\frac{d}{10} - m|Y|$. Therefore, it suffices to show that $\frac{|X||Y|}{2n} - \frac{1}{10}\left(|X| + |Y| - \frac{n}{2}\right) \ge 0$. For fixed $|X|$, consider this as a function of $|Y|$. At $|Y| = |X|$, the left hand side factors as $\frac{1}{2}\left(\frac{|X|}{\sqrt{n}} - \frac{\sqrt{n}}{5}\right)^2 + \frac{3n}{100}$, so the inequality holds. At $|Y| = \frac{n}{2}$, the inequality holds trivially. Since it holds at both endpoints of $\left[|X|, \frac{n}{2}\right]$, and the left hand side is linear in $|Y|$ for fixed $|X|$, it follows that it holds for the whole interval.

In either case, we have that Equation 2 holds. So we can apply Lemma 4.6 as desired, obtaining for each i a subset $E'_i \subset E_{H_i}$ such that the subgraph induced by $(E_G(A_i,B_i)\backslash E_{H_i}) \cup E'_i$, which we denote H'_i, is regular. Then $\{H_i\}_{i=1}^{\lceil 20\log d\rceil} \cup \{H'_i\}_{i=1}^{\lceil 20\log d\rceil}$ is our desired collection, because every edge in G is contained in $E_G(A_i,B_i)$ for some i, and H_i and H'_i together cover every edge of $E_G(A_i,B_i)$.

□

5 Conclusions and Future Work

We gave an alternate proof of the upper bound $2^{O(\sqrt{\log n})}$ for the first problem. To improve the upper bound, one can try to improve our coloring above. This would mean choosing a c that uses fewer colors but still ensures that there is no bad K_4 of the first type. It is possible one could do this by considering c to be a *good* coloring (for some meaning of the word good) and then showing that this will produce a good coloring in the K_n, to recursively produce a bound. It is not enough to say that *good* means that c contains no monochromatic triangles or K_4's with at most 2 distinct colors, as the last part of our argument will not work with only such a coloring (we could have xz and yw crossing between different pairs of sets).

Another direction would be to attempt to prove the lower bound. We believe the correct value is closer to the upper bound (though it is not clear whether the upper bound leaves room for improvement). We hope that working on the upper bound can provide intuition for what properties optimal colorings need, which one could use to come up with an argument that strengthens the lower bound (finding a way to ensure that any coloring with at most k colors has a bad K_4).

For the second problem, we found a covering of random d-regular graphs with $c \log d$ regular bipartite subgraphs that works asymptotically almost surely, for an absolute constant c.

6 Acknowledgements

I sincerely thank my mentor, Dr. Asaf Ferber, for introducing me to this area of research and guiding me throughout the entire process. I would also like to thank Dr. Tanya Khovanova for her advice on writing this paper. Finally, I would like to thank the PRIMES-USA program and the MIT Math Department for this wonderful opportunity.

References

[1] D. Conlon, J. Fox, and B. Sudakov, "Recent developments in graph Ramsey theory," *Surveys in Combinatorics*, pp. 49–118, Jul. 2015. DOI: 10.1017/CBO9781316106853.003.

[2] D. Conlon, J. Fox, C. Lee, and Benny, "The Erdős-Gyárfás problem on generalized Ramsey numbers," *Proceedings of the London Mathematical Society*, vol. 110, no. 1, pp. 1–18, Oct. 2014. DOI: 10.1112/plms/pdu049.

[3] P. Erdős and A. Gyárfás, "A variant of the classical Ramsey problem," *Combinatorica*, vol. 17, no. 4, pp. 459–467, 1997. DOI: 10.1007/bf01195000.

[4] P. C. Fishburn and P. L. Hammer, "Bipartite dimensions and bipartite degrees of graphs," *Discrete Mathematics*, vol. 160, no. 1-3, pp. 127–148, 1996.

[5] D. Mubayi, "Note – edge-coloring cliques with three colors on all 4-cliques," *Combinatorica*, vol. 18, no. 2, pp. 293–296, 1998. DOI: 10.1007/pl00009822.

[6] J. Fox and B. Sudakov, "Ramsey-type problem for an almost monochromatic K_4," *SIAM Journal on Discrete Mathematics*, vol. 23, no. 1, pp. 155–162, 2009. DOI: 10.1137/070706628.

[7] D. Mubayi, "An explicit construction for a Ramsey problem," *Combinatorica*, vol. 24, no. 2, pp. 313–324, Jan. 2004. DOI: 10.1007/s00493-004-0019-6.

[8] S. Hoory, N. Linial, and A. Wigderson, "Expander graphs and their applications," *Bulletin of the American Mathematical Society*, vol. 43, no. 4, pp. 439–561, 2006.

[9] A. Ferber and V. Jain, "1-factorizations of pseudorandom graphs," *arXiv preprint arXiv:1803.10361*, 2018.

[10] L. Mirsky, "Combinatorial theorems and integral matrices," *Journal of Combinatorial Theory*, vol. 5, no. 1, pp. 30–44, 1968.

[11] N. Alon and J. H. Spencer, *The probabilistic method.* John Wiley & Sons, 2004.

[12] D. Kühn, D. Osthus, and A. Treglown, "Hamilton decompositions of regular tournaments," *Proceedings of the London Mathematical Society*, vol. 101, no. 1, pp. 303–335, 2010.

[13] K. Tikhomirov, P. Youssef, *et al.*, "The spectral gap of dense random regular graphs," *The Annals of Probability*, vol. 47, no. 1, pp. 362–419, 2019.