Edge-disjoint rainbow spanning trees in complete graphs

James M. Carraher∗ Stephen G. Hartke† Paul Horn‡

May 9, 2013

Abstract

Let $G$ be an edge-colored copy of $K_n$, where each color appears on at most $n/2$ edges (the edge-coloring is not necessarily proper). A rainbow spanning tree is a spanning tree of $G$ where each edge has a different color. Brualdi and Hollingsworth [4] conjectured that every properly edge-colored $K_n$ ($n \geq 6$ and even) using exactly $n-1$ colors has $n/2$ edge-disjoint rainbow spanning trees, and they proved there are at least two edge-disjoint rainbow spanning trees. Kaneko, Kano, and Suzuki [13] strengthened the conjecture to include any proper edge-coloring of $K_n$, and they proved there are at least three edge-disjoint rainbow spanning trees. Akbari and Alipouri [1] showed that each $K_n$ that is edge-colored such that no color appears more than $n/2$ times contains at least two rainbow spanning trees. 

We prove that if $n \geq 1,000,000$ then an edge-colored $K_n$, where each color appears on at most $n/2$ edges, contains at least $\left\lfloor n/(1000 \log n) \right\rfloor$ edge-disjoint rainbow spanning trees.

Keywords: rainbow spanning trees
AMS classification: Primary: 05C15; Secondary: 05C05, 05C70

1 Introduction

Let $G$ be an edge-colored copy of $K_n$, where each color appears on at most $n/2$ edges (the edge-coloring is not necessarily proper). A rainbow spanning tree is a spanning tree of $G$ such that each edge has a different color. Brualdi and Hollingsworth [4] conjectured that every properly edge-colored $K_n$ ($n \geq 6$ and even) where each color class is a perfect matching has a decomposition of the edges of $K_n$ into $n/2$ edge-disjoint rainbow spanning trees. They proved there are at least two edge-disjoint rainbow spanning trees in such an edge-colored $K_n$. Kaneko, Kano, and Suzuki [13] strengthened the conjecture to say that for any proper edge-coloring of $K_n$ ($n \geq 6$) contains at least $\left\lfloor n/2 \right\rfloor$ edge-disjoint rainbow spanning trees, and they proved there are at least three edge-disjoint rainbow spanning trees. Akbari and Alipouri [1] showed that each $K_n$ that is an edge-colored such that no color appears more than $n/2$ times contains at least two rainbow spanning trees.

Our main result is

Theorem 1. Let $G$ be an edge-colored copy of $K_n$, where each color appears on at most $n/2$ edges and $n \geq 1,000,000$. The graph $G$ contains at least $\left\lfloor n/(1000 \log n) \right\rfloor$ edge-disjoint rainbow spanning trees.

The strategy of the proof of Theorem 1 is to randomly construct $\left\lfloor n/(1000 \log n) \right\rfloor$ edge-disjoint subgraphs of $G$ such that with high probability each subgraph has a rainbow spanning tree. This result is the best known for the conjecture by Kaneko, Kano, and Suzuki. Horn [12] has shown that if the edge-coloring is a

∗Department of Mathematics, University of Nebraska–Lincoln, s-jcarrah1@math.unl.edu. Partially supported by NSF grant DMS-0914815.
†Department of Mathematics, University of Nebraska–Lincoln, hartke@math.unl.edu. Partially supported by NSF grant DMS-0914815.
‡Department of Mathematics, Harvard University, phorn@math.harvard.edu
proper coloring where each color class is a perfect matching then there are at least $cn$ rainbow spanning trees for some positive constant $c$, which is the best known result for the conjecture by Brualdi and Hollingsworth.

There have been many results in finding rainbow subgraphs in edge-colored graphs; Kano and Li [14] surveyed results and conjecture on monochromatic and rainbow (also called heterochromatic) subgraphs of an edge-colored graph. Related work includes Brualdi and Hollingsworth [5] finding rainbow spanning trees and forests in edge-colored complete bipartite graphs, and Constantine [8] showing that for certain values of $n$ there exists a proper coloring of $K_n$ such that the edges of $K_n$ decompose into isomorphic rainbow spanning trees.

The existence of rainbow cycles has also been studied. Albert, Frieze, and Reed [2] showed that for an edge-colored $K_n$ where each color appears at most $\lceil cn \rceil$ times then there is a rainbow hamiltonian cycle if $c < 1/64$ (Rue (see [11]) provided a correction to the constant). Frieze and Krivelevich [11] proved that there exists a $c$ such that if each color appears at most $\lceil cn \rceil$ times then there are rainbow cycles of all lengths.

This paper is organized as follows. Section 2 includes definitions and results used throughout the paper. Section 3, 4, and 5 contains lemmas describing properties of the random subgraphs we generate. The final section provides the proof of our main result.

2 Definitions

First we establish some notation that we will use throughout the paper. Let $G$ be a graph and $S \subseteq V(G)$. Let $G[S]$ denote the induced subgraph of $G$ on the vertex set $S$. Let $[S, \overline{S}]_G$ be the set of edges between $S$ and $\overline{S}$ in $G$. For natural numbers $q$ and $k$, $[q]$ represents the set $\{1, \ldots, q\}$, and $\binom{[q]}{k}$ is the collection of all $k$-subsets of $[q]$. Throughout the paper the logarithm function used has base $e$. One inequality that we will use often is the union sum bound which states that for events $A_1, \ldots, A_r$ that

$$\mathbb{P}\left[\bigcup_{i=1}^{r} A_i \right] \leq \sum_{i=1}^{r} \mathbb{P}[A_i].$$

Throughout the rest of the paper let $G$ be an edge-colored copy of $K_n$, where the set of edges of each color has size at most $n/2$, and $n \geq 1,000,000$. We assume $G$ is colored with $q$ colors, where $n-1 \leq q \leq \binom{n}{2}$. Let $C_j$ be the set of edges of color $j$ in $G$. Define $c_j = |C_j|$, and without loss of generality assume $c_1 \geq c_2 \geq \cdots \geq c_q$. Note that $1 \leq c_j \leq n/2$ for all $j$.

Let $t = \lfloor n/(C \log n) \rfloor$ where $C = 1000$. Note that we have not optimized the constant $C$, and it can be slightly improved at the cost of more calculation. Since $\frac{n}{C \log n} - 1 \leq t \leq \frac{n}{C \log n}$, we have

$$-\frac{1}{t} \leq -\frac{C \log n}{n} \quad \text{and} \quad \frac{C \log n}{n} \leq \frac{1}{t} \leq \left(\frac{n}{n - C \log n}\right) \frac{C \log n}{n}.$$  \hfill (*)

We will frequently use these bounds on $t$.

We construct edge-disjoint subgraphs $G_1, \ldots, G_t$ of $G$ in the following way: independently and uniformly select each edge of $G$ to be in $G_i$ with probability $1/t$. Each $G_i$ (considered as an uncolored graph) is distributed as an Erdős-Rényi random graph $G(n, 1/t)$. Note that the subgraphs are not independent. We will show that with high probability each of the subgraphs $G_1, \ldots, G_t$ simultaneously contain a rainbow spanning tree.

To prove that a graph has a rainbow spanning tree we will use Theorem 2 below that gives necessary and sufficient conditions for the existence of a rainbow spanning tree. Broersma and Li [3] showed that determining the largest rainbow spanning forest of $H$ can be solved by applying the Matroid Intersection Theorem [10] (see Schrijver [15, p. 700]), to the graphic matroid and the partition matroid on the edge set of $H$ defined by the color classes. Schrijver [15] translated the conditions of the Matroid Intersection Theorem into necessary and sufficient conditions for the existence of a rainbow spanning tree. Suzuki [16] and Carraher and Hartke [6] gave graph-theoretical proofs of this same theorem.
Theorem 2. A graph $G$ has a rainbow spanning tree if and only if, for every partition $\pi$ of $V(G)$, at least $s-1$ different colors are represented between the parts of $\pi$, where $s$ is the number of parts of $\pi$.

We show that for every partition $\pi$ of $V(G)$ into $s$ parts, that there are at least $s-1$ colors between the parts for each $G_i$. Sections 3, 4 and 5 describe properties of the subgraphs $G_1, \ldots, G_\ell$ for certain partitions $\pi$ of $V(G)$ into $s$ parts. Many of our proofs use the following variant of Chernoff’s inequality [7], frequently attributed to Bernstein (see [9]).

Lemma 3 (Bernstein’s Inequality). Suppose $X_i$ are independently identically distributed Bernoulli random variables, and $X = \sum X_i$. Then

$$\Pr\{X \geq \mathbb{E}[X] + \lambda\} \leq \exp\left(-\frac{\lambda^2}{2(\mathbb{E}[X] + \lambda/3)}\right)$$

and

$$\Pr\{X \leq \mathbb{E}[X] - \lambda\} \leq \exp\left(-\frac{\lambda^2}{2\mathbb{E}[X]}\right).$$

In several places in the paper we use Jensen’s inequality.

Lemma 4 (Jensen’s Inequality (see [17])). Let $f(x)$ be a real-valued convex function defined on an interval $I = [a, b]$. If $x_1, \ldots, x_n \in I$ and $\lambda_1, \ldots, \lambda_n \geq 0$ with $\sum_{i=1}^n \lambda_i = 1$, then

$$f\left(\sum_{i=1}^n \lambda_ix_i\right) \leq \sum_{i=1}^n \lambda_if(x_i).$$

We also make use of the following upper bounds for binomial coefficients:

$$\binom{n}{k} \leq \left(\frac{en}{k}\right)^k = \exp(k \log n - k \log k + k) \leq n^k.$$ 

3 Partitions with $n$ or $n-1$ parts

In this section we show that a partition $\pi$ of $V(G)$ into $n$ or $n-1$ parts has enough colors between the parts. Since color classes can have small size, there might not be any edges of a given color in a subgraph $G_i$. Therefore, we group small color classes together to form larger pseudocolor classes. Recall that $c_j$ is the size of the color class $C_j$, and $c_1 \geq c_2 \geq \cdots \geq c_q$. Define the pseudocolor classes $D_1, \ldots, D_{n-1}$ of $G'$ recursively as follows:

$$D_k = \left(\bigcup_{j=1}^\ell C_j\right) \setminus \left(\bigcup_{i=1}^{k-1} D_i\right),$$

where $\ell$ is the smallest integer such that $\left|\left(\bigcup_{j=1}^\ell C_j\right) \setminus \left(\bigcup_{i=1}^{k-1} D_i\right)\right| \geq n/4$. Note that the $n-1$ pseudocolor classes might not contain all the edges of $G$.

Lemma 5. Each of the $n-1$ pseudocolor classes $D_1, \ldots, D_{n-1}$ have size at least $n/4$ and at most $n/2$.

Proof. Consider the pseudocolor class $D_k$, for $1 \leq k \leq n-1$. Since each of the pseudocolor classes $D_1, \ldots, D_{k-1}$ has size at most $n/2$, there are at least $\frac{n}{2}(n-k)$ edges not in $\bigcup_{i=1}^{k-1} D_i$. Therefore there exists $\ell'$ and $\ell$ such that $D_k = \bigcup_{i=\ell'}^{\ell} C_i$, where $|D_k| = \sum_{i=\ell'}^{\ell} c_i \geq n/4$.

If $\ell' = \ell$ then $|D_k| = |C_\ell| \leq n/2$. Otherwise, we know $c_{\ell'} \leq c_{\ell-1} \leq c_\ell \leq n/4$. So,

$$|D_k| = \sum_{i=\ell'}^{\ell-1} c_i + c_\ell \leq \frac{n}{4} + c_\ell \leq \frac{n}{4} + \frac{n}{4} = \frac{n}{2},$$

which proves that the pseudocolor class $D_k$ has size at most $\frac{n}{2}$. 

\[\square\]
Lemma 6. For a fixed subgraph $G_i$ and pseudocolor class $D_j$,
\[
\mathbb{P} \left[ |E(G_i) \cap D_j| \leq \frac{|D_j|}{t} - \sqrt{\frac{3n}{t} \log n} \right] \leq \frac{1}{n^3}.
\]
As a consequence, with probability at least $1 - \frac{1}{n}$ every subgraph $G_i$ has at least one edge from each of the pseudocolor classes $D_1, \ldots, D_{n-1}$.

Proof. Fix a subgraph $G_i$ and a pseudocolor class $D_j$. The expected number of edges in $G_i$ from the pseudocolor class is $\frac{|D_j|}{t}$. By Bernstein’s inequality where $\lambda = \sqrt{\frac{3n}{t} \log n}$, we have
\[
\mathbb{P} \left[ |E(G_i) \cap D_j| \leq \frac{|D_j|}{t} - \sqrt{\frac{3n}{t} \log n} \right] \leq \exp \left( \frac{-\lambda^2}{2|D_j|n/4} \right) \leq \exp \left( \frac{-\lambda^2}{2|D_j|n/4} \right) = \frac{1}{n^3}.
\]
Since $|D_j| \geq n/4$, $n \geq 1,000$, and $C \geq 50$,
\[
\frac{|D_j|}{t} - \sqrt{\frac{3n}{t} \log n} \geq \frac{n}{4t} - \sqrt{\frac{3n}{t} \log n} \geq 1.
\]
The second statement follows from the previous inequalities by using the union sum bound for the $n - 1$ pseudocolor classes and $t$ subgraphs and recalling that $t < n$. \square

Lemma 6 shows that if we consider a partition $\pi$ of $V(G)$ into $s$ parts, where $s = n$ there must be at least $n - 1$ colors in $G_i$ between the parts of $\pi$. In the case when the partition has $s = n - 1$ parts there is at most one edge inside the parts of $\pi$, so there are at least $n - 2$ colors in $G_i$ between the parts of $\pi$.

4 Partitions where $\left(1 - \frac{14}{\sqrt{C}}\right)n \leq s \leq n - 2$

In this section we consider partitions $\pi$ of $V(G)$ into $s$ parts where $\left(1 - \frac{14}{\sqrt{C}}\right)n \leq s \leq n - 2$. First, we introduce a new function that will help with our calculations. The function $f$ will be used to bound the probability that $q - (s - 2)$ colors do not appear between the parts of $\pi$ in $G_i$.

Lemma 7. For an integer $\ell$ and real numbers $c_1, \ldots, c_q$, define
\[
f(c_1, \ldots, c_q; \ell) = \sum_{I \subseteq \{q\}} \exp \left( -\frac{1}{\ell} \sum_{j \in I} c_j \right).
\]
If $1 \leq c_j \leq \frac{n}{2}$ for each $j$, $\sum_{I=1}^q c_j = \binom{n}{2}$, and $\frac{n}{2} \leq \ell \leq n - 4$, then
\[
f(c_1, \ldots, c_q; \ell) \leq \exp \left( -\frac{49C}{200} (n - \ell) \log n \right).
\]
Proof.
For convenience we define $w(I) = \sum_{j \in I} c_j$ for a subset $I \subseteq \{q\}$.

Claim 1. $f(c_1, \ldots, c_q; \ell) \leq f\left(1,1,\ldots,1,x^*,\frac{n}{2},\ldots,\frac{n}{2};\ell\right)$, where $1 \leq x^* < \frac{n}{2}$, and where $k$ and $x^*$ are so that $(k - 1) + (q - k)\frac{n}{2} + x^* = \binom{n}{2}$.
Proof of Claim 1. Since $f(c_1, \ldots, c_q; \ell)$ is a symmetric function in the $c_j$'s, it suffices to show that when $c_2 \geq c_1$,

$$f(c_1, c_2, \ldots, c_q; \ell) \leq f(c_1 - \epsilon, c_2 + \epsilon, \ldots, c_q; \ell),$$

where $\epsilon = \min\{c_1 - 1, \frac{q}{2} - c_2\}$.

$$f(c_1, c_2, \ldots, c_q; \ell) = \sum_{I \in ([q] \setminus \{1, 2\})} \exp\left(-\frac{w(I)}{t}\right) + \sum_{I \in ([q] \setminus \{1, 2\})} \exp\left(-\frac{c_1 - c_2}{t} - \frac{w(I)}{t}\right) + \sum_{I \in ([q] \setminus \{1, 2\})} \left(\exp\left(-\frac{c_1}{t} - \frac{w(I)}{t}\right) + \exp\left(-\frac{c_2}{t} - \frac{w(I)}{t}\right)\right)$$

The first two summations are unchanged in $f(c_1 - \epsilon, c_2 + \epsilon, \ldots, c_q; \ell)$, and hence it suffices to show that for every $I \in ([q] \setminus \{1, 2\})$,

$$\exp\left(-\frac{c_1}{t} - \frac{w(I)}{t}\right) + \exp\left(-\frac{c_2}{t} - \frac{w(I)}{t}\right) \leq \exp\left(-\frac{(c_1 - \epsilon)}{t} - \frac{w(I)}{t}\right) + \exp\left(-\frac{(c_2 + \epsilon)}{t} - \frac{w(I)}{t}\right).$$

This follows immediately by Jensen's inequality and the convexity of $\exp(\alpha x + \beta)$ as a function in $x$. 

Claim 2. 

$$f\left(1, 1, \ldots, 1, x^*, \frac{n}{2}, \ldots, \frac{n}{2}; \ell\right) \leq f\left(1, \ldots, 1, \frac{n}{2}, \ldots, \frac{n}{2}; \ell\right),$$

where \(\frac{n(n-2)}{2} \leq k + (q-k)\frac{n}{2} \leq \binom{n}{2}\).

Proof of Claim 2. The function $f$ is decreasing in each $c_j$, and in particular $c_k$. 

Now consider

$$f\left(1, \ldots, 1, \frac{n}{2}, \ldots, \frac{n}{2}; \ell\right) = \sum_{I \in ([q] \setminus \{0, \ell\})} \exp\left(-\frac{1}{t}w(I)\right)$$

$$\leq \sum_{r=\max\{0, \ell-(q-k)\}}^{\min\{\ell, k\}} k^r(q-k)^{q-k-(\ell-r)} \exp\left(-\frac{1}{t}\left(n(n-2)/2 - (\ell-r)(n/2 - r)\right)\right)$$

$$\leq \sum_{r=\max\{0, \ell-(q-k)\}}^{\min\{\ell, k\}} k^r(q-k)^{q-k-(\ell-r)} \exp\left(-\frac{1}{t}\left(\frac{n}{2}(n-\ell+r-2) - r\right)\right)$$

$$\leq \sum_{r=\max\{0, \ell-(q-k)\}}^{\min\{\ell, k\}} \exp\left(\log n - \frac{1}{t}\left(n - \ell + 2r + C\right) \right) \text{ by (*)}$$

$$\leq n \exp\left(\log n \left(q - k + 2r + C\right) + r\left(2 - C + C/n\right) + C\right)$$

5
\[ \leq \exp \left( \log n \left( (n - \ell) \left( \frac{1 - C}{2} \right) + C + 1 \right) \right) \]

Since \( n - \ell \geq 4 \) and \( C \geq 250 \), we have
\[ 1 + \frac{1}{n - \ell} \leq \frac{C}{200} \leq C \left( \frac{1}{2} - \frac{1}{n - \ell} \right) - 49 \frac{C}{200}. \]
Thus the sum above is bounded by
\[ \exp \left( -\frac{49C}{200} (n - \ell) \log n \right) . \]

**Lemma 8.** Let \( \Pi \) be the set of partitions of \( V(G) \) into \( s \) parts, where \( \left( 1 - \frac{14}{\sqrt{C}} \right) n \leq s \leq n - 2 \). For a partition \( \pi \in \Pi \), let \( B_{\pi,i} \) be the event that there are less than \( s - 1 \) colors between the parts of \( \pi \) in \( G_i \). Then
\[ \mathbb{P} \left[ \bigcup_{i=1}^{t} \bigcup_{\pi \in \Pi} B_{\pi,i} \right] \leq \frac{1}{n}. \]

**Proof.** Fix a subgraph \( G_i \) and a partition \( \pi \in \Pi \). Recall that \( C_1, \ldots, C_q \) are the color classes of \( G \) with sizes \( c_1, \ldots, c_q \), respectively. Let \( I_{\pi,i} \) be the set of colors that do not appear on edges of \( G_i \) between the parts of \( \pi \).

The total number of edges in \( G \) that have a color indexed by \( I_{\pi,i} \) is \( \sum_{j \in I_{\pi,i}} c_j \). By convexity of \( \binom{n}{s} \), there are at most \( \binom{n - s + 1}{2} \) edges inside the parts of \( \pi \). Note that if \( I_{\pi,i} \) does not have size \( q - (s - 2) \), then it contains a set \( I' \subseteq I_{\pi,i} \) of size \( q - (s - 2) \), and the event that no edges of \( G_i \) between the parts of \( \pi \) have colors in \( I_{\pi,i} \) is contained in the event that no edges of \( G_i \) between the parts have colors in \( I' \). Thus,
\[ \mathbb{P} [ B_{\pi,i} ] \leq \sum_{I \in \binom{[n]}{q} - \binom{[n]}{q - (s - 2)}} \left( 1 - \frac{1}{\ell} \right)^{\sum_{j \in I} c_j} \binom{n - s + 1}{2} \]
\[ \leq f(c_1, c_2, \ldots, c_q; s - 2) \left( 1 - \frac{1}{\ell} \right)^{\binom{n - s + 1}{2}} \]
\[ \leq f(c_1, c_2, \ldots, c_q; s - 2) \exp \left( \frac{1}{\ell} \binom{n - s + 1}{2} \right) \]
\[ \leq \exp \left( -\frac{49C}{200} (n - (s - 2)) \log n + \frac{(n - s + 1)^2}{2} \right) \text{ by Lemma 7.} \]

Since \( s \geq \left( 1 - \frac{14}{\sqrt{C}} \right) n \), we know \( n - s + 1 \leq \frac{14n}{\sqrt{C}} + 1 \). Thus we can bound the previous line by
\[ \leq \exp \left( (n - s + 1) \left( -\frac{49C}{200} \log n + \frac{1}{2\ell} \left( \frac{14}{\sqrt{C}} \right) n + 1 \right) \right) \]
\[ \leq \exp \left( (n - s + 1) \log n \left( -\frac{49C}{200} + \frac{n}{n - C \log n} \left( \frac{14\sqrt{C}}{2} + \frac{C}{2n} \right) \right) \right) \text{ by (s).} \]

We now perform a union bound over all partitions \( \pi \in \Pi \). The number of partitions of \( V(G) \) into \( s \) nonempty parts is at most
\[ \binom{n}{s} s^{n-s} \leq \binom{n}{n-s} n^{n-s} \leq n^{2(n-s)} = \exp(2(n-s) \log n) \leq \exp(2(n-s+1) \log n) . \]
Therefore,
\[
\mathbb{P} \left[ \bigcup_{\pi \in \Pi \text{ with } s \text{ parts}} B_{\pi,i} \right] \leq \exp \left( (n - s + 1) \log n \left( 2 - \frac{49C}{200} + \frac{n}{n - C \log n} \left( \frac{14\sqrt{C}}{2} + \frac{C}{2n} \right) \right) \right)
\]

Since \( C = 1000 \) and \( n \geq 1,000,000 \), we have
\[
2 - \frac{49C}{200} + \frac{n}{n - C \log n} \left( \frac{14\sqrt{C}}{2} + \frac{C}{2n} \right) \leq -1,
\]
and since \((n - s + 1) \geq 3\),
\[
\mathbb{P} \left[ \bigcup_{\pi \in \Pi \text{ with } s \text{ parts}} B_{\pi,i} \right] \leq \exp (-3 \log n) = \frac{1}{n^3}.
\]

This gives a bound on the probability for a fixed partition size \( s \). Using the union sum bound over all partition sizes \( s \), where \((1 - \frac{14}{\sqrt{C}}) n \leq s \leq n - 2\), and over all \( t \) subgraphs completes the proof.

This proves when \( s \) is large there are enough colors between the parts.

## 5 Partitions where \( 2 \leq s \leq \left( 1 - \frac{14}{\sqrt{C}} \right) n \)

Next, we prove several results that will be used to show there are enough colors in \( G_i \) between the parts of the partition when the number of parts is small. Our goal is to show that for a partition \( \pi \) of \( V(G) \) into \( s \) parts, the number of edges between the parts in \( G_i \) is so large that there must be at least \( s - 1 \) colors between the parts.

**Lemma 9.** For a fixed subgraph \( G_i \) and color \( j \),
\[
\mathbb{P} \left[ |E(G_i) \cap C_j| \geq \frac{n}{2t} + 4\sqrt{\frac{n}{t} \log n} \right] \leq \frac{1}{n^3}.
\]

As a consequence, with probability at least \( 1 - \frac{1}{n} \), every color appears at most \( \frac{n}{2t} + 4\sqrt{\frac{n}{t} \log n} \) times in every \( G_i \).

**Proof.** Fix a color \( j \) and a subgraph \( G_i \). Order the edges of \( C_j \) as \( e_1, \ldots, e_{c_j} \). For \( 1 \leq k \leq c_j \), let \( X_k \) be the indicator random variable for the event \( e_k \in E(G_i) \). For a color class with size less than \( \frac{n}{2t} \) we introduce dummy random variables, so we can apply Bernstein’s inequality. For \( c_j + 1 \leq k \leq n/2 \), let \( X_k \) be a random variable distributed independently as a Bernoulli random variable with probability \( 1/t \).

By construction, \( |E(G_i) \cap C_j| \leq X = \sum_{k=1}^{n/2} X_k \) and \( \mathbb{E}[X] = \frac{n}{2t} \). By Bernstein’s Inequality where \( \lambda = 4\sqrt{\frac{n}{t} \log n} \), we have
\[
\mathbb{P} \left[ |E(G_i) \cap C_j| \geq \frac{n}{2t} + 4\sqrt{\frac{n}{t} \log n} \right] \leq \mathbb{P} \left[ X \geq \frac{n}{2t} + 4\sqrt{\frac{n}{t} \log n} \right] \leq \exp \left( -\frac{16n}{t} \log n \right)
\]
\[
\leq \exp \left( -\frac{16n}{t} \log n / \left( \frac{n}{2t} + 4\sqrt{\frac{n}{t} \log n} \right) \right)
\]
\[
\sum_{\text{inequality 4. Since sum for all partitions. If}} \sum_{\text{for all partitions}} \text{to find a lower bound on the number of edges between the parts for a partition } \pi
\]

Applying the union sum bound for the

\[\text{So}
\]

\[\text{Fix a subgraph } G
\]

\[\text{Proof. Fix a subgraph } G_i \text{ and a set of vertices } S \subseteq V(G). \text{ Let } r = |S|. \text{ The expected number of edges in } G_i \text{ between } S \text{ and } \overline{S} \text{ is } r(n-r)/t. \text{ By Bernstein’s inequality with } \lambda = \sqrt{6\frac{r(n-r)}{t}} \min\{r, n-r\} \log n, \text{ we have}
\]

\[\Pr[\mathcal{B}_{S,i}] \leq \exp\left(-\frac{6(n-r)}{t} \min\{r, n-r\} \log n \right) = n^{-3 \min\{r, n-r\}}.
\]

So

\[\Pr\left[\bigcup_{S \subseteq V(G)} \mathcal{B}_{S,i}\right] \leq \sum_{r=1}^{n/2} \binom{n}{r} n^{-3r} + \sum_{r=n/2}^{n} \binom{n}{n-r} n^{-3(n-r)} = 2 \sum_{r=1}^{n/2} \binom{n}{r} n^{-3r}
\]

\[\leq 2 \sum_{r=1}^{n/2} n^{-2r} \leq 2n^{-2} + 2 \left(\sum_{r=2}^{n} n^{-4}\right) \leq \frac{2}{n^2} + \frac{2}{n^3} \leq \frac{4}{n^2}.
\]

Applying the union sum bound for the \(t\) subgraphs gives the final statement of the lemma. \(\square\)

The previous lemma gives a lower bound on the number of edges between \(S\) and \(\overline{S}\). We use this lemma to find a lower bound on the number of edges between the parts for a partition \(\pi = \{P_1, \ldots, P_s\}\) of \(V(G)\).

**Definition 11.** For \(x \in [0, n]\), let

\[f(x) = \frac{x(n-x)}{t} - \sqrt{\frac{6x(n-x)}{t}} \min\{x, n-x\} \log n.
\]

If none of the bad events \(\mathcal{B}_{S,i}\) from Lemma 10 occur, then the sum \(\frac{1}{2} \sum_{\pi=(P_1, \ldots, P_s)} f(|P_i|)\), where \(\sum_{i=1}^{s} |P_i| = n\), is a lower bound on the number of edges between the parts of the partition \(\pi\). We bound this sum for all partitions. If \(-f(x)\) is not convex then we could immediately find a lower bound by using Jensen’s inequality 4. Since \(-f(x)\) is not convex, we bound it with a function that is convex.
Let $h(x)$ be a function with domain $[a, b]$. We say a function $h$ is concave down if for $x, y \in [a, b]$ and $\lambda \in [0, 1]$, then $h(\lambda x + (1 - \lambda)y) \geq \lambda h(x) + (1 - \lambda)h(y)$. First, we present two basic results about concave down functions.

**Lemma 12.** Let $h(x)$ be a differentiable function with domain $[a, b]$. Suppose that $h$ is concave down on $[z, b]$, where $z \in (a, b)$. Let $\ell(x)$ be the line tangent to $h$ at the point $(z, h(z))$. Then the function

$$h_1(x) = \begin{cases} \ell(x) & \text{if } a \leq x \leq z, \\ h(x) & \text{if } z < x \leq b \end{cases}$$

is concave down.

**Proof.** Let $y_1, y_2 \in [a, b]$ where $y_1 \leq y_2$, and $\lambda \in [0, 1]$. If $y_1$ and $y_2$ are both in $[a, z]$ or $[z, b]$ then

$$h_1(\lambda y_1 + (1 - \lambda)y_2) \geq \lambda h_1(y_1) + (1 - \lambda)h_1(y_2),$$

since $\ell$ and $h$ are both concave down.

Consider the case when $y_1 \in [a, z]$ and $y_2 \in (z, b]$. Let $\lambda \in [0, 1]$ and $w = \lambda y_1 + (1 - \lambda)y_2$. Let $b$ be the $y$-intercept of the line $\ell(x)$, i.e. $\ell(x) = h'(z)x + b$. Since $h$ is concave down on the interval $[z, b]$, we know that $h$ lies below the tangent line $\ell(x)$ on the interval $[a, b]$. In particular, $h'(z)y_2 + b \geq h(y_2)$. Let $\epsilon = h'(z)y_2 + b - h(y_2) \geq 0$.

If $w \leq z$, then we want to show that $h_1(w) = h'(z)(w - y_1) + \ell(y_1) \geq \frac{h(y_2) - h(y_1)}{y_2 - y_1}(w - y_1) + \ell(y_1)$. Note that it is enough to show that $h'(z) \geq \frac{h(y_2) - h(y_1)}{y_2 - y_1}$. Since $\epsilon \geq 0$, we have

$$h'(z) \geq h'(z) \frac{y_2 - y_1 - \epsilon}{y_2 - y_1} = \frac{(h'(z)y_2 + b - \epsilon) - (h'(z)y_1 + b)}{y_2 - y_1} = \frac{h(y_2) - \ell(y_1)}{y_2 - y_1}.$$

Suppose $w > z$. The line between $z$ and $y_2$ is given by $\frac{h(y_2) - h(z)}{y_2 - z}(w - z) + h(z)$, and by the concavity of $h(x)$ on $[z, b]$ we know $h_1(w) \geq \frac{h(y_2) - h(z)}{y_2 - z}(w - z) + h(z)$. We want to show $\frac{h(y_2) - h(z)}{y_2 - z}(w - z) + h(z) \geq \frac{h(y_2) - h(z)}{y_2 - y_1}(w - y_1) + h(z)$. It is enough to show that $\frac{h(y_2) - h(z)}{y_2 - z} \geq \frac{h(y_2) - h(z)}{y_2 - y_1}$. We know $y_1 < z$ and $\epsilon \geq 0$. Thus

$$\epsilon y_1 \leq \epsilon z \leq (h'(z)y_2 - h'(z)y_2z - h'(z)y_1y_2 + h'(z)zy_1 - \epsilon y_2) + \epsilon y_1 \leq (h'(z)y_2 - h'(z)y_2z - h'(z)y_1y_2 + h'(z)zy_1 - \epsilon y_2) + \epsilon z \leq (h'(z)y_2 + b - \epsilon) - (h'(z)y_1 + b) \leq \frac{h(y_2) - h(z)}{y_2 - z} \leq \frac{h(y_2) - \ell(y_1)}{y_2 - y_1}.$$

**Lemma 13.** If $h_1$ and $h_2$ are concave down functions, then $h(x) = \min\{h_1(x), h_2(x)\}$ is concave down.

**Proof.** For every $x, y$ and $\lambda \in [0, 1]$ we have

$$h(\lambda x + (1 - \lambda)y) = \min\{h_1(\lambda x + (1 - \lambda)y), h_2(\lambda x + (1 - \lambda)y)\} \leq \lambda \min\{h_1(x), h_2(x)\} + (1 - \lambda) \min\{h_1(y), h_2(y)\} = \lambda h_1(x) + (1 - \lambda)h_2(x).$$
We next define several functions that will lead to a concave down lower bound for the function $f$. Define on $[0,n]$ the functions

$$f_1(p) = \frac{x(n-x)}{t} - x\sqrt{\frac{6(n-x)}{t} \log n},$$

$$f_2(p) = \frac{x(n-x)}{t} - (n-x)\sqrt{\frac{6x}{t} \log n}.$$ 

Note that

$$f(x) = \begin{cases} f_1(x) & 0 \leq x \leq n/2, \\ f_2(x) & n/2 < x \leq n. \end{cases}$$

Let $\ell(x) = f_2'(x)(x-n/2) - f_2(n/2)$ be the tangent line of $f_2(x)$ at the point $\left(\frac{n}{2}, \frac{n^2}{14} - \frac{n}{2} \sqrt{\frac{3n}{14} \log n}\right)$. Let $c$ be the point such that $f_1(x)$ achieves its maximum value on the interval $[0,n]$. Define

$$f_3(p) = \begin{cases} \ell(x) & 0 \leq x \leq n/2, \\ f_2(x) & n/2 < x \leq n. \end{cases}$$

and

$$f_4(x) = \begin{cases} f_1(x) & 0 \leq x \leq c, \\ f_1(c) & c < x \leq n. \end{cases}$$

By Lemma 12 the functions $f_3$ and $f_4$ are concave down.

On the interval $[0,n]$ define $f_5(x) = \min\{f_3(x), f_4(x)\}$. The function $f_5(x)$ is concave down by Lemma 13, where $f(x) \geq f_5(x)$ for all $x \in [0,n]$. Figure 1 shows the functions $f(x)$ and $\ell(x)$ used to create $f_5(x)$.

**Lemma 14.** The sum $\sum_{i=1}^s f(x_i)$, where $\sum_{i=1}^s x_i = n$ and $x_i \geq 1$ for all $i$, is bounded below by

$$\sum_{i=1}^s f(x_i) \geq (s-1)f(1) + f(n-s+1).$$

**Proof.** The proof is broken up into two cases based on whether $s \leq n/2$, or $s > n/2$.

When $s \leq n/2$ the function $f(x) \geq f_5(x)$, so $\sum_{i=1}^s f(x_i) \geq \sum_{i=1}^s f_5(x_i)$. Since the function $f_5(x)$ is concave down the sum $\sum_{i=1}^s f_5(x)$ is minimized when there is one part of size $n-s+1$ and all the other parts are of size 1. Since $n-s+1 \geq n/2$, we have $f_5(n-s+1) = f(n-s+1)$. Note that $\ell(1) \geq f_1(1)$, which implies $f_5(1) = f(1)$. Thus

$$\sum_{i=1}^s f(x_i) \geq \sum_{i=1}^s f_5(x_i) \geq (s-1)f_5(1) + f_5(n-s+1) = (s-1)f(1) + f(n-s+1).$$

When $s > n/2$, we have $x_i \leq n/2$ for all $i$. Therefore $f(x_i) = f_1(x_i)$ for all $i$. Since $f_1(x)$ is concave down the sum is minimized when one parts has size $n-s+1$ and the rest have size 1.

**Lemma 15.** Let $\pi$ be a partition of the vertices of $G$ into $s$ parts. Suppose none of the events $B_{S,i}$ from Lemma 10 hold for all $S \subseteq V(G)$ and $1 \leq i \leq t$. Then in each of the subgraphs $G_1, \ldots, G_t$, the number of edges between the parts of $\pi$ is at least

$$\frac{1}{2} \left( (s-1) \left( \frac{n-1}{t} - \sqrt{\frac{6(n-1)}{t} \log n} \right) + \frac{(n-s+1)(s-1)}{t} - (s-1)\sqrt{\frac{6(n-s+1)}{t} \log n} \right)$$

when $s \leq n/2$, and

$$\frac{1}{2} \left( (s-1) \left( \frac{n-1}{t} - \sqrt{\frac{6(n-1)}{t} \log n} \right) + \frac{(n-s+1)(s-1)}{t} - (n-s+1)\sqrt{\frac{6(s-1)}{t} \log n} \right)$$

when $s > n/2$. 

10
Figure 1: The function \( f(x) \), along with the line \( \ell(x) \).

**Proof.** If none of the events \( B_{S,i} \) hold then the sum \( \frac{1}{2} \sum_{\pi=(P_1,\ldots,P_s)} f(x) \) where \( \sum_{i=1}^{s} |P_i| = n \) is a lower bound on the number of edges between the parts of \( \pi \). By Lemma 14 we know this sum is bounded below by \( \frac{1}{2} ((s-1)f(1) + f(n-s+1)) \).

**Lemma 16.** Let \( \pi \) be a partition of the vertices of \( G \) into \( s \) parts, where \( 2 \leq s \leq \left( 1 - \frac{14}{\sqrt{C}} \right) n \). Suppose none of the events \( B_{S,i} \) from Lemma 10 hold for all \( S \subseteq V(G) \) and \( 1 \leq i \leq t \), and every color appears in each \( G_i \) at most \( \frac{n}{2t} + 4\sqrt{\frac{n}{t}} \log n \) times (as in Lemma 9). Then in each of the subgraphs \( G_1,\ldots,G_t \), the number of colors between the parts of \( \pi \) is at least \( s-1 \).

**Proof.** Suppose there exists a subgraph \( G_i \) and a partition \( \pi \) into \( s \) parts where there are at most \( s-2 \) colors between the parts in \( G_i \). Then by assumption there are at most

\[
(s-2) \left( \frac{n}{2t} + 4\sqrt{\frac{n}{t}} \log n \right)
\]

edges in \( G_i \) between the parts of \( \pi \). We will show that the number of edges between the parts of \( \pi \) can not be this small, giving a contradiction.

Suppose \( \frac{n}{2} < s \leq \left( 1 - \frac{14}{\sqrt{C}} \right) n \). By Lemma 15 there are at least

\[
\frac{1}{2} \left( (s-1) \left( \frac{n-1}{t} - \sqrt{6(n-1) \frac{\log n}{t}} \right) + \frac{(n-s+1)(s-1)}{t} - (n-s+1) \sqrt{6(s-1) \frac{\log n}{t}} \right)
\]
edges in $G_i$ between the parts of $\pi$. If $\pi$ has at most $s - 2$ colors in $G_i$ between the parts, then

$$(s - 2) \left( \frac{n}{2t} + 4 \sqrt{\frac{n}{t} \log n} \right) \geq \frac{s - 1}{2} \left( \frac{n - 1}{t} - \sqrt{6(n - 1) \log n \over t} + \frac{(n - s + 1)}{t} - (n - s + 1) \sqrt{6 \log n \over (s - 1)t} \right).$$

Rearranging we have

$$\frac{s - 2}{s - 1} \left( \frac{n}{t} + 8 \sqrt{\frac{n}{t} \log n} \right) + \frac{1}{t} + \sqrt{6(n - 1) \log n \over t} + (n - s + 1) \sqrt{6 \log n \over (s - 1)t} \geq \frac{n}{t} + \frac{(n - s + 1)}{t}.$$

We will give an upper bound to the left side and a lower bound to the right side that give a contradiction.

Since $s$ is an integer and $n/2 < s$, we have

$$(n - s + 1) \sqrt{6 \over n(s - 1)} \leq \sqrt{12 \over n^2} = \sqrt{3}. \quad (\dagger)$$

Therefore

$$\frac{s - 2}{s - 1} \left( \frac{n}{t} + 8 \sqrt{\frac{n}{t} \log n} \right) + \frac{1}{t} + \sqrt{6(n - 1) \log n \over t} + (n - s + 1) \sqrt{6 \log n \over (s - 1)t}$$

$$\leq \sqrt{C} \log n \left( \frac{n}{n - C \log n} \right) \left( \sqrt{C} + \sqrt{C \over n} + \sqrt{n - C \log n \over n} \left( 8 + \sqrt{6(n - 1) \over n} + (n - s + 1) \sqrt{6 \over n(s - 1)} \right) \right)$$

$$\leq \sqrt{C} \log n \left( \frac{n}{n - C \log n} \right) \left( \sqrt{C} + \sqrt{C \over n} + \sqrt{n - C \log n \over n} \left( 8 + \sqrt{6 + \sqrt{3}} \right) \right) \quad \text{by} \ (\dagger)$$

Since $C = 1000$ and $n \geq 1,000,000$, $n \log n \over n - C \log n \leq 1.02$ and $n \log n \over n - C \log n \leq 1.01$. Thus the term above is bounded above by

$$\sqrt{C} \log n \left( 1.02 \sqrt{C} + \frac{1.02 \sqrt{C}}{n} + 1.01(8 + \sqrt{6 + \sqrt{3}}) \right) \leq \sqrt{C} \log n \left( 1.02 \sqrt{C} + 12.31 \right)$$

We next bound the right side. By $(\ast)$ we have $\frac{1}{t} \geq \frac{1}{C \log n \over n}$, and since $s \leq \left( 1 - \frac{14}{\sqrt{C}} \right) n$, so

$$\frac{n}{t} + \frac{(n - s + 1)}{t} \geq C \log n + C \log n \frac{n - s + 1}{n} \geq C \log n + C \log n \frac{14}{\sqrt{C}} = \sqrt{C} \log n (\sqrt{C} + 14).$$

When $C = 1000$ and $n \geq 1,000,000$ we have $\sqrt{C} + 14 > 1.02 \sqrt{C} + 12.31$, which gives a contradiction. So, there must be at least $s - 1$ colors in $G_i$ between the parts of $\pi$ when $n/2 < s \leq \left( 1 - \frac{14}{\sqrt{C}} \right) n$.

Suppose $2 \leq s \leq n/2$. By Lemma 15 there are at least

$$\frac{1}{2} \left( (s - 1) \left( \frac{n - 1}{t} - \sqrt{6(n - 1) \log n \over t} \right) + \frac{(n - s + 1)(s - 1)}{t} - (s - 1) \sqrt{6(n - s + 1) \log n \over t} \right)$$

edges in $G_i$ between the parts of $\pi$. If $\pi$ has at most $s - 2$ colors in $G_i$ between the parts then

$$(s - 2) \left( \frac{n}{2t} + 4 \sqrt{\frac{n}{t} \log n} \right) \geq \frac{(s - 1)}{2} \left( \frac{n - 1}{t} - \sqrt{6(n - 1) \log n \over t} + \frac{(n - s + 1)}{t} - \sqrt{6(n - s + 1) \log n \over t} \right).$$
Rearranging we have
\[
\frac{s - 2}{s - 1} \left( \frac{n}{t} + 8 \sqrt{\frac{n}{t} \log n} \right) + \frac{1}{t} \geq \frac{1}{s} \left( \frac{n}{t} + 6(n - 1) \frac{\log n}{t} + 6(n - s + 1) \frac{\log n}{t} \right) \geq \frac{n}{t} + \frac{(n - s + 1)}{t}.
\]

Using \( \frac{1}{t} \leq \frac{C \log n}{n - C \log n} \) from (*), we have
\[
\frac{s - 2}{s - 1} \left( \frac{n}{t} + 8 \sqrt{\frac{n}{t} \log n} \right) + \frac{1}{t} + \frac{6(n - 1) \log n}{t} + \frac{6(n - s + 1) \log n}{t} \\
\leq \sqrt{C} \log n \left( \frac{n}{n - C \log n} \right) \left( \sqrt{C} + \frac{\sqrt{C}}{n} + \frac{n - C \log n}{n} \left( 8 + \sqrt{\frac{6(n - 1)}{n} + \frac{6(n - s + 1)}{n}} \right) \right).
\]

Since \( C = 1000 \) and \( n \geq 1,000,000 \), \( \frac{n}{n - C \log n} \leq 1.02 \) and \( \sqrt{\frac{n}{n - C \log n}} \leq 1.01 \). Thus the term above is bounded above by
\[
\sqrt{C} \log n \left( 1.02 \sqrt{C} + \frac{1.02 \sqrt{C}}{n} + 1.01 \left( 8 + 2 \sqrt{6} \right) \right) \leq \sqrt{C} \log n \left( 1.02 \sqrt{C} + 13.1 \right).
\]

Bounding the right side using \( \frac{1}{t} \geq \frac{C \log n}{n} \) from (*), and \( s \leq \frac{n}{2} \), we have
\[
\frac{n}{t} + \frac{(n - s + 1)}{t} \geq C \log n + C \log n \frac{(n - s + 1)}{n} \geq C \log n + C \log n \frac{n}{n} = \sqrt{C} \log n \left( \frac{3 \sqrt{C}}{2} \right).
\]

Again, when \( C = 1000 \) and \( n \geq 1,000,000 \) we have \( \frac{3 \sqrt{C}}{2} > 1.02 \sqrt{C} + 13.1 \) which leads to a contradiction. Thus, there must be at least \( s - 1 \) colors in \( G_i \) between the parts of \( \pi \) when \( 2 \leq s \leq \frac{n}{2} \).

\[\square\]

6 Main Result

**Theorem 1.** Let \( G \) be an edge-colored copy of \( K_n \), where each color appears on at most \( n/2 \) edges and \( n \geq 1,000,000 \). The graph \( G \) contains at least \( \lfloor n/(1000 \log n) \rfloor \) edge-disjoint rainbow spanning trees.

**Proof.** Recall that \( t = n/(C \log n) \) where \( C = 1000 \). We perform the random experiment of decomposing the edges of \( G \) into \( t \) edge-disjoint subgraphs \( G_i \) by independently and uniformly selecting each edge of \( G \) to be in the subgraph \( G_i \) with probability \( 1/t \). With probability at least \( 1 - \frac{2}{n} \) none of the bad events from Lemmas 6, 8, 9, and 10 occur in any of the subgraphs \( G_i \). Henceforth let \( G_1, \ldots, G_t \) be fixed subgraphs where none of these bad events occur.

We want to show that each \( G_i \) has a rainbow spanning tree. By Theorem 2 it is enough to show that for every partition \( \pi \) of \( V(G) \) into \( s \) parts, there are at least \( s - 1 \) different colors appearing on the edges of \( G_i \) between the parts of \( \pi \).

By Lemma 6, every \( G_i \) has at least one edge from each of the \( n - 1 \) pseudocolor classes. When \( s = n \) there must be at least \( n - 1 \) colors in \( G_i \) between the parts of \( \pi \). When \( s = n - 1 \) there is at most one edge inside the parts of \( \pi \), so there are at least \( n - 2 \) colors in \( G_i \) between the parts of \( \pi \).

If \( \left( 1 - \frac{1}{\sqrt{2}} \right) n \leq s \leq n - 2 \), then by Lemma 8 every partition \( \pi \) of \( V(G) \) into \( s \) parts has at least \( s - 1 \) colors in \( G_i \) between the parts, for every subgraph \( G_1, \ldots, G_t \).
Finally, we assume that $s \leq \left(1 - \frac{14}{\sqrt{\pi}}\right) n$. When $s = 1$ there are zero colors between the parts, so the condition is vacuously true. So suppose $2 \leq s \leq \left(1 - \frac{14}{\sqrt{\pi}}\right) n$. Since Lemmas 9 and 10 hold, by Lemma 16 the number of colors between the parts of $\pi$ is at least $s - 1$ for every subgraph $G_1, \ldots, G_t$.

Therefore all of the subgraphs $G_1, \ldots, G_t$ contain a rainbow spanning tree, and so $G$ contains at least $t = \lfloor n/(1000 \log n) \rfloor$ edge-disjoint rainbow spanning trees.

\section*{Acknowledgements}

The authors thank Douglas B. West and the Research Experience for Graduate Students (REGS) at the University of Illinois at Urbana-Champaign for their support and hospitality. Funding for the authors’ visit in the summer 2012 was provided by the UIUC Department of Mathematics through National Science Foundation grant DMS 08-38434, “EMSW21-MCTP: Research Experience for Graduate Students”.

\section*{References}


