AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **86(1)** (2023), Pages 1–23

An extremal problem on rainbow spanning trees in graphs

Matthew DeVilbiss[1] Bradley Fain[2] Amber Holmes[3]

Paul Horn[4] Sonwabile Mafunda[5] K.E. Perry[6]

Abstract

A subgraph of an edge-colored graph is rainbow provided that no color appears on more than one edge. In this paper we consider the natural extremal problem of maximizing and minimizing the number of rainbow spanning trees in a graph G. Such a question clearly needs restrictions on the colorings to be meaningful. For edge-colorings using $n-1$ colors and without rainbow cycles, known in the literature as JL-colorings, there turns out to be a particularly nice way of counting the rainbow spanning trees and we solve this problem completely for JL-colored complete graphs K_n and complete bipartite graphs $K_{n,m}$. In both cases, we find tight upper and lower bounds; the lower bound for K_n, in particular, proves to have an unexpectedly chaotic and interesting behavior. We further investigate this question for JL-colorings of general graphs and prove several results including characterizing graphs which have JL-colorings achieving the lowest possible number of rainbow spanning trees. We establish other results for general $n-1$ colorings, including providing an analogue of Kirchoff's matrix tree theorem which yields a way of counting rainbow spanning trees in a general graph G.

[1]Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, U.S.A. devilbiss.16@osu.edu. This author's research supported in part by NSF grant DMS-1343651.
[2]Department of Mathematics, University of Delaware, U.S.A. bfain@udel.edu. This author's research supported in part by NSF grant DMS-1343651.
[3]Department of Mathematics, University of Kentucky, U.S.A. Amber.Holmes@uky.edu. This author's research partially supported by NSF grant DMS-1343651.
[4]Department of Mathematics, University of Denver, U.S.A. paul.horn@du.edu. This author's research partially by NSF grant DMS-1343651, Simons Collaboration Grant 525039, and a University of Denver Internationalization Grant.
[5]Department of Mathematics and Applied Mathematics, University of Johannesburg, South Africa. smafunda@uj.ac.za. This author's research partially supported by NSF grant DMS-1343651 and the National Research Foundation in South Africa.
[6]Mathematics, Soka University of America, U.S.A. kperry@soka.edu. This author's research partially supported by NSF grant DMS-1343651 and a University of Denver Internationalization Grant.

ISSN: 2202-3518

1 Introduction

Let G be an edge-colored simple graph with $|V(G)| = n$ and note that the edge-coloring need not be proper. A rainbow spanning tree (RST) in G is an acyclic, connected, spanning subgraph such that no color appears on more than one edge. Given a coloring $\varphi : E(G) \to \mathbb{N}$ let

$$\mathcal{R}(G, \varphi) = \{T \subseteq E(G) : T \text{ is a rainbow spanning tree}\}.$$

The study of rainbow spanning trees in complete graphs, and more general graphs, has attracted a great deal of attention lately, especially on work related to the Brualdi-Hollingsworth conjecture which posits that if the edges of K_{2n} are colored via a one-factorization then the edge set can be partitioned into edge-disjoint RSTs. See [1, 3, 4, 5, 9, 10, 14, 15] for the conjecture and some recent developments along these lines.

In this paper we are concerned with a natural extremal problem regarding rainbow spanning trees: maximizing and minimizing $|\mathcal{R}(G, \varphi)|$ over a collection of colorings. One immediately notes that the problem, without restrictions on the colorings, is not interesting: any coloring with fewer than $n-1$ colors cannot possibly contain a rainbow spanning tree so for such a coloring, $|\mathcal{R}(G, \varphi)| = 0$. On the other hand, if all edge colors are distinct, the number of RSTs is simply the number of trees in the graph. This can be easily computed by the matrix tree theorem of Kirchoff (see [13]) for a general graph G and is n^{n-2} by Cayley's formula for the special case where $G = K_n$ (see [2]).

To make the problem interesting and non-trivial, and in the spirit of anti-Ramsey results, we consider this extremal problem on a certain class of colorings, known in the literature as *JL-colorings* [6, 8, 11, 12]. A coloring $\varphi : E(G) \to [n-1]$ is a JL-coloring if it is surjective and rainbow cycle free. Note that these properties are rather delicately balanced with respect to an interplay between RSTs and cycles: if n colors appear in an edge coloring, then G necessarily contains a rainbow cycle, but if fewer than $n-1$ colors appear in an edge coloring, then no RSTs can exist.

Given a JL-coloring φ, let $\mathcal{C}_1, \ldots, \mathcal{C}_{n-1}$ denote the color classes of φ. If a single edge of each color is selected, this gives $n-1$ edges of distinct colors; further, since φ is a rainbow cycle free coloring, this collection of edges yields a rainbow spanning tree. This simple observation means that for a JL-coloring φ,

$$|\mathcal{R}(G, \varphi)| = \prod_{i=1}^{n-1} |\mathcal{C}_i|. \tag{i}$$

Further, since $\sum |C_i| = |E(G)|$, convexity immediately implies that

$$(|E(G)| - (n-2)) \cdot 1^{n-2} \leq |\mathcal{R}(G, \varphi)| \leq \left(\frac{|E(G)|}{n-1}\right)^{n-1}. \tag{ii}$$

How good are these particular estimates? While both can be tight (simultaneously, in the case where G is itself a tree), for the interesting special case where $G = K_n$ they are both far from tight. In particular, we prove that

Theorem 1.1. *Let $\varphi : E(K_n) \to [n-1]$ be a JL-coloring. Then,*

$$2^{2n-O(\log n)} = \frac{\mu(n)}{n} \leq |\mathcal{R}(K_n, \varphi)| \leq (n-1)!$$

where $n > 1$, $\mu(n)$ has the defining property that if s is the unique integer power of 2 such that $\frac{n}{3} \leq s < \frac{2n}{3}$ then,

$$\mu(n) = n \cdot \mu(s) \cdot \mu(n-s)$$

and $\mu(1) = 1$. For each inequality there is a coloring φ for which equality holds.

As we shall see, this gives a surprisingly (to us) chaotic lower bound for $|\mathcal{R}(G, \varphi)|$ (cf. Figure 3 in Section 3.2 and the surrounding discussion), which grows exponentially in n, as opposed to the trivial linear lower bound in the inequality in (ii). For $n \leq 14$, this evaluates to the lower bounds given below:

$n =$	2	3	4	5	6	7	8	9	10	11	12	13	14
$\lvert\mathcal{R}(K_n, \varphi)\rvert \geq$	1	2	4	12	32	96	256	960	3072	10752	32768	122880	393216

We further study the extremal problem on complete bipartite graphs, proving

Theorem 1.2. *Let $\varphi : E(K_{n,m}) \to [n+m-1]$ be a JL-coloring. Then for $n \geq m$,*

$$(n-1)(m-1) + 1 \leq |\mathcal{R}(K_{n,m}, \varphi)| \leq m^{n-m+1}((m-1)!)^2.$$

Both inequalities have colorings φ for which they are tight.

Particularly interesting here, to us, is the stark difference between this case and the case of K_n in terms of the proof mechanics: in particular, the lower bound, difficult in K_n, is now the trivial bound, while the upper bound, quite easy in the K_n case, is comparatively more difficult.

Finally, we consider some related problems: What happens if we work with more general graphs and/or more general colorings? Here, we are able to characterize graphs with JL-colorings for which the trivial lower bound from (*ii*) is tight and we prove an analogue of the matrix tree theorem counting rainbow spanning trees in general graphs that may be of interest in future investigations along these lines for non-JL colorings (cf. Theorem 5.4).

The remainder of the paper is organized as follows: In the next section we introduce a particularly nice way of thinking about JL-colorings which allows us to derive our bounds. We then turn our attention to the proof of Theorem 1.1 in Section 3 before proving Theorem 1.2 in Section 4. We conclude with results concerning general graphs and colorings, some open questions, and directions for future work.

2 The Structure of JL-Colorings

Recall that a JL-coloring φ is a rainbow cycle free $(n-1)$-edge-coloring of a connected graph G with order n. There is a representation of JL-colorings as labeled binary trees which we will use to count RST's in a given JL-colored graph. The key to this approach, which has appeared in a series of papers of Johnson and collaborators (see [6, 8, 11, 12]), is the following proposition:

Proposition 2.1. *Suppose φ is a JL-coloring of a connected graph G. Then there is a partition of $V(G)$ into sets A and $\bar{A}$, so that $e(A, \bar{A})$, the set of edges between vertices in A and $\bar{A}$, is monochromatic in φ, and both $\varphi|_A$ and $\varphi|_{\bar{A}}$ are JL-colorings of the graphs induced on A and $\bar{A}$ respectively. Furthermore, both subsets A and $\bar{A}$ of $V(G)$ induce connected subgraphs of G.*

This was originally proved for complete graphs in [6], complete bipartite graphs in [11], and finally for complete multipartite graphs in [12]. Recently, it has been established for arbitrary connected graphs in [8]. Iterating Proposition 2.1 on the induced subgraphs gives iteratively nested subsets so that each non-trivial subset A is partitioned into two subsets A' and $\bar{A}'$ where the edges between the subsets are monochromatic and the coloring induced on each is a JL-coloring.

This allows us to create a rooted binary tree with $n-1$ internal vertices from every JL-coloring. Here, each vertex is labeled with sets: the root is labeled with $V(G)$ and the children of a vertex labeled A are the two subsets A' and $\bar{A}'$ guaranteed by Proposition 2.1.

It is easy to see that this construction of a tree from a JL-coloring actually gives a correspondence between JL-colorings of a graph and subgraph-labeled binary trees with $n-1$ internal vertices, where each subgraph is connected and the label of any internal vertex is partitioned by the labels of its two children. The colors of the corresponding JL-coloring can be associated with the internal vertices so that the edges of a color are exactly the edges between the two children of the associated internal vertex.

These representations can be simplified for the two main graph classes considered in this paper: complete graphs and complete bipartite graphs, and we do so below.

2.1 JL-Colorings of K_n

For the case where $G = K_n$, the exact sets labeling the vertices in the associated tree make no difference when enumerating rainbow spanning trees: only the number of vertices in each label matters. Thus, a JL-coloring of K_n is equivalent (up to vertex labeling) to a rooted binary tree with $n-1$ internal vertices, so that the root is labeled by n and the two children of a vertex labeled $r \geq 2$ are labeled p and q with $r = p + q$, $p, q \geq 1$, and all n leaves are labeled 1. We call such a tree a *JL-tree*. Equivalently, a JL-tree is a rooted binary tree in which every vertex is labeled with the number of leaves below (or including) itself. Further, there is a bijection between JL-trees and JL-colored K_n's.

As a clarifying example, we illustrate a JL-coloring and its respective JL-tree for K_5 in Figure 1.

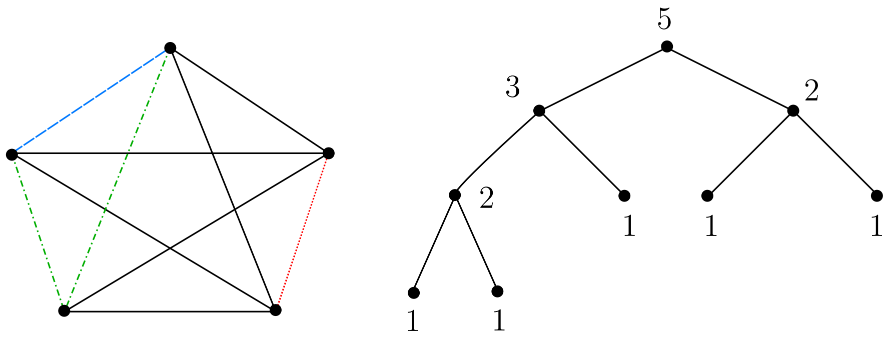

Figure 1: A JL-coloring of K_5 and its associated JL-tree.

2.2 JL-Colorings of $K_{n,m}$

If $G = K_{n,m}$, then the trees described above can also be simplified. In this case, the connected subgraphs A and $\bar{A}$ are smaller complete bipartite graphs. The tree is determined by the size of each bipartite label and from which part in the parent label each smaller part originates.

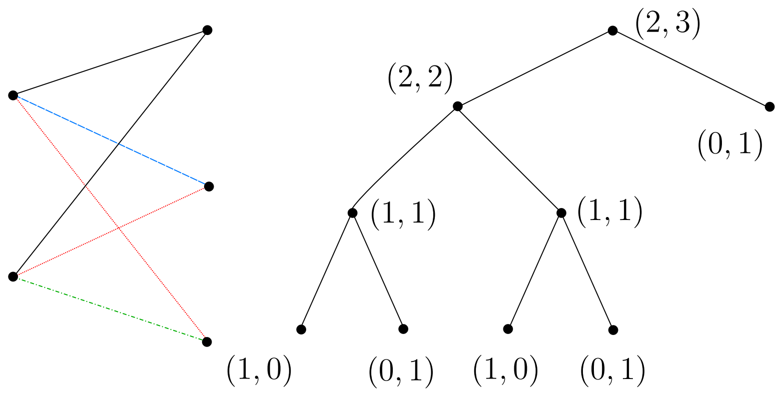

Figure 2: A JL-coloring of $K_{2,3}$ and its associated JL-tree.

In light of this, a JL-coloring of $K_{n,m}$ is equivalent to a rooted full binary tree with $n+m-1$ internal vertices so that the root is labeled (n, m) and the children of a vertex labeled (r_1, r_2) are labeled (p_1, p_2) and (q_1, q_2) so that $p_1+q_1 = r_1$, $p_2+q_2 = r_2$ with the p_i, q_i non-negative and if $p_1 = 0$ then $p_2 = 1$ (respectively if $p_2 = 0$, then $p_1 = 1$). This last restriction is because a single vertex – $K_{1,0}$ – is connected, but $K_{p_1,0}$, $p_1 > 1$ is not. Note that the vertices labeled $(1, 0)$ or $(0, 1)$ are exactly the leaves of the tree. We again observe that there is a bijection between the JL-trees and JL-colored $K_{m,n}$'s.

An example of a JL-tree for $K_{2,3}$ is given in Figure 2.

3 Rainbow spanning trees in K_n

We begin by considering the case where the graph G is complete. In this instance, we observe that the JL-tree (introduced above in Section 2.1) captures not only the structure of the JL-coloring, but also the number of rainbow spanning trees in the coloring.

Since the graph is complete, the number of edges with a given color associated with an inner vertex r is the product of the sizes of its two children, p and q. It follows that multiplying the sizes of all color classes together in a JL-coloring of K_n (as in (i)) is equivalent to taking the product of all non-root labels of its associated JL-tree (or, equivalently, finding the product of all labels of the associated tree and dividing by n).

3.1 The Upper Bound

We first turn our attention to the upper bound in Theorem 1.1. This turns out to be relatively simple after the discussion above. We prove that the JL-tree maximizing the product is the one where the two children of a vertex labeled r are labeled $r-1$ and 1.

Proof of the upper bound in Theorem 1.1. We prove that the JL-tree maximizing the product is as described in the paragraph above: a tree where the vertex labeled r has children labeled $r-1$ and 1. Such a tree has product $n!$ and hence, describes a coloring with $(n-1)!$ RSTs. We proceed by induction on n, noting that it is trivially true for $n=1$. Now, suppose that in a maximizing tree, a vertex r is split as p and q with $p, q \geq 1$: By the inductive hypothesis, the labels below the vertex labeled r have product at most $p!q!$, but it is easy to see that $p!q! \leq (r-1)!$ if $p+q=r$ as this is equivalent to the statement that $\binom{r}{p} \geq r$ for $1 \leq p \leq r-1$. Thus, the optimal split is $p=1$ and $q=r-1$, and the result follows. □

3.2 The Lower Bound

We now turn to the significantly harder case of the lower bound. Since the upper bound was obtained by taking the splits in the JL-tree to be as unbalanced as possible one might expect, or hope, that the lower bound would be achieved by taking the split to be as balanced as possible, namely a vertex labeled n should split as $\lfloor \frac{n}{2} \rfloor$ and $\lceil \frac{n}{2} \rceil$. While this holds for powers of 2, it turns out to be false in general: one part of the optimal split is always a power of 2, specifically the unique power of 2 between $\frac{n}{3}$ and $\frac{2n}{3}$. To show this we study the following function.

For $n \in \mathbb{N}$, let

$$\mu(n) = \min_{1 \leq p \leq n-1} n \cdot \mu(p) \cdot \mu(n-p), \tag{iii}$$

and let $\mu(1) = 1$.

This function corresponds to n times the minimum number of RSTs. This can be seen by noticing that if one takes an interior vertex of a JL-tree, as well as the vertices below it, one obtains a JL-tree for a smaller complete graph. Thus, $\mu(n)$ is taking the product of all of the labels of the vertices of our 'minimum' JL-tree recursively.

In light of this, we are interested in proving the following theorem, which is the lower bound of Theorem 1.1:

Theorem 3.1. *Let $n > 1$ and s denote the unique integer power of 2 so that $\frac{n}{3} \leq s < \frac{2n}{3}$. Then,*

$$\mu(n) = n\mu(s)\mu(n-s).$$

Remark: This does not quite finish the stated bound in Theorem 1.1 that

$$\frac{\mu(n)}{n} = 2^{2n-O(\log n)};$$

the final step in this equality is recorded in Proposition 3.1 at the conclusion of this section.

In order to prove Theorem 3.1, we first introduce the following continuous analogue of μ. For $x \geq 1$, consider the function

$$\tau(x) = \frac{2^{2x-2}}{x}. \tag{iv}$$

It is not necessarily obvious that τ is, in any sense, a continuous analogue of μ. To this end, note that

$$\log_2 \tau(x) = 2x - \log_2(x) - 2$$

is a convex function of x. This log convexity means that for $x \geq 2$,

$$\begin{aligned}\min_{1\leq p\leq x-1} x \cdot \tau(p) \cdot \tau(x-p) &= x\tau(x/2)^2 \\ &= x\frac{2^{2x-4}}{(x/2)^2} \\ &= \frac{2^{2x-2}}{x} = \tau(x),\end{aligned}$$

so that $\tau(x)$ satisfies the defining property (iii) of μ while extending the minimization to all real numbers as opposed to merely integers, and $\tau(1) = \mu(1) = 1$. We now make some elementary observations.

Claim 1: For all integers $n \geq 1$, $\mu(n) \geq \tau(n)$.

Proof. To see this, proceed by induction. Equality holds if $n = 1$, and for $n \geq 2$ note that for some $1 \leq p \leq n-1$,

$$\mu(n) = n \cdot \mu(p) \cdot \mu(n-p) \geq n \cdot \tau(p) \cdot \tau(n-p) \geq n \cdot \tau(n/2) \cdot \tau(n/2) = \tau(n).$$

□

Claim 2: For all integers $i \geq 0$, $\mu(2^i) = \tau(2^i)$.

Proof. This is also shown by induction. Equality holds for $i = 0$ and for $i \geq 1$ observe that,

$$\mu(2^i) \geq \tau(2^i) = 2^i \tau(2^{i-1})^2 = 2^i \mu(2^{i-1})^2 \geq \mu(2^i).$$

Here, the first inequality is Claim 1 and the final inequality is from the definition of μ (iii). Combined, the inequalities force equality and complete the inductive step. $\square$

We remark here that Claims 1 and 2, in fact, prove the lower bound in Theorem 1.1 is achieved since $\frac{\mu(n)}{n} \geq \frac{\tau(n)}{n} = 2^{2n-2\log_2 n-2}$ and $\mu(n) = \tau(n)$ when n is a power of 2. It remains to show that $\mu(n)$ is always of the order $2^{2n-O(\log n)}$. This is done in Proposition 3.1 at the end of this section.

Ultimately, we are interested in the relationship between $\mu(n)$ and $\tau(n)$; to this end let

$$\beta(n) = \frac{\mu(n)}{\tau(n)}. \tag{v}$$

To get a sense of values of β, μ, and τ, we include some values of them below:

$n =$	2	3	4	5	6	7	8	9	10	11	12	13
$\mu(n)$	2	6	16	60	192	672	2048	8640	30720	118272	393216	1597440
$\tau(n)$	2	$5\frac{1}{3}$	16	$51\frac{1}{5}$	$170\frac{2}{3}$	$585\frac{1}{7}$	2048	$7281\frac{7}{9}$	$26214\frac{2}{5}$	$95325\frac{1}{11}$	$349525\frac{1}{3}$	$129055\frac{1}{13}$
$\beta(n)$	1	$\frac{9}{8}$	1	$\frac{75}{64}$	$\frac{9}{8}$	$\frac{147}{128}$	1	$\frac{1215}{1024}$	$\frac{75}{64}$	$\frac{2541}{2048}$	$\frac{9}{8}$	$\frac{2535}{2048}$

By Claim 1, we know that $\beta(n) \geq 1$ for all n. A straightforward calculation reveals that if $\mu(n) = n\mu(p)\mu(n-p)$, then

$$\beta(n) = \frac{\mu(n)}{\tau(n)} = \frac{n^2 \cdot \mu(p)\mu(n-p)}{2^{2n-2}} = \frac{n^2}{4p(n-p)}\beta(p)\beta(n-p), \tag{vi}$$

and finding the minimizing split that defines $\mu(n)$ is equivalent to finding the value of p that minimizes (vi). To emphasize the above observations and allow the reader an easy way to refer back, we collect the above statements into the following Lemma, a proof of which is left to the reader.

Lemma 3.1. *For all integers $n \geq 1$, let $\beta(n) = \frac{\mu(n)}{\tau(n)}$. For $p \in \{1, 2, \ldots n-1\}$, if $\mu(n) = n\mu(p)\mu(n-p)$, then $\beta(n) = \frac{n^2}{4p(n-p)}\beta(p)\beta(n-p)$ and finding the value of p that minimizes $n\mu(p)\mu(n-p)$ is equivalent to finding the value of p that minimizes $\frac{n^2}{4p(n-p)}\beta(p)\beta(n-p)$.*

We now proceed with the proof of Theorem 3.1 which will show that if s is the unique power of 2 so that $\frac{n}{3} \leq s < \frac{2n}{3}$, then s is the value of p that minimizes (vi). As used earlier, a *split* of n is a pair (p, q) of positive integers such that $p+q = n$. We will say that an *optimal split* of n is a split (p, q) of n such that $\mu(n) = n\mu(p)\mu(q)$.

We remark here that working with β proves to be a bit simpler than dealing with μ directly, as we at least have some information (and some clue as to why the powers of 2 occur): $\beta(n) \geq 1$ with $\beta(2^n) = 1$, so minimizing the product (vi) 'prefers' powers of 2. Unfortunately, this is not enough to complete the proof as the β function is quite chaotic and $\limsup \beta(n) = \infty$.

We present in Figure 3 the values of $\beta(n)$ for $1 \leq n \leq 256$.

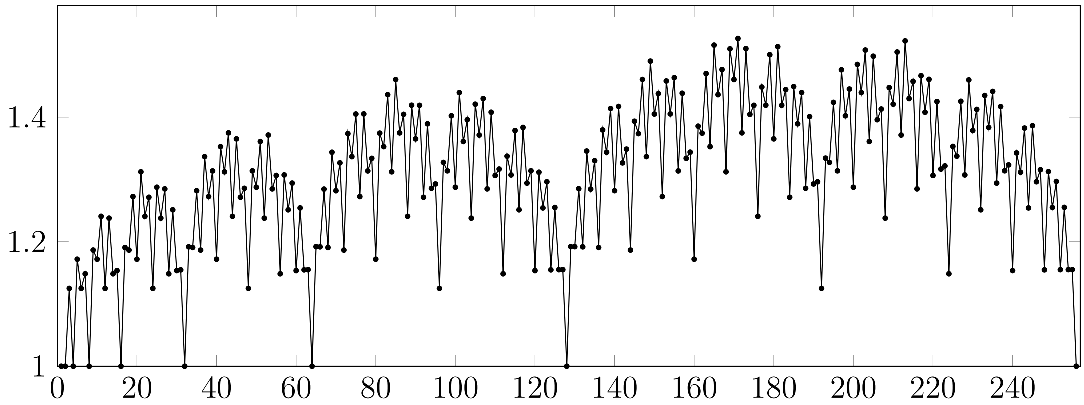

Figure 3: A plot of $\beta(n)$ for $1 \leq n \leq 256$ (linearly interpolating between points).

Several striking features of $\beta(n)$ appear in Figure 3: for instance it appears that $\beta(n)$ has some self-similarity properties, and alternates between increasing and decreasing. Both of these turn out to be true: it is not difficult to verify that $\beta(2n) = \beta(n)$, and a more involved argument shows that $\beta(n)$ for even n is smaller than $\beta(n-1)$ and $\beta(n+1)$. These facts, however, turn out to be not important for solving the recurrence, so we do not record their (rather laborious, in the second case) proofs here.

Proof of Theorem 3.1. We proceed by induction. Theorem 3.1 holds for $n = 2$ so let us assume that $n > 2$ and Theorem 3.1 holds for all integers $1 < k < n$. We shall prove that it holds for n.

For any positive integers p, q with $p + q = n$, let

$$\beta(p, q) = \frac{n^2}{4 \cdot p \cdot q} \beta(p)\beta(q).$$

We observe here that by Lemma 3.1, $\beta(n) \leq \beta(p, q)$ when $p + q = n$. We want to show that $\beta(p, q) \geq \beta(s, n - s)$ where s is unique power of 2 with $\frac{n}{3} \leq s < \frac{2n}{3}$.

We first prove the following lemma:

Lemma 3.2. *Let p, q be positive integers with $p + q = n$ and $p \leq q$. Let (p_1, p_2) and (q_1, q_2) be optimal minimizing splits of p and q respectively, with the two numbers in the splits ordered arbitrarily. Then:*

(a)

$$\beta(p, q) \geq \frac{pq}{(p_1 + q_1)(p_2 + q_2)} \beta(p_1 + q_1, p_2 + q_2)$$

(b)

$$\beta(p,q) \geq \frac{q}{p+q_1}\beta(p+q_1,q_2)$$

Proof. To prove (a), we observe:

$$\begin{aligned}
\beta(p,q) &= \frac{n^2}{4\cdot p\cdot q}\beta(p)\beta(q)\\
&= \frac{n^2\cdot p\cdot q}{4^3\cdot p_1\cdot p_2\cdot q_1\cdot q_2}\beta(p_1)\beta(p_2)\beta(q_1)\beta(q_2)\\
&= \frac{n^2\cdot p\cdot q(p_1+q_1)^2(p_2+q_2)^2}{4^3\cdot p_1\cdot p_2\cdot q_1\cdot q_2\cdot(p_1+q_1)^2(p_2+q_2)^2}\beta(p_1)\beta(p_2)\beta(q_1)\beta(q_2)\\
&= \frac{n^2\cdot p\cdot q}{4\cdot(p_1+q_1)^2(p_2+q_2)^2}\cdot\frac{(p_1+q_1)^2}{4\cdot p_1\cdot q_1}\beta(p_1)\beta(q_1)\cdot\frac{(p_2+q_2)^2}{4\cdot p_2\cdot q_2}\beta(p_2)\beta(q_2)\\
&\geq \frac{n^2\cdot p\cdot q}{4\cdot(p_1+q_1)^2(p_2+q_2)^2}\beta(p_1+q_1)\beta(p_2+q_2)\\
&= \frac{p\cdot q}{(p_1+q_1)(p_2+q_2)}\cdot\frac{n^2}{4\cdot(p_1+q_1)(p_2+q_2)}\beta(p_1+q_1)\beta(p_2+q_2)\\
&= \frac{p\cdot q}{(p_1+q_1)(p_2+q_2)}\beta(p_1+q_1,p_2+q_2)
\end{aligned}$$

where the inequality comes from the fact that (p_1, q_1) and (p_2, q_2) might be suboptimal splits for $p_1 + q_1$ and $p_2 + q_2$, respectively.

The proof of (b) follows in the same manner, only splitting q instead of both q and p. □

We now proceed by comparing an arbitrary split $p + q = n$, $p \leq q$, to our conjectured optimal split, $s + t = n$, where s is the unique power of 2 so that $\frac{n}{3} \leq s < \frac{2n}{3}$. (For the remainder of this section, let s and t be defined as such.) We consider two cases: the first in which (p, q) is a more balanced split than (s, t) and Case 2 in which (p, q) is less balanced. In both cases, we show $\beta(p, q) \geq \beta(s, t)$, thus proving the theorem. To clarify the mechanics of this section, we give a brief example of the balancing and then unbalancing in Case 1 after the conclusion of the proof for the benefit of the reader.

Case 1 (More Balanced Split): $\min(s, t) < p \leq q < \max(s, t)$.

Suppose that $\min(s, t) < p \leq q < \max(s, t)$. Let (p_1, p_2) and (q_1, q_2) be the optimal splits of p and q, respectively. By induction, p and q split as conjectured; let p_1 and q_1 denote the guaranteed powers of 2, respectively. We note that as $p \leq q$, $p_1 \leq q_1$. We now establish the following claim.

Claim: The following inequalities hold:

$$\frac{s}{4} \leq p_1 \leq \frac{s}{2}; \qquad \frac{s}{2} \leq q_1 \leq s; \qquad p_1 \leq q_1 \leq 2p_1.$$

We begin with the first inequality. Note that,

$$p_1 \geq \frac{p}{3} > \min\{s,t\}/3.$$

In particular, if $s = \min\{s,t\}$ then this means that $p_1 > \frac{s}{3}$, and thus, $\frac{s}{2}$ is a lower bound for p_1 (as p_1 is a power of 2). Otherwise, if $t = \min\{s,t\}$, note that $s < 2t$ so that $\frac{p}{3} > \frac{t}{3} > \frac{s}{6}$ and the smallest power of 2 greater than $\frac{s}{6}$ is $\frac{s}{4}$. In either case, $p_1 \geq \frac{s}{4}$.

On the other hand, $p_1 \leq \frac{s}{2}$ as

$$p \leq \frac{(s+t)}{2} \leq \frac{3s}{2};$$

and $p_1 < \frac{2}{3}p$, so that $p_1 < s$. This establishes the first inequality and we note that we have actually shown that if $s = \min\{s,t\}$, then $p_1 = \frac{s}{2}$.

We now turn to the second inequality. First, note that

$$q_1 < \frac{2}{3}q < \frac{2}{3}\max\{s,t\}.$$

By definition s is the unique power of 2 such that $\frac{n}{3} \leq s < \frac{2n}{3}$. Thus, $\frac{s}{2}$ is the unique power of 2 between $\frac{n}{6}$ and $\frac{n}{3}$.

If $s = \max\{s,t\}$, observe that $\frac{q}{3} < \frac{s}{3} < \frac{s}{2}$ while $\frac{q}{3} \geq \frac{n}{6}$ (since $q \geq \frac{n}{2}$) so that $q_1 \geq \frac{s}{2}$. We also observe that $q_1 < q < s$ and note that since q_1 and s are powers of 2, we have that $q_1 \leq \frac{s}{2}$. This establishes the second inequality and we note that we have actually shown that if $s = \max\{s,t\}$, then $q_1 = \frac{s}{2}$.

Otherwise, if $t = \max\{s,t\}$, note that $q_1 \geq \frac{s}{2}$ since $\frac{q}{2} > \frac{s}{2}$ whilst $q_1 < s$ since $2s \geq q$.

The third inequality follows from our already extant work; we have already observed that $p_1 \leq q_1$. If $s = \min\{s,t\}$ we have that $p_1 = \frac{s}{2}$ while $q_1 \leq s$ so $q_1 \leq 2p_1$.

If $s = \max\{s,t\}$ we have that $q_1 = \frac{s}{2}$, and as $p_1 \geq \frac{s}{4}$ so again, $q_1 \leq 2p_1$, and the third inequality holds.

Combining, we have three possibilities: $p_1 = q_1 = \frac{s}{2}$, $p_1 = \frac{s}{4}$ while $q_1 = \frac{s}{2}$, and $p_1 = \frac{s}{2}$ while $q_1 = s$.

Balancing:

If $p_1 = q_1 = \frac{s}{2}$ then we apply Lemma 3.2 to see that

$$\beta(p,q) \geq \frac{pq}{(p_1+q_1)(p_2+q_2)}\beta(p_1+q_1, p_2+q_2) = \frac{pq}{st}\beta(s,t) > \beta(s,t).$$

Here we use the fact that $p_1 + q_1 = s$ so $p_2 + q_2 = t$, and the earlier observation that $pq > st$.

We use the same argument if $p_2 = q_1 = \frac{s}{2}$. This happens if $p = 3 \cdot 2^k$ for some integer k. In this case, $p_1 = \frac{s}{4}$; however, letting $p_2 = \frac{s}{2}$ play the role of p_1 in the above computation similarly shows that the less balanced partition (s,t) is preferred.

Balancing then Unbalancing :

Now suppose that either $p_1 = \frac{s}{2}$ while $q_1 = s$ or $p_1 = \frac{s}{4}$ while $q_1 = \frac{s}{2}$. In the case where $p_1 = \frac{s}{4}$, we may further assume that $p_2 < \frac{s}{2}$ as otherwise the preceding argument applies.

We proceed by first comparing (p,q) to a *more* balanced split (p',q') such that $p' = p_1 + q_1$ and $q' = p_2 + q_2$.

If $p_1 = \frac{s}{2}$ and $q_1 = s$, observe that $p' = \frac{3s}{2} \geq \frac{n}{2}$ by definition of s. However, since $p_1 \neq q_1$, $p \neq q$, so we must have that $p < \frac{n}{2}$ and thus, $p' > p$. If $p_1 = \frac{s}{4}$ while $q_1 = \frac{s}{2}$, then as shown above, we may assume $p_2 < \frac{s}{2}$ which implies $p_2 < 2p_1$. It follows that $p' > p$ since $p' = p_1 + q_1 = \frac{s}{4} + \frac{s}{2} = p_1 + 2p_1 > p_1 + p_2 = p$.

Furthermore, $p' < q$ since $p_1 = \frac{q_1}{2} < q_2$ so that this split is truly more balanced.

$$\beta(p,q) \geq \frac{pq}{(p_1+q_1)(p_2+q_2)}\beta(p_1+q_1, p_2+q_2) = \frac{pq}{p'q'}\beta(p',q').$$

By construction, p' splits optimally as $(p_1', p_2') = (p_1, q_1)$. Note that one of the two terms is necessarily $\frac{s}{2}$—this may be either p_1 or q_1; by possibly reordering, we assume $p_1' = \frac{s}{2}$. We claim that the optimal split of q', (q_1', q_2') also contains $\frac{s}{2}$ as its unique power of 2 (which we denote by q_1').

There are two possibilities. First, if $q_1 = s$, this means that $s = \min\{s,t\}$ (since if $s = \max\{s,t\}$, then $q_1 < q < s$, a contradiction). In this case $q' > s$ since $s < p \leq p', q' < q < t$ so that $\frac{q'}{2} > \frac{s}{2}$, so that $q_1' > \frac{s}{2}$. In the other direction, note that

$$q' = p_2 + q_2 = (p - s/2) + (q - s) = n - 3s/2,$$

so that

$$\frac{2q'}{3} = \frac{2n}{3} - s \leq \frac{n}{3} \leq s,$$

where the last inequality holds as $s \geq \frac{n}{3}$ by definition of s as the power of 2 in the (conjectural) optimal (s,t) split of n. Since q_1 is the power of 2 in the split of q' and is less than $2q'/3$, it hence is at most $\frac{s}{2}$. Thus, if $q_1 = s$, we have that $q_1' = \frac{s}{2}$.

Otherwise, $q_1 = \frac{s}{2}$ and $p_1 = \frac{s}{4}$ so that $s = \max\{s,t\}$ (since, as we showed above, $p_1 = \frac{s}{2}$ when $s = \min\{s,t\}$). In this case, $\frac{q'}{2} < \frac{s}{2}$ and hence $q_1' \leq \frac{s}{2}$. In the other direction, note that

$$q' = p_2 + q_2 = (p - s/4) + (q - s/2) = n - 3s/4$$

whence

$$\frac{q'}{3} = \frac{n}{3} - \frac{s}{4} > \frac{s}{2} - \frac{s}{4} = \frac{s}{4},$$

where we used that $s < \frac{2n}{3}$ to derive that $\frac{s}{2} < \frac{n}{3}$. But this means $q_1' > \frac{s}{4}$, and hence $q_1' \geq \frac{s}{2}$—and so again, $q_1' = \frac{s}{2}$. Now we combine, obtaining

$$\begin{aligned}\beta(p,q) \geq \frac{pq}{p'q'}\beta(p',q') &\geq \frac{pq}{p'q'}\frac{p'q'}{(s/2+s/2)(p_2'+q_2')}\beta(s/2+s/2, p_2'+q_2') \\ &= \frac{pq}{st}\beta(s,t) > \beta(s,t).\end{aligned}$$

Case 2 (Less Balanced Split): $p < \min(s,t) \leq \max(s,t) < q$.

This case works much like Case 1, with somewhat of an opposite feel since the split we are considering is less balanced than our conjectured optimal split. To that end, suppose that $p < \min(s,t) \le \max(s,t) < q$. As above, let (q_1, q_2) be the optimal split of q and note that, by induction, q splits as conjectured; let q_1 denote the power of 2. It follows that $q_1 = s$ or $\frac{s}{2}$ since $n > q > s$.

Balancing:

If $q_1 = s$, this is quite easy. We apply Lemma 3.2 directly to see that

$$\beta(p,q) \ge \frac{q}{p+q_2}\beta(p+q_2, q_1) = \frac{q}{t}\beta(t,s) > \beta(s,t)$$

where we use the fact that $q > \max(s,t)$ and also the fact that $n = s+t = q_1+p+q_2$ implies $t = p + q_2$.

Unbalancing then Balancing:

If $q_1 = \frac{s}{2}$ we first note that

$$\beta(p,q) \ge \frac{q}{p+q_2}\beta(q_1, p+q_2).$$

Now, consider the optimal split r_1, r_2 of $p + q_2 = n - \frac{s}{2}$, where r_1 denotes the power of 2. Observe that since $s < \frac{2n}{3}$, then $\frac{s}{2} < \frac{n}{3}$, and hence $p + q_2 > \frac{2n}{3} \ge \max(s,t)$. Thus, this split is also less balanced than s, t (and potentially less balanced than p, q so that the ratio $\frac{q}{p+q_2}$ appearing above may be less than one.) Nonetheless, we proceed noting that since $n > p + q_2 > s$, either $r_1 = s$ or $r_1 = \frac{s}{2}$.
If $r_1 = s$, then we again apply Lemma 3.2 to see that:

$$\beta(p,q) \ge \frac{q}{p+q_2}\beta(q_1, p+q_2) \ge \frac{q}{p+q_2} \cdot \frac{p+q_2}{q_1+r_2}\beta(s,t) = \frac{q}{t}\beta(s,t) > \beta(s,t).$$

Otherwise, if $r_1 = s/2$, then we balance slightly differently:

$$\beta(p,q) \ge \frac{q}{p+q_2}\beta(q_1, p+q_2) \ge \frac{q}{p+q_2} \cdot \frac{p+q_2}{q_1+r_1}\beta(q_1+r_1, r_2) = \frac{q}{s}\beta(s,t) > \beta(s,t).$$

In both cases, we see that $\beta(p,q) \ge \beta(s,t)$, thus proving the theorem. □

Example 3.2. To better understand the mechanics of the proof above, it is rather helpful to work through an example. We will show an example of the balancing then unbalancing of Case 1, where we compare the optimal split to a more balanced split.

To that end, consider the case $n = 187$, whose optimal split is $s = 64$ and $n - s = 123$. To show this, we want to compare $(64, 123)$ to an arbitrary split of 187. Suppose we start with a more balanced split: $(90, 97)$. To compare these splits, we first compare the $(90, 97)$ split to the more balanced split $(91, 96)$, and then ultimately to the (less balanced, but optimal) $(64, 123)$ split. Figures 4, 5, and 6 illustrate the optimal split for $n = 187$, along with the splits we compare them to which, by induction, we know split optimally below the first step.

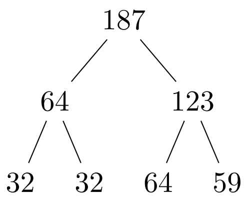

Figure 4: The optimal split.

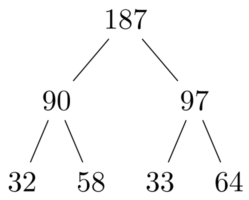

Figure 5: The $(90, 97)$ split.

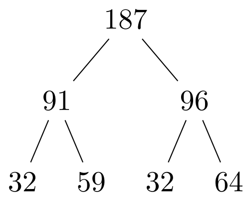

Figure 6: The $(91, 96)$ split.

$$
\begin{aligned}
\beta(90,97) &= \frac{187^2}{4 \cdot 90 \cdot 97}\beta(90)\beta(97) \\
&= \frac{187^2 \cdot 90 \cdot 97}{4^3 \cdot 32 \cdot 58 \cdot 33 \cdot 64}\beta(32)\beta(58)\beta(33)\beta(64) \\
&\geq \frac{187^2 \cdot 90 \cdot 97}{4 \cdot 91^2 \cdot 96^2}\beta(91)\beta(96) \\
&= \frac{90 \cdot 97}{91 \cdot 96}\left[\frac{187^2}{4 \cdot 91 \cdot 96}\beta(91)\beta(96)\right] \\
&= \frac{90 \cdot 97}{91 \cdot 96}\left[\frac{187^2 \cdot 91 \cdot 96}{4^3 \cdot 32^2 \cdot 59 \cdot 64}\beta(32)^2\beta(59)\beta(64)\right] \\
&= \frac{90 \cdot 97}{64 \cdot 123}\left[\frac{187^2}{4 \cdot 64 \cdot 123}\beta(64)\beta(123)\right] \\
&> \frac{187^2}{4 \cdot 64 \cdot 123}\beta(64)\beta(123) = \beta(64, 123).
\end{aligned}
$$

Hence, the $(90, 97)$ split of $n = 187$ is not an optimal split since there exists a split of $(64, 123)$ with smaller β.

We conclude this section with a final proposition, completing the claimed bound that $\frac{\mu(n)}{n} = 2^{2n - O(\log n)}$ from Theorem 1.1 for all n.

Proposition 3.1. *$\beta(n) \leq n^{O(1)}$ and $\beta(n) \geq \frac{9}{8}$ for all n which are not powers of 2.*

Proof. Both statements follow from (vi) and Theorem 3.1 which together show

$$\beta(n) = \frac{n^2}{4s(n-s)}\beta(s)\beta(n-s),$$

where s is the unique integer power of 2 satisfying $\frac{n}{3} \leq s < \frac{2n}{3}$. Since $\beta(s) = 1$, one obtains that

$$\beta(n) \leq \frac{9}{8}\beta(n-s).$$

Now we iterate, letting $n - s = n_1$ and noting that $n_1 \leq \frac{2}{3}n$. Then (again letting s_1 be the unique power of 2 in the decomposition of n_1) we obtain $\beta(n) \leq \frac{9}{8}\beta(n_1) \leq \left(\frac{9}{8}\right)^2 \beta(n_1 - s_1)$. We continue to iterate, letting $n_2 = n_1 - s_1 \leq \frac{2}{3}n_1 \leq (\frac{2}{3})^2 n$, until

n_i is a power of 2; since it decreases exponentially this takes at most $\log_{3/2}(n)$ steps. Hence $\beta(n) \leq (9/8)^{\log_{3/2}(n)} = n^{\log(9/8)/\log(3/2)}$, and this proves the first statement.

The second statement follows by strong induction, noting that it is true for $n = 2^i + 2^{i+1}$ from the fact that for such integers, $\frac{n^2}{4 \cdot s \cdot (n-s)} = \frac{9}{8}$, and for other integers, at least one of the terms appearing in the decomposition of $\beta(n)$ is not a power of 2 and hence is at least $\frac{9}{8}$. □

Remark: The fact that $\beta(n) \leq n^{O(1)}$ completes the claimed bound on $\frac{\mu(n)}{n}$ from Theorem 1.1 as

$$\frac{\mu(n)}{n} = \beta(n)\frac{\tau(n)}{n} = \beta(n)2^{2n-2\log_2(n)} = 2^{2n-O(\log n)}.$$

4 Rainbow Spanning Trees in $K_{n,m}$

We now consider the case where the graph G is complete bipartite. As with the JL-tree associated with a complete graph, the JL-tree (introduced in Section 2.2) associated with a complete bipartite graph captures both the structure of the JL-coloring and the number of RSTs in that coloring.

In this instance, the number of edges with color $\mathcal{C}_1$ associated with an inner vertex (p, q) and children (p_1, p_2) and (q_1, q_2) would be the sum $p_1q_2 + p_2q_1$.

Now, we turn our attention to the proof of Theorem 1.2. As used earlier, a *split* of (n, m) is a split (n_1, m_1), (n_2, m_2) such that $n_1 + n_2 = n$ and $m_1 + m_2 = m$. An optimal minimizing [maximizing] split of (n, m) is a split that minimizes [maximizes] $\nu(n, m)$. When it causes no confusion to the reader, we will simply refer to an optimal minimizing or maximizing split as an optimal split.

We begin by proving the lower bound, followed by the upper bound.

4.1 The Lower Bound

We first consider the lower bound in Theorem 1.2. We prove that $|\mathcal{R}(K_{n,m}, \varphi)| \geq (n-1)(m-1)+1$ and further, that there exists a coloring achieving this lower bound.

Proof of the lower bound in Theorem 1.2. Let $G = K_{n,m}$ be a complete bipartite graph with partitions N and M, respectively. By (ii), for a graph G of order n, $|\mathcal{R}(G, \varphi)| \geq |E(G)| - (n-2)$, so it follows that for $K_{n,m}$,

$$|\mathcal{R}(K_{n,m}, \varphi)| \geq (n-1)(m-1) + 1.$$

We construct a coloring achieving this bound as follows. Fix one vertex $a \in N$ and $b \in M$ from each partite set. Color the edges incident to a and b with distinct colors, and color all other edges the same as the ab edge, so that all color classes except for one have size one. This coloring has $n+m-1$ colors, it has $n+m-2$ color classes class of size one, and the remaining class has size $nm - (n+m-2) = (n-1)(m-1)+1$. This coloring is also rainbow cycle free, as any cycle must use two edges of the ab edge's color. This realizes the bound of (ii) and proves the theorem. □

Note that the coloring described above is represented by the JL-tree where the children of the root vertex labeled (n, m) are $(1, m-1)$ and $(n-1, 1)$, respectively.

4.2 The Upper Bound

We now turn our attention to proving the upper bound in Theorem 1.2. To that end, we let the function $\nu(n, m)$ for $n, m \in \mathbb{N}$ be the maximum number of rainbow spanning trees occurring in any JL-coloring of $K_{n,m}$. We are interested in proving the following theorem.

Theorem 4.1. *Let $n \geq m \in \mathbb{N}$. Then*

$$\nu(n, m) = m^{n-m+1}((m-1)!)^2.$$

Observe that proving Theorem 4.1 proves the upper bound in Theorem 1.2.

Proof of the upper bound in Theorem 1.2. The proof proceeds by induction. Observe that the upper bound in Theorem 1.2 holds for the base case $K_{1,1}$. We shall prove it holds for $K_{n,m}$.

Now, we first claim that for a vertex (a, b) with $a \geq b$ in the JL-tree, the optimal split for producing the most RSTs is the two vertices $(1, 0)$ and $(a-1, b)$. Notice that for $K_{n,m}$ with $n \geq m$, these splits, applied starting with the root (n, m), yield $n-m+1$ color classes of size m and two color classes of each size 1 through $m-1$. By the observations made above, this split produces $m^{n-m+1}((m-1)!)^2$ RSTs.

Now, suppose to the contrary that the split described above does not maximize $|\mathcal{R}(K_{n,m}, \varphi)|$. Then there exists some split, (n_1, m_1) and (n_2, m_2) with $n_1 + n_2 = n$, $m_1 + m_2 = m$, of (n, m) that produces more RSTs. We claim this cannot be the case.

To that end, notice that either $n_1 \geq m_1$ or $n_2 \geq m_2$. Without loss of generality, suppose $n_1 \geq m_1$ and observe that by induction, (n_1, m_1) splits in the conjectured optimal way. Thus, the number of RSTs produced by this (n_1, m_1) and (n_2, m_2) split is the following:

$$\begin{aligned}(n_1m_2 + n_2m_1)\nu(n_1, m_1)\nu(n_2, m_2) &= (n_1m_2 + n_2m_1)m_1\nu(1, 0)\nu(n_1 - 1, m_1)\nu(n_2, m_2) \\ &= m_1(n_1m_2 + n_2m_1)\nu(n_1 - 1, m_1)\nu(n_2, m_2)\end{aligned}$$

Now, the number of RSTs produced by the conjectured optimal split, $(1, 0)$ and $(n-1, m)$, is $m\nu(n-1, m)$. Thus, proving our claim is equivalent to showing $m\nu(n-1, m) \geq m_1(n_1m_2 + n_2m_1)\nu(n_1 - 1, m_1)\nu(n_2, m_2)$.

To that end, observe that

$$\begin{aligned}m\nu(n-1, m) &= m\nu((n_1 - 1) + n_2, m_1 + m_2) \\ &\geq m[(n_1 - 1)m_2 + n_2m_1]\nu(n_1 - 1, m_1)\nu(n_2, m_2)\end{aligned}$$

where the inequality comes from the fact that $(n_1 - 1, m_1)$ and (n_2, m_2) might be suboptimal splits for $(n-1, m)$.

Therefore, it is enough to show that

$$m_1(n_1m_2+n_2m_1)\nu(n_1-1,m_1)\nu(n_2,m_2)\leq m[(n_1-1)m_2+n_2m_1]\nu(n_1-1,m_1)\nu(n_2,m_2).$$

Using the fact that $m = m_1 + m_2$ and rearranging, this is equivalent to showing that

$$0 \leq (n_1 - 1)(m_2)^2 + (n_2 - 1)m_1m_2.$$

If $n_1, n_2 > 0$, then $(n_1 - 1)(m_2)^2 + (n_2 - 1)m_1m_2 \geq 0$. Now, observe that $n_1 \neq 0$ because we assumed $m_1 \leq n_1$ and $(0, 0)$ is not a valid vertex in a JL-tree. Thus, it remains to consider the case where $n_2 = 0$. If $n_2 = 0$ then $m_2 = 1$ and thus,

$$\begin{aligned}(n_1 - 1)(m_2)^2 + (n_2 - 1)m_1m_2 &= n - 1 - (m - 1)\\ &= n - m\\ &\geq 0\end{aligned}$$

It follows that $(n_1 - 1)(m_2)^2 + (n_2 - 1)m_1m_2 \geq 0$, thus completing the proof. □

5 General Graphs, General Colorings, and Further Questions

In this section we briefly investigate a few related questions: How do the results above generalize to arbitrary graphs? How do these results generalize to other $n-1$ colorings, when rainbow cycles are allowed? We note that there are a myriad of interesting open questions in these areas, some of them raised below, that will likely require new ideas to address.

5.1 General Graphs with JL-colorings

As noted in the introduction, the number of rainbow spanning trees in a JL-coloring of a general graph is the product of the sizes of the color classes. In (ii) we observed that by convexity,

$$|E(G)| - (n - 2) \leq |\mathcal{R}(G, \varphi)| \leq \left(\frac{|E(G)|}{n - 1}\right)^{n-1},$$

when φ is a JL-coloring of G.

We have seen that, in the case of a complete bipartite graph, the lower bound is actually achievable. Furthermore, as also observed in the introduction, a rainbow coloring of any tree meets both bounds. The following are natural questions which arise when considering the strength of these bounds.

For the remainder of Section 5.1, assume that G is connected and all colorings are JL-colorings.

(1) **Sharpness of the lower bound:** For which graphs is there a coloring so that the lower bound is sharp? Can they be characterized?

(2) **Sharpness of the upper bound:** Are there any non-trivial examples of the sharpness of the upper bound? The trivial upper bound given above can be strengthened, somewhat, as the sizes of color classes must be integral. Let $a_1, \ldots, a_{n-1}$ be positive integers so that $\sum a_i = |E(G)|$ and $|a_i - a_j| \leq 1$, for $1 \leq i < j \leq n-1$. Then (applying convexity more carefully),

$$|\mathcal{R}(G, \varphi)| \leq \prod_{i=1}^{n-1} a_i. \tag{vii}$$

For which graphs is there a coloring so that (vii) is sharp?

(3) **Graphs maximizing rainbow spanning trees:** Note that for the complete graph, the upper bound (vii) is *not* satisfied and a coloring maximizing $|\mathcal{R}(K_n, \varphi)|$ does not have all color classes the same size. (In fact, at least one color appears on only one edge; this holds true for any connected graph on at least two vertices, by the main result of [8].) This leaves open the possibility that some other connected n-vertex graph G has a coloring so that $\max_\varphi |\mathcal{R}(K_n, \varphi)| \leq \max_\varphi |\mathcal{R}(G, \varphi)|$. Does such a graph exist?

We give brief answers, partial in some cases, to these questions. The first question, in (1), we can answer precisely and we obtain the following.

Theorem 5.1. *The lower bound*

$$|E(G)| - (n-2) \leq |\mathcal{R}(G, \varphi)|$$

is realized for some coloring iff $V(G)$ can be partitioned into two parts (X, Y) so that $G[X]$ and $G[Y]$ are trees and $|E(X, Y)| \geq 1$.

Remark: This is not the traditional presentation of $K_{n,m}$, where we have already observed this bound to be tight. We note, however, that $K_{n,m}$ can also be thought of as two stars, $K_{1,m-1}$ and $K_{n-1,1}$, along with a complete bipartite graph between the leaves and a single edge connecting the roots.

Proof. If G has the desired form, then one colors each of the trees in a rainbow way, with each color used once and each tree using disjoint sets of colors, and then the bipartite graph on (X, Y) a distinct color. Then the coloring has no rainbow cycle (as any cycle must use multiple edges of the bipartite graph (X, Y), uses $(|X| - 1) + (|Y| - 1) + 1 = n - 1$ colors, and furthermore, only one color class has size larger than one so the lower bound is realized.

In the other direction, suppose G has a JL-coloring realizing the lower bound. Such a coloring has at most one color class of size larger than one. If each is of size one, G is a tree, which is of the desired form with X and Y being any partition into connected subtrees. So suppose G is not a tree. Since the coloring is rainbow cycle free, the $n-2$ color classes of size one induce a forest with two components (X and Y). The remaining (larger) color class cannot have any edges within X or Y without forming a rainbow cycle, and hence forms a bipartite graph between them, as desired. $\square$

In the complete graph, however, the lower bound is exponential and this leaves many related open questions. In particular, can one characterize graphs for which this number grows exponentially (or polynomially)? Is it true, for instance, that in a non-bipartite expander graph $|\mathcal{R}(G,\varphi)|$ is necessarily exponential in the number of vertices?

We answer (2), the second question, as follows:

Theorem 5.2. *Let G be a connected graph and let $a_1, a_2, \ldots, a_n$ denote positive integers so that $\sum_{i=1}^{n-1} a_i = |E(G)|$ and, for all indices i and j, $|a_i - a_j| \leq 1$. Then, as noted above, convexity implies that*

$$|\mathcal{R}(G,\varphi)| \leq \prod_{i=1}^{n-1} a_i \leq \left(\frac{|E(G)|}{n-1}\right)^{n-1}.$$

If $|E(G)| \geq 2(n-1)$, then the first inequality is strict.

Remark: When G is a tree both inequalities above are equalities, but it is also easily seen that the first (lower) inequality is equality when G is unicyclic (that is, when $|E(G)| = n$). An interesting open question would be to find the largest $|E(G)|$ for an n-vertex graph G where this inequality can be tight.

Proof. If $|E(G)| \geq 2(n-1)$, then the values a_i satisfying the hypothesis of the theorem are all at least 2. On the other hand, the tree decomposition of a JL-coloring described in Section 2, by iteratively partitioning the graph, ends with two parts of size one — and hence, with a color class of size 1. Thus, in any JL-coloring $|C_i| = 1$ for some i and the bound on the product given is never sharp. □

This leaves open the rather interesting question of whether there is a general improvement to (vii).

Finally we answer the third question, (3), completely with the following.

Theorem 5.3. *If G is an n-vertex non-complete graph, then*

$$\max_{\varphi} |\mathcal{R}(K_n,\varphi)| > \max_{\varphi} |\mathcal{R}(G,\varphi)|.$$

Proof. This follows immediately from the decomposition of JL-colored graphs given in Section 2. Given a graph G and cut $(A, \bar{A})$ in the decomposition of G guaranteed by Proposition 2.1, increasing the number of edges in such a cut gives a JL-colored graph with more edges in the color class (and hence, more rainbow spanning trees). Iterating eventually gives a JL-colored complete graph. This has more rainbow spanning trees than in G, as not all of the cuts augmented were originally complete (as G is not complete). □

5.2 General Colorings

Another interesting set of questions deals with the case where instead of JL-colorings, one considers general colorings. As noted in the introduction, if too general colorings are allowed, the question of counting RSTs can become trivial. To this end, for an n vertex graph G, let

$$\begin{aligned}\mathcal{J}(G) &= \{\varphi : E(G) \to [n-1] : \varphi \text{ is a JL-coloring}\}, \text{ and}\\ \mathcal{C}(G) &= \{\varphi : E(G) \to [n-1]\}\end{aligned}$$

denote the set of JL-colorings and set of general colorings, possibly with rainbow cycles, but restricted to having no more than $n-1$ colors. It is easy to see that when $V(G) > 2$

$$0 = \min_{\varphi \in \mathcal{C}(G)} |\mathcal{R}(G, \varphi)| < \min_{\varphi \in \mathcal{J}(G)} |\mathcal{R}(G, \varphi)|.$$

The question of *maximizing* the number of rainbow spanning trees, however, seems quite interesting. In particular we raise the following question.

Question: Is the following true?

$$\max_{\varphi \in \mathcal{C}(K_n)} |\mathcal{R}(K_n, \varphi)| = \max_{\varphi \in \mathcal{J}(K_n)} |\mathcal{R}(K_n, \varphi)| = (n-1)!$$

The inequality $\max_{\varphi \in \mathcal{J}} |\mathcal{R}(K_n, \varphi)| \leq \max_{\varphi \in \mathcal{C}} |\mathcal{R}(K_n, \varphi)|$ is trivial, as the first maximization is over a smaller set. The inequality in the other direction, that $\max_{\varphi \in \mathcal{J}} |\mathcal{R}(K_n, \varphi)| \geq \max_{\varphi \in \mathcal{C}} |\mathcal{R}(K_n, \varphi)|$ initially appeared unlikely to us, but after some experimentation and thought, it seems plausible. We can show, at least, that colorings with more rainbow spanning trees than the maximizing JL-coloring are quite rare.

Theorem 5.4. *Let $\mathcal{C}'(K_n) \subseteq \mathcal{C}(K_n)$ denote the set of colorings φ of $E(K_n)$ satisfying $\mathcal{R}(K_n, \varphi) \geq (n-1)!$. Then*

$$\lim_{n \to \infty} \frac{|\mathcal{C}'(K_n)|}{|\mathcal{C}(K_n)|} = 0.$$

Proof. Let φ denote a uniform random coloring of the edges of K_n so that the color of each edge is independently and uniformly chosen from $[n-1]$. For a fixed spanning tree T, the probability that T is rainbow is $(n-1)!/(n-1)^{n-1}$. Then by Cayley's formula and linearity of expectation

$$\mathbb{E}\Big[|\mathcal{R}(K_n, \varphi)|\Big] = n^{n-2} \frac{(n-1)!}{(n-1)^{n-1}} = \left(\frac{n}{n-1}\right)^{n-2} \cdot \frac{1}{n-1} \cdot (n-1)! \leq \frac{e}{n-1} \cdot (n-1)!.$$

The result then follows by Markov's inequality, as

$$\frac{|\mathcal{C}'(K_n)|}{|\mathcal{C}(K_n)|} = \mathbb{P}\Big(|\mathcal{R}(K_n, \varphi)| \geq (n-1)!\Big) \leq \frac{e}{n-1} \to 0.$$

$\square$

In general, understanding $|\mathcal{R}(G,\varphi)|$ for an arbitrary $\varphi \in \mathcal{C}$ seems difficult. It is clear that, if $\mathcal{C}_1, \ldots, C_{n-1}$ are the color classes of φ, then

$$|\mathcal{R}(G,\varphi)| \leq \prod_{i=1}^{n-1} |\mathcal{C}_i|.$$

The inequality is strict when collections of $n-1$ edges, one of each color, include cycles. Understanding these collections in a simple way, however, seems difficult.

As a first step in this direction, we observe that we can prove an analogue of the matrix tree theorem of Kirchoff, which gives a way of counting rainbow spanning trees in a general graph.

Recall that the combinatorial Laplacian matrix of a graph is the matrix

$$L = D - A,$$

where D is a diagonal matrix consisting of vertex degrees and A is the adjacency matrix. Note that both A and D above are formed with reference to the same ordering of the vertices of a graph. Then the matrix tree theorem states that the determinant of any cofactor of L is the number of spanning trees in this graph.

We generalize this result to colored graphs. Because we deal with $n-1$ edge colored graphs, and because the statement is cleaner in this case, we focus on the $n-1$ colored case. Given a graph G and an edge coloring $\varphi : E(G) \to [n-1]$, we define the colored graph Laplacian L_φ of G so that

$$[L_\varphi]_{ij} = \begin{cases} 0 & \text{if } i \neq j,\ v_i \not\sim v_j \\ -c_{\varphi(v_i v_j)} & \text{if } i \neq j \text{ and } v_i \sim v_j \\ \sum\limits_{k: v_i \sim v_k} c_{\varphi(v_i v_k)} & \text{if } i = j \end{cases},$$

where c_i for $i = 1, \ldots, n-1$ are indeterminates. Note that if one sets $c_i = 1$, for all i, then one recovers the ordinary graph Laplacian, as above.

Theorem 5.5 (Matrix Tree Theorem for Rainbow Spanning Trees)**.** *Let G be a graph and $\varphi : E(G) \to [n-1]$ an edge coloring of G. Let L_φ of G be the colored graph Laplacian defined above. Let L' denote a principle cofactor of $L_\varphi(G)$ and*

$$f(c_1, \ldots, c_{n-1}) = \det L'.$$

Then

$$|\mathcal{R}(G,\varphi)| = [f(c_1, \ldots, c_{n-1})]_{c_1 c_2 \ldots c_{n-1}} = \frac{\partial}{\partial c_1} \frac{\partial}{\partial c_2} \cdots \frac{\partial}{\partial c_{n-1}} \det L'.$$

Remark: The proof, a simple modification of the usual proof of the matrix tree theorem, actually shows that $\det L'$ is a generating function for different colorings of spanning trees. This remains true for colorings with more than $n-1$ colors. Rainbow spanning trees, in this setting, are counted by the coefficients of squarefree terms. The advantage in stating the $n-1$ color case is that there is only one such term.

Proof. The proof largely follows that of the standard matrix tree theorem.

Let B_φ be a $|V| \times |E|$ matrix, indexed by vertices and edges respectively. The column indexed by edge v_iv_j has non-zero entries only in the v_i and v_j positions: one of these is set to be $\sqrt{c_{\varphi(v_iv_j)}}$ and the other $-\sqrt{c_{\varphi(v_iv_j)}}$, with the signing chosen arbitrarily. Then it is easy to check that

$$L_\varphi = B_\varphi B_\varphi^T,$$

just as with the standard Laplacian. If the v_ith row and column of the Laplacian are removed, then $L' = B'(B')^T$, where B' is obtained by removing the v_ith row of B_φ.

Then, by the Cauchy-Binet formula,

$$\begin{aligned} f(c_1, \ldots, c_{n-1}) = \det L' &= \det(B')(B')^T \\ &= \sum_{\substack{A \subset E(G) \\ |A|=n-1}} \det(B'|_A) \det((B')^T|_A) \\ &= \sum_{\substack{A \subset E(G) \\ |A|=n-1}} \det(B'|_A)^2, \end{aligned}$$

and it is straightforward to verify that

$$\det(B'|_A) = \begin{cases} 0 & \text{if the edges in } A \text{ contain a cycle} \\ \pm \prod_{e \in A} \sqrt{c_{\varphi(e)}} & \text{if the edges in } A \text{ form a spanning tree} \end{cases}$$

Thus,

$$f(c_1, \ldots, c_{n-1}) = \sum_{\substack{T \ \textit{spanning} \\ \textit{tree of } G}} \prod_{e \in T} c_{\varphi(e)}.$$

Then the number of rainbow spanning trees is exactly the coefficient of the monomial where each of the c_is has degree one, as claimed. As this polynomial is homogeneous of degree $n-1$ in the variables c_i, the coefficient can be recovered by iteratively taking derivatives.

As a quick observation, often in such a result one would then evaluate at $c_i = 0$ for all i to remove unwanted contributions. In this case, however, this is unnecessary – all non-rainbow trees will be be missing some variable c_j, and will be destroyed when taking that derivative so that after taking the partial derivatives only the contribution from rainbow trees remains. □

Acknowledgements

The work described in this article is a result of a collaboration made possible by the National Science Foundation (DMS-134365) and the SAMSA-Masamu Program. The work for this article began during a 10 day workshop, hosted by the Masamu Advanced Study Institute in November 2019. We also acknowledge the Simons Foundation (Grant 52503), National Research Foundation in South Africa, University of Denver, and Auburn University. We also generously thank the reviewers for their thoughtful comments and careful read of this paper.

References

[1] R. Brualdi and S. Hollingsworth, Multicolored trees in complete graphs, *J. Combin. Theory Ser. B* 68(2) (1996), 310–313.

[2] A. Cayley, A theorem on trees, *Quart. J. Pure Appl. Math.* 23 (1889), 376–378.

[3] J. Carraher, S. Hartke, and P. Horn, Edge-disjoint rainbow spanning trees in complete graphs, *European J. Combin.* 57 (2016), 71–84.

[4] H. L. Fu, Y. H. Lo, K. E. Perry and C. A. Rodger, On the number of rainbow spanning trees in edge-colored complete graphs, *Discrete Math.* 341 (8) (2018), 2343–2352.

[5] S. Glock, D. Kühn, R. Mongtgomery and D. Osthus, Decompositions into isomorphic rainbow spanning trees, *J. Combin. Theory Ser. B* 146 (2021), 439–484.

[6] A. Gouge, D. Hoffman, P. Johnson, L. Nunley and L. Paben, Edge colorings of K_n which forbid rainbow cycles, *Utilitas Math.* 83 (2010), 219–232.

[7] R. L. Graham and H. O. Pollak, On embedding graphs in squashed cubes, Graph theory and applications, *Lec.Notes in Math.* 303; Springer, Berlin, (1972), 99–110.

[8] D. Hoffman, P. Horn, P. Johnson and A. Owens, On rainbow-cycle-forbidding edge colorings of finite graphs, *Graphs Combin.* 35 (6) (2019), 1585–1596.

[9] P. Horn, Rainbow spanning trees in complete graphs colored by one factorizations, *J. Graph Theory* 87 (3) (2018), 333–346.

[10] P. Horn and L. Nelsen, Many edge-disjoint rainbow spanning trees in general graph, (preprint).

[11] P. Johnson and C. Zhang, Edge colorings of $K_{n,m}$ with $m + n - 1$ colors which forbid rainbow cycles, *Theory and Applicns. of Graphs* 4 (1) Art. 1 (2017), 17.

[12] P. Johnson and A. Owens, Edge colorings of complete multipartite graphs forbidding rainbow cycles, *Theory and Applicns. of Graphs* 4 (2) Art. 2 (2017), 9.

[13] G. Kirchhoff, Über die Auflösung der Gleichungen, auf welche man bei der untersuchung der linearen verteilung galvanischer Ströme geführt wird, *Ann. Phys. Chem.* 72 (1847), 497–508.

[14] R. Montgomery, A. Pokrovskiy and B. Sudakov, Decompositions into spanning rainbow structures, *Proc. Lond. Math. Soc.* (3) 119 (4) (2019), 899–959.

[15] A. Pokrovskiy and B. Sudakov, Linearly many rainbow trees in properly edge-coloured complete graphs, *J. Combin. Theory Ser. B* 132 (2018), 134–156.

(Received 16 June 2021; revised 6 May 2022, 23 Jan 2023)