

A proof of Tomescu's graph coloring conjecture

Jacob Fox*,

Xiaoyu He†

Freddie Manners‡

October 23, 2018

Abstract

In 1971, Tomescu conjectured that every connected graph G on n vertices with chromatic number $k \geq 4$ has at most $k!(k-1)^{n-k}$ proper k -colorings. Recently, Knox and Mohar proved Tomescu's conjecture for $k = 4$ and $k = 5$. In this paper, we complete the proof of Tomescu's conjecture for all $k \geq 4$, and show that equality occurs if and only if G is a k -clique with trees attached to each vertex.

1 Introduction

Let k be a positive integer and $G = (V, E)$ be a graph¹. A *proper k -coloring*, or simply a *k -coloring*, of G is a function $c : V \rightarrow [k]$ (here $[k] = \{1, \dots, k\}$) such that $c(u) \neq c(v)$ whenever $uv \in E$. The chromatic number $\chi(G)$ is the minimum k for which there exists a k -coloring of G . Let $P_G(k)$ denote the number of k -colorings of G . This function is a polynomial in k and is thus called the *chromatic polynomial* of G . In 1912, Birkhoff [5] introduced the chromatic polynomial for planar graphs in an attempt to solve the Four Color Problem using tools from analysis. The chromatic polynomial was later defined and studied for general graphs by Whitney [41].

Despite a great deal of attention over the past century, our understanding of the chromatic polynomial is still quite poor. In particular, proving general bounds on the chromatic polynomial remains a major challenge. The first result of this type, due to Birkhoff [6], states that

$$P_G(k) \geq k(k-1)(k-2)(k-3)^{n-3}$$

for any planar graph G on n vertices and any real number $k \geq 5$. Birkhoff and Lewis [7] later conjectured this to be true for all real $k \geq 4$. The case $k = 4$ is equivalent to the celebrated Four Color Theorem, which was resolved much later.

In 1971, Tomescu [36] conjectured that

$$P_G(k) \leq k!(k-1)^{n-k} \tag{1.1}$$

for every connected graph G on n vertices with chromatic number $k \geq 4$. This inequality can be easily shown if G contains a k -clique. It is also tight if the 2-core of G is exactly a k -clique, that is, if after repeatedly deleting vertices of degree at most one, the remaining graph is a k -clique. Equivalently, such a graph G is formed by attaching a tree to each vertex of the complete graph K_k . Tomescu further conjectured that such graphs are the only examples for which inequality (1.1) is tight.

Tomescu's conjecture has received considerable attention over the past 46 years. Tomescu observed that if $k = 2$, inequality (1.1) holds as an equality for every connected bipartite graph, whereas it is false for $k = 3$ when G is an odd cycle of length at least five. This shows that the condition $k \geq 4$ is necessary; see [21, 37, 39, 40] for details. Tomescu [38] also proved his conjecture for 4-chromatic planar graphs. Further partial results were obtained in [4, 14, 15], and the book [13] gives an overview of the area. Recently, Knox and Mohar solved the cases $k = 4$ [21] and $k = 5$ [22] of Tomescu's conjecture.

In this paper, we prove the full conjecture.

*Department of Mathematics, Stanford University, Stanford, CA 94305, USA. Email: jacobfox@stanford.edu. Research supported by a Packard Fellowship, and by NSF Career Award DMS-1352121.

†Department of Mathematics, Stanford University, Stanford, CA 94305, USA. Email: alkjash@stanford.edu.

‡Department of Mathematics, Stanford University, Stanford, CA 94305, USA. Email: fmanners@stanford.edu.

¹All graphs in this paper are simple and undirected.

Theorem 1. *If G is a connected graph on n vertices with $\chi(G) = k \geq 4$, then*

$$P_G(k) \leq k!(k-1)^{n-k},$$

with equality if and only if the 2-core of G is a k -clique.

In order to prove Theorem 1, we introduce a method for bounding the number of k -colorings of a graph using what we call the Overprediction Lemma. The idea is to construct a probability distribution on the set of colorings by randomly coloring one vertex at a time. From this distribution one obtains an upper bound on $P_G(k)$ using the fact that the cross-entropy between two random variables is at least the entropy of either (this is equivalent to the nonnegativity of Kullback–Leibler divergence). We expect this lemma to have further applications in graph coloring. The argument closely resembles what is sometimes called the “entropy method” for combinatorial upper bounds, first developed by Radhakrishnan [34] to give an alternative proof of Brégman’s Theorem, and applied subsequently in various forms, see e.g. [19, 20, 29].

In the next section, we state and prove the Overprediction Lemma using the entropy inequality mentioned above and develop probabilistic tools for applying it.

In Section 3, we collect some basic results on graphs with chromatic number k , and set up the pieces of our inductive proof of Theorem 1. There, we also show that Theorem 1 holds for small graphs G in the range $|V| \leq 2k - 2$, which serve as base cases for the induction.

In Section 4, we describe a simple linear program whose optimum describes an upper bound for the number of proper k -colorings of a possible counterexample to Tomescu’s conjecture, and complete the argument in Section 5. Together with the Overprediction Lemma, this proves Theorem 1 for large graphs G , in the range $|V| > 2k - 2$ when $k \geq 5$.

Finally, in Section 6 we use a more delicate calculation to handle the case $k = 4$.

2 The Overprediction Lemma

Fix a graph $G = (V, E)$ of chromatic number at most k . Given a k -coloring c of G and a (linear) ordering π of V , we can predict the total number $P_G(k)$ of k -colorings of G as follows.

Define a prefix of π to be a set $U \subseteq V$ with the property that if $v \in U$ then any vertex before v lies in U . Define a π -precoloring of G to be a k -coloring of an induced subgraph of G whose vertices form a prefix of π . Define a *maximal* π -precoloring to be a π -precoloring which cannot be extended to the next vertex in π . In particular, all k -colorings are also maximal π -precolorings.

Define the *backneighborhood* of a vertex v with respect to π , denoted $N_\pi^-(v)$, to be the set of neighbors of v coming before v in the ordering π . If c is a maximal π -precoloring of G , let $X_\pi(c, v)$ be the number of colors in $[k]$ that do not appear among the colored vertices in $N_\pi^-(v)$. If U is the set of vertices colored by c , let

$$X_\pi(c) = \prod_{v \in U} X_\pi(c, v).$$

If the vertices of G are greedily colored in the order of π , the number of choices of the color for v , having already colored all the preceding vertices according to c , is exactly $X_\pi(c, v)$. It is thus natural to suspect that if c is a k -coloring of G , then $X_\pi(c)$ is a good estimate for $P_G(k)$. For example, if $G = K_k$, then for any ordering π and any k -coloring c , $P_G(k) = X_\pi(c) = k!$.

We will be taking expectations of random variables over random choices of a vertex, a coloring, and an ordering of V . The subscripts in the expectation operator $\mathbb{E}[\cdot]$ denote the random choice(s) we are taking expectation over, and choices are made uniformly and independently from among all possible ones.

For example, we write $\mathbb{E}_{\pi, c}[X]$ for the expectation of the random variable X over a uniform random choice of a vertex ordering π and a k -coloring c of G .

The Overprediction Lemma. *For any graph $G = (V, E)$ with $\chi(G) \leq k$ and any ordering π of V ,*

$$P_G(k) \leq \exp(\mathbb{E}_c[\log X_\pi(c)]).$$

Proof. Let C be the family of all k -colorings of G and C_π be the family of all maximal π -precolorings of G , so $C \subseteq C_\pi$ and $|C| = P_G(k)$. We wish to show

$$\mathbb{E}_{c \in C}[\log X_\pi(c)] \geq \log |C|. \quad (2.1)$$

Define two probability distributions p, q over C_π as follows. The distribution p is uniform over colorings $c \in C$. The distribution q picks each $c \in C_\pi$ with probability $X_\pi(c)^{-1}$. Equivalently, q constructs a random maximal pre-coloring c by greedily coloring G one vertex at a time in the order of π , making a uniform choice for the color of v at each step out of all $X_\pi(c, v)$ possible colors. This process stops at a maximal π -precoloring when it runs out of choices or vertices.

By the nonnegativity of Kullback–Leibler divergence from p to q , we find that

$$-\sum_{c \in C_\pi} p(c) \log q(c) \geq -\sum_{c \in C_\pi} p(c) \log p(c). \quad (2.2)$$

Note that p is supported on C . Thus, the left hand side of the inequality is

$$-\sum_{c \in C_\pi} p(c) \log q(c) = \mathbb{E}_{c \in C}[\log X_\pi(c)],$$

while the right hand side is

$$-\sum_{c \in C_\pi} p(c) \log p(c) = \log |C|,$$

so (2.2) is exactly equivalent to the desired inequality (2.1). \square

Let

$$T(c, v) = \exp(\mathbb{E}_\pi[\log X_\pi(c, v)]) \quad (2.3)$$

be the geometric mean of $X_\pi(c, v)$ over a uniform random permutation π , and define

$$T = \prod_{v \in V} \exp(\mathbb{E}_c[\log T(c, v)]).$$

By linearity of expectation,

$$\begin{aligned} \exp(\mathbb{E}_{\pi, c}[\log X_\pi(c)]) &= \prod_{v \in V} \exp(\mathbb{E}_{\pi, c}[\log X_\pi(c, v)]) \\ &= \prod_{v \in V} \exp(\mathbb{E}_c[\log T(c, v)]) \\ &= T. \end{aligned}$$

After averaging over orderings π , we can upper bound $P_G(k)$ by the quantity T .

Given a k -coloring c , define $c_i(v)$ to be the number of neighbors of v colored with color i . The value of $T(c, v)$ is completely determined by the k -tuple $(c_i(v))_{i \leq k}$, and we can obtain an upper bound for $T(c, v)$ via the AM–GM inequality. Define

$$W(c, v) = \sum_{i \leq k} \frac{1}{c_i(v) + 1}. \quad (2.4)$$

Lemma 2. *With $T(c, v)$ and $W(c, v)$ defined by (2.3) and (2.4) respectively, we have*

$$T(c, v) \leq W(c, v)$$

for every k -coloring c and vertex v .

Proof. Note that $X_\pi(c, v) = k - Z_\pi(c, v)$, where $Z_\pi(c, v)$ is the number of colors of c that appear in $N_\pi^-(v)$. By the AM–GM inequality and linearity of expectation,

$$\begin{aligned} T(c, v) &= \exp(\mathbb{E}_\pi[\log X_\pi(c, v)]) \\ &\leq \mathbb{E}_\pi[X_\pi(c, v)] \\ &= k - \mathbb{E}_\pi[Z_\pi(c, v)] \\ &= k - \sum_{i \leq k} \mathbb{E}_\pi[Z_\pi^i(c, v)], \end{aligned}$$

where $Z_\pi^i(c, v)$ is the indicator random variable for the event that color i appears in $N_\pi^-(v)$.

But for each individual color i , $\Pr[Z_\pi^i(c, v)]$ is the probability that v comes after at least one of the $c_i(v)$ vertices of color i in its neighborhood. In a uniform random ordering π , the probability that v appears before all $c_i(v)$ of these vertices is just $\frac{1}{c_i(v)+1}$. Thus,

$$T(c, v) \leq k - \sum_{i \leq k} \left(1 - \frac{1}{c_i(v) + 1}\right) = \sum_{i \leq k} \frac{1}{c_i(v) + 1} = W(c, v),$$

as desired. □

3 Setup

We outline the proof of Theorem 1. Our strategy is to control the size of a minimal counterexample to Tomescu’s conjecture (equivalently, to induct on the number of vertices n in G).

Suppose Theorem 1 is false for a certain $k \geq 4$. Fix this k henceforth. Let n be the minimum number of vertices in a counterexample to Theorem 1, so that $n > k$. Note that such a counterexample G either satisfies (1.1) strictly, or else

$$P_G(k) = k!(k-1)^{n-k}$$

and the 2-core of G is not a k -clique.

Define a graph $G = (V, E)$ to be *bad* if $\chi(G) = k$, $|V| = n$, and G is a minimal counterexample to Theorem 1, i.e. every proper subgraph of G satisfies Theorem 1.

Our proof breaks into three parts. First, we use a straightforward edge-counting argument to show that there are no bad graphs for $n \leq 2k - 2$. Then, we show using the Overprediction Lemma that any bad graph with $k \geq 5$ satisfies $n \leq 2k - 2$. Finally, for $k = 4$ we modify the argument to show that $n \leq 7$, which leaves exactly the case $k = 4$, $n = 7$ to check by hand.

Recall that a graph G is *k-critical* if $\chi(G) = k$ and deleting any vertex or edge reduces its chromatic number.

Lemma 3. *Every bad graph is k-critical.*

Proof. Suppose G is bad but not k -critical, and there is an edge e for which $G \setminus e$ is still k -chromatic. Since G is a minimal counterexample to Theorem 1 with respect to edge deletion, if $G \setminus e$ is connected, then it satisfies

$$P_{G \setminus e}(k) \leq k!(k-1)^{n-k}.$$

But every k -coloring of G is a k -coloring of $G \setminus e$, so G also satisfies Theorem 1, which is a contradiction.

If $G \setminus e$ is disconnected, let the two components be the induced subgraphs $G[V_1]$ and $G[V_2]$ for some vertex partition $V = V_1 \sqcup V_2$. Since e is the only edge between these two components in G and $\chi(G) = k$, one of the two components (say $G[V_1]$) has chromatic number k . Since $G[V_2]$ is connected, every k -coloring of $G[V_1]$ extends to at most $(k-1)^{|V_2|}$ k -colorings of G , with equality if and only if V_2 is a tree. Also, by the minimality of G ,

$$P_{G_1}(k) \leq k!(k-1)^{|V_1|-k},$$

with equality if and only if the 2-core of $G[V_1]$ is a k -clique. Thus,

$$P_G(k) \leq P_{G_1}(k) \cdot (k-1)^{|V_2|} \leq k!(k-1)^{n-k},$$

with equality if and only if the 2-core of $G[V_1]$ is a k -clique and $G[V_2]$ is a tree. But then the 2-core of G is a k -clique, so G is not a counterexample to Theorem 1. This is a contradiction and completes the proof. \square

We will need some well-known structural results about k -critical graphs. It is a basic fact that the minimum degree of a k -critical graph G is at least $k - 1$. We will also need a classical result of Brooks.

Lemma 4 (Brooks' Theorem [9]). *If $k \geq 4$, a k -critical graph that is not the k -clique has at least one vertex of degree at least k .*

The structure of k -critical graphs has been studied extensively, and much more is now known – see for example Gallai [16, 17], Dirac [12], and Kostochka and Yancey [23, 24].

We also use a general fact about k -colorings of graphs using the minimum possible number of colors.

Lemma 5. *For every graph $G = (V, E)$ of chromatic number k and for every proper k -coloring c of G , there are k vertices $v_1, \dots, v_k \in V$ such that, for each $i \leq k$, $c(v_i) = i$ and $\{v_i\} \cup N(v_i)$ contains a vertex of every color.*

Proof. Suppose there is no vertex of some color i with every color in $[k] \setminus i$ appearing in its neighborhood. Then, for every vertex v with $c(v) = i$, there is some color $j \neq i$ not appearing in $N(v)$, and we can change the color of v to j . Since the vertices of color i in k -coloring c form an independent set, these changes can be made simultaneously to create a proper $(k - 1)$ -coloring of G not using color i . This shows $\chi(G) \leq k - 1$, which is a contradiction. \square

Thus, to every k -coloring c of G , we can associate a set of k vertices. There is one of each color, and each sees all the other colors in its neighborhood. We call these k vertices the *radiant* vertices of k -coloring c . When there are multiple vertices of the same color with this property, we arbitrarily fix a single one of them to call radiant. If a vertex v is radiant in k -coloring c , then we say v is *c -radiant*.

Finally, for bad graphs specifically, we show that each pair of vertices receive the same color in less than a fraction $1/(k - 1)$ of all k -colorings.

Lemma 6. *If G is a bad graph, and u, v are any two distinct vertices of G , then*

$$\Pr[c(u) = c(v)] < \frac{1}{k - 1},$$

where the probability is over a uniform choice of a k -coloring c of G .

Proof. Suppose this is not true for some pair u, v . Note that u, v are not adjacent as otherwise u and v must receive different colors and $\Pr[c(u) = c(v)] = 0$. Let G' be the graph obtained from G by contracting u and v together. That is, the two vertices u and v are replaced by a single vertex whose neighborhood is the union $N(u) \cup N(v)$. The k -colorings of G' are in bijection with the k -colorings of G that assign u and v the same color. Thus,

$$P_{G'}(k) = \Pr[c(u) = c(v)] P_G(k) \geq \frac{1}{k - 1} P_G(k) \geq k!(k - 1)^{|V| - k - 1}$$

since G itself contradicts Theorem 1. Note that G' is connected because G is connected.

Thus, either the 2-core of G' is a k -clique or G' is a counterexample to Theorem 1 with fewer vertices. The latter assumption contradicts the minimality of G , so the 2-core of G' is a k -clique.

Since G is bad, G is k -critical and has minimum degree at least $k - 1 \geq 3$. But a single contraction reduces the degree of each vertex by at most 1, so the minimum degree of G' is at least 2. Thus, G' is its own 2-core and must be exactly K_k .

This is impossible. Indeed, it would have to be the case that $G \setminus \{u, v\}$ is a $(k - 1)$ -clique, and since $\chi(G) = k$ at least one of u and v must be complete to this $(k - 1)$ -clique. But then G would contain a k -clique as a proper subgraph, and would not be k -critical. \square

The next lemma is a straightforward consequence of Lemma 6.

Lemma 7. *Let $G = (V, E)$ be a bad graph. For any subset $U \subset V$, if $P(c, U)$ is the number of pairs u, v of distinct vertices in U for which $c(u) = c(v)$, then*

$$\mathbb{E}_c[P(c, U)] < \frac{1}{k-1} \binom{|U|}{2}.$$

Proof. For every pair (u, v) of vertices in U , the probability they are the same color in a random k -coloring of G is less than $\frac{1}{k-1}$ by Lemma 6. The stated inequality follows by linearity of expectation. \square

In general, we will apply Lemma 7 when $U = V$ or when $U \subseteq N(v)$ for some vertex v and $|U| = k - 1$. We describe an immediate application of Lemma 6, showing there are no small bad graphs.

Lemma 8. *If there exists a bad graph on n vertices, then $n \geq 2k - 1$.*

Proof. Suppose there is a bad graph $G = (V, E)$ with $\chi(G) = k \geq 4$ and $|V| = n \leq 2k - 2$. Define $P(c) = P(c, V)$ to be the total number of unordered pairs $u \neq v \in V$ for which $c(u) = c(v)$.

Applying Lemma 6 to every pair $u \not\sim v$ of nonadjacent vertices in G and summing over all such pairs, we find by linearity of expectation that

$$\mathbb{E}_c[P(c)] < \frac{1}{k-1} \left(\binom{n}{2} - |E| \right). \quad (3.1)$$

On the other hand, $P(c) = \sum_i \binom{c_i}{2}$, where c_i is the size of the i -th color class in k -coloring c . As $\sum_i c_i = n \leq 2k - 2$ and each c_i is a nonnegative integer, convexity of the binomial coefficient implies

$$P(c) = \sum_{i \leq k} \binom{c_i}{2} \geq (n - k) \binom{2}{2} + (2k - n) \binom{1}{2} = n - k. \quad (3.2)$$

Comparing (3.1) to (3.2), we have

$$n - k < \frac{1}{k-1} \left(\binom{n}{2} - |E| \right),$$

which simplifies to

$$|E| < \binom{n}{2} - (n - k)(k - 1) = \frac{n^2 + n}{2} + k^2 - kn - k. \quad (3.3)$$

Since G is bad, it has at least $k + 1$ vertices. By an old result of Gallai [17], the number of edges of any k -critical graph on n vertices with $k < n \leq 2k - 1$ is at least

$$\frac{k-1}{2}n + \frac{(n-k)(2k-n)}{2} - 1. \quad (3.4)$$

Comparing (3.3) with (3.4), it follows that

$$n^2 + (1 - 3k)n + (2k^2 - k + 1) > 0,$$

and it is easy to check that this inequality does not hold whenever $k \geq 4$ and $n \leq 2k - 2$. \square

4 A Linear Program

We reduced upper bounding the number of k -colorings of G to estimating the geometric mean of the quantity $W(c, v)$ defined by (2.4), which depends only on the integer partition $\deg(v) = c_1(v) + c_2(v) + \cdots + c_k(v)$ of $\deg(v)$ into the sizes of the color classes within $N(v)$. We now solve a linear program that we will use to bound $W(c, v)$ in the next section.

Lemma 9. Suppose $x \geq k \geq 4$ are positive integers, $S \in [k-1, 3k-3]$ is a real number, and a_j , for $j \geq 0$, are nonnegative real numbers for which

$$\sum_{j \geq 0} a_j = x \quad (4.1)$$

$$\sum_{j \geq 0} j a_j = k-1 \quad (4.2)$$

$$\sum_{j \geq 0} j^2 a_j \leq S. \quad (4.3)$$

Then,

$$\sum_{j \geq 0} \frac{1}{j+1} a_j \leq \begin{cases} x - \frac{4(k-1)-S}{6} & \text{if } S \leq 2k-2 \\ x - \frac{6(k-1)-S}{12} & \text{if } S > 2k-2. \end{cases} \quad (4.4)$$

Proof. Write $W = \sum_{j \geq 0} \frac{1}{j+1} a_j$.

The maximum value of W will always be attained when (4.3) achieves equality. Otherwise, subtract 2ε from some nonzero $a_j, j \geq 1$, and add ε to each of a_{j-1}, a_{j+1} . For a small enough $\varepsilon > 0$, this operation maintains the constraints (4.1), (4.2), and (4.3) while increasing W .

It is a standard fact that for a linear program with n variables whose set of feasible solutions is nonempty and bounded, there exists an optimal solution for which some set of n linearly independent constraints are exactly equality. Such solutions are called *basic feasible solutions* (see for example Theorems 2.3 and 2.7 of [3]). For the linear program (4.1)-(4.3) in question, there are only three constraints other than the nonnegativity constraints $a_j \geq 0$. Thus, some optimal solution has at most three nonzero a_j . We assume henceforth that there are at most three distinct values of j with $a_j \neq 0$.

Subtracting (4.2) from (4.1), we see that $a_0 > 0$, so there are at most two nonzero a_j with $j > 0$. Suppose $0 < j_0 < j_1$ are two indices such that if $j \notin \{0, j_0, j_1\}$ then $a_j = 0$. The three constraints are linearly independent and uniquely determine the values of a_0, a_{j_0}, a_{j_1} :

$$\begin{aligned} a_0 + a_{j_0} + a_{j_1} &= x \\ j_0 a_{j_0} + j_1 a_{j_1} &= k-1 \\ j_0^2 a_{j_0} + j_1^2 a_{j_1} &= S. \end{aligned}$$

The unique solution to this system is

$$(a_0, a_{j_0}, a_{j_1}) = \left(x - a_{j_0} - a_{j_1}, \frac{(k-1)j_1 - S}{j_0 j_1 - j_0^2}, \frac{S - (k-1)j_0}{j_1^2 - j_0 j_1} \right).$$

As we require $a_{j_0}, a_{j_1} \geq 0$, the solution exists whenever $j_0 \leq \frac{S}{k-1} \leq j_1$. In terms of j_0, j_1 , the final objective function is

$$\begin{aligned} W &= a_0 + \frac{1}{j_0+1} a_{j_0} + \frac{1}{j_1+1} a_{j_1} \\ &= x - \frac{j_0}{j_0+1} a_{j_0} - \frac{j_1}{j_1+1} a_{j_1} \\ &= x - \frac{(k-1)(j_0+j_1+1) - S}{(j_0+1)(j_1+1)}. \end{aligned} \quad (4.5)$$

We see that W is at most the maximum value of (4.5) over all choices of integers $0 < j_0 < j_1$ for which $j_0 \leq \frac{S}{k-1} \leq j_1$. It is easy to check that when $\frac{S}{k-1} \in [1, 2]$, the best choice is $j_0 = 1, j_1 = 2$, and if $\frac{S}{k-1} \in [2, 3]$, the best choice is $j_1 = 2, j_2 = 3$. Plugging these choices into (4.5), we obtain exactly the desired bound (4.4). \square

Although we will only use the case $x = k$, we state Lemma 9 for general x because it may be useful to tackle the problem of counting x -colorings of graphs with chromatic number k even when $x > k$.

5 Main Theorem

We are almost ready to apply the Overprediction Lemma to bad graphs. We make one final simplification which allows us to treat G as if it is $(k-1)$ -regular.

Suppose $G = (V, E)$ is a bad graph, so it has minimum degree at least $k-1$. For each vertex $v \in V$, fix an arbitrary set $N^*(v) \subseteq N(v)$ of exactly $k-1$ neighbors. This specified set of neighbors will take the place of $N(v)$ in some of the proceeding calculations. Define $N_\pi^{*-}(v) = N_\pi^-(v) \cap N^*(v)$, and $X_\pi^*(c, v)$ to be the number of distinct colors of k -coloring c missing from $N_\pi^{*-}(v)$.

Similarly, define $c_i^*(v)$ to be the number of vertices of color i in $N^*(v)$,

$$T^*(c, v) = \exp(\mathbb{E}_\pi[\log X_\pi^*(c, v)])$$

and

$$W^*(c, v) = \sum_{i \leq k} \frac{1}{c_i^*(v) + 1}.$$

Then, it follows that $X_\pi(c, v) \leq X_\pi^*(c, v)$, $c_i(v) \geq c_i^*(v)$, $T(c, v) \leq T^*(c, v)$, and $W(c, v) \leq W^*(c, v)$. It will suffice to bound the quantity $W^*(c, v)$ in place of $W(c, v)$.

Theorem 10. *If there exists a bad graph on n vertices, then either $k = 4$ and $n \leq 12$, or $k \geq 5$ and $n \leq 2k - 2$.*

Proof. Let G be a bad graph. By applying Lemma 7 to $N^*(v)$ and averaging over all v ,

$$\mathbb{E}_{c,v} \left[\sum_{i \leq k} \binom{c_i^*(v)}{2} \right] < \frac{1}{k-1} \binom{k-1}{2} = \frac{k-2}{2}. \quad (5.1)$$

It is possible to bound $T(c, v)$ on average with (5.1), but we can obtain a stronger bound on $T(c, v)$ for the radiant vertices. Thus, we bound $T(c, v)$ when v is c -radiant and when v is not c -radiant separately.

If v is c -radiant, let u_1, \dots, u_{k-1} be $k-1$ neighbors of v representing every color except $c(v)$. In a uniform random ordering of V , the number of u_i which appear before v is uniformly distributed among $0, \dots, k-1$. Thus, the probability that $N_\pi^-(v)$ contains vertices of at least i different colors is at least $\frac{i}{k}$. Under these conditions,

$$T(c, v) = \exp \mathbb{E}_\pi[\log X_\pi(c, v)] = \prod_{i \leq k-1} (k-i)^{\Pr[X_\pi(c, v) = k-i]} \leq (k!)^{\frac{1}{k}}. \quad (5.2)$$

This is our bound for $T(c, v)$ when v is c -radiant. Now we seek to control $T(c, v)$ when v is not c -radiant using (5.1). Define

$$\mathbb{E}_{c,v}^*[Y] = \mathbb{E}_c[\mathbb{E}_v[Y | v \text{ is not } c\text{-radiant}]]$$

to be the expectation of $Y = Y(c, v)$ under a uniform random choice of a pair (c, v) of a k -coloring c and a vertex v which is not c -radiant.

In a uniform random choice of a pair (c, v) of a k -coloring and a vertex, the probability that v is c -radiant is $\frac{k}{n}$. Since the sum of $\binom{c_i^*(v)}{2}$ is always nonnegative, this observation together with (5.1) implies

$$\mathbb{E}_{c,v}^* \left[\sum_{i \leq k} \binom{c_i^*(v)}{2} \right] \leq \left(1 - \frac{k}{n}\right)^{-1} \mathbb{E}_{c,v} \left[\sum_{i \leq k} \binom{c_i^*(v)}{2} \right] < \frac{n(k-2)}{2(n-k)}. \quad (5.3)$$

For $j \geq 0$, let $a_j = \mathbb{E}_{c,v}^*[\{i : c_i^*(v) = j\}]$. Then, we claim that these a_j satisfy the conditions

$$\sum_{j \geq 0} a_j = k \quad (5.4)$$

$$\sum_{j \geq 0} j a_j = k - 1 \quad (5.5)$$

$$\sum_{j \geq 0} j^2 a_j < S \quad (5.6)$$

of Lemma 9 with

$$S = (k-1) + \frac{n(k-2)}{n-k}.$$

The first constraint (5.4) is the fact that there are k colors in total. The second constraint (5.5) comes from the fact that $|N^*(v)| = k-1$ for each v . Adding (5.5) to twice (5.3), we get the third constraint (5.6).

The corresponding objective function $\sum_{j \geq 0} \frac{1}{j+1} a_j$ in Lemma 9 is exactly $\mathbb{E}_{c,v}^*[W^*(c, v)]$. We apply Lemma 9 and obtain

$$\mathbb{E}_{c,v}^*[W^*(c, v)] \leq \begin{cases} \frac{7k+5}{12} + \frac{n(k-2)}{12(n-k)} & \text{if } 2k-1 \leq n \leq k^2-k \\ \frac{k+1}{2} + \frac{n(k-2)}{6(n-k)} & \text{if } n > k^2-k. \end{cases}$$

From Lemma 2, $T(c, v) \leq W(c, v)$, and we noted that $W(c, v) \leq W^*(c, v)$. Hence, by the Overprediction Lemma,

$$\begin{aligned} P_G(k) &\leq \prod_{v \in V} \exp(\mathbb{E}_c[\log T(c, v)]) \\ &= \exp(n \mathbb{E}_{c,v}[\log T(c, v)]) \\ &= \exp(k \mathbb{E}_{c,v}[\log T(c, v) | v \text{ is } c\text{-radiant}] + (n-k) \mathbb{E}_{c,v}^*[\log T(c, v)]) \\ &\leq k! \mathbb{E}_{c,v}^*[W^*(c, v)]^{n-k} \\ &\leq \begin{cases} k! \left(\frac{7k+5}{12} + \frac{n(k-2)}{12(n-k)} \right)^{n-k} & \text{if } 2k-1 \leq n \leq k^2-k \\ k! \left(\frac{k+1}{2} + \frac{n(k-2)}{6(n-k)} \right)^{n-k} & \text{if } n > k^2-k. \end{cases} \end{aligned}$$

The factor of $k!$ comes from (5.2). Note that

$$\frac{7k+5}{12} + \frac{n(k-2)}{12(n-k)} < k-1$$

when $k \geq 5$ and $2k-1 \leq n \leq k^2-k$, while

$$\frac{k+1}{2} + \frac{n(k-2)}{6(n-k)} < k-1$$

when $k \geq 5$ and $n > k^2-k$ or $k=4$ and $n \geq 13$.

Plugging these bounds in, we get

$$P_G(k) < k!(k-1)^{n-k}$$

when $k \geq 5$ and $n \geq 2k-1$ or $k=4$ and $n \geq 13$. This completes the proof. \square

Theorem 10 and Lemma 8 together show that there are no bad graphs except possibly in the case $k=4$ and $7 \leq n \leq 12$. In the next section we improve the above argument to handle $8 \leq n \leq 12$ and then check $n=7$ by hand.

6 Small 4-Critical Graphs

In our application of the Overprediction Lemma in Theorem 10, we bounded $T(c, v)$ by the larger but simpler quantity $W(c, v)$ via the AM–GM Inequality. When $k=4$ it is possible to directly bound the average value of $T(c, v)$, improving Theorem 10. We use the same notations as in Section 5.

Lemma 11. *If there exists a bad graph on n vertices with chromatic number $k=4$, then $n \leq 7$.*

Proof. Let G be the bad graph in question. We use a very similar argument to the proof of Theorem 10. Inequalities (5.2) and (5.3) still hold for G , showing (respectively) that if v is c -radiant,

$$T(c, v) \leq (4!)^{\frac{1}{4}}, \tag{6.1}$$

and that conditional on v being non-radiant,

$$\mathbb{E}_{c,v}^* \left[\sum_{i \leq 4} \binom{c_i^*(v)}{2} \right] < \frac{n}{n-4}. \quad (6.2)$$

Recall that $\mathbb{E}_{c,v}^*$ is the expectation over a uniform choice of a coloring c and a vertex v which is not c -radiant.

Now, observe that $|N^*(v)| = k - 1 = 3$ for every vertex v . For each $j \in 1, 2, 3$, define S_j to be the set of non-radiant vertices v for which $N^*(v)$ contains vertices of exactly j distinct colors. The exact value of $T^*(c, v)$ is determined as follows:

$$T^*(c, v) = \begin{cases} 4^{\frac{1}{4}} \cdot 3^{\frac{3}{4}} & \text{if } v \in S_1 \\ 4^{\frac{1}{4}} \cdot 3^{\frac{1}{3}} \cdot 2^{\frac{5}{12}} & \text{if } v \in S_2 \\ (4!)^{\frac{1}{4}} & \text{if } v \in S_3. \end{cases}$$

The first case, for example, comes from the fact that if the three elements of $N^*(v)$ are all the same color, then there is a $\frac{1}{4}$ chance that v precedes all of $N^*(v)$ in the random ordering π so that $X_\pi(c, v) = 4$, and otherwise $X_\pi(c, v) = 3$.

Let $s_j = \mathbb{E}_c[|S_j|]$. Conditioning on the value of j for which $v \in S_j$, we are left with the problem of upper bounding

$$(n-4)\mathbb{E}_{c,v}^*[\log T^*(c, v)] = s_1 \log(4^{\frac{1}{4}} \cdot 3^{\frac{3}{4}}) + s_2 \log(4^{\frac{1}{4}} \cdot 3^{\frac{1}{3}} \cdot 2^{\frac{5}{12}}) + s_3 \log((4!)^{\frac{1}{4}}). \quad (6.3)$$

Using the fact that there are 4 radiant vertices, we see that

$$s_1 + s_2 + s_3 = n - 4. \quad (6.4)$$

Next, consider the size of

$$P(v) = \sum_{i \leq 4} \binom{c_i^*(v)}{2}$$

when $v \in S_1, S_2$, and S_3 . When $v \in S_1$, there is a single color for which $c_i^*(v) = 3$, so $P(v) = 3$. When $v \in S_2$, the sizes of $c_i^*(v)$ are 2, 1, 0, 0, so $P(v) = 1$. When $v \in S_3$, the sizes of $c_i^*(v)$ are 1, 1, 1, 0, so $P(v) = 0$. Thus, by breaking up the expectation in (6.2) conditional on which S_j each vertex lies in, we get

$$3s_1 + s_2 = (n-4)\mathbb{E}_{c,v}^* \left[\sum_{i \leq 4} \binom{c_i^*(v)}{2} \right] \leq n. \quad (6.5)$$

It remains to solve another linear program. Assuming $n \geq 6$, we find that $(s_1, s_2, s_3) = (2, n-6, 0)$ is the unique triple maximizing the quantity (6.3) subject to $s_j \geq 0$ and the two constraints (6.4) and (6.5).

Putting all this together with the Overprediction Lemma,

$$\begin{aligned} P_G(4) &\leq \exp(n\mathbb{E}_{c,v}[\log T(c, v)]) \\ &= \exp(4\mathbb{E}_{c,v}[\log T(c, v)|v \text{ is } c\text{-radiant}] + (n-4)\mathbb{E}_{c,v}^*[\log T(c, v)]) \\ &\leq 4!\mathbb{E}_{c,v}^*[T^*(c, v)]^{n-4} \\ &\leq 4! \cdot (4^{\frac{1}{4}} \cdot 3^{\frac{3}{4}})^2 \cdot (4^{\frac{1}{4}} \cdot 3^{\frac{1}{3}} \cdot 2^{\frac{5}{12}})^{n-6} \\ &= 4! \cdot 3^{(2n-3)/6} \cdot 2^{(11n-54)/12}. \end{aligned}$$

The factor of $4!$ comes from (6.1). The last quantity above is smaller than $4! \cdot 3^{n-4}$ if $n \geq 8$, so there are no bad graphs when $n \geq 8$, as desired. \square

We can now finish the proof of Tomescu's conjecture.

Proof of Theorem 1. Combining Theorem 10 with Lemma 8, we see that there are no bad graphs except possibly if $k = 4$ and $7 \leq n \leq 12$. In this case, Lemma 11 rules out $n \geq 8$, which leaves the sole possibility of $k = 4$ and $n = 7$.

Up to isomorphism, there are exactly two 4-critical graphs on 7 vertices. One of them is the so-called Mycielskian of a triangle, which can be described explicitly as the graph on vertices $u_1, u_2, u_3, v_1, v_2, v_3, w$,

where the edges are $u_i \sim u_j$ for all $i \neq j$, $v_i \sim u_j$ for all $i \neq j$, and $w \sim v_i$ for all i . This graph has $4! \cdot 13$ colorings by 4 colors.

The second 4-critical graph is the Moser spindle, which is obtained by gluing two diamonds at a vertex and connecting their opposite vertices. Explicitly, it is the graph on vertices $u_1, u_2, u_3, v_1, v_2, v_3, w$ where $u_i \sim u_j$ if $i \neq j$, $v_i \sim v_j$ if $i \neq j$, $u_1 \sim v_1$, and w is adjacent to u_2, u_3, v_2, v_3 . This graph has $4! \cdot 16$ colorings by 4 colors.

Both graphs satisfy $P_G(4) \leq 4! \cdot 3^3$, so we are done. \square

7 Concluding Remarks

In 1990, Tomescu [38] extended his original conjecture to a general number of colors.

Conjecture 12. *If $x \geq k \geq 4$, and G is a graph on n vertices with chromatic number k , then*

$$P_G(x) \leq (x)_k (x-1)^{n-k},$$

with equality if and only if the 2-core of G is a k -clique.

Here $(x)_k$ is the falling factorial $x(x-1)\cdots(x-k+1)$. Knox and Mohar [21, 22] proved this more general conjecture for $k = 4$ and 5 , but our methods meet an obstruction to resolving the general case $x > k$. We relied on the existence of radiant vertices in Lemma 5 to recover the factor of $k!$ when $x = k$, and such vertices do not necessarily exist in x -colorings of graphs of chromatic number k when $x > k$.

Our proof of Tomescu's conjecture began with the fact that a minimal counterexample would have to be k -critical. However, while the bound in Tomescu's conjecture is tight for chromatic number k , it is no longer tight when we restrict our attention to k -critical graphs. We are thus led to the following question.

Question 13. *What is the maximum of $P_G(k)$ over all k -critical graphs on n vertices?*

Knox and Mohar [21] recently announced the bound $P_G(k) \leq (k-1)^{n-c \log n} (k-2)^{c \log n}$ holds for every graph G on n vertices of minimum degree at least three and with no twins (two vertices with identical neighborhoods), where $c > 0$ is an absolute constant. This same bound holds for every k -critical graph with $k \geq 4$ because critical graphs have no twins, and improves on Theorem 1 when n is superexponentially large in k^2 . Our methods suggest much stronger bounds than Theorem 1 may hold in Question 13 and we plan to explore this problem further.

There are many other attractive optimization problems in the study of graph colorings. Linial [28] studied the problem of minimizing $P_G(k)$ over all graphs with n vertices and m edges, and found that the complete graph K_s with one additional vertex adjacent to l vertices of the clique and $n-s-1$ isolated vertices, where $m = \binom{s}{2} + l$ and $0 \leq l < s$, is a minimizer for every integer k .

Linial also proposed the related problem of maximizing $P_G(k)$ over graphs G with n vertices and m edges, which is more challenging because the extremal graphs no longer have such a simple structure. Bender and Wilf [2, 42] arrived at the same maximization problem independently, via the study of the backtrack algorithm for enumerating all proper k -colorings of a graph G . Lazebnik [25] was the first to make substantial progress on this problem, and conjectured that the Turán graphs are always extremal graphs for the coloring-maximization problem (see [27]). Many cases of Lazebnik's conjecture have been verified [26, 30, 33, 35], but the full conjecture was recently disproved by Ma and Naves [31].

More recently, Galvin and Tetali [18] considered the problem of maximizing $P_G(k)$ over d -regular graphs G with n vertices. They conjectured that for n a multiple of $2d$, the optimal G consists of $\frac{n}{2d}$ disjoint copies of the complete bipartite graph $K_{d,d}$. Using entropy methods, they proved this conjecture under the additional assumption that G is bipartite. This conjecture was later resolved using a linear programming relaxation for 3-regular graphs by Davies et al. [11] and for 4-regular graphs when $k \geq 5$ by Davies [10].

Acknowledgements. We would like to thank Aysel Erey, János Pach and the referees for many helpful comments.

After we submitted this paper, we learned [32] that Bojan Mohar and Fiachra Knox have also announced a proof of Tomescu's conjecture. We also learned from Persi Diaconis that the counting method in the Overprediction Lemma can be viewed as an instance of a general estimation method known as sequential importance sampling (see for example [8]).

References

- [1] N. Alon and J. H. Spencer, **The Probabilistic Method**, 3rd ed., Wiley, 2008.
- [2] E. Bender and H. S. Wilf, A theoretical analysis of backtracking in the graph coloring problem, *Journal of Algorithms* **6** (1985), 275–282.
- [3] D. Bertsimas and J. N. Tsitsiklis, **Introduction to linear optimization**, Athena Scientific, 1997.
- [4] J. Brown and A. Erey, New bounds for chromatic polynomials and chromatic roots, *Discrete Math.* **338** (2015), 1938–1946.
- [5] G. D. Birkhoff, A determinant formula for the number of ways of coloring a map, *Ann. of Math.* **14** (1912), 42–46.
- [6] G. D. Birkhoff, On the number of ways of colouring a map, *Proc. Edinb. Math. Soc.* **2** (1930), 83–91.
- [7] G. D. Birkhoff and D. Lewis, Chromatic polynomials, *Trans. Amer. Math. Soc.* **60** (1946), 355–451.
- [8] J. Blitzstein and P. Diaconis, A sequential importance sampling algorithm for generating random graphs with prescribed degrees, *Internet Math.* **6** (2010), 489–522.
- [9] R. L. Brooks, On colouring the nodes of a network, *Proc. Cambridge Philosophical Society, Math. Phys. Sci.* **37** (1941), 194–197.
- [10] E. Davies, Counting proper colourings in 4-regular graphs via the Potts Model, preprint arXiv:1801.07547.
- [11] E. Davies, M. Jenssen, W. Perkins, B. Roberts, Extremes of the internal energy of the Potts model on cubic graphs, *Random Structures Algorithms* **53** (2018), 59–75.
- [12] G. Dirac, A theorem of R. L. Brooks and a conjecture of H. Hadwiger, *Proc. London Math. Soc.* **7** (1957), 161–195.
- [13] F. M. Dong, K. M. Koh, and K. L. Teo, Chromatic polynomials and chromaticity of graphs. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2005.
- [14] A. Erey, Maximizing the number of x -colorings of 4-chromatic graphs, *Discrete Math.* **341** (2018), 1419–1431.
- [15] A. Erey, On the maximum number of colorings of a graph, *J. Combinatorics*, to appear.
- [16] T. Gallai, Kritische Graphen I, *Publ. Math. Inst. Hungar. Acad. Sci.* **8** (1963), 165–192.
- [17] T. Gallai, Kritische Graphen II, *Publ. Math. Inst. Hungar. Acad. Sci.* **8** (1963), 373–395.
- [18] D. Galvin and P. Tetali, On weighted graph homomorphisms, *Graphs, morphisms and statistical physics, DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, vol. 63, Amer. Math. Soc., Providence, RI, 2004, pp. 97–104.
- [19] R. Glebov and Z. Luria, On the maximum number of Latin transversals, *J. Combin. Theory Ser. A* **141** (2016), 136–146.
- [20] J. Kahn, An entropy approach to the hard-core model on bipartite graphs, *Combin. Probab. Comput.*, **10** (2001), 219–237.
- [21] F. Knox and B. Mohar, Maximum number of colourings, I. 4-chromatic graphs, preprint arXiv:1708.01781.
- [22] F. Knox and B. Mohar, Maximum number of colourings, II. 5-chromatic graphs, preprint arXiv:1710.06535.

- [23] A. Kostochka and M. Yancey, Ore’s conjecture is almost true, *J. Combin. Theory Ser. B* **109** (2014), 73–101.
- [24] A. Kostochka and M. Yancey, Ore’s conjecture for $k = 4$ and Grötzsch theorem, *Combinatorica* **34** (2014), 323–329.
- [25] F. Lazebnik, On the greatest number of 2 and 3 colorings of a (V, E) -graph, *J. Graph Theory* **13** (1989), 203–214.
- [26] F. Lazebnik and S. Tofts, An extremal property of Turán graphs, *Electron. J. Comb.* **17** (2010), 1–11.
- [27] F. Lazebnik, O. Pikhurko and A. Woldar, Maximum number of colorings of $(2k, k^2)$ -graphs, *J. Graph Theory* **56** (2007), 135–148.
- [28] N. Linial, Legal coloring of graphs, *Combinatorica* **6** (1986), 49–54.
- [29] N. Linial and Z. Luria, An upper bound on the number of Steiner triple systems, *Random Structures Algorithms* **43** (2013), 399–406.
- [30] P.-S. Loh, O. Pikhurko and B. Sudakov, Maximizing the number of q -colorings, *Proc. Lond. Math. Soc.* **101** (2010), 655–696.
- [31] J. Ma and H. Naves, Maximizing proper colorings on graphs, *J. Combin. Theory Ser. B* **115** (2015), 236–275.
- [32] B. Mohar, personal communication.
- [33] S. Norine, Turán graphs and the number of colorings, *Siam J. Discrete Math.* **25** (2011), 260–266.
- [34] J. Radhakrishnan, An entropy proof of Bregman’s theorem, *J. Comb. Theory, Ser. A* **77** (1997), 161–164.
- [35] S. Tofts, An extremal property of Turán graphs, II, *J. Graph Theory*, **75** (2014), 275–283.
- [36] I. Tomescu, Le nombre des graphes connexes k -chromatiques minimaux aux sommets étiquetés, *C. R. Acad. Sci. Paris* **273** (1971), 1124–1126.
- [37] I. Tomescu, The maximum number of 3-colorings of a connected graph, *Discrete Math.* **4** (1972), 351–356.
- [38] I. Tomescu, Maximal chromatic polynomials of connected planar graphs, *Journal of Graph Theory* **14** (1990), 101–110.
- [39] I. Tomescu, Maximum chromatic polynomials of 2-connected graphs, *J. Graph Theory* **18** (1994), 329–336.
- [40] I. Tomescu, Maximum chromatic polynomial of 3-chromatic blocks, *Discrete Math.* **172** (1997), 131–139.
- [41] H. Whitney, A logical expansion in mathematics, *Bull. Amer. Math. Soc.* **38** (1932), 572–579.
- [42] H. S. Wilf, Backtrack: An $O(1)$ expected time algorithm for the graph coloring problem, *Information Processing Letters* **18** (1984), 119–121.