The extremality of 2-partite Turán graphs with respect to the number of colorings

Melissa M Fuentes *
Department of Mathematics and Statistics
Villanova University
Villanova, PA 19085

Submitted: December 31, 2021

Abstract

We consider a problem proposed by Linial and Wilf to determine the structure of graphs which allows the maximum number of q-colorings among graphs with n vertices and m edges. Let $T_r(n)$ denote the Turán graph - the complete r-partite graph on n vertices with partition sizes as equal as possible. We prove that for all odd integers $q \geq 5$ and sufficiently large n, the Turán graph $T_2(n)$ has at least as many q-colorings as any other graph G with the same number of vertices and edges as $T_2(n)$, with equality holding if and only if $G = T_2(n)$. Our proof builds on methods by Norine and methods by Loh, Pikhurko, and Sudakov, which reduces the problem to a quadratic program.

1 Introduction

For a positive integer q, let $[q] = \{1, 2, \ldots, q\}$. A function $f : V(G) \to [q]$ such that $f(x) \neq f(y)$ for every edge xy of a graph G is called a *proper vertex coloring of G in at most q colors*, or simply a *q-coloring* of G. The set $[q]$ is often referred to as the *set of colors*.

Let $P_G(q)$ denote the number of all q-colorings of a given graph G. This number was introduced and studied by Birkhoff [3], who proved that it is always a polynomial in q, and is known as the *chromatic polynomial* of G. Birkhoff's original motivation to consider the chromatic polynomial was to use it to solve the famous four-color conjecture (now a theorem), which asserts that if G is a planar graph then $P_G(4) > 0$, i.e., a 4-coloring of G exists.

Linial [12] and Wilf [2, 22, 20] had independently posed the following problem.

Problem 1. *Let n, m, and q be positive integers. What is the greatest number of q-colorings that a graph with n vertices and m edges can have, and which graphs attain this maximum number of colorings?*

Linial studied the problem of minimizing the chromatic polynomial over the family of graphs with n vertices and m edges by studying the worst-case computational complexity of a certain

*This document is the results of the author's doctoral dissertation. This research was supported in part by the Graduate Scholar's Award Fellowship from the University of Delaware and also in part by funds provided by the National Science Foundation Grant: 1855723.

algorithm. At the end of his paper, Linial posed Problem 1. Around the same time, Wilf [22, 20] and Bender and Wilf [2] studied the backtrack algorithm for the decision problem on the existence of q-coloring of a given graph G, particularly for the family of graphs with n vertices and m edges. This prompted Wilf to consider Problem 1.

Problem 1 remains largely open in general and has been the topic of extensive research. Several upper bounds on $P_G(q)$ over the family of graphs with n vertices and m edges have been obtained (see, for instance, [5, 9, 13, 16]). For a survey of numerous results on Problem 1, see Lazebnik [11].

For a positive integer r, let $T_r(n)$ denote the *r-partite Turán graph*, that is, the complete r-partite graph of order n with all parts nearly equal in size (each part is of size $\lfloor n/r \rfloor$ or $\lceil n/r \rceil$). Let $t_r(n)$ denote the number of edges of $T_r(n)$. Our main motivation is the following conjecture made by Lazebnik in 1987, although it first appeared in print in [7].

Conjecture 1 ([7]). *For all $n \geq r \geq 2$ and $q \geq r$, the Turán graph $T_r(n)$ is the only graph on n vertices and $t_r(n)$ edges that attains the maximum number of q-colorings.*

When $q = r$, the statement follows from the celebrated Turán Theorem [18], since any graph with n vertices and $t_r(n)$ edges that is not $T_r(n)$ does not have a r-coloring. Lazebnik proved Conjecture 1 when $r = 2$ and $q \geq (n/2)^5$ in [9] and when n is a positive integer divisible by r and $q \geq 2\binom{t_r(n)}{3}$ in [10]. For $r = 2$ and $q = 3$, Lazebnik, Pikhurko, and Woldar [7] proved the conjecture when n is even, as well as an asymptotic version when $q = 4$ for even n, as long as n is sufficiently large. Their result for $q = 4$ was extended by Tofts [17] to all $n \geq 4$. Loh, Pikhurko, and Sudakov proved the conjecture for $q = r + 1$ for large n in [14], and their result was later extended to all $n \geq r$ by Lazebnik and Tofts in [6]. This was greatly improved by Norine [16], who developed further powerful techniques from [14], and showed that for any positive integers q and r such that $2 \leq r < q$ and r divides q, Conjecture 1 is true, provided that n is sufficiently large. The most recent result was by Ma and Naves [15], who showed that Conjecture 1 is true for all $q \geq 100r^2/(\log(r))$, provided that n is sufficiently large.

Conjecture 1 was widely believed to be true, especially since there are many results confirming it in several cases. However, Ma and Naves [15] constructed counterexamples in some ranges of r and q. For example, if $r + 3 \leq q \leq 2r - 7$ and $r \geq 10$ then Conjecture 1 is false. Also, for all integers $r \geq 50000$ and q_0 such that $20r \leq q_0 \leq \frac{r^2}{200\log(r)}$, there exists an integer q within distance at most r from q_0, such that Conjecture 1 is false for r and q.

Nevertheless, Conjecture 1 is still believed to be true for integers r and q, where $2 \leq r \leq 9$ and $q \geq r$. The first case for which there are no explicit or asymptotic results is for $r = 2$ and odd $q \geq 5$. This motivated the research done in this paper. The main result is the following theorem.

Theorem 1. *Let $q \geq 5$ be an odd integer. Then for all sufficiently large n, the Turán graph $T_2(n)$ has more q-colorings than any other graph with the same number of vertices and edges.*

Note that $T_2(n)$ is isomorphic to the complete bipartite graph, $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$.

In a breakthrough paper, Loh, Pikhurko, and Sudakov [14] developed a new approach which allowed them to solve Problem 1 for $q = 3$, some $m = m(n)$, and sufficiently large n, by reducing it to a quadratically constrained linear optimization problem. In this work we use ideas and results from [14], [16], and [15], as well as new techniques, to prove Theorem 1.

1.1 Notation

All graphs in this article are finite, undirected, and have neither loops nor multiple edges. For all missing definitions and basic facts which are mentioned but not proved, we refer the reader to Bollobás [4].

For a graph G, let $V = V(G)$ and $E = E(G)$ denote the vertex set of G and the edge set of G, respectively. Let $|A|$ denote the size of a set A. Let $e(G) = |E(G)|$ denote the number of edges of G. For $A \subseteq V(G)$, let $G[A]$ denote the subgraph of G *induced by* A, which means that $V(G[A]) = A$, and $E(G[A])$ consists of all edges xy of G with both x and y in A. For a vertex v of G, let $d_A(v)$ denote the *degree* of v in A, the number of vertices in A that are adjacent to v. For two disjoint subsets $A, B \subseteq V(G)$, by $G[A, B]$ we denote the bipartite subgraph of G induced by A and B, which means that $V(G[A, B]) = A \cup B$, and $E(G[A, B])$ consists of all edges of G with one end-vertex in A and the other in B. Let k be a positive integer. A *k-partition* of a set S is a collection of disjoint subsets $A_1, A_2, \ldots, A_k$ (possibly empty) such that $S = A_1 \cup A_2 \cup \ldots \cup A_k$.

1.2 Organization

In section 2, the aforementioned approach from [14] is introduced, and two important graph constructions and powerful related results from [14] and [15] are presented. Section 3 is dedicated to making two new important observations which will facilitate solving the linear optimization problem from [14]. In section 4, new techniques are used to solve the relevant instances of the linear optimization problem by Loh et al. In section 5, an approximate version of Theorem 1 is proved. The main result, Theorem 1, is derived in section 6, using the results of the previous sections. Finally, several open problems for future investigation are mentioned in section 7.

2 The linear optimization problem and associated graph constructions

In their breakthrough paper, Loh, Pikhurko, and Sudakov [14] developed the optimization problem **OPT** for future researchers to use to determine the graphs that maximize the number of q-colorings among all graphs with the same numbers of vertices and edges. They remark that in solving Problem 1, "the remaining challenge is to find analytic arguments which solve the optimization problem for general q." We will solve particular cases of the optimization problem in Section 4.

For the remainder of this paper we will think of the integer $q \geq 2$ as being fixed.

2.1 The linear optimization problem by Loh, Pikhurko, and Sudakov

It is shown in [14] that solving Problem 1 for large n reduces to a quadratically constrained linear program, which we will now define.

Let $\mathbb{R}$ denote the set of real numbers. Let $\boldsymbol{\alpha} = (\alpha_A)_{A \subseteq [q], A \neq \emptyset}$ be a vector with $2^q - 1$ components $\alpha_A \in \mathbb{R}$ that are indexed by the nonempty subsets $A \subseteq [q]$. When all components of $\boldsymbol{\alpha}$ are nonnegative we will write $\boldsymbol{\alpha} \geq 0$. The logarithms below and in the rest of the paper are natural.

Fix a real number γ that satisfies $0 < \gamma \le (q-1)/(2q)$. Consider the following functions of $\boldsymbol{\alpha}$:

$$\mathrm{OBJ}_q(\boldsymbol{\alpha}) := \sum_{A \neq \emptyset} \alpha_A \log|A|; \quad \mathrm{V}_q(\boldsymbol{\alpha}) := \sum_{A \neq \emptyset} \alpha_A; \quad \mathrm{E}_q(\boldsymbol{\alpha}) := \sum_{\substack{\{A,B\}: A \cap B = \emptyset \\ A \neq \emptyset, B \neq \emptyset}} \alpha_A \alpha_B.$$

The sums in $\mathrm{OBJ}_q(\boldsymbol{\alpha})$ and $V_q(\boldsymbol{\alpha})$ run over all nonempty subsets of $[q]$, and the sum in $\mathrm{E}_q(\boldsymbol{\alpha})$ runs over all *unordered* pairs of disjoint nonempty subsets of $[q]$. In the remainder of this paper, we will suppress mentioning that the sets over which the sums above are taken are nonempty.

Let $\mathrm{FEAS}_q(\gamma)$ be defined by

$$\mathrm{FEAS}_q(\gamma) := \{\boldsymbol{\alpha} \in \mathbb{R}^{2^q-1} : \boldsymbol{\alpha} \ge 0,\ \mathrm{V}_q(\boldsymbol{\alpha}) = 1, \text{ and } \mathrm{E}_q(\boldsymbol{\alpha}) \ge \gamma\}.$$

The elements of $\mathrm{FEAS}_q(\gamma)$ will be referred to as *feasible vectors*. Our goal is to maximize $\mathrm{OBJ}_q(\boldsymbol{\alpha})$ over $\mathrm{FEAS}_q(\gamma)$.

Optimization Problem (OPT). *Find*

$$\mathrm{OPT}_q(\gamma) := \max_{\boldsymbol{\alpha} \in \mathrm{FEAS}_q(\gamma)} \mathrm{OBJ}_q(\boldsymbol{\alpha}).$$

As noted in [14], a solution of **OPT** exists by continuity of $\mathrm{OPT}_q(\gamma)$ and by compactness of the set $\mathrm{FEAS}_q(\gamma)$. We say that $\boldsymbol{\alpha}$ *solves* $\mathrm{OPT}_q(\gamma)$ (or just **OPT**) if $\boldsymbol{\alpha} \in \mathrm{FEAS}_q(\gamma)$ and $\mathrm{OBJ}_q(\boldsymbol{\alpha}) = \mathrm{OPT}_q(\gamma)$.

Our objective in solving **OPT** is to obtain the approximate structure of a graph on n vertices and at least γn^2 edges that has the most number of q-colorings, provided that n is sufficiently large. We define such a graph in the next section.

Loh et al. [14] had solved **OPT** in for all $q \ge 3$ when γ satisfies $0 \le \gamma \le \kappa_q$, where

$$\kappa_q := \left(\sqrt{\frac{\log(q/(q-1))}{\log(q)}} + \sqrt{\frac{\log(q)}{\log(q/(q-1))}} \right)^{-2} \approx \frac{1}{q \log(q)}.$$

Norine [16] presented an argument that solves **OPT** when $\gamma = (r-1)/(2r)$, where r is the number of parts in the Turán graph, $T_r(n)$, and q is divisible by r. In particular, Norine completely solved **OPT** for $\gamma = 1/4$ (that is, $r = 2$) and all even integers $q \ge 2$. In Section 4 we extend his solution to all γ within a closed interval of real numbers $[1/4 - \epsilon, 1/4]$ when $\epsilon > 0$ is sufficiently small and $q \ge 5$.

2.2 Graph constructions based on feasible vectors

We begin with the following simple fact that we will be used to define a graph construction.

Proposition 2. *Let γ be a real number, n be a positive integer, and $\boldsymbol{\alpha} \in \mathrm{FEAS}_q(\gamma)$. Then there exists a set of integers $\{n_A : A \subseteq [q], A \neq \emptyset\}$ such that $n_A \in \{\lfloor n\alpha_A \rfloor, \lceil n\alpha_A \rceil\}$ and $\sum_{A \neq \emptyset} n_A = n$.*

The following is a graph construction from [14] that is based on an arbitrary feasible vector $\boldsymbol{\alpha}$. We point out to the reader that this construction may not result in a unique graph.

Construction ($G_{\boldsymbol\alpha}(n)$). *Let γ be a real number that satisfies $0 < \gamma \le (q-1)/(2q)$, $n \ge 1$, and $\boldsymbol\alpha \in \mathrm{FEAS}_q(\gamma)$. Using Proposition 2 we partition n into a sum of nonnegative integers $n_A \in \{\lfloor n\alpha_A \rfloor, \lceil n\alpha_A \rceil\}$ such that $\sum_{A\neq\emptyset} n_A = n$. Let V be a set of n vertices. For every non-empty $A \subseteq [q]$, let $V_A \subseteq V$ such that $|V_A|= n_A$ and the sets V_A partition V. Let V be the set of vertices of the graph $G_{\boldsymbol\alpha}(n)$. We form the edge set of $G_{\boldsymbol\alpha}(n)$ by taking every pair of non-empty $A, B \subseteq [q]$ with $A \cap B = \emptyset$ and joining every vertex in V_A to every vertex of V_B by an edge.*

Let us consider an example of a graph $G_{\boldsymbol\alpha}(n)$.

Example 1. Let $n = 16$, $q = 4$, and $\gamma = 1/4$. Define $\boldsymbol\alpha \in \mathbb{R}^{15}$ by $\alpha_{\{1\}} = \alpha_{\{2\}} = \alpha_{\{1,2\}} = \alpha_{\{3,4\}} = 1/4$, and all other entries are 0. Then we have

$$\mathrm{V}_4(\boldsymbol\alpha) = \alpha_{\{1\}} + \alpha_{\{2\}} + \alpha_{\{1,2\}} + \alpha_{\{3,4\}} = 1,$$

$$\mathrm{E}_4(\boldsymbol\alpha) = \alpha_{\{1\}}\alpha_{\{2\}} + \alpha_{\{1\}}\alpha_{\{3,4\}} + \alpha_{\{2\}}\alpha_{\{3,4\}} + \alpha_{\{1,2\}}\alpha_{\{3,4\}} = 4(1/4)^2 = 1/4,$$

and thus, $\boldsymbol\alpha \in \mathrm{FEAS}_4(1/4)$. The vertex set V of $G_{\boldsymbol\alpha}(16)$ consists of the partition $V = V_{\{1\}} \cup V_{\{2\}} \cup V_{\{1,2\}} \cup V_{\{3,4\}}$ and $|V_{\{1\}}|= |V_{\{2\}}|= |V_{\{1,2\}}|= |V_{\{3,4\}}|= 4$. The graph $G_{\boldsymbol\alpha}(16)$ is shown below. Gray areas between partition sets depict all possible edges between the sets.

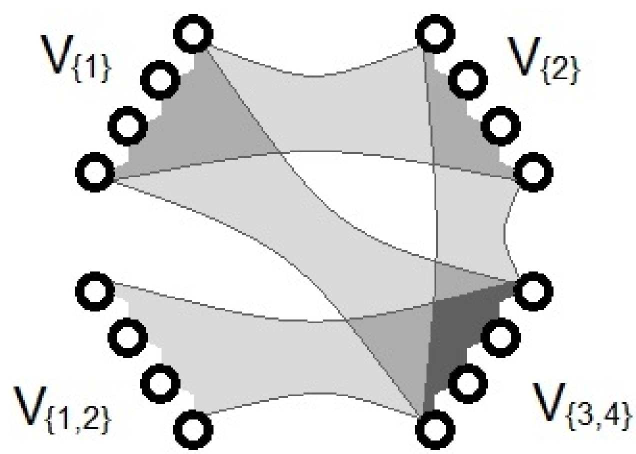

Figure 1: The graph $G_{\boldsymbol\alpha}(16)$ for the vector $\boldsymbol\alpha$ defined in Example 1.

If all $n\alpha_A$ happened to be integers, as in Example 1, the graph $G_{\boldsymbol\alpha}(n)$ would be unique and $G_{\boldsymbol\alpha}(n)$ would have precisely $\mathrm{E}_q(\boldsymbol\alpha)n^2$ edges and at least

$$\prod_{A\neq\emptyset} |A|^{n\alpha_A} = e^{\mathrm{OBJ}_q(\boldsymbol\alpha)n}$$

q-colorings, since any coloring in which the vertices in V_A are only colored with colors in A results in a q-coloring of $G_{\boldsymbol\alpha}(n)$. However, if this is not the case, the choice in size of each V_A may result in a graph with fewer than γn^2 edges (recall that $\mathrm{E}_q(\boldsymbol\alpha)n^2 \ge \gamma n^2$). Fortunately, we may use the Proposition 3 below in order to obtain another graph on n vertices which is "close" (to be defined next) to $G_{\boldsymbol\alpha}(n)$, and which has at least γn^2 edges.

Given graphs G and H with the same set of vertices, their *edit distance* is the minimum number of edges that need to be added or deleted from one graph to obtain a graph isomorphic to the other. Proposition 3 provides bounds for the edit distance and the difference in the number of edges between any two graphs as defined in Construction $\mathbf{G_{ff}(n)}$.

Proposition 3 ([14]). *For any feasible vector $\boldsymbol{\alpha}$, the number of edges in any graph $G_{\boldsymbol{\alpha}}(n)$ differs from $\mathrm{E}_q(\boldsymbol{\alpha})n^2$ by less than $2^q n$. Also, for any other feasible vector $\boldsymbol{\nu}$, the edit distance between $G_{\boldsymbol{\alpha}}(n)$ and $G_{\boldsymbol{\nu}}(n)$ is at most $\|\boldsymbol{\alpha}-\boldsymbol{\nu}\|_1 n^2 + 2^{q+1}n$, where $\|\cdot\|_1$ is the L^1-norm on $\mathbb{R}^{2^q-1}$.*

Proposition 3 and Theorem 4 (shown below) will be important in the proof of our main result, Theorem 1, in Section 6.

Theorem 4 essentially states that if G is a graph that has at least as many q-colorings as any other graph with the same number of vertices and edges, then G is close to a graph $G_{\boldsymbol{\alpha}}(n)$ with respect to their edit distance. For $\delta > 0$ we say that the graphs G and H, each with n vertices, are *δn^2-close* if their edit distance is at most δn^2.

Theorem 4 ([14]). *For any $\delta, \kappa > 0$, the following holds for all sufficiently large n. Let G be an n-vertex graph with m edges, where $m \leq \kappa n^2$, which has at least as many q-colorings as any other graph with the same number of vertices and edges. Then G is ϵn^2-close to a graph $G_{\boldsymbol{\alpha}}(n)$ for some feasible vector $\boldsymbol{\alpha}$ which solves $\mathrm{OPT}_q(\gamma)$ for some γ, where $|\gamma - m/n^2| < \epsilon$ and $\gamma \leq \kappa$.*

The *support* of a feasible vector $\boldsymbol{\alpha} = (\alpha_A)_{A\subseteq[q],A\neq\emptyset}$, denoted by $\mathrm{supp}_q(\boldsymbol{\alpha})$, is the collection of sets $A \subseteq [q]$ such that $\alpha_A > 0$. The following result holds only for feasible vectors which are solutions of **OPT**.

Proposition 5 ([15]). *For any solution $\boldsymbol{\alpha}$ of **OPT**, we have*

$$\bigcup_{A\in\mathrm{supp}_q(\boldsymbol{\alpha})} A = [q].$$

The graph construction below is from [15].

Construction ($\mathbf{SUPP}_q(\boldsymbol{\alpha})$). *Let $\boldsymbol{\alpha} \in \mathrm{FEAS}_q(\gamma)$. The set of vertices of the graph $\mathrm{SUPP}_q(\boldsymbol{\alpha})$ is $\mathrm{supp}_q(\boldsymbol{\alpha})$ and the edge set is formed by connecting pairs of disjoint sets.*

The graph $\mathrm{SUPP}_q(\boldsymbol{\alpha})$ is called the *support graph* of a feasible vector $\boldsymbol{\alpha}$. Let us consider an example of a support graph of a particular feasible vector.

Example 2. Let $n = 16$, $q = 4$, and $\gamma = 1/4$. Consider the feasible vector $\boldsymbol{\alpha}$ defined in Example 1. The graph $\mathrm{SUPP}_q(\boldsymbol{\alpha})$ is shown below.

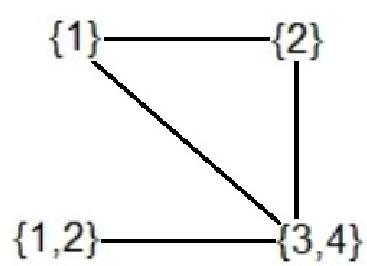

Figure 2: The graph $\mathrm{SUPP}_q(\boldsymbol{\alpha})$ for the vector $\boldsymbol{\alpha}$ defined in Example 1.

We define two classes of graphs $\mathrm{SUPP}_q(\boldsymbol{\alpha})$ as follows. Let $\mathcal{P}_k$ be the class of all graphs $\mathrm{SUPP}_q(\boldsymbol{\alpha})$ for which $\mathrm{supp}_q(\boldsymbol{\alpha})$ forms a k-partition $A_1, A_2, \ldots, A_k$ of $[q]$, with only nonempty sets. Let $\mathcal{Q}_k$ be the class of all graphs $\mathrm{SUPP}_q(\boldsymbol{\alpha})$ for which $\mathrm{supp}_q(\boldsymbol{\alpha})$ consists of a k-partition

$A_1, A_2, \ldots, A_k$ of $[q]$, with only nonempty sets, together with the set $A_1 \cup A_2$. Note that by Construction $\mathbf{SUPP_q}(\boldsymbol{\alpha})$, any graph in the class $\mathcal{P}_k$ is a complete graph on k vertices and any graph in the class $\mathcal{Q}_k$ is a complete graph on k vertices with two edges incident to the same vertex missing.

Fix a positive integer k and let $\boldsymbol{\alpha}$ be a vector such that the graph $\mathrm{SUPP}_q(\boldsymbol{\alpha})$ is in $\mathcal{P}_k$ or $\mathcal{Q}_k$. Consider a restricted version of **OPT** for such vectors $\boldsymbol{\alpha}$:

Optimization Problem (OPT 2). *Maximize*

$$\sum_{i=1}^{k} \alpha_{A_i} \log(|A_i|) + \alpha_{A_1 \cup A_2} \log(|A_1|+|A_2|),$$

subject to

$$\begin{aligned}
&\sum_{i=1}^{k} \alpha_{A_i} + \alpha_{A_1 \cup A_2} = 1, \\
&\sum_{i=1}^{k} \alpha_{A_i}^2 + \alpha_{A_1 \cup A_2}^2 + 2\alpha_{A_1 \cup A_2}(\alpha_{A_1} + \alpha_{A_2}) \leq 1 - 2\gamma \qquad (1) \\
&\alpha_{A_i} \geq 0, \alpha_{A_1 \cup A_2} \geq 0, \ \textit{fo}\ i = 1, \ldots, k.
\end{aligned}$$

The conditions of **OPT 2** are consistent with the conditions of **OPT**, when restricted to vectors with support in $\mathcal{Q}_k \cup \mathcal{P}_k$. The following observation was made in ([15]) using the Cauchy-Schwarz inequality.

Observation 1. *We have $k \geq 1/\lceil 1 - 2\gamma \rceil$ in **OPT 2**, with a strict inequality holding if $\alpha_{A_1} + \alpha_{A_2} > 0$ and $\alpha_{A_1 \cup A_2} > 0$.*

Define

$$\mathcal{P} = \bigcup_{\lceil \frac{1}{1-2\gamma} \rceil \leq k \leq q} \mathcal{P}_k.$$

The next theorem determines the structure of the graph $\mathrm{SUPP}_q(\boldsymbol{\alpha})$ when $\boldsymbol{\alpha}$ solves **OPT**.

Theorem 6 ([15])**.** *For an integer q and real γ which satisfy $0 < \gamma \leq (q-1)/(2q)$, all solutions to **OPT** are such that $\mathrm{SUPP}_q(\boldsymbol{\alpha})$ is in either $\mathcal{P}$ or $\mathcal{Q}_{\lceil 1/(1-2\gamma) \rceil}$. When $\gamma < (q-1)/(2q)$, we have $\mathrm{SUPP}_q(\boldsymbol{\alpha}) \notin \mathcal{P}_q$.*

3 Some observations on the support of a feasible vector

In this section, we present two new useful lemmas which simplify the problem of solving **OPT** by showing that we can make additional assumptions about feasible vectors. Two sets are called *intersecting* if they have a non-empty intersection.

Lemma 1. *Let $\gamma \in \mathbb{R}$. If $\boldsymbol{\alpha} \in \mathrm{FEAS}_q(\gamma)$ such that $\mathrm{supp}_q(\boldsymbol{\alpha})$ contains three pairwise intersecting sets then there exists a vector $\boldsymbol{\alpha}^* \in \mathrm{FEAS}_q(\gamma)$ such that $\mathrm{OBJ}_q(\boldsymbol{\alpha}^*) = \mathrm{OBJ}_q(\boldsymbol{\alpha})$, $|\mathrm{supp}_q(\boldsymbol{\alpha}^*)| < |\mathrm{supp}_q(\boldsymbol{\alpha})|$, and $\boldsymbol{\alpha}^*$ does not contain three pairwise intersecting sets.*

Proof. Let $A, B, C \in \operatorname{supp}_q(\boldsymbol{\alpha})$ be three sets that are pairwise intersecting. The constraint $\mathrm{E}_q(\boldsymbol{\alpha}) \geq \gamma$ can be expressed as

$$\mathrm{E}_q(\boldsymbol{\alpha}) = a\alpha_A + b\alpha_B + c\alpha_C + d \geq \gamma,$$

where a, b, c, and d are nonnegative and depend solely on the entries α_S such that S is distinct from each of A, B, and C. Then

$$\alpha_A + \alpha_B + \alpha_C = 1 - \sum_{S \notin \{A,B,C\}} \alpha_S$$

and

$$\log|A|\alpha_A + \log|B|\alpha_B + \log|C|\alpha_C = \mathrm{OBJ}_q(\boldsymbol{\alpha}) - \sum_{S \notin \{A,B,C\}} \log|S|\alpha_S.$$

Let us show that the components α_A, α_B, and α_C can be modified in such a way to obtain another feasible vector with support of smaller size than $\boldsymbol{\alpha}$. For this, let us consider a system of linear homogeneous equations

$$\begin{cases} x + y + z = 0 \\ \log|A|x + \log|B|y + \log|C|z = 0. \end{cases}$$

Since we have a system of two homogeneous linear equations with three unknowns, this system has a non-trivial solution $(x_A, x_B, x_C) \in \mathbb{R}^3$. For any real t, let $x_A(t) := \alpha_A + tx_A$, $x_B(t) := \alpha_B + tx_B$, and $x_C(t) := \alpha_C + tx_C$. Define the vector $\boldsymbol{\alpha}(t) = (\alpha(t)_S)_{S \subseteq [q], S \neq \emptyset}$ by replacing α_A, α_B, and α_C in the vector $\boldsymbol{\alpha}$ by $x_A(t), x_B(t)$, and $x_C(t)$, respectively, and letting $\alpha(t)_S = \alpha_S$ for all other $S \notin \{A, B, C\}$.

Note that $\mathrm{E}_q(\boldsymbol{\alpha}(t))$ can be rewritten as

$$\mathrm{E}_q(\boldsymbol{\alpha}(t)) = (ax_A + bx_B + cx_C)t + \mathrm{E}_q(\boldsymbol{\alpha}).$$

Suppose that $ax_A + bx_B + cx_C \geq 0$. As the vector (x_A, x_B, x_C) is not the zero vector, and $x_A + x_B + x_C = 0$, then without loss of generality we can assume that $x_A > 0$ and $x_B < 0$. If $x_C \geq 0$, let $t^* = -\alpha_B/x_B > 0$. Then $x_B(t^*) = 0$, $x_A(t^*) \geq 0$, and $x_C(t^*) \geq 0$. If $x_C < 0$ then let $t^* = \min\{-\alpha_B/x_B, -\alpha_C/x_C\} \geq 0$. Then we have $x_A(t^*) \geq 0$, $x_B(t^*) \geq 0$, and $x_C(t^*) \geq 0$, and at least one of $x_B(t^*)$ or $x_C(t^*)$ is equal to zero.

Suppose that $ax_A + bx_B + cx_C < 0$. Then similarly, we may assume that $x_A > 0$ and $x_B < 0$. If $x_C \geq 0$, we let $t^* = \min\{-\alpha_A/x_A, \alpha_C/x_C\} \leq 0$, and if $x_C < 0$ we let $t^* = -\alpha_A/x_A < 0$. Then, in either case, we have $x_A(t^*) \geq 0$, $x_B(t^*) \geq 0$, and $x_C(t^*) \geq 0$, and at least one of $x_B(t^*)$ or $x_C(t^*)$ is equal to zero.

Summarizing, we have $\boldsymbol{\alpha}(t^*) \geq 0$, $V_q(\boldsymbol{\alpha}(t^*)) = 1$, $E_q(\boldsymbol{\alpha}(t^*)) \geq E_q(\boldsymbol{\alpha}) \geq \gamma$, and $\mathrm{OBJ}_q(\boldsymbol{\alpha}(t^*)) = \mathrm{OBJ}_q(\boldsymbol{\alpha})$. Thus, $\boldsymbol{\alpha}(t^*) \in \mathrm{FEAS}_q(\gamma)$ and $|\operatorname{supp}_q(\boldsymbol{\alpha}(t^*))| < |\operatorname{supp}_q(\boldsymbol{\alpha})|$.

If $\operatorname{supp}_q(\boldsymbol{\alpha}(t^*))$ contains another triple of pairwise intersecting sets then we may continue to apply this same process until we obtain a feasible vector $\boldsymbol{\alpha}^*$ as desired in the statement of the lemma. ■

The proof of Lemma 2 below is analogous to the proof of Lemma 1.

Lemma 2. *Let $\gamma \in \mathbb{R}$. If $\boldsymbol{\alpha} \in \mathrm{FEAS}_q(\gamma)$ such that $\operatorname{supp}_q(\boldsymbol{\alpha})$ contains a pair of intersecting sets of the same size then there exists a vector $\boldsymbol{\alpha}^* \in \mathrm{FEAS}_q(\gamma)$ such that $\mathrm{OBJ}_q(\boldsymbol{\alpha}^*) = \mathrm{OBJ}_q(\boldsymbol{\alpha})$, $|\operatorname{supp}_q(\boldsymbol{\alpha}^*)| < |\operatorname{supp}_q(\boldsymbol{\alpha})|$, and $\boldsymbol{\alpha}^*$ does not contain a pair of intersecting sets of the same size.*

4 Relevant solutions of OPT

This section is the most original part of the paper. We use new methods to provide an analytic solution to **OPT** for $q \geq 5$ for a particular range of γ. We mention that although a computer solution can be found for fixed γ, we emphasize the fact that an analytic solution is necessary to solve **OPT** for all γ within a closed interval of real numbers $[1/4 - \epsilon, 1/4]$ for small $\epsilon > 0$.

The main result of this section is the following theorem. For any set S, we denote its *complement* by S^c.

Theorem 7. *The following holds for all γ sufficiently close to $1/4$. Any solution $\boldsymbol{\alpha}$ of **OPT** for $q \geq 5$ has*

$$\operatorname{supp}_q(\boldsymbol{\alpha}) = \{A, A^c\},$$

where $|A| = \lceil q/2 \rceil$, and

$$\alpha_A = \frac{1 + \sqrt{1-4\gamma}}{2} \quad \text{and} \quad \alpha_{A^c} = 1 - \alpha_A,$$

which gives

$$\mathrm{OPT}_q(\gamma) = \frac{1}{2}\log\left(\frac{q^2-1}{4}\right) + \frac{\sqrt{1-4\gamma}}{2}\log\left(\frac{q+1}{q-1}\right).$$

Note that the vector $\boldsymbol{\alpha}$ defined in Theorem 7 satisfies

$$\mathrm{V}_q(\boldsymbol{\alpha}) = \alpha_A + \alpha_{A^c} = \alpha_A + (1 - \alpha_A) = 1$$

and

$$\mathrm{E}_q(\boldsymbol{\alpha}) = \alpha_A \alpha_{A^c} = \alpha_A(1-\alpha_A) = \left(\frac{1+\sqrt{1-4\gamma}}{2}\right)\left(\frac{1-\sqrt{1-4\gamma}}{2}\right) = \gamma.$$

Therefore, $\boldsymbol{\alpha} \in \mathrm{FEAS}_q(\gamma)$.

We will use a constraint that is equivalent to $\mathrm{E}_q(\boldsymbol{\beta}) \geq \gamma$ for all $\boldsymbol{\beta} \in \mathrm{FEAS}(\gamma)$, specifically, $1 - 2\,\mathrm{E}_q(\boldsymbol{\beta}) \leq 1 - 2\gamma$. In addition, note that

$$1 - 2\,\mathrm{E}_q(\boldsymbol{\beta}) = \sum_{(B,S):B\cap S \neq \emptyset} \beta_B \beta_S = \sum_S P_S(q)\beta_S, \tag{2}$$

where

$$P_S(q) := \sum_{B: B\cap S\neq\emptyset} \beta_B.$$

Therefore,

$$\sum_S P_S(q)\beta_S = 1 - 2\,\mathrm{E}_q(\boldsymbol{\beta}) \leq 1 - 2\gamma. \tag{3}$$

Recall that q is the number of colors. We continue with two lemmas which hold for all integers $q \geq 2$ that will be used in the proof of Theorem 7. For $\boldsymbol{\beta} = (\beta_S)_{S\subseteq[q], S\neq\emptyset} \in \mathrm{FEAS}_q(\gamma)$ and every $S \in \operatorname{supp}_q(\boldsymbol{\beta})$, let

$$Q_S(q) := \frac{|S|}{q}.$$

Lemma 3 below is a technical result that will be used in the proof of Lemma 4.

Lemma 3. *Let* γ *satisfy* $0 < \gamma \leq 1/4$ *and let* $\boldsymbol{\beta} \in \mathrm{FEAS}_q(\gamma)$*. Then*

$$\sum_S \left(2\sqrt{2Q_S(q)} - (3-4\gamma)\right)\beta_S \leq 0. \tag{4}$$

Proof. By a comparison of the geometric and arithmetic mean of two numbers (AMGM) we obtain

$$2\sqrt{2Q_S(q)} \leq (2-4\gamma)\frac{Q_S(q)}{P_S(q)} + \frac{1}{1-2\gamma}P_S(q). \tag{5}$$

Thus,

$$\sum_S \left(2\sqrt{2Q_S(q)} - (3-4\gamma)\right)\beta_S \leq \sum_S \left((2-4\gamma)\frac{Q_S(q)}{P_S(q)} + \frac{1}{1-2\gamma}P_S(q) - (3-4\gamma)\right)\beta_S. \tag{6}$$

Now we show that

$$\sum_S \left((2-4\gamma)\frac{Q_S(q)}{P_S(q)} + \frac{1}{1-2\gamma}P_S(q) - (3-4\gamma)\right)\beta_S \leq 0. \tag{7}$$

Let us prove that

$$\sum_S \frac{Q_S(q)}{P_S(q)}\beta_S \leq 1. \tag{8}$$

Let $\emptyset \neq S \subseteq [q]$, and $\chi_S : [q] \to \{0,1\}$ be defined by

$$\chi_S(x) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \notin S. \end{cases}$$

Take any $x \in [q]$. Then

$$\sum_S \frac{\chi_S(x)}{P_S(q)}\beta_S = \sum_{S:x\in S}\frac{1}{P_S(q)}\beta_S = \sum_{S:x\in S}\left(\frac{1}{\sum\limits_{B:B\cap S\neq\emptyset}\beta_B}\right)\beta_S \leq \sum_{S:x\in S}\left(\frac{1}{\sum\limits_{B:x\in B}\beta_B}\right)\beta_S = 1, \tag{9}$$

where for $x \in S$, the last inequality in (9) is due to the fact that $x \in B$ implies $B \cap S \neq \emptyset$.

Note that $\sum_{x\in[q]}\chi_S(x) = |S|$. Then, using (9), we have

$$q = \sum_{x\in[q]} 1 \geq \sum_{x\in[q]}\sum_S \frac{\chi_S(x)}{P_S(q)}\beta_S = \sum_S \frac{\sum_{x\in[q]}\chi_S(x)}{P_S(q)}\beta_S = \sum_S \frac{|S|}{P_S(q)}\beta_S = q\sum_S \frac{Q_S(q)}{P_S(q)}\beta_S,$$

which implies (8).

Using (8), (3), and $\mathrm{V}_q(\boldsymbol{\beta}) = 1$, we obtain (7). Therefore, (6) and (7) imply that (4) holds. ∎

Next we make a key observation that will be used throught the proof of Theorem 7. We show that the sum of the components corresponding to sets of size roughly $q/2$ in the support of a feasible vector $\boldsymbol{\beta}$ with $\mathrm{OBJ}_q(\boldsymbol{\beta})$ being at least as large as the optimal value claimed in Theorem 7 carry more "weight" than the sum of components corresponding to sets of other sizes (recall that $\mathrm{V}_q(\boldsymbol{\beta}) = 1$ and all components of feasible vectors are nonnegative).

Lemma 4. *Let $\boldsymbol{\beta} \in \mathrm{FEAS}_q(\gamma)$, where $q \geq 5$. If*

$$\mathrm{OBJ}_q(\boldsymbol{\beta}) \geq \mathrm{OBJ}_q(\boldsymbol{\alpha}) = \frac{1}{2}\log\left(\frac{q^2-1}{4}\right) + \frac{\sqrt{1-4\gamma}}{2}\log\left(\frac{q+1}{q-1}\right), \tag{10}$$

then for all γ sufficiently close to $1/4$,

$$\sum_{S:|S|=\lceil q/2\rceil,\lfloor q/2\rfloor} \beta_S \geq \begin{cases} 0.76 & \text{if } q=5 \\ 0.80 & \text{if } q \geq 7. \end{cases}. \tag{11}$$

Proof. Let $\boldsymbol{\beta} = (\beta_S) \in \mathrm{FEAS}_q(\gamma)$. Then by Lemma 3, we have

$$\sum_S \left(2\sqrt{2Q_S(q)} - (3-4\gamma)\right)\beta_S \leq 0.$$

Then

$$\begin{aligned}
&\sum_S \left(2\sqrt{2Q_S(q)} - (3-4\gamma)\right)\beta_S \\
&\leq \mathrm{OBJ}_q(\boldsymbol{\beta}) - \frac{1}{2}\log\left(\frac{q^2-1}{4}\right) - \frac{\sqrt{1-4\gamma}}{2}\log\left(\frac{q+1}{q-1}\right) \qquad \text{by (10)}, \\
&= \sum_S \log(|S|)\,\beta_S - \frac{1}{2}\log\left(\frac{q^2-1}{4}\right) + \frac{\sqrt{1-4\gamma}}{2}\log\left(\frac{q-1}{q+1}\right) \\
&= \sum_S \left(\log(2Q_S(q)) + \frac{1}{2}\log\left(\frac{q^2}{q^2-1}\right) + \frac{\sqrt{1-4\gamma}}{2}\log\left(\frac{q-1}{q+1}\right)\right)\beta_S,
\end{aligned}$$

and therefore,

$$\begin{aligned}
0 &\leq \sum_S \left(\log(2Q_S(q)) + \frac{1}{2}\log\left(\frac{q^2}{q^2-1}\right) + \frac{\sqrt{1-4\gamma}}{2}\log\left(\frac{q-1}{q+1}\right) - 2\sqrt{2Q_S(q)} + 3 - 4\gamma)\right)\beta_S \\
&= \sum_{m\in\{2,4,6,\ldots,2q\}} \sum_{S:|S|=m/2} f(m,q,\gamma)\,\beta_S, \qquad (12)
\end{aligned}$$

where we define

$$f(m,q,\gamma) = \log\left(\frac{m}{q}\right) + \frac{1}{2}\log\left(\frac{q^2}{q^2-1}\right) + \frac{\sqrt{1-4\gamma}}{2}\log\left(\frac{q-1}{q+1}\right) - 2\sqrt{\frac{m}{q}} + 3 - 4\gamma. \tag{13}$$

The following claim will help us find maximum values of $f(m,q,\gamma)$ over certain ranges of m.

Claim 1. *Let $q \geq 5$ and $m \in [2, 2q]$ be real numbers. For fixed q, define*

$$h_q(m) = \log\left(\frac{m}{q}\right) - 2\sqrt{\frac{m}{q}}.$$

Then $h_q(m)$ is increasing on $[2,q)$ and is decreasing on $(q,2q]$. Moreover, for a real number $k \in [1, q-2]$,

$$h_q(q-k) < h_q(q+k). \tag{14}$$

Proof of Claim 1. Observe that

$$\frac{d}{dm} h_q(m) = \frac{1 - \sqrt{m/q}}{m}.$$

Since $\sqrt{m/q} > 1$ whenever $q < m \le 2q$ and $\sqrt{m/q} < 1$ whenever $2 \le m < q$, then the first part of the claim holds.

Let us prove the inequality (14). If we substitute $t = k/q$, the inequality (14) is equivalent to

$$\log\left(\frac{1+t}{1-t}\right) > 2\left(\sqrt{1+t} - \sqrt{1-t}\right) \tag{15}$$

for $1/q \le t \le 1 - 2/q$.

Let us show that (15) holds for any $t \in (0,1)$. Since $[1/q, 1-2/q] \subseteq (0,1)$, it will imply (14). Consider the function

$$L(t) := \log\left(\frac{1+t}{1-t}\right) - 2\left(\sqrt{1+t} - \sqrt{1-t}\right) .$$

For $t \in (0,1)$, we have

$$L'(t) = \frac{2}{1-t^2} - \left(\frac{1}{\sqrt{1+t}} + \frac{1}{\sqrt{1-t}}\right) = \frac{2 - \sqrt{1-t^2}\left(\sqrt{1+t} + \sqrt{1-t}\right)}{1-t^2}.$$

It can be easily verified that

$$\sqrt{1-t^2} \cdot (\sqrt{1+t} + \sqrt{1-t}) < \sqrt{1+t} + \sqrt{1-t} < 2.$$

Therefore, $L'(t) > 0$ for all $t \in (0,1)$. As$L(0) = 0$ and L is continuous at 0, $L(t) > 0$ on $(0,1)$, which proves the inequality (15). Therefore,(14) holds. ■

Let $g(q,\gamma) := \frac{1}{2}\log\left(\frac{q^2}{q^2-1}\right) + \frac{\sqrt{1-4\gamma}}{2}\log\left(\frac{q-1}{q+1}\right) + 3 - 4\gamma$. Claim 1 implies that

$$\begin{aligned}
\max_{m \in \{q-1,q+1\}} f(m,q,\gamma) &= g(q,\gamma) + \max_{m \in \{q-1,q+1\}} h_q(m) \\
&= g(q,\gamma) + h_q(q+1) \\
&= f(q+1,q,\gamma)
\end{aligned}$$

and

$$\begin{aligned}
\max_{m \notin \{q-1,q+1\}} f(m,q,\gamma) &= g(q,\gamma) + \max_{m \notin \{q-1,q+1\}} h_q(m) \\
&= g(q,\gamma) + \max_{m \in \{q-3,q+3\}} h_q(m) \\
&= g(q,\gamma) + h_q(q+3) \\
&= f(q+3,q,\gamma).
\end{aligned}$$

By combining these observations with (13) we obtain

$$\begin{aligned}
0 &\le \sum_{m\in\{2,4,6,\dots,2q\}} \sum_{S:|S|=m/2} f(m,q,\gamma)\,\beta_S \\
&= \sum_{m\in\{q-1,q+1\}} \sum_{S:|S|=m/2} f(m,q,\gamma)\,\beta_S + \sum_{m\notin\{q-1,q+1\}} \sum_{S:|S|=m/2} f(m,q,\gamma)\,\beta_S \\
&\le f(q+1,q,\gamma) \sum_{S:|S|\in\{\lfloor q/2\rfloor,\lceil q/2\rceil\}} \beta_S + f(q+3,q,\gamma) \sum_{S:|S|\notin\{\lfloor q/2\rfloor,\lceil q/2\rceil\}} \beta_S \\
&= f(q+1,q,\gamma) \sum_{S:|S|\in\{\lfloor q/2\rfloor,\lceil q/2\rceil\}} \beta_S + f(q+3,q,\gamma) \left(1 - \sum_{S:|S|\in\{\lfloor q/2\rfloor,\lceil q/2\rceil\}} \beta_S\right) \\
&= \left(f(q+1,q,\gamma) - f(q+3,q,\gamma)\right) \sum_{S:|S|\in\{\lfloor q/2\rfloor,\lceil q/2\rceil\}} \beta_S + f(q+3,q,\gamma).
\end{aligned}$$

By Claim 1, we have $f(q+1,q,1/4) - f(q+3,q,1/4) > 0$ for all $q \ge 5$. Since for fixed q and m, $f(m,q,\gamma)$ is continuous as a function of γ on $[0,1/4]$, then

$$f(q+1,q,\gamma) - f(q+3,q,\gamma) > 0$$

for all γ sufficiently close to $1/4$. Therefore,

$$\begin{aligned}
\sum_{S:|S|\in\{\lceil q/2\rceil,\lfloor q/2\rfloor\}} \beta_S &\ge \frac{-f(q+3,q,\gamma)}{-f(q+3,q,\gamma)+f(q+1,q,\gamma)} \\
&= \frac{\log\left(\frac{q+3}{\sqrt{q^2-1}}\right) - 2\sqrt{\frac{q+3}{q}} - \frac{\sqrt{1-4\gamma}}{2}\log\left(\frac{q+1}{q-1}\right) - 4\gamma + 3}{\log\left(\frac{q+3}{q+1}\right) + 2\left(\sqrt{\frac{q+1}{q}} - \sqrt{\frac{q+3}{q}}\right)} =: \mu(q,\gamma).
\end{aligned}$$

Let

$$A = \log\left(\frac{q+3}{q+1}\right) + 2\left(\sqrt{\frac{q+1}{q}} - \sqrt{\frac{q+3}{q}}\right).$$

Then

$$\frac{d}{d\gamma}\mu(q,\gamma) = \frac{1}{A}\left(\frac{\log\left(\frac{q+1}{q-1}\right)}{\sqrt{1-4\gamma}} - 4\right).$$

Since for all $q \ge 5$ we have $A < 0$ and

$$\frac{\log\left(\frac{q+1}{q-1}\right)}{\sqrt{1-4\gamma}} - 4 > 0,$$

for $\gamma \le 1/4$ sufficiently close to $1/4$, then $\frac{d}{d\gamma}\mu(q,\gamma) < 0$ and

$$\mu(q,\gamma) \ge \mu(q,1/4) = \frac{\log\left(\frac{q+3}{\sqrt{q^2-1}}\right) - 2\sqrt{\frac{q+3}{q}} + 2}{\log\left(\frac{q+3}{q+1}\right) + 2\left(\sqrt{\frac{q+1}{q}} - \sqrt{\frac{q+3}{q}}\right)} \tag{16}$$

for all $q \ge 5$ and $\gamma \le 1/4$ sufficiently close to $1/4$.

Note that $\mu(5,1/4) \approx 0.768$. We prove the following claim to show that $\mu(q,1/4) \ge 0.8$ for all $q \ge 7$.

Claim 2. *For all $q \geq 7$, $\mu(q, 1/4) \geq 4/5$.*

Proof. We will prove that, for all $q \geq 7$,

$$\frac{\log\left(\frac{q+3}{\sqrt{q^2-1}}\right) - 2\sqrt{\frac{q+3}{q}} + 2}{\log\left(\frac{q+3}{q+1}\right) + 2\left(\sqrt{\frac{q+1}{q}} - \sqrt{\frac{q+3}{q}}\right)} \geq \frac{4}{5}. \tag{17}$$

Note that since $f(x) = x - \log(x)$ is strictly increasing for all $x > 1$, then

$$\sqrt{\frac{q+1}{q}} - \log\left(\sqrt{\frac{q+1}{q}}\right) < \sqrt{\frac{q+3}{q}} - \log\left(\sqrt{\frac{q+3}{q}}\right),$$

and hence, the denominator of the left side of (17) is negative for all $q \geq 7$. Therefore, proving (17) is equivalent to showing that

$$2\sqrt{\frac{q+3}{q}} - 2\log\sqrt{\frac{q+3}{q}} + 8\sqrt{\frac{q+1}{q}} - 3\log\sqrt{\frac{q+1}{q}} - 5\log\sqrt{\frac{q}{q-1}} - 10 \geq 0.$$

With the substitution $q = \frac{1}{x}$, it suffices to prove that, for all $x \in [0, 1/7]$,

$$F(x) := 2\sqrt{1+3x} - \ln(1+3x) + 8\sqrt{1+x} - \frac{3}{2}\ln(1+x) + \frac{5}{2}\ln(1-x) - 10 \geq 0.$$

We have

$$F'(x) = \frac{3}{\sqrt{1+3x}} - \frac{3}{1+3x} + \frac{4}{\sqrt{1+x}} - \frac{3}{2(1+x)} - \frac{5}{2(1-x)},$$

$$F''(x) = -\frac{9}{2(1+3x)^{3/2}} + \frac{9}{(1+3x)^2} - \frac{2}{(1+x)^{3/2}} + \frac{3}{2(1+x)^2} - \frac{5}{2(1-x)^2},$$

and

$$F'''(x) = \frac{81}{4(1+3x)^{5/2}} - \frac{54}{(1+3x)^3} + \frac{3}{(1+x)^{5/2}} - \frac{3}{(1+x)^3} - \frac{5}{(1-x)^3}.$$

We have, for all $x \in [0, 1/7]$,

$$\begin{aligned} F'''(x) &\leq \left(\frac{81}{4(1+3x)^2} - \frac{54}{(1+3x)^3}\right) + \left(\frac{3}{(1+x)^2} - \frac{3}{(1+x)^3} - \frac{5}{(1+x)^3}\right) \\ &= \frac{27(9x-5)}{4(1+3x)^3} + \frac{3x-5}{(1+x)^3} \\ &< 0. \end{aligned}$$

Note that $F''(0) > 0$ and $F''(1/7) < 0$. Thus, there exists $x_0 \in (0, 1/7)$ such that $F''(x_0) = 0$, $F''(x) > 0$ for $x \in [0, x_0)$, and $F''(x) < 0$ for $x \in (x_0, 1/7]$. Since $F'(0) = 0$ and $F'(1/7) < 0$, there exists $x_1 \in (x_0, 1/7)$ such that $F'(x_1) = 0$, $F'(x) > 0$ on $(0, x_1)$, and $F'(x) < 0$ on $(x_1, 1/7)$. Note that $F(0) = 0$ and $F(1/7) > 0$. Thus, $F(x) \geq 0$ on $[0, 1/7]$ and (17) holds. ■

Therefore, by (16) and (17), we have

$$\sum_{S:|S|=\lfloor q/2\rfloor, \lceil q/2\rceil} \beta_S \geq \begin{cases} 0.76 & \text{if } q = 5 \\ 0.80 & \text{if } q \geq 7. \end{cases},$$

as long as $\boldsymbol{\beta} \in \mathrm{FEAS}_q(\gamma)$ for $\gamma \leq 1/4$ sufficiently close to $1/4$, as desired. ■

We are ready to embark on the proof of Theorem 7. We use the following approach:

(i) Assume that $\boldsymbol{\beta} \in \mathrm{FEAS}_q(\gamma)$ is a solution of **OPT** for odd $q \geq 5$. Then since the vector $\boldsymbol{\alpha}$ defined in the statement of Theorem 7 is also a feasible vector, we must have

$$\mathrm{OBJ}_q(\boldsymbol{\beta}) \geq \mathrm{OBJ}_q(\boldsymbol{\alpha}). \tag{18}$$

(ii) By Lemma 4, (11) holds. Out of all such $\boldsymbol{\beta}$ we pick the one such that $|\mathrm{supp}_q(\boldsymbol{\beta})|$ is minimum. By Lemma 2, $\mathrm{supp}_q(\boldsymbol{\beta})$ cannot contain two intersecting sets of the same size. Therefore, $\mathrm{supp}_q(\boldsymbol{\beta})$ contains at most one set of size $\lceil q/2 \rceil$ and at most two sets of size $\lfloor q/2 \rfloor$. In addition, (11) implies that $\mathrm{supp}_q(\boldsymbol{\beta})$ must have at least one set of size $\lceil q/2 \rceil$ or $\lfloor q/2 \rfloor$. Therefore, $\mathrm{supp}_q(\boldsymbol{\beta})$ must fall into one of the following disjoint cases.

The support of $\boldsymbol{\beta}$ contains

Case 1: no set of size $\lceil q/2 \rceil$ and exactly one set of size $\lfloor q/2 \rfloor$, or vice versa.

Case 2: no set of size $\lceil q/2 \rceil$ and exactly two sets of size $\lfloor q/2 \rfloor$.

Case 3: exactly one set A of size $\lceil q/2 \rceil$, a set B of size $\lfloor q/2 \rfloor$, but not A^c.

Case 4: exactly one set A of size $\lceil q/2 \rceil$ and A^c, and $|\mathrm{supp}_q(\boldsymbol{\beta})| > 2$.

Case 5: exactly one set A of size $\lceil q/2 \rceil$ and A^c only.

We show that cases 1, 2, 3, and 4 are impossible, and conclude that $\boldsymbol{\beta}$ must fall into Case 5. From this, we easily obtain that $\boldsymbol{\beta}$ must be of the same form as $\boldsymbol{\alpha}$ in Theorem 7.

Proof of Theorem 7. Throughout the proof we will be making a series of claims which hold for all γ sufficiently close to $1/4$. Suppose that $\boldsymbol{\beta} \in \mathrm{FEAS}_q(\gamma)$ is a solution of **OPT** for odd $q \geq 5$ such that $|\mathrm{supp}_q(\boldsymbol{\beta})|$ is minimum. Then

$$\mathrm{OBJ}_q(\boldsymbol{\beta}) \geq \mathrm{OBJ}_q(\boldsymbol{\alpha}),$$

where $\boldsymbol{\alpha}$ is a vector as defined in the statement of Theorem 7, and by Lemma 4,

$$\sum_{S:|S|=\lfloor q/2 \rfloor, \lceil q/2 \rceil} \beta_S \geq \psi(q) = \begin{cases} 0.76 & \text{if } q = 5 \\ 0.80 & \text{if } q \geq 7, \end{cases} \tag{19}$$

Let us recall our constraints for feasible vectors. We have

$$\mathrm{V}_q(\boldsymbol{\beta}) = \sum_S \beta_S = 1 \tag{20}$$

and

$$1 - 2\gamma \geq 1 - 2\,\mathrm{E}_q(\boldsymbol{\beta}) = \sum_{(B,S):B \cap S \neq \emptyset} \beta_B \beta_S \tag{21}$$

by (2).

As was explained above, we may proceed with our five cases.

Case 1: *The support of $\boldsymbol{\beta}$ contains no set of size $\lceil q/2 \rceil$ and exactly one set of size $\lfloor q/2 \rfloor$, or vice versa.*

Let S be the only set of size $\lfloor q/2 \rfloor$ in $\mathrm{supp}_q(\boldsymbol{\beta})$ and suppose that no sets of size $\lceil q/2 \rceil$ are present in $\mathrm{supp}_q(\boldsymbol{\beta})$. Then by (19), $\beta_S \geq \psi(q)$, and using (21), we have

$$\psi(q)^2 \leq \beta_S^2 \leq 1 - 2\,\mathrm{E}_q(\boldsymbol{\beta}) \leq 1 - 2\gamma.$$

However, as $\gamma \to 0.25^-$, note that $\psi(5)^2 \to 0.76^2 \geq 0.57$, $\psi(q)^2 \to 0.8^2 = 0.64$ for $q \geq 7$, and $1 - 2\gamma \to 0.5$. Therefore, the inequality above cannot hold for all γ sufficiently close to $1/4$. Thus, $\boldsymbol{\beta} \notin \mathrm{FEAS}_q(\gamma)$, which is a contradiction.

Exactly the same argument holds if S is the only set of size $\lceil q/2 \rceil$ in $\mathrm{supp}_q(\boldsymbol{\beta})$ and there are no sets of size $\lfloor q/2 \rfloor$ in $\mathrm{supp}_q(\boldsymbol{\beta})$. In either subcase we obtain a contradiction and therefore, this entire case is impossible.

For Cases 2, 3, and 4 we apply the results from [15] that were stated in Section 3.

Case 2: *The support of $\boldsymbol{\beta}$ contains no set of size $\lceil q/2 \rceil$ and exactly two sets of size $\lfloor q/2 \rfloor$.*

Let S_1 and S_2 be the only two sets of size $\lfloor q/2 \rfloor$ in $\mathrm{supp}_q(\boldsymbol{\beta})$. By Lemma 2, we have $S_1 \cap S_2 = \emptyset$. Then Theorem 6 implies that we have only two possibilies for $\mathrm{supp}_q(\boldsymbol{\beta})$, which we consider below.

Subcase 2.1: $\mathrm{supp}_q(\boldsymbol{\beta}) = \{S_1, S_2, (S_1 \cup S_2)^c\}$.

Since $\beta_{S_1} + \beta_{S_2} + \beta_{(S_1 \cup S_2)^c} = 1$, we have

$$\begin{aligned}
\mathrm{OBJ}_q(\boldsymbol{\beta}) &= \sum_S \log(|S|)\beta_S \\
&= \log\left(\lfloor q/2 \rfloor\right)\left(\beta_{S_1} + \beta_{S_2}\right) \\
&\leq \log\left(\frac{q-1}{2}\right) \\
&< \frac{1}{2}\log\left(\frac{q^2-1}{4}\right) \quad \text{for all } q \geq 5 \\
&\leq \frac{1}{2}\log\left(\frac{q^2-1}{4}\right) + \frac{\sqrt{1-4\gamma}}{2}\log\left(\frac{q+1}{q-1}\right) \\
&= \mathrm{OBJ}_q(\boldsymbol{\alpha}).
\end{aligned}$$

However, this contradicts (18), and thus, this case is impossible.

Subcase 2.2: $\mathrm{supp}_q(\boldsymbol{\beta}) = \{S_1, S_2, S_1 \cup S_2, (S_1 \cup S_2)^c\}$.

In this case, the entries of $\boldsymbol{\beta}$ are subject to the following constraints

$$\beta_{S_1} + \beta_{S_2} + \beta_{S_1 \cup S_2} + \beta_{(S_1 \cup S_2)^c} = 1, \tag{22}$$

$$\left(\beta_{S_1 \cup S_2} + \beta_{S_1} + \beta_{S_2}\right)^2 - 2\beta_{S_1}\beta_{S_2} < 1 - 2\gamma, \tag{23}$$

$$\beta_{S_1} + \beta_{S_2} \geq \psi(q), \tag{24}$$

where $\beta_{S_1} > 0$, $\beta_{S_2} > 0$, $\beta_{S_1 \cup S_2} > 0$, and $\beta_{(S_1 \cup S_2)^c} > 0$. Solving (23) for $\beta_{S_1 \cup S_2}$ we obtain

$$\beta_{S_1 \cup S_2} \leq \sqrt{1 - 2\gamma + 2\beta_{S_1}\beta_{S_2}} - \left(\beta_{S_1} + \beta_{S_2}\right),$$

noting that $1-2\gamma+2\beta_{S_1}\beta_{S_2}\geq 0$ for any $\gamma\leq 1/4$. Let $t=\beta_{S_1}+\beta_{S_2}$. Then by (22) and (23), we have $\psi(q)\leq t\leq 1$. Since $2\beta_{S_1}\beta_{S_2}\leq t^2/2$, we have

$$\beta_{S_1\cup S_2}\leq\sqrt{1-2\gamma+t^2/2}-t.$$

Thus,

$$\begin{aligned}\mathrm{OBJ}_q(\boldsymbol{\beta})&=\log\left(\lfloor q/2\rfloor\right)t+\log(q-1)\beta_{S_1\cup S_2}\\&\leq\log\left(\frac{q-1}{2}\right)t+\log(q-1)(\sqrt{1-2\gamma+t^2/2}-t)\\&\leq-\log(2)t+\log(q-1)\sqrt{1-2\gamma+t^2/2}=:\lambda(t,q,\gamma).\end{aligned}$$

We will maximize $\lambda(t,q,\gamma)$ with respect to t. Note that $2t^2-8\gamma+4\geq 0$ for $\gamma\leq 1/4$, and hence,

$$\frac{d\lambda}{dt}=-\log(2)+\frac{\log(q-1)\,t}{\sqrt{2t^2-8\gamma+4}}\geq 0$$

when

$$t\geq\sqrt{\frac{\log(2)\,(4-8\gamma)}{\log(q-1)^2-2\log(2)^2}}=:\tau(q,\gamma).$$

Let us first consider when $q=5$. Note that

$$\tau(5,\gamma)>0.76=\psi(5)$$

for all $\gamma\leq 1/4$. Thus, $\lambda(t,5,\gamma)$ is decreasing with respect to t on $[\psi(5),\tau(5,\gamma))$ and is increasing with respect to t on $[\tau(5,\gamma),1]$. Thus,

$$\mathrm{OBJ}_5(\boldsymbol{\beta})\leq\max\{\lambda(\psi(5),5,\gamma),\lambda(1,5,\gamma)\}.$$

Note that

$$\lim_{\gamma\to\frac{1}{4}^-}\lambda(\psi(5),5,\gamma)=\lambda(\psi(5),5,1/4)=-\log(2)\cdot 0.76+\log(4)\sqrt{1/2+(0.76)^2/2}\approx 0.70$$

and

$$\lim_{\gamma\to\frac{1}{4}^-}\lambda(1,5,\gamma)=\lambda(1,5,1/4)=\log(2)\approx 0.69.$$

Since $\lambda(t,5,\gamma)$ is continuous with respect to γ, then

$$\begin{aligned}\mathrm{OBJ}_5(\boldsymbol{\beta})&\leq\lambda(\psi(5),5,\gamma)\\&=-\log(2)\cdot 0.76+\log(4)\sqrt{1-2\gamma+(0.76)^2/2}\end{aligned}$$

for γ sufficiently close to $1/4$. Since

$$\frac{1}{2}\log\left(\frac{5^2-1}{4}\right)=\frac{\log(6)}{2}\approx 0.89,$$

then

$$\begin{aligned}
\mathrm{OBJ}_5(\boldsymbol{\beta}) &\le \lambda(\psi(5), 5, \gamma) \\
&< \frac{1}{2}\log\left(\frac{5^2-1}{4}\right) \quad \text{for } \gamma \text{ sufficiently close to } 1/4 \\
&\le \frac{1}{2}\log\left(\frac{5^2-1}{4}\right) + \frac{\sqrt{1-4\gamma}}{2}\log\left(\frac{3}{2}\right) \\
&= \mathrm{OBJ}_5(\boldsymbol{\alpha}).
\end{aligned}$$

We have contradicted (18), and hence, this subcase is impossible for $q = 5$.

Now suppose that $q \ge 7$. Note that

$$\lim_{\gamma \to \frac{1}{4}^-} \tau(q,\gamma) = \tau(q, 1/4) = \sqrt{\frac{2\log(2)}{\log(q-1)^2 - 2\log(2)^2}} \le 0.80 = \psi(q)$$

for all $q \ge 7$. Since $\tau(q,\gamma)$ is decreasing with respect to γ as $\gamma \to 0.25^-$, then

$$\tau(q,\gamma) \le \psi(q)$$

for all $\gamma \le 1/4$ and $q \ge 7$. Thus, $\lambda(t, q, \gamma)$ is increasing on the interval $[\psi(q), 1]$ with respect to t, and hence,

$$\mathrm{OBJ}_q(\boldsymbol{\beta}) \le \lambda(1, q, \gamma).$$

Notice that

$$\lim_{\gamma \to \frac{1}{4}^-} \lambda(1,q,\gamma) = \lambda(1, q, 1/4) = \frac{\log(q-1)}{2} < \frac{1}{2}\log\left(\frac{q^2-1}{4}\right),$$

where the last inequality holds for all $q \ge 7$. Thus,

$$\begin{aligned}
\mathrm{OBJ}_q(\boldsymbol{\beta}) &\le \lambda(1, q, \gamma) \\
&< \frac{1}{2}\log\left(\frac{q^2-1}{4}\right) \quad \text{for } \gamma \text{ sufficiently close to } 1/4 \\
&\le \frac{1}{2}\log\left(\frac{q^2-1}{4}\right) + \frac{\sqrt{1-4\gamma}}{2}\log\left(\frac{q+1}{q-1}\right) \\
&= \mathrm{OBJ}_q(\boldsymbol{\alpha}).
\end{aligned}$$

However, this contradicts (18), and thus, this subcase is also impossible for $q \ge 7$.

Case 3: *The support of $\boldsymbol{\beta}$ contains exactly one set A of size $\lceil q/2 \rceil$, a set B of size $\lfloor q/2 \rfloor$, but not A^c.*

Since $A^c \notin \mathrm{supp}_q(\boldsymbol{\beta})$, then $A \cap B \neq \emptyset$. If B is the only set of size $\lfloor q/2 \rfloor$ in $\mathrm{supp}_q(\boldsymbol{\beta})$, then $\beta_A + \beta_B \ge \psi(q)$, and hence, by (21), we have

$$\psi(q)^2 \le (\beta_A + \beta_B)^2 \le 1 - 2\,\mathrm{E}_q(\boldsymbol{\beta}) \le 1 - 2\gamma.$$

However, as was shown in Case 1, the inequality above cannot hold for all γ sufficiently close to $1/4$. Thus, $\boldsymbol{\beta} \notin \mathrm{FEAS}_q(\gamma)$, a contradiction.

We may assume that there exists another set $C \in \operatorname{supp}_q(\boldsymbol{\beta}) \setminus \{B\}$ that satisfies $|C|= \lfloor q/2 \rfloor$. By Lemma 2, $B \cap C = \emptyset$ and B and C are the only sets of size $\lfloor q/2 \rfloor$ in $\operatorname{supp}_q(\boldsymbol{\beta})$. Since B and C are disjoint, but they each intersect with the set A, Theorem 6 implies that $A = B \cup C$ and that

$$\operatorname{supp}_q(\boldsymbol{\beta}) = \{A, B, C, (B \cup C)^c\}.$$

However, then we would have

$$\frac{q+1}{2} = |A|= |B|+|C|= q-1,$$

which is impossible. Therefore, this entire case is impossible.

Case 4: *The support of $\boldsymbol{\beta}$ contains exactly one set A of size $\lceil q/2 \rceil$ and A^c, and $|\operatorname{supp}_q(\boldsymbol{\beta})|> 2$.*

Since $A, A^c \in \operatorname{supp}_q(\boldsymbol{\beta})$ and $|\operatorname{supp}_q(\boldsymbol{\beta})|> 2$, then any other set in $\operatorname{supp}_q(\boldsymbol{\beta})$ must intersect A or A^c. Thus, Theorem 6 implies that one set in $\operatorname{supp}_q(\boldsymbol{\beta})$ must be the union of precisely two other set in $\operatorname{supp}_q(\boldsymbol{\beta})$. Therefore, by Observation 1 we have

$$|\operatorname{supp}_q(\boldsymbol{\beta})|> \lceil 1/(1-2\gamma) \rceil + 1 = 3,$$

since $\lceil 1/(1-2\gamma) \rceil = 2$ for γ sufficiently close to $1/4$. Thus, the only possibility is that

$$\operatorname{supp}_q(\boldsymbol{\beta}) = \{A, A^c, A_1, A_2\},$$

where A_1 and A_2 are disjoint sets which form a 2-partition of A or A^c.

Without loss of generality, assume that A_1 and A_2 form a 2-partition of A. Then the entries of $\boldsymbol{\beta}$ are subject to the following constraints

$$\beta_A + \beta_{A^c} + \beta_{A_1} + \beta_{A_2} = 1, \tag{25}$$

$$\beta_A^2 + \beta_{A^c}^2 + \beta_{A_1}^2 + \beta_{A_2}^2 + 2\beta_A(\beta_{A_1} + \beta_{A_2}) \le 1 - 2\gamma, \tag{26}$$

$$\beta_A + \beta_{A^c} \ge \psi(q), \tag{27}$$

where $\beta_A > 0$, $\beta_{A^c} > 0$, $\beta_{A_1} > 0$, and $\beta_{A_2} > 0$. We can combine (26) and (26) by substituting $\beta_{A^c} = 1 - \beta_A - \beta_{A_1} - \beta_{A_2}$ into (26) to obtain

$$\begin{aligned} 1 - 2\gamma &\ge \beta_A^2 + \beta_{A^c}^2 + \beta_{A_1}^2 + \beta_{A_2}^2 + 2\beta_A(\beta_{A_1} + \beta_{A_2}) \\ &= \frac{1}{2}(2\beta_A + 2\beta_{A_1} + 2\beta_{A_2} - 1)^2 + \frac{1}{2} - \beta_{A_1}^2 - \beta_{A_2}^2 - 2\beta_{A_1}\beta_{A_2}. \end{aligned} \tag{28}$$

By solving for β_A in (28) we obtain

$$\beta_A \le \frac{\sqrt{1 - 4\gamma + 2(\beta_{A_1}^2 + \beta_{A_2}^2 + 2\beta_{A_1}\beta_{A_2})} + 1}{2} - (\beta_{A_1} + \beta_{A_2}).$$

Let $t = \beta_{A_1} + \beta_{A_2}$. Then $0 < t \le 1 - \psi(q)$ by (25) and (27). Note that

$$\beta_{A_1}^2 + \beta_{A_2}^2 + 2\beta_{A_1}\beta_{A_2} \le t^2,$$

and hence, we have

$$\beta_A \le \frac{\sqrt{1 - 4\gamma + 2t^2} + 1}{2} - t.$$

Thus,

$$\begin{aligned}\mathrm{OBJ}_q(\boldsymbol{\beta}) &\le \log\left(\lfloor q/2\rfloor\right)\left(\beta_{A^c}+\beta_{A_1}+\beta_{A_2}\right)+\log\left(\lceil q/2\rceil\right)\beta_A\\ &= \log\left(\frac{q-1}{2}\right)+\log\left(\frac{q+1}{q-1}\right)\beta_A, \quad \text{by (25)}\\ &\le \log\left(\frac{q-1}{2}\right)+\log\left(\frac{q+1}{q-1}\right)\left(\frac{\sqrt{1-4\gamma+2t^2}+1}{2}-t\right) =: \rho(t,q,\gamma).\end{aligned}$$

We will maximize $\rho(t,q,\gamma)$ with respect to t. Note that $\sqrt{1-4\gamma+2t^2}\ge 0$ for $\gamma\le 1/4$, and hence,

$$\frac{d\rho}{dt}=\frac{t}{\sqrt{1-4\gamma+2t^2}}+1<0$$

for

$$t>\sqrt{4\gamma-1}.$$

Since $4\gamma-1\le 0<t$, then $\rho(t,q,\gamma)$ is always decreasing with respect to t on $[0,1-\psi(q)]$. Thus,

$$\begin{aligned}\mathrm{OBJ}_q(\boldsymbol{\beta}) &< \rho(0,q,\gamma)\\ &= \log\left(\frac{q-1}{2}\right)+\log\left(\frac{q+1}{q-1}\right)\frac{\sqrt{1-4\gamma}+1}{2}\\ &= \frac{1}{2}\log\left(\frac{q^2-1}{4}\right)+\frac{\sqrt{1-4\gamma}}{2}\log\left(\frac{q+1}{q-1}\right)\\ &= \mathrm{OBJ}_q(\boldsymbol{\alpha}),\end{aligned}$$

which contradicts (18). Therefore, this case is impossible.

Case 5: *The support of* $\boldsymbol{\beta}$ *contains exactly one set* A *of size* $\lceil q/2\rceil$ *and* A^c *only.*

For any $\gamma\le 1/4$, define the roots of the equation $x(1-x)=\gamma$ by

$$M(\gamma)^+=\frac{1+\sqrt{1-4\gamma}}{2}$$

and

$$M(\gamma)^-=\frac{1-\sqrt{1-4\gamma}}{2}.$$

In this case, we have $\mathrm{supp}_q(\boldsymbol{\beta})=\{A,A^c\}$. Then by (20), we have

$$\beta_A+\beta_{A^c}=1,$$

and we also have

$$\mathrm{E}_q(\boldsymbol{\beta})=\beta_A\beta_{A^c}=\beta_A(1-\beta_A)\ge\gamma,$$

which implies that

$$M(\gamma)^-\le\beta_A\le M(\gamma)^+.$$

Similarly, we can also show that

$$M(\gamma)^-\le\beta_{A^c}\le M(\gamma)^+.$$

Therefore, it is clear that in this case, we have

$$\begin{aligned}\mathrm{OBJ}_q(\boldsymbol{\beta}) &= \log\left(\frac{q-1}{2}\right)\beta_{A^c} + \log\left(\frac{q+1}{2}\right)\beta_A \\ &\leq \log\left(\frac{q-1}{2}\right)M(\gamma)^- + \log\left(\frac{q+1}{2}\right)M(\gamma)^+ \\ &= \mathrm{OBJ}_q(\boldsymbol{\alpha}),\end{aligned}$$

where equality holds everywhere if and only if $\beta_{A^c} = M(\gamma)^-$ and $\beta_A = M(\gamma)^+$. That is, equality holds if and only if $\boldsymbol{\beta}$ is of the same form as $\boldsymbol{\alpha}$ in the statement of Theorem 7.

We have considered all possible cases for the solution vector $\boldsymbol{\beta}$ and have shown that the only possibility is that $\boldsymbol{\beta}$ must be of the same form as the vector $\boldsymbol{\alpha}$ defined in the statement of Theorem 7.

Our proof is complete. ■

Corollary 1. *Let n be a positive integer. If $\boldsymbol{\alpha}$ solves $\mathrm{OPT}_q(1/4)$ for $q \geq 5$, then the graph $G_{\boldsymbol{\alpha}}(n)$ is isomorphic to $T_2(n)$.*

Proof. If $\boldsymbol{\alpha}$ is a solution of $\mathrm{OPT}_q(1/4)$ then according to Theorem 7, $\mathrm{supp}_q(\boldsymbol{\alpha}) = \{A, A^c\}$, where $A \subseteq [q]$ such that $|A| = \lceil q/2 \rceil$, and $\alpha_A = \alpha_{A^c} = 1/2$. Then by Construction $\boldsymbol{G_{\alpha}(n)}$, the graph $G_{\boldsymbol{\alpha}}(n)$ is a complete 2-partite graph with parts V_A and V_{A^c}, where $|V_A| = \lfloor n/2 \rfloor$ and $|V_{A^c}| = \lceil n/2 \rceil$, or vice versa. Therefore, $G_{\boldsymbol{\alpha}}(n) \cong K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$, which is isomorphic to $T_2(n)$. ■

5 Approximate Version of Theorem 1

Recall that the Turán graph, $T_2(n)$, is the complete 2-partite graph with parts of size $\lfloor n/2 \rfloor$ and $\lceil n/2 \rceil$. We denote the number of edges in $T_2(n)$ by $t_2(n)$. For positive integers n and m, we call a graph G a *(n, m)-graph* if it has n vertices and m edges.

For the reader's convenience, we remind them that for $\delta > 0$, two graphs G and H, each with n vertices, are *δn^2-close* if their *edit distance*, the minimum number of edges that need to be added or deleted from one graph to obtain a graph isomorphic to the other, is at most δn^2.

This section is dedicated to proving an "approximate", or "local", version of Theorem 1. This version is nearly the same as Theorem 1, but has an additional requirement: namely, that a $(n, t_2(n))$-graph G must be δn^2-close to $T_2(n)$ for sufficiently small $\delta > 0$. That is, $T_2(n)$ "locally maximizes" the number of q-colorings among the class of $(n, t_2(n))$-graphs for $q \in \{5, 7\}$.

The main result of this section is as follows.

Theorem 8. *There exists a $\delta > 0$ such that the following holds for sufficiently large n. Let $q \geq 2$ be an odd integer and let G be a $(n, t_2(n))$-graph such that G is δn^2-close to $T_2(n)$. Then G has at most as many q-colorings as $T_2(n)$, with equality holding if and only if G is isomorphic to $T_2(n)$.*

Intuitively, Theorem 8 states that if a graph is "close", with respect to edit distance, in structure to $T_2(n)$ then the number of its q-colorings is at most $P_{T_2(n)}(q)$ for $q \in \{5, 7\}$.

The following four results will be referenced throughout the proof of Theorem 8. The first result, Lemma 5, states the existence of a particular partition of the vertex set of a $(n, t_2(n))$-graph that is δn^2-close $T_2(n)$.

Lemma 5. *Let $\delta > 0$. If G is a $(n, t_2(n))$-graph that is δn^2-close to $T_2(n)$ then there exists a partition $A_1 \cup A_2 = V(G)$ such that $e(G[A_1, A_2]) \geq t_2(n) - \frac{\delta}{2} n^2$.*

Proof. Suppose that $T_2(n)$ has vertex set partition $A_1 \cup A_2 = V(T_2(n))$. Since G is δn^2-close to $T_2(n)$, then by definition, $V(G) = V(T_2(n))$. We claim that

$$e(G[A_1, A_2]) \geq t_2(n) - \frac{\delta}{2}n^2.$$

First observe that

$$E(G[A_1, A_2]) = E(G) \cap E(T_2(n)) = (E(G) \cup E(T_2(n))) \setminus (E(G) \triangle E(T_2(n))), \tag{29}$$

where the first equality holds since $T_2(n)$ has all possible edges between the parts A_1 and A_2. By taking the cardinalities of each set in (29) we have

$$\begin{aligned} e(G[A_1, A_2]) &= |E(G) \cup E(T_2(n))| - |E(G) \triangle E(T_2(n))| \\ &= e(G) + e(T_2(n)) - |E(G) \cap E(T_2(n))| - |E(G) \triangle E(T_2(n))| \\ &\geq 2t_2(n) - e(G[A_1, A_2]) - \delta n^2. \end{aligned}$$

Therefore,

$$e(G[A_1, A_2]) \geq t_2(n) - \frac{\delta}{2}n^2.$$

■

The following result is Proposition 5.13 from Babai and Frankl in [1].

Lemma 6 ([1])**.** *For positive integers r and t such that $t \leq r/4$ we have*

$$\sum_{i=0}^{t} \binom{r}{i} < 2\binom{r}{t}.$$

The next result, from [9], provides an upper bound for the number of q-colorings of a graph G. This bound will be used in some of the bounding arguments in the proof of Theorem 8. Recall that the number of q-colorings of a graph G is denoted by $P_G(q)$, and is called the *chromatic polynomial* of G.

Lemma 7 ([9])**.** *Let G be an (n, m)-graph and let $q \geq 2$ be an integer. Then*

$$P_G(q) \leq \left(1 - \frac{1}{q}\right)^{\lceil(\sqrt{1+8m}-1)/2} q^n \leq \left(1 - \frac{1}{q}\right)^{\lceil(\sqrt{m}-1)/2} q^n.$$

Let us now examine the asymptotic behavior of $P_{T_2(n)}(q)$ for a fixed odd integer $q \geq 2$ as $n \to \infty$. Since $T_2(n)$ is a 2-partite graph, one can view its chromatic polynomial as the sum of the number of q-colorings for which one part is colored with precisely i colors and the other part with any nonempty subset of the remaining $q - i$ colors, where $1 \leq i \leq q - 1$. In the following lemma, we extract the terms of highest magnitude of this sum, specifically those which correspond to the q-colorings for which both parts are colored with roughly $q/2$ colors. All other colorings in the sum constitute little-o of this number of q-colorings as $n \to \infty$.

Lemma 8. *Let $q \geq 3$ be an odd integer and $n = 2k + s$, where $k \geq 1$ is an integer and $s \in \{0, 1\}$. Then as $n \to \infty$,*

$$P_{T_2(n)}(q) = \left((qs + 2(1-s))\binom{q}{\lfloor q/2 \rfloor} + o(1)\right)(\lfloor q/2 \rfloor \lceil q/2 \rceil)^{\lfloor n/2 \rfloor}.$$

Proof. Since $T_2(n)$ is isomorphic to the complete 2-partite graph $K_{\lfloor n/2\rfloor,\lceil n/2\rceil}$, we let $V(T_2(n)) = A\cup B$, where $|A|=\lfloor n/2\rfloor = k$ and $|B|=\lceil n/2\rceil = k+s$. We can properly color the vertices of $T_2(n)$ by first coloring the vertices of A and then of B as follows:

(i) Select i colors from $[q]$ with which to color the vertices of A with. There are $\binom{q}{i}$ possible choices.

(ii) Color the vertices of A using *exactly* i colors. Select a partition of A into i non-empty sets. The number of ways to partition a set of k objects into i non-empty subsets is denoted by $S(k,i)$, a Stirling number of the second kind. For each partition of A into i parts, we can distribute the i colors we have selected in $i!$ ways.

(iii) Color the vertices of B with any nonempty subset of the $q-i$ remaining colors that were not used to color the vertices of A. There are $(q-i)^{k+s}$ possibilities.

Therefore, by summing $\binom{q}{i}S(k,i)i!\,(q-i)^{k+s}$ over all $i\in[q-1]$ we obtain

$$P_{T_2(n)}(q) = \sum_{i=1}^{q}\binom{q}{i}S(k,i)i!\,(q-i)^{k+s}.$$

It is well-known that (see Wilf, [21])

$$S(k,i) := \frac{1}{i!}\sum_{j=0}^{i}(-1)^j\binom{i}{j}(i-j)^k.$$

Note that the term of largest magnitude in $S(k,i)$ is $i^k/i!$. Therefore, since i is fixed, we can write that

$$\begin{aligned}
S(k,i)i! &= \left(\sum_{j=0}^{i}(-1)^j\binom{i}{j}\left(\frac{i-j}{i}\right)^k\right)i^k\\
&= \left(1 - i\left(\frac{i-1}{i}\right)^k + \binom{i}{2}\left(\frac{i-2}{i}\right)^k - \ldots + (-1)^{i-1}i\right)i^k\\
&= (1+o(1))i^k, \quad \text{as } k\to\infty.
\end{aligned}$$

Therefore,

$$\begin{aligned}
P_{T_2(n)}(q) &= \sum_{i=1}^{q}\binom{q}{i}S(k,i)i!\,(q-i)^{k+s}\\
&= \binom{q}{\lfloor q/2\rfloor}\cdot S(k,\lfloor q/2\rfloor)\cdot\lfloor q/2\rfloor!\cdot\lceil q/2\rceil^{k+s}\\
&+ \binom{q}{\lceil q/2\rceil}\cdot S(k,\lceil q/2\rceil)\cdot\lceil q/2\rceil!\cdot\lfloor q/2\rfloor^{k+s} + \sum_{\substack{i=1\\ i\notin\{\lfloor q/2\rfloor,\lceil q/2\rceil\}}}^{q-1}\binom{q}{i}S(k,i)i!\,(q-i)^{k+s}\\
&= \binom{q}{\lfloor q/2\rfloor}(1+o(1))\lfloor q/2\rfloor^k\lceil q/2\rceil^{k+s} + \binom{q}{\lceil q/2\rceil}(1+o(1))\lceil q/2\rceil^k\lfloor q/2\rfloor^{k+s}\\
&+ \sum_{\substack{i=1\\ i\notin\{\lfloor q/2\rfloor,\lceil q/2\rceil\}}}^{q-1}\binom{q}{i}(1+o(1))i^k(q-i)^{k+s}
\end{aligned}$$

$$= \binom{q}{\lfloor q/2 \rfloor} \left(\lceil q/2 \rceil^s + \lfloor q/2 \rfloor^s + o(1)\right) \left(\lfloor q/2 \rfloor \lceil q/2 \rceil\right)^k$$
$$= \left((qs + 2(1-s))\binom{q}{(q-1)/2} + o(1)\right) \left(\frac{q^2-1}{4}\right)^{\lfloor n/2 \rfloor},$$

where the last equality holds since q is odd and $s \in \{0, 1\}$. The proof is complete. ■

Note that the term $\left((qs + 2(1-s))\binom{q}{(q-1)/2} + o(1)\right) \left(\frac{q^2-1}{4}\right)^{\lfloor n/2 \rfloor}$ appears only in the enumeration of those q-colorings in which one partition of $T_2(n)$ is colored with precisely $\lfloor q/2 \rfloor$ colors and the other partition is colored with precisely $\lceil q/2 \rceil$ colors.

We will now embark on the proof of Theorem 8. Our approach is very similar to the one used by Norine [16] in his proof of Lemma 4.2, with some adjustments, since q is not an even integer.

The approach is as follows:

We apply Lemma 5 to obtain a partition $A_1 \cup A_2$ of the vertices of a $(n, t_2(n))$-graph G that is δn^2-close to $T_2(n)$ such that the number of edges between parts A_1 and A_2 is "nearly" $t_2(n)$.

For a q-coloring f of G and for each $i \in [2]$, we define $\mathcal{R}_f(i)$ to be the subset of colors in $[q]$ which occur relatively frequently in A_i, that is the colors for which a nontrivial number of vertices in A_i are colored with.

The number of q-colorings of G can be bounded above by the number of q-colorings f for which the sets $\mathcal{R}_f(1)$ and $\mathcal{R}_f(2)$ consist of nonempty disjoint subsets of $[q]$. We then consider two cases: the q-colorings where at most one of $|\mathcal{R}_f(1)|$ or $|\mathcal{R}_f(2)|$ is roughly $q/2$ and those for which both $|\mathcal{R}_f(1)|$ and $|\mathcal{R}_f(2)|$ equal roughly $q/2$.

Proof of Theorem 8. Throughout the proof we will be making a series of claims which hold for positive δ, sufficiently small as a function of q, and for n, sufficiently large as a function of q and δ. The eventual choice of δ and n will be implicitly made so that all of our claims are valid.

Let G be a $(n, t_2(n))$-graph, where $n \geq 2$. Suppose, to the contrary, that G has more q-colorings than $T_2(n)$. By Lemma 5, there exists a partition $A_1 \cup A_2 = V(G)$ such that $e(G[A_1, A_2]) \geq t_2(n) - \frac{\delta}{2} n^2$. Assume that the pair (A_1, A_2) is chosen to maximize $e(G[A_1, A_2])$. Let $\delta' := 2n^{-1} \left| |A_1| - \frac{n}{2} \right| = 2n^{-1} \left| |A_2| - \frac{n}{2} \right|$. Then

$$\begin{aligned} e(G[A_1, A_2]) &\leq |A_1| \cdot |A_2| \\ &= \left(\frac{n}{2} + \left||A_1| - \frac{n}{2}\right|\right) \left(\frac{n}{2} - \left||A_2| - \frac{n}{2}\right|\right) \\ &= \left(\frac{n}{2}\right)^2 - \left(\frac{n\delta'}{2}\right)^2 \\ &\leq t_2(n) + \frac{1}{4} - \frac{1}{2}\left(\frac{n\delta'}{2}\right)^2. \end{aligned}$$

Then for sufficiently small δ and sufficiently large n, $(\delta' n/2)^2 \leq 1/2 + \delta n^2 \leq (q/2)\delta n^2$, and thus, $\delta' \leq \sqrt{(q/2)\delta}$. Therefore, it suffices to show that the conclusion of the lemma holds as long as not only δ, but $\max\{\delta, \delta'\}$ is sufficiently small. To simplify the notation at the expense of overloading it, we will use δ in the remainder of the proof to denote $\max\{\delta, \delta'\}$. In particular, we have

$$||A_i|-n/2| \leq \frac{\delta}{2}n \tag{30}$$

for all $i \in [2]$ for sufficiently small δ.

Let $\epsilon := \sqrt{\delta}$. We say that a vertex $v \in V(G)$ is *good* if $d_{A_i}(v) \geq (1-\epsilon)|A_i|$ for the $i \in [2]$ such that $v \notin A_i$; that is, v has "many" neighbors in the part A_i that does not contain v. Otherwise, we say that v is *bad*. Let B denote the set of bad vertices of G. By counting the edges in $\overline{G}[A_1, A_2]$, where $\overline{G}$ denotes the complementary graph of G, we obtain

$$\begin{aligned}
\epsilon(1-\delta)\frac{n}{2}|B| &\leq \sum_{i=1}^{2}|A_i \cap B|\,\epsilon\,|A_{3-i}| \\
&\leq \sum_{i=1}^{2}\sum_{v \in A_i \cap B}\big(|A_{3-i}| - d_{A_{3-i}}(v)\big) \\
&\leq e(\overline{G}[A_1, A_2]) \\
&= t_2(n) - e(G[A_1, A_2]) \\
&\leq t_2(n) - \left(t_2(n) - \frac{\delta}{2}n^2\right) \\
&= \frac{\delta}{2}n^2,
\end{aligned}$$

and hence, $|B| \leq \frac{\epsilon}{1-2\epsilon^2}n \leq 2\epsilon n$ for sufficiently small ϵ (and hence, δ).

Let $f : V(G) \to [q]$ be a q-colorings of G. For each $i \in [2]$ define

$$\mathcal{R}_f(i) := \left\{c \in [q] : \left|f^{-1}(c) \cap A_i\right| > \epsilon|A_i|\right\},$$

i.e., $\mathcal{R}_f(i)$ is the set of colors which occur relatively frequently in A_i under the coloring f. We say that each color in $\mathcal{R}_f(i)$ is an *essential color in* A_i.

We make two observations about $\mathcal{R}_f(i)$. First, note that we can ensure that $\mathcal{R}_f(i) \neq \emptyset$ for each $i \in [2]$ by ensuring that $|A_i| = \sum_{c \in [q]}\left|f^{-1}(c) \cap A_i\right| > q\epsilon|A_i|$ by choosing $\epsilon < 1/q$. Secondly, the sets $\mathcal{R}_f(1)$ and $\mathcal{R}_f(2)$ are disjoint. Note that for every essential color $c \in \mathcal{R}_f(i)$ we have $f^{-1}(c) \subseteq A_i \cup B$. Otherwise, if say, $i = 1$, and there were some $v \in A_2 \setminus B$ such that $f(v) = c$, then since $d_{A_1}(v) \geq (1-\epsilon)|A_1|$, the vertex v must be adjacent to some vertex in A_1 with color c, contradicting that f is a q-coloring. Therefore, we see that there exists at least one essential color in each A_i, and that A_1 and A_2 cannot share essential colors.

Let us define the vector $\boldsymbol{\mathcal{R}}_f := (\mathcal{R}_f(1), \mathcal{R}_f(2))$. Given another vector $\boldsymbol{\mathcal{R}} = (\mathcal{R}_1, \mathcal{R}_2)$ such that the components $\mathcal{R}_1$ and $\mathcal{R}_2$ are disjoint nonempty subsets of $[q]$, let

$$\mathcal{P}_G(\boldsymbol{\mathcal{R}}) := |\{f : V(G) \to [q] : \mathcal{R}_i \text{ is the set of essential colors in } A_i \text{ for each } i \in [2]\}|.$$

We will bound $\mathcal{P}_G(\boldsymbol{\mathcal{R}})$ in two distinct cases, each of which is based upon a comparison of the cardinalities of the sets $\mathcal{R}_1$ and $\mathcal{R}_2$.

Case 1: The components of $\boldsymbol{\mathcal{R}} = (\mathcal{R}_1, \mathcal{R}_2)$ which satisfy

$$(|\mathcal{R}_1|, |\mathcal{R}_2|) \notin \{(\lfloor q/2\rfloor, \lceil q/2\rceil), (\lceil q/2\rceil, \lfloor q/2\rfloor)\}.$$

We can estimate $\mathcal{P}_G(\boldsymbol{\mathcal{R}})$ in the following way:

(i) Allow the vertices of B to be colored arbitrarily. There are $q^{|B|}$ possibilities.

(ii) Allow $|R_i|$ choices of colors for each of the vertices in A_i. There are $|\mathcal{R}_i|^{|A_i|}$ possibilities for each $i \in [2]$.

(iii) Account for the number of subsets of A_i which will *not* be colored with any of the $|\mathcal{R}_i|$ essential colors. Every subset of A_i that is colored with a non-essential color of A_i is of size at most $\epsilon |A_i|$. Thus, there are most $(q - |\mathcal{R}_i|)\epsilon |A_i|$ vertices in A_i colored with a non-essential color. We may choose ϵ (and hence, δ) sufficiently small so that $(q - |\mathcal{R}_i|)\epsilon |A_i| \le |A_i| /4$. Then by Proposition 6, we have at most

$$\sum_{0 \le \ell \le (q-|\mathcal{R}_i|)\epsilon|A_i|} \binom{|A_i|}{\ell} \le 2\binom{|A_i|}{(q - |\mathcal{R}_i|)\epsilon |A_i|}$$

non-essentially colored subset of A_i for each $i \in [2]$.

(iv) Color the subset of A_i chosen in (iii). There are at most $(q - |\mathcal{R}_i|)^{(q-|\mathcal{R}_i|)\epsilon|A_i|}$ possibilities.

By estimating $\mathcal{P}_G(\boldsymbol{\mathcal{R}})$ this way we obtain

$$\begin{aligned}
\mathcal{P}_G(\boldsymbol{\mathcal{R}}) &\le q^{|B|} \left(\prod_{i=1}^{2} |\mathcal{R}_i|^{|A_i|} \cdot 2\binom{|A_i|}{(q - |\mathcal{R}_i|)\epsilon |A_i|} \cdot (q - |\mathcal{R}_i|)^{(q-|\mathcal{R}_i|)\epsilon|A_i|} \right) \\
&\le 4 \cdot q^{|B|} \left(\prod_{i=1}^{2} |\mathcal{R}_i|^{|A_i|} \cdot \left(\frac{e\,|A_i|}{(q - |\mathcal{R}_i|)\epsilon\,|A_i|} \right)^{(q-|\mathcal{R}_i|)\epsilon|A_i|} \cdot (q - |\mathcal{R}_i|)^{(q-|\mathcal{R}_i|)\epsilon|A_i|} \right) \\
&\le 4 \cdot q^{2\epsilon n} \cdot ((\lceil q/2 \rceil + 1)(\lfloor q/2 \rfloor - 1))^{(1+\delta)\frac{n}{2}} (e/\epsilon)^{q\epsilon n} \\
&= 4 \cdot q^{2\epsilon n} \cdot \left(\left(\frac{q+3}{2} \right) \left(\frac{q-3}{2} \right) \right)^{(1+\delta)\frac{n}{2}} (e/\epsilon)^{q\epsilon n} \\
&= 4 \left(\left(\frac{q^2-9}{4} \right) \right)^{n/2} \cdot \exp\left(\left(\frac{\delta}{2} \log\left(\frac{q^2-9}{4} \right) + 2\epsilon \log(q) + q\epsilon \log(e/\epsilon) \right) n \right) \\
&< \frac{1}{3^q} \left(\frac{q^2-1}{4} \right)^{(n-2)/2},
\end{aligned}$$

for ϵ (and hence, δ) sufficiently small and n sufficiently large, since

$$\exp\left(\left(\frac{\delta}{2} \log\left(\frac{q^2-9}{4} \right) + 2\epsilon \log(q) + q\epsilon \log(e/\epsilon) \right) n \right) \to 1$$

as $\epsilon = \sqrt{\delta} \to 0$. It follows that

$$\sum_{\boldsymbol{\mathcal{R}}} \mathcal{P}_G(\boldsymbol{\mathcal{R}}) < \sum_{\boldsymbol{\mathcal{R}}} \frac{1}{3^q} \left(\frac{q^2-1}{4} \right)^{(n-2)/2} = \left(\frac{q^2-1}{4} \right)^{(n-2)/2}, \tag{31}$$

where the summation is taken over all $\boldsymbol{\mathcal{R}} = (\mathcal{R}_1, \mathcal{R}_2)$ such that

$$(|\mathcal{R}_1|, |\mathcal{R}_2|) \notin \{(\lfloor q/2 \rfloor, \lceil q/2 \rceil), (\lceil q/2 \rceil, \lfloor q/2 \rfloor)\},$$

and

$$\begin{aligned}
&\left|\{\boldsymbol{\mathcal{R}} : (|\mathcal{R}_1|, |\mathcal{R}_2|) \notin \{(\lfloor q/2 \rfloor, \lceil q/2 \rceil), (\lceil q/2 \rceil, \lfloor q/2 \rfloor)\}\}\right| \\
&\leq |\{(S_1, S_2) : S_1, S_2 \subseteq [q], S_1 \cap S_2 = \emptyset, S_i \neq \emptyset\}| \\
&= \sum_{j=0}^{q} \binom{q}{j} 1^j \cdot 2^{q-j} \\
&= 3^q.
\end{aligned}$$

Case 2: The components of $\boldsymbol{\mathcal{R}} = (\mathcal{R}_1, \mathcal{R}_2)$ which satisfy

$$(|\mathcal{R}_1|, |\mathcal{R}_2|) \in \{(\lfloor q/2 \rfloor, \lceil q/2 \rceil), (\lceil q/2 \rceil, \lfloor q/2 \rfloor)\}.$$

In this case, we will bound $\mathcal{P}_G(\boldsymbol{\mathcal{R}})$ when $\boldsymbol{\mathcal{R}}$ corresponds to a partition of $[q]$ into two parts, one of which is roughly of size $q/2$. Note that under any such q-coloring f, all of the vertices in $A_i \setminus B$ are only colored with colors from $\mathcal{R}_f(i)$. Otherwise, if there is a q-coloring f such that $\boldsymbol{\mathcal{R}}_f = \boldsymbol{\mathcal{R}}$ and we have, for example, a vertex $v \in A_1 \setminus B$ such that $f(v) = c$ for some $c \in \mathcal{R}_f(2)$, then since $d_{A_2}(v) \geq (1-\epsilon)|A_2|$ and there exist more than ϵk vertices with color c in G_2, then v would be adjacent to some vertex of color c in G_2, a contradiction.

Suppose first that there exists a vertex $v \in V(G)$ such that $d_{A_i}(v) \geq \delta^{2/5}|A_i|$ for every $i \in [2]$. We can estimate $\mathcal{P}_G(\boldsymbol{\mathcal{R}})$ in the following way:

(i) Arbitrarily color the vertices of B with any of the q colors. There are $q^{|B|}$ possibilities.

(ii) If $f(v) \in \mathcal{R}_f(j)$ for some $j \in [2]$ then arbitrarily color the neighbors of v in $A_j \setminus B$ with any of the available $|\mathcal{R}_f(j)| - 1$ colors in $\mathcal{R}_f(j) \setminus \{f(v)\}$. We see that there are $2 \cdot (|\mathcal{R}_f(j)| - 1)^{d_{A_j}(v)}$ possibilities.

(iii) Arbitrarily color the vertices in $A_i \setminus B$ which are *not* neighbors of v using any of the colors in $\mathcal{R}_f(j)$. There are $|\mathcal{R}_f(j)|^{|A_j| - d_{A_j}(v)}$ possibilities.

(iv) Arbitrarily color the vertices in A_{3-j} using any of the colors in $\mathcal{R}_f(3-j)$. There are $|\mathcal{R}_f(3-j)|^{A_{3-j}}$ possibilities.

By estimating $\mathcal{P}_G(\boldsymbol{\mathcal{R}})$ this way we obtain

$$\begin{aligned}
\mathcal{P}_G(\boldsymbol{\mathcal{R}}) &\leq 2 \cdot q^{|B|}(|\mathcal{R}_f(j)| - 1)^{d_{A_j}(v)} |\mathcal{R}_f(j)|^{|A_j| - d_{A_j}(v)} |\mathcal{R}_f(3-j)|^{|A_{3-j}|} \\
&= 2 \cdot q^{|B|} \left(\frac{|\mathcal{R}_f(j)| - 1}{|\mathcal{R}_f(j)|}\right)^{d_{A_j}(v)} |\mathcal{R}_f(j)|^{|A_j|} |\mathcal{R}_f(3-j)|^{|A_{3-j}|} \\
&\leq 2 \cdot q^{2\delta^{1/3}} \left(\frac{\lceil q/2 \rceil - 1}{\lceil q/2 \rceil}\right)^{\delta^{2/5}(1-\delta)n/2} (\lfloor q/2 \rfloor \lceil q/2 \rceil)^{n/2 + \delta n} \\
&= 2 \cdot q^{2\delta^{1/3}} \left(\frac{q-1}{q+1}\right)^{\delta^{2/5}(1-\delta)n/2} (\lfloor q/2 \rfloor \lceil q/2 \rceil)^{n/2 + \delta n} \\
&= 2 \cdot q^{2\delta^{1/3}} \left(\frac{q-1}{q+1}\right)^{\delta^{2/5}(1-\delta)n/2} \left(\frac{q^2 - 1}{4}\right)^{n/2 + \delta n} \\
&< \frac{1}{2^q} \left(\frac{q^2 - 1}{4}\right)^{n/2},
\end{aligned}$$

for sufficiently small δ and sufficiently large n. Recall that from (31) in Case 1, we have

$$\sum_{\substack{\boldsymbol{\mathcal{R}}:(|\mathcal{R}_1|,|\mathcal{R}_2|)\notin \\ \{(\lfloor q/2\rfloor,\lceil q/2\rceil),(\lceil q/2\rceil,\lfloor q/2\rfloor)\}}} \mathcal{P}_G(\boldsymbol{\mathcal{R}}) < \left(\frac{q^2-1}{4}\right)^{(n-2)/2}.$$

Thus,

$$\begin{aligned}
P_G(q) &\leq \sum_{\boldsymbol{\mathcal{R}}} \mathcal{P}_G(\boldsymbol{\mathcal{R}}) \\
&= \sum_{\substack{\boldsymbol{\mathcal{R}}:(|\mathcal{R}_1|,|\mathcal{R}_2|)\notin \\ \{(\lfloor q/2\rfloor,\lceil q/2\rceil),(\lceil q/2\rceil,\lfloor q/2\rfloor)\}}} \mathcal{P}_G(\boldsymbol{\mathcal{R}}) + \sum_{\substack{\boldsymbol{\mathcal{R}}:(|\mathcal{R}_1|,|\mathcal{R}_2|)\in \\ \{(\lfloor q/2\rfloor,\lceil q/2\rceil),(\lceil q/2\rceil,\lfloor q/2\rfloor)\}}} \mathcal{P}_G(\boldsymbol{\mathcal{R}}) \\
&< \left(\frac{q^2-1}{4}\right)^{(n-2)/2} + \sum_{\substack{\boldsymbol{\mathcal{R}}:(|\mathcal{R}_1|,|\mathcal{R}_2|)\in \\ \{(\lfloor q/2\rfloor,\lceil q/2\rceil),(\lceil q/2\rceil,\lfloor q/2\rfloor)\}}} \frac{1}{2^q}\left(\frac{q^2-1}{4}\right)^{n/2} \\
&= \left(\frac{q^2-1}{4}\right)^{(n-2)/2} + \frac{\binom{q}{\lfloor q/2\rfloor}}{2^q}\left(\frac{q^2-1}{4}\right)^{n/2} \\
&\leq 2\left(\frac{q^2-1}{4}\right)^{n/2},
\end{aligned}$$

which is less than the number of q-colorings of $T_2(n)$ (computed in Lemma 8), a contradiction. Therefore, a vertex v as above does not exist. It follows from the choice of the partition (A_1, A_2) that for every $i \in [2]$ the subgraph $G[A_i]$ of G has maximum degree at most $\delta^{2/5}n$. Let $e_i := e(G[A_i \setminus B])$ for each $i \in [2]$. Then

$$\sum_{i=1}^{2}\left(e_i + \delta^{2/5}n\,|B \cap A_i|\right) \geq \sum_{i=1}^{2} e(G[A_i]) = e(G) - e(G[A_1, A_2]) \geq \epsilon\,(1-\delta)\,\frac{n}{2}\,|B|\,.$$

It follows that

$$e_1 + e_2 \geq \left(\delta^{1/3}(1-\delta)\frac{1}{2} - \delta^{2/5}\right)|B|\,n \geq \delta^{2/5}\,|B|\,n \tag{32}$$

for sufficiently small δ. Using Lemma 7 we obtain

$$\begin{aligned}
&\mathcal{P}_G(\boldsymbol{\mathcal{R}}) \\
&\le q^{|B|} \prod_{i=1}^{2} P_{G[A_i]}(\mathcal{R}_i) \\
&\le q^{|B|} \prod_{i=1}^{2} |\mathcal{R}_i|^{|A_i|} \left(\frac{|\mathcal{R}_i|-1}{|\mathcal{R}_i|}\right)^{\sqrt{e_i}} \\
&\le q^{|B|} \prod_{i=1}^{2} |\mathcal{R}_i|^{(1+\delta)\frac{n}{2}} \left(\frac{\lceil q/2\rceil-1}{\lceil q/2\rceil}\right)^{\sqrt{e_i}} \\
&= q^{|B|} \prod_{i=1}^{2} |\mathcal{R}_i|^{(1+\delta)\frac{n}{2}} \left(\frac{q-1}{q+1}\right)^{\sqrt{e_i}} \\
&\le q^{|B|} \left(\frac{q-1}{q+1}\right)^{\sqrt{\delta^{2/5}|B|n}} \left(\lfloor q/2\rfloor\lceil q/2\rceil\right)^{(1+\delta)\frac{n}{2}} \\
&= \left(\frac{q^2-1}{4}\right)^{n/2} \exp\left(\log(q)\,|B| - \log\left(\frac{q+1}{q-1}\right)\sqrt{\delta^{2/5}\,|B|\,n} + \log\left(\frac{q^2-1}{4}\right)\frac{\delta}{2}n\right) \\
&\le \left(\frac{q^2-1}{4}\right)^{n/2} \exp\left(\left(\log(q) - 10^{-1/2}\delta^{-1/20}\log\left(\frac{q+1}{q-1}\right)\right)|B| + \log\left(\frac{q^2-1}{4}\right)\frac{\delta}{2}n\right),
\end{aligned}$$

where the last inequality holds since $|B| \le 2\delta^{1/3}n$. If $|B| \neq 0$ then

$$\mathcal{P}_G(\boldsymbol{\mathcal{R}}) \le \frac{1}{2^q}\left(\frac{q^2-1}{4}\right)^{n/2},$$

since

$$\exp\left(\left(\log(q) - 10^{-1/2}\delta^{-1/20}\log\left(\frac{q+1}{q-1}\right)\right)|B| + \log\left(\frac{q^2-1}{4}\right)\frac{\delta}{2}n\right) \to 0$$

as $\delta \to 0$. Then again, together with the inequality (31), we have

$$\begin{aligned}
P_G(q) &\le \sum_{\boldsymbol{\mathcal{R}}} \mathcal{P}_G(\boldsymbol{\mathcal{R}}) \\
&= \sum_{\substack{\boldsymbol{\mathcal{R}}:(|\mathcal{R}_1|,|\mathcal{R}_2|)\notin \\ \{(\lfloor q/2\rfloor,\lceil q/2\rceil),(\lceil q/2\rceil,\lfloor q/2\rfloor)\}}} \mathcal{P}_G(\boldsymbol{\mathcal{R}}) + \sum_{\substack{\boldsymbol{\mathcal{R}}:(|\mathcal{R}_1|,|\mathcal{R}_2|)\in \\ \{(\lfloor q/2\rfloor,\lceil q/2\rceil),(\lceil q/2\rceil,\lfloor q/2\rfloor)\}}} \mathcal{P}_G(\boldsymbol{\mathcal{R}}) \\
&< \left(\frac{q^2-1}{4}\right)^{(n-2)/2} + \sum_{\substack{\boldsymbol{\mathcal{R}}:(|\mathcal{R}_1|,|\mathcal{R}_2|)\in \\ \{(\lfloor q/2\rfloor,\lceil q/2\rceil),(\lceil q/2\rceil,\lfloor q/2\rfloor)\}}} \frac{1}{2^q}\left(\frac{q^2-1}{4}\right)^{n/2} \\
&= \left(\frac{q^2-1}{4}\right)^{(n-2)/2} + \frac{\binom{q}{\lfloor q/2\rfloor}}{2^q}\left(\frac{q^2-1}{4}\right)^{n/2} \\
&\le 2\left(\frac{q^2-1}{4}\right)^{n/2},
\end{aligned}$$

which is less than the number of q-colorings of $T_2(n)$ (computed in Lemma 8), a contradiction.

We may assume that $|B| = 0$. It suffices to assume that $G[A_1, A_2]$ is a complete bipartite graph. Indeed, if this were not the case, then there exist nonadjacent vertices $v_1 \in A_1$ and $v_2 \in A_2$ and adjacent vertices x and y, either both in A_1 or both in A_2. Let the graph $G' := G - xy + v_1v_2$. If f is a q-coloring in Case 2, then $f(x) \neq f(y)$ and $f(v_1) \neq f(v_2)$. Thus, we can apply the coloring f to the graph G' to obtain a proper q-coloring of G' that satisfies the conditions of Case 2. Therefore, $\mathcal{P}_G(\boldsymbol{\mathcal{R}}) \leq \mathcal{P}_{G'}(\boldsymbol{\mathcal{R}})$. We may repeat this process until all possible edges between A_1 and A_2 are present. Then $G \cong K_{|A_1|,|A_2|}$, and since $t_2(n) = e(G)$ then G must be isomorphic to $T_2(n)$. Therefore, we have shown that any $(n, t_2(n))$-graph G which is δn^2 close to $T_2(n)$ has at most as many q-colorings as $T_2(n)$, with equality holding if and only if G is isomorphic to $T_2(n)$, provided that δ is sufficiently small and n is sufficiently large. ∎

6 Proof of Theorem 1

We are now ready to prove our main result, Theorem 1. The proof combines our solution of **OPT** for odd $q \geq 5$ in Theorem 7, Theorem 8, and uses a similar approach to Norine's proof of his main result, Theorem 1.1, in [16].

Proposition 9 ([14])**.** *If* $\boldsymbol{\alpha} \in \mathrm{FEAS}_q(\gamma)$ *solves* $\mathrm{OPT}_q(\gamma)$*, then* $\mathrm{E}_q(\boldsymbol{\alpha}) = \gamma$.

Proof of Theorem 1. We proceed by contradiction. Assume there exists an increasing sequence of positive integers $\{n_i\}_{i=1}^{\infty}$ and a sequence of graphs $\{H_i\}_{i=1}^{\infty}$ with the following properties.

(i) H_i is a $(n_i, t_2(n_i))$-graph.

(ii) H_i is not isomorphic to $T_2(n_i)$.

(iii) H_i has at least as many q-colorings as any other $(n_i, t_2(n_i))$-graph.

Choose $\epsilon > 0$ so that a real number $\gamma \in [1/4 - \epsilon, 1/4]$ and the conclusion of Theorem 7 holds. We apply Theorem 4 for $\kappa = 1/4$ and a sequence of positive real numbers $\{\delta_i\}_{i=1}^{\infty}$ with $0 < \delta_i \leq \epsilon$ and $\lim_{i\to\infty} \delta_i = 0$. By possibly restricting $\{n_i\}_{i=1}^{\infty}$ to a subsequence, we obtain a sequence $\{\boldsymbol{\alpha}_i\}_{i=1}^{\infty}$ such that H_i is $\delta_i n_i^2$-close to the graph $G_{\boldsymbol{\alpha}_i}(n_i)$, $\boldsymbol{\alpha}_i$ solves $\mathrm{OPT}_q(\gamma_i)$ for some real number γ_i such that $\gamma_i \in [1/4 - \epsilon, 1/4]$, and $\lim_{i\to\infty} \gamma_i = 1/4$.

Since $\boldsymbol{\alpha}_i \in \mathrm{FEAS}_q(\gamma_i)$, and $\mathrm{FEAS}_q(\gamma_i)$ is a compact set, we may further restrict our sequence $\{\boldsymbol{\alpha}_i\}_{i=1}^{\infty}$ (and hence, the sequences $\{n_i\}$ and $\{\gamma_i\}$) by assuming that the $\boldsymbol{\alpha}_i$'s converge in the L^1-norm to a vector $\boldsymbol{\alpha}^*$ with $\mathrm{E}_q(\boldsymbol{\alpha}^*) = \gamma^*$. Then by Theorem 7, for odd $q \geq 5$, we have

$$\begin{aligned}
&\mathrm{OBJ}_q(\boldsymbol{\alpha}^*) - \frac{1}{2}\log\left(\frac{q^2-1}{4}\right) \\
&= \lim_{i\to\infty}\left[\mathrm{OBJ}_q(\boldsymbol{\alpha}_i) - \left(\frac{1}{2}\log\left(\frac{q^2-1}{4}\right) + \frac{\sqrt{1-4\gamma_i}}{2}\log\left(\frac{q+1}{q-1}\right)\right)\right] \\
&= \lim_{i\to\infty}\left[\mathrm{OPT}_q(\gamma_i) - \left(\frac{1}{2}\log\left(\frac{q^2-1}{4}\right) + \frac{\sqrt{1-4\gamma_i}}{2}\log\left(\frac{q+1}{q-1}\right)\right)\right] \\
&= 0.
\end{aligned}$$

Therefore,

$$\mathrm{OBJ}_q(\boldsymbol{\alpha}^*) = \frac{1}{2}\log\left(\frac{q^2-1}{4}\right).$$

Since $\boldsymbol{\alpha}_i$ solves $\mathrm{OPT}_q(\gamma_i)$, by Proposition 9, $\mathrm{E}_q(\boldsymbol{\alpha}_i) = \gamma_i$. Since $\{\boldsymbol{\alpha}_i\}_{i=1}^{\infty}$ converges to $\boldsymbol{\alpha}^*$ in the L^1-norm, we have

$$1/4 = \lim_{i\to\infty} \gamma_i = \lim_{i\to\infty} \mathrm{E}_q(\boldsymbol{\alpha}_i) = \mathrm{E}_q(\boldsymbol{\alpha}^*) = \gamma^*.$$

Then $\gamma^* = 1/4$, and hence, by Theorem 7, $\boldsymbol{\alpha}^*$ solves $\mathrm{OPT}_q(1/4)$. Then Corollary 1 tells us that the graph $G_{\boldsymbol{\alpha}^*}(n) = T_2(n)$ for every n.

Let $\delta > 0$ be chosen so that the conclusion of Theorem 8 holds. By Proposition 3, $G_{\boldsymbol{\alpha}_i}(n_i)$ is $\delta n_i^2/2$-close to $G_{\boldsymbol{\alpha}^*}(n_i) = T_2(n_i)$ for sufficiently large i, since $\boldsymbol{\alpha}_i \to \boldsymbol{\alpha}^*$. We may assume that $\delta_i \le \delta/2$ for sufficiently large i since $\delta_i \to 0$ as $i \to \infty$. Consequently, as each H_i is $\delta_i n_i^2$-close to $G_{\boldsymbol{\alpha}_i}(n_i)$, then the edit distance between H_i and $G_{\boldsymbol{\alpha}^*}(n_i) = T_2(n_i)$ is at most

$$\frac{\delta}{2} n_i^2 + \delta_i n_i^2 \le \frac{\delta}{2} n_i^2 + \frac{\delta}{2} n_i^2 = \delta n_i^2.$$

That is, H_i is δn_i^2-close to $T_2(n_i)$ for sufficiently large i. However, this contradicts Theorem 8, finishing the proof of the theorem. ∎

7 Concluding remarks and open problems

In Section 1, we presented Conjecture 1 and stated that although several cases of the conjecture had been solved for various ranges of r, q, and n, it was not true in general, as counterexamples were discovered in [15]. Nonetheless, several cases of the conjecture remain open, one of them being the following conjecture.

Conjecture 2. *Let r and q be integers such that $2 \le r \le 9$ and $r \le q$. Then for all $n \ge r$, the Turán graph $T_r(n)$ has more q-colorings than any other graph with the same number of vertices and edges.*

It seems that it may be difficult to resolve Conjecture 2 for all $n \ge r$. However, asymptotic versions (for n sufficiently large), may be more attainable by finding solutions of **OPT** for $2 \le r \le 9$, $q \ge r$, and positive γ which satisfy $\gamma \le (q-1)/(2q)$.

8 Acknowledgements

I would like to thank Felix Lazebnik, my former advisor, for all of his help and encouragement on this work while I was a doctoral student at the University of Delaware.

References

[1] L. Babai and P. Frankl, *Linear Algebra Methods in Combinatorics, Part 1*, University of Chicago, Department of Computer Science, 1988.

[2] E. Bender and H. Wilf, "A theoretical analysis of backtracking in the graph coloring problem", *J. Algorithms*, **6** (1985), 275-282.

[3] G. Birkhoff, "A determinant formula for the number of ways of coloring a map", *Adv. Math.*, **14**, 1912, 42-46.

[4] B. Bollobás, *Modern Graph Theory*, Springer-Verlag, Berlin, 1998.

[5] O. D. Byer, "Some new bounds for the maximum number of vertex colorings of a (v, e)-graph", *J. Graph Theory*, **28**, 1998, 115-128.

[6] F. Lazebnik, S.N. Tofts, "An extremal property of Turán's Graphs", *Electron. J. Combin.*, **17 (1)**, 2010, 1-11.

[7] F. Lazebnik, O. Pikhurko, and A. Woldar, "Maximum number of colorings of $(2k, k^2)$-graphs", *J. Graph Theory*, **56**, 2007, 135-148.

[8] F. Lazebnik, "New upper bounds for the greatest number of proper colorings of a (V, E)-graph", *J. Graph Theory*, **14**, 1990, 25-29.

[9] F. Lazebnik, "On the greatest number of 2 and 3 colorings of a (v, e)-graph", *J. Graph Theory*, **13**, 1989, 203-214.

[10] F. Lazebnik, "Some corollaries of a theorem of Whitney on the chromatic polynomial", *Discrete Math.*, **87**, 1991, 53-64.

[11] F. Lazebnik, "The maximum number of colorings of graphs of given order and size: A survey", *Discrete Math.*, **342 (10)**, 2019, 2783-2791.

[12] N. Linial, "Legal colorings of graphs", *Combinatorica*, **6**, 1986, 49-54.

[13] R. Y. Liu, "The maximum number of proper 3-colorings of a graph", *Math. Appl. (Wuhan)*, **6**, 1993, 88-91.

[14] P.-S. Loh, O. Pikhurko, and B. Sudakov, "Maximizing the number of q-colorings", *Proc. London Math. Soc.*, **101**, 2010, 655-696.

[15] J. Ma and H. Naves, "Maximizing proper colorings on graphs", *J. Comb. Theory Ser. B*, **115**, 2015, 236–275.

[16] S. Norine, "Turán graphs and the number of colorings", *SIAM J Discrete*, **25**, 2011, 260-266.

[17] S. N. Tofts, "An extremal property of Turán's Graphs, II", *J. Graph Theory*, **75 (3)**, 2014, 275-283.

[18] P. Turán, "On an extremal problem in graph theory", *Mat. Fiz. Lapok*, **48**, 1941, 436-452.

[19] H. Whitney, "A logical expansion in mathematics", *Bull. Amer. Math. Soc.*, **38**, 1932, 572-579.

[20] H.S. Wilf, "Backtrack: an $O(1)$ expected time algorithm for the graph coloring problem", *Inform. Process. Lett.*, **18**, 1984, 119-121.

[21] H.S. Wilf, *Generatingfunctionology*, Elsevier Science, United States, 2013, 19.

Wilf, Herbert S.. Generating Functionology. United States, Elsevier Science, 2013.

[22] H.S. Wilf. Personal communication with Felix Lazebnik, 1982.