Decay of correlation for edge colorings when $q > 3\Delta$

Zejia Chen[1], Yulin Wang[1], Chihao Zhang[1], and Zihan Zhang[2]

[1] *Shanghai Jiao Tong University*
[2] *Graduate Institute for Advanced Studies, SOKENDAI*

September 14, 2025

Abstract

We examine various perspectives on the decay of correlation for the uniform distribution over proper q-edge colorings of graphs with maximum degree Δ.

First, we establish the coupling independence property when $q \geq 3\Delta$ for general graphs. Together with the recent work of Chen, Feng, Guo, Zhang and Zou (2024), this result implies a fully polynomial-time approximation scheme (**FPTAS**) for counting the number of proper q-edge colorings.

Next, we prove the strong spatial mixing property on *trees*, provided that $q > (3 + o(1))\Delta$. The strong spatial mixing property is derived from the spectral independence property of a version of the weighted edge coloring distribution, which is established using the matrix trickle-down method developed in Abdolazimi, Liu and Oveis Gharan (FOCS, 2021) and Wang, Zhang and Zhang (STOC, 2024).

Finally, we show that the weak spatial mixing property holds on trees with maximum degree Δ if and only if $q \geq 2\Delta - 1$.

Contents

1 Introduction

Sampling from the uniform distribution of proper edge colorings received lots of attention recently, with the advent of new tools in analyzing high-dimensional distributions [DHP20, ALG21a, WZZ24, CCFV25]. A proper edge coloring of an undirected graph with maximum degree Δ is an assignment of each edge with one of q colors so each adjacent edges receive different color. Clearly a proper *edge* coloring can be viewed as a proper *vertex* coloring on its line graph.The sampling problem is to draw a proper edge coloring uniformly. Most of previous work focuses on the mixing time of Glauber dynamics. The work of [WZZ24] established the spectral independence property, a notion to measure the correlation in high-dimensional distributions [ALG21b], whenever $q > (2 + o(1))\Delta$ for general graphs and the work of [CCFV25] establish approximate tensorization of variance on trees when $q \geq \Delta + 1$. Note that Glauber dynamics is known to be reducible when $q < 2\Delta$ on general graphs [HJNP19] and therefore the $q > (2 + o(1))\Delta$ condition for Glauber dynamics is asymptotically tight. However, it is still open whether $q \approx 2\Delta$ is the threshold for efficient sampling proper edge colorings and there are some recent attempts to design other sampling algorithm under better conditions [DMKLP25]. A (almost) uniformly sampling algorithm can be turned into fully polynomial-time randomized scheme (**FPRAS**) for counting the number of proper colorings using standard reduction [JVV86].

In this work we examine some other aspects for the correlation of the uniform distribution on edge colorings. We first established the *coupling independence* property on general graphs when $q \geq 3\Delta$. Coupling independence [CZ23] is a notion stronger than spectral independence, and thanks to recent work of [CFG+24], building on the machinery of Moitra [Moi19], it (together with some other properties) implies a fully polynomial-time approximate scheme (**FPTAS**) for counting proper edge colorings. This is the first *determinstic* approximate counting algorithm in this range of parameters. Moreover, our proof of the coupling independence property differs from those directly derived from contractive couplings for establishing rapid mixing results.

Theorem 1 (Informal). *If $q \geq 3\Delta$, then there exists an **FPTAS** for counting the number of proper q-edge colorings on any graph G with maximum degree Δ.*

Before our work, there is no similar results tailored for counting edge colorings. The best **FPTAS** for counting edge coloring is the same as the one for general vertex coloring, which requires $q > 3.618\Delta$ [CV25, CFG+24] for sufficiently large Δ.

We then study the strong spatial mixing (SSM) property for edge colorings on *trees*, an important notion to measure the correlation between sites in Gibbs distributions whose definition is in Section 2.3.

Theorem 2 (Informal). *Let T be a tree with maximum degree Δ. If $q \geq (3+o(1))\Delta$, then the uniform distribution over q-colorings on T exhibits strong spatial mixing with exponential decay rate.*

Similar strong spatial mixing bounds on trees have been thoroughly studied for vertex colorings (e.g. [EGH+19, CLMM23]). It is conjectured that SSM holds whenever $q \geq \Delta + 1$ and in [CLMM23], a $q \geq \Delta + 3$ condition was established, almost resolving the conjecture. However, much less is known for edge colorings. The $q > (3 + o(1))\Delta$ bound established in this work is by no means tight. We also discuss the limit of our approach and possible further improvement. On the other hand, we show that one cannot expect the strong spatial mixing property to hold when $q < 2\Delta$, as we prove that $2\Delta - 1$ is the threshold for *weak spatial mixing*. Therefore, we conjecture that SSM holds for edge colorings on graphs with maximum degree Δ whenever $q \geq 2\Delta + \gamma$ for certain constant γ. We also show that the best bound one can expect using the analysis in this paper cannot be better than $q \approx 2.618\Delta$.

Theorem 3 (Informal)**.** *If $q \geq 2\Delta - 1$, then the uniform distribution over q-colorings on any tree with maximum degree Δ exhibits weak spatial mixing with exponential decay rate. Otherwise, there exists a tree with maximum degree Δ such that the uniform distribution over q-coloring on it does not satisfy the weak spatial mixing property.*

In the following, we give an overview with our technique, with an emphasis on the novelty.

1.1 Technical contribution

A new coupling strategy The coupling independence property is established via a new local coupling for edge colorings. Our coupling can somehow be viewed as a multi-spin version of Chen and Gu's coupling [CG24] for Holant problems with boolean domain. Their coupling, using the problem of b-matching as an example, begins with two instances differing at one pendant edge, or equivalently, two instances with a single constraint discrepancy. During the coupling process, the number of discrepancies can never increase but has a nonzero probability of decreasing to zero. Therefore, the coupling process terminates in expected constant number of steps.

In the problem of edge coloring, we can design a local coupling starting from a single discrepancy so that the number of discrepancies can either increase to two, decrease to zero, or remain unchanged. We then use marginal probability bounds to control the probability of discrepancy increasing, while ensuring that the number of discrepancies decreases in expectation.

Dimension reduction We establish the strong spatial mixing property on trees by analyzing marginal recursions, which is similar to [CLMM23]. However, unlike previous work for spin systems where the marginal on a single site is considered, we study the recursion for marginals on a "broom", namely all edges incident to the root. For each partial coloring on the broom, we can represent its marginal as a function of marginal probabilities of partial colorings in subtrees. However, the Jacobian matrix of this system can be as large as $q^{\Delta} \times \Delta q^{\Delta}$, and is technically very hard to analyze. Our key observation is that the Jacobian matrix is of low rank, and therefore we can apply a trace trick to bound its 2-norm by the norm of a much smaller matrix. We call this step *dimension reduction*.

From spectral independence to strong spatial mixing It is still challenging to directly bound the 2-norm of the small matrix. We then observe that it can be written as the product of certain covariance matrices of marginal distributions on brooms. Therefore, ideally we can apply the known bounds for these covariance matrices, or equivalently the spectral independence bound for these marginals. However, these marginal probabilities are from the distribution of certain "weighted edge colorings" and one cannot directly apply previous spectral independence result for edge colorings. As a result, we apply the machinery of matrix trickle-down developed in [ALG21a] in the way of [WZZ24] to establish the desired spectral independence result.

Top-down analysis of recursion There is another subtle technical issue in the above approach. When analyzing the contraction of marginal recursion, one needs to analyze the gradient / Jacobian at certain "midpoint" between two boundary conditions due to the application of fundamental theorem of calculus or mean-value theorem in the analysis. These midpoints, however, are not necessarily probabilities because the recursion may involve a potential function[1]. In previous work, only certain marginal bounds are used to prove the contraction, and these bounds are also satisfied by the midpoints. However, in our case, we require these midpoints to satisfy refined properties, such as spectral independence, which does not hold in general. Therefore, we cannot apply the recursion in the traditional bottom-to-top manner, where one fixes the boundary value at level L, analyze the contraction at level $L-1$, then fixes boundary value at level $L-1$ and analyze the contraction at level $L-2$, and so on. Instead, we only fix boundary value at the leaves and analyze the composition of the recursion at each level as a whole. Therefore, we need to take "midpoints" only at the leaves, which defines our "weighted edge coloring" instance. Its spectral independence property can be established by the matrix trickle-down method, as discussed earlier.

1.2 Organization of the paper

After introducing the necessary preliminaries in Section 2, we give the marginal recursions and prove useful marginal bounds in Section 3. Then we present our proof of coupling independence, which implies the **FPTAS**, in Section 4. Strong spatial mixing on trees is proved in Section 5 and the results of weak spatial mixing are proved in Section 7. In Section 6, we prove the spectral independence property for weighted edge coloring that will be used in the proof of strong spatial mixing.

2 Preliminaries

We use the following notations. For any $a, b \in \mathbb{R}$, let $a \wedge b := \min\{a, b\}$, $a \vee b := \max\{a, b\}$. For any two non-negative integers $a \geq b$, let $a^{\underline{b}}$ be the falling factorial, i.e. $a^{\underline{b}} = \prod_{i=a-b+1}^{a} i$. Let Id denote the identity matrix. For a function $f: \Omega \to \mathbb{R}$ defined on a finite domain Ω, we use $\left[f(x)\right]_{x \in \Omega}$ to denote the corresponding (column) vector in $\mathbb{R}^{\Omega}$. For any set S and an element $x \in S$, we write $S - x$ for $S \setminus \{x\}$.

For two probability measures μ, ν on the same probability space Ω, we define $\|\mu - \nu\|_{\mathrm{TV}} = \frac{1}{2}\sum_{\omega \in \Omega} |\mu(\omega) - \nu(\omega)| = \sup_{A \subseteq \Omega} |\mu(A) - \nu(A)|$ for the total variation distance between μ, ν. If μ, ν are two probability distributions on finite state spaces Ω_1, Ω_2 respectively, then we say ω is a coupling of μ, ν when it is a joint distribution on $\Omega_1 \times \Omega_2$ with μ, ν as its marginals, i.e. $\mu(x) = \sum_{y \in \Omega_2} \omega(x, y)$ and $\nu(y) = \sum_{x \in \Omega_1} \omega(x, y)$ for every $x \in \Omega_1, y \in \Omega_2$.

Given a graph $G = (V, E)$, for any vertex $v \in V$, let $E(v) = \{e \in E \mid v \in e\}$ and $\deg(v) = |E(v)|$ be the degree of v; for any edge $e \in E$, let $\deg(e)$ be the degree of e, which is the number of edges adjacent to e. Moreover, we write $\deg(G)$ for the maximum (vertex) degree in G. For two edges $e, e' \in E$, we write $\mathrm{dist}_G(e, e')$ for the length of the shortest path between them in G (not containing e, e') and $\mathrm{dist}_G(e, e') = \infty$ if e and e' is disconnected. Similarly, for vertices $v, v' \in V$, $\mathrm{dist}_G(v, v')$ is the shortest path between them in G and $\mathrm{dist}_G(v, v') = \infty$ if v and v' is disconnected.

[1] Theses quantities are referred to as "subdistributions" in [CLMM23]

2.1 List edge coloring

Fix a color set $[q] = \{1, 2, \ldots, q\}$ where $q \in \mathbb{N}$. Let $G = (V, E)$ be an undirected graph and $\mathcal{L} = \{\mathcal{L}(e) \subseteq [q] : e \in E\}$ be a collection of color lists associated with each edge in E. The pair $(G, \mathcal{L})$ is an instance of list edge coloring.

If $\mathcal{L}(e) = [q]$ for any $e \in E$, we say $(G, \mathcal{L})$ is a q-edge coloring instance. If $|\mathcal{L}(e)| \geq \deg(e) + \beta$ for any $e \in E$, we say $(G, \mathcal{L})$ is a β-extra edge coloring instance. We say $\sigma : E \to [q]$ is a proper edge coloring if $\sigma(e) \in \mathcal{L}(e)$ for any $e \in E$ and $\sigma(e_1) \neq \sigma(e_2)$ for any $e_1 \cap e_2 \neq \varnothing$. Let Ω denote the set of all proper edge colorings and μ be the uniform distribution on Ω.

Let $\Lambda \subseteq E$ and $\tau \in [q]^{\Lambda}$. We say τ is a proper partial edge coloring on Λ if it is a proper coloring on $(G[\Lambda], \mathcal{L}|_{\Lambda})$ where $G[\Lambda]$ is the subgraph of G induced by Λ and $\mathcal{L}|_{\Lambda} = \{\mathcal{L}(e) \in \mathcal{L} : E \in \Lambda\}$. Let Ω^{τ} be the set of all proper edge colorings on E that is compatible with τ, i.e. $\Omega^{\tau} = \{\sigma \in \Omega \mid \tau \subset \sigma\}$. We also define μ^{τ} on Ω which is supported on Ω^{τ} as $\mu^{\tau}(\cdot) = \mathbb{P}_{\sigma \sim \mu}[\sigma = \cdot \mid \tau \subset \sigma]$. For a subset $S \subseteq E \setminus \Lambda$ and a partial coloring ω on S, define Ω_S^{τ} as the set of all proper partial edge colorings on S that is compatible with τ and $\mu_S^{\tau}(\omega) = \mathbb{P}_{\sigma \sim \mu}[\omega \subset \sigma \mid \tau \subset \sigma]$. Especially, when $\Lambda = \{i\}$ and $\tau(i) = c$, we write the conditional distribution and the conditional marginal distribution by $\mu^{i \leftarrow c}$ and $\mu_S^{i \leftarrow c}$. Besides, we define the color lists after pinning τ by $\mathcal{L}^{\tau}$ such that for any $e \in E \setminus \Lambda$, $\mathcal{L}^{\tau}(e) = \{c \in \mathcal{L}(e) \mid \mu_e^{\tau}(c) > 0\}$ and the degree after pinning by $\deg^{\tau}(e) = |\{e \cup f \neq \varnothing \mid f \in E \setminus \Lambda\}|$.

For a given list edge coloring instance $(G, \mathcal{L})$, let $Z_{G,\mathcal{L}}(M)$ denote the number of proper colorings with the condition M satisfied, (or event M happens) and $\mathbb{P}_{G,\mathcal{L}}[M]$ denote the probability that the condition M is satisfied when a proper coloring is drawn uniformly at random. For an edge set $F \subseteq E$, we usually use $c(F)$ to denote the partial coloring on F. With a little abuse of notation, $c(F)$ is sometimes referred to as the set of colors used on F. For a color a, we write $a \in F, a \notin F$ as shorthands for $a \in c(F), a \notin c(F)$ respectively.

2.2 The Wasserstein distance

In this work, we restrict our discussions and terminologies to finite probability spaces without invoking general measure theory.

Definition 4 (Wasserstein distance). *Let μ, ν be two distributions defined on the same finite set Ω equipped with a metric $d(\cdot, \cdot)$. We define $\Gamma(\mu, \nu)$ as the set of couplings of μ and ν. Then the* Wasserstein (1-)distance *is defined by*

$$\mathcal{W}_1(\mu, \nu) := \inf_{\tau \in \Gamma(\mu, \nu)} \mathbf{E}_{(x,y) \sim \tau}[d(x, y)].$$

In this paper, our metric d is always the Hamming distance. For two configurations σ, τ on $[q]^V$, their Hamming distance is defined as $d(\sigma, \tau) = |\{v \in V \mid \sigma(v) \neq \tau(v)\}|$. We define the notion of coupling independence for Gibbs distribution here.

Definition 5 (Coupling independence). *We say a Gibbs distribution μ over $[q]^V$ satisfies C-coupling independence if for any two partial configurations $\sigma, \tau \in [q]^{\Lambda}$ on $\Lambda \subseteq V$ such that $d(\sigma, \tau) = 1$,*

$$\mathcal{W}_1(\mu^{\sigma}, \mu^{\tau}) \leq C$$

where μ^{σ} and μ^{τ} denote the Gibbs distribution conditional on σ and τ, respectively.

We will use the following inequality w.r.t Wasserstein distance in later proof.

Proposition 6. *Let μ, ν be arbitrary distributions on a common finite metric space (Ω, d). If there exists non-negative constants $\lambda_i, 1 \le i \le k$ and distributions $\{\mu_i\}_{1\le i\le k}$, $\{\nu_i\}_{1\le i\le k}$ on Ω such that*

$$\mu - \nu = \sum_{i=1}^{k} \lambda_i \cdot (\mu_i - \nu_i),$$

where we regard both sides as functions on Ω. Then

$$\mathcal{W}_1(\mu, \nu) \le \sum_{i=1}^{k} \lambda_i \mathcal{W}_1(\mu_i, \nu_i).$$

Proof. The Kantotrovich-Rubinstein duality theorem (see Theorem 1.14 in [Vil21] for the proof) states a equivalent form of Wasserstein distance:

$$\mathcal{W}_1(\mu, \nu) = \sup_{f \in L^1(\Omega)} \langle f, \mu - \nu \rangle,$$

where $\langle f, \mu - \nu \rangle := \sum_{x\in\Omega} f(x)(\mu(x) - \nu(x))$ and $L^1(\Omega) := \{f : \Omega \to \mathbb{R} \mid \forall x, y \in \Omega : f(x) - f(y) \le d(x, y)\}$ is the space of 1-Lipschiz functions. Then

$$\mathcal{W}_1(\mu, \nu) = \sup_{f\in L^1(\Omega)} \langle f, \mu - \nu \rangle = \sup_{f\in L^1(\Omega)} \langle f, \sum_{i=1}^{k} \lambda_i(\mu_i - \nu_i) \rangle \le \sum_{i=1}^{k} \lambda_i \sup_{f\in L^1(\Omega)} \langle f, \mu_i - \nu_i \rangle = \sum_{i=1}^{k} \lambda_i \mathcal{W}_1(\mu_i, \nu_i).$$

□

2.3 Correlation Decay

Correlation decay refers to the phenomenon that the correlation between the color assignments of edges diminishes as their distance in the graph increases. Specifically, there are two primary notions of correlation decay: strong spatial mixing and weak spatial mixing. These two notions differ in how they measure the "distance" over which the correlation should decay.

Definition 7 (Strong spatial mixing)**.** *The Gibbs distribution μ of the list edge coloring instance $(G = (V, E), \mathcal{L})$ satisfies strong spatial mixing (SSM) with exponential decay rate $1 - \delta$ and constant $C = C(q, \Delta)$ if for any $e \in E$, every subset $\Lambda \subseteq E \setminus \{e\}$ and every pair of feasible pinning τ_1, τ_2 on Λ which differ on $\partial_{\tau_1,\tau_2} = \{e \in \Lambda \mid \tau_1(e) \ne \tau_2(e)\}$, we have that*

$$\|\mu_e^{\sigma} - \mu_e^{\tau}\|_{\mathrm{TV}} \le C(1 - \delta)^K$$

where $K = \min_{e'\in\partial_{\tau_1,\tau_2}} \mathrm{dist}_G(e, e')$.

Definition 8 (Weak spatial mixing)**.** *The Gibbs distribution μ of the list edge coloring instance $(G = (V, E), \mathcal{L})$ satisfies weak spatial mixing (WSM) with exponential decay rate $1 - \delta$ and constant $C = C(q, \Delta)$ if for any $e \in E$, every subset $\Lambda \subseteq E \setminus \{e\}$ and every pair of feasible pinning σ, τ on Λ, we have that*

$$\|\mu_e^{\sigma} - \mu_e^{\tau}\|_{\mathrm{TV}} \le C(1 - \delta)^K$$

where $K = \min_{e'\in\Lambda} \mathrm{dist}_G(e, e')$.

Remark. *When defining Strong Spatial Mixing (SSM) and Weak Spatial Mixing (WSM) for individual problem instances, it is technically possible to choose a constant C large enough to satisfy the decay inequality since the instance is finite. For SSM/WSM to meaningfully imply algorithmic tractability, these constants must hold universally for an entire class of instances, independent of the size of graph. Therefore, when we informally state that* we are establishing weak / strong spatial mixing property *in this context, we refer to the existence of uniform constants C and δ for a certain class of edge coloring instances even as the graph grows arbitrarily large.*

3 Recursion and marginals

In this section, we analyze the marginal bounds on the set of edges adjacent to a given vertex in a β-extra list-edge-coloring instance under various conditions.

3.1 Marginal bounds for general edge coloring

We begin by examining the marginal bounds on general graphs.

Lemma 9. *Given a list-edge-coloring instance $(G,\mathcal{L})$ where $G=(V,E)$, and a vertex $v\in V$ such that $\forall e\in E_v : |\mathcal{L}(e)|-\deg(e)\geq\beta\geq 2$. Then for any $v\in V$, $F\subseteq E(v)$ and color a:*

$$\mathbb{P}_{G,\mathcal{L}}\left[a\in c(F)\right]\leq\frac{|F|}{\beta-1+|F|}\quad\textit{and}\quad\frac{\mathbb{P}_{G,\mathcal{L}}\left[a\in c(F)\right]}{\mathbb{P}_{G,\mathcal{L}}\left[a\notin c(E(v))\right]}\leq\frac{|F|}{\beta-1}.$$

Proof. We denote the set of all proper colorings on $(G,\mathcal{L})$ by Ω, and define

$$A:=\{\omega\in\Omega\mid\exists e\in F:\omega(e)=a\}.$$

Then define a function $\iota:(\Omega\setminus A)\times A\to\mathbb{R}_{\geq 0}$ such that for any $\omega'\in A$:

$$\sum_{\omega\in\Omega\setminus A}\iota(\omega,\omega')=1. \tag{1}$$

To continue with the construction of ι, for any $\omega'\in A$:

$$B(\omega'):=\left\{\omega\in\Omega\setminus A:d(\omega,\omega')=1\right\}.$$

where d is the Hamming distance, i.e., the number of edges colored differently between the two colorings. Then

$$\iota(\omega,\omega'):=\begin{cases}\frac{1}{|B(\omega')|}, & \omega\in B(\omega')\\ 0, & \text{otherwise}\end{cases}.$$

It is easy to verify eq. (1) holds. Notice that if $\omega\in B(\omega')$ such that $\omega'(e)=a$ for some $e\in F$, then ω is obtained from ω' by recoloring e to another color other than a. There are at least $\beta-1$ choices according to the assumption, so $|B(\omega')|\geq\beta-1$ for any $\omega'\in A$.

Similarly, if $\omega'\in B(\omega)$, then ω' is obtained from ω by recoloring an edge $e\in F$ to a, there are at most $|F|$ choices, so we have $\sum_{\omega'}\iota(\omega,\omega')\leq\frac{|F|}{\beta-1}$ for all $\omega\in\Omega\setminus A$.

Then

$$\begin{aligned}\mathbb{P}_{G,\mathcal{L}}\left[a\in c(F)\right]&=\frac{\sum_{\omega'\in A}1}{\sum_{\omega\in\Omega\setminus A}1+\sum_{\omega'\in A}1}\\&=\frac{\sum_{\omega'\in A}\sum_{\omega\in\Omega\setminus A}\iota(\omega,\omega')}{\sum_{\omega\in\Omega\setminus A}1+\sum_{\omega'\in A}\sum_{\omega\in\Omega\setminus A}\iota(\omega,\omega')}\\&=\frac{\sum_{\omega\in\Omega\setminus A}\sum_{\omega'\in A}\iota(\omega,\omega')}{\sum_{\omega\in\Omega\setminus A}\left(1+\sum_{\omega'\in A}\iota(\omega,\omega')\right)}\\&\leq\frac{|F|/(\beta-1)}{1+|F|/(\beta-1)}=\frac{|F|}{\beta-1+|F|},\end{aligned}$$

proving the first part of the lemma.

Similarly, we define $\Omega_0:=\{\omega\in\Omega\mid\nexists\, e\in E(v):\omega(e)=a\}$. Note that by definition, $\iota(\cdot,\cdot)$ is supported on $\Omega_0\times A$ and for any ω', $B(\omega')\subseteq\Omega_0$, so we have

$$\frac{\mathbb{P}_{G,\mathcal{L}}\left[a\in c(F)\right]}{\mathbb{P}_{G,\mathcal{L}}\left[a\notin c(E(v))\right]}=\frac{\sum_{\omega'\in A}1}{\sum_{\omega\in\Omega_0}1}=\frac{\sum_{\omega'\in A}\sum_{\omega\in\Omega\setminus A}\iota(\omega,\omega')}{\sum_{\omega\in\Omega_0}1}.$$

This can be rewritten as

$$\frac{\sum_{\omega\in\Omega_0}\sum_{\omega'\in A}\iota(\omega,\omega')}{\sum_{\omega\in\Omega_0}1} \le \frac{|F|}{\beta-1}.$$

□

Especially, when $|F| = 1$, we have the following corollary.

Corollary 10. *Let $(G = (V, E), \mathcal{L})$ be a β-extra list-edge-coloring instance. Then for any $e \in E, v \in V, a \in \mathcal{L}(e)$,*

$$\frac{\mathbb{P}_{G,\mathcal{L}}[c(e) = a]}{\mathbb{P}_{G,\mathcal{L}}[a \notin c(E_v)]} \le \frac{1}{\beta-1}.$$

3.2 Tree recursion for edge coloring

We now turn our attention to a more specialized structure: trees. Throughout this section, we fix a β-extra edge-coloring instance $(T = (V, E), \mathcal{L})$ with root r. For any vertex $v \in V$, let T_v be the sub-tree of T rooted at v and E_{T_v} be the edge set of T_v. Let C_v be the set of proper partial colorings on $E_{T_v}(v)$ and $\mathcal{D}(C_v)$ be the set of all distributions on C_v.

Assume r has d children $v_1, v_2, \dots, v_d$. For any $i \in [d]$, let $e_i = (r, v_i)$ and $T_i = T_{v_i}$. For brevity, since the color lists are fixed, we omit the color lists $\mathcal{L}$ in the subscript in $\mathbb{P}_{T,\mathcal{L}}[\cdot]$ and $Z_{T,\mathcal{L}}(\cdot)$. For any $i \in [d]$, we also write $\mathbb{P}_{T_i}[\cdot]$ for $\mathbb{P}_{T_i,\mathcal{L}|_{T_i}}[\cdot]$ and $Z_{T_i}(\cdot)$ for $Z_{T_i,\mathcal{L}|_{T_i}}(\cdot)$.

We introduce a tree recursion on the marginal distributions of partial colorings on "brooms", where a "broom" is referred to as the edge set $E_{T_v}(v)$ for a vertex $v \in V$. This recursion demonstrates how the marginal distributions on brooms propagate through the tree structure as in Figure 1.

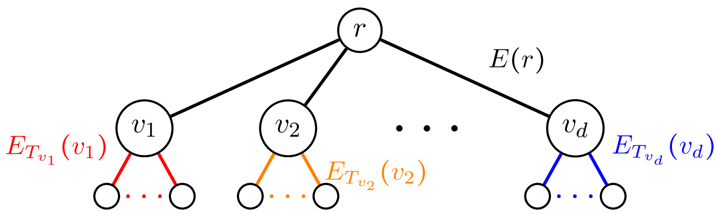

Figure 1: Brooms on a tree

Lemma 11. *Given distributions $\mathbf{p}_i = (\mathbb{P}_{T_i}[c(E_{T_i}(v_i)) = \tau])_{\tau\in C_{v_i}}$ on C_{v_i} for each v_i, we can compute the marginal distribution $\mathbf{p}_r = (\mathbb{P}_T[c(E(r)) = \pi])_{\pi\in C_r}$ on C_r:*

$$\mathbf{p}_r(\pi) = f_\pi(\{\mathbf{p}_i\}_{i\in[d]}) = \frac{\prod_i \sum_{\tau\in C_{v_i}:\pi(e_i)\notin\tau}\mathbf{p}_i(\tau)}{\sum_{\rho\in C_r}\prod_i\sum_{\tau\in C_{v_i}:\rho(e_i)\notin\tau}\mathbf{p}_i(\tau)}.$$

Proof. By the definition of marginal probabilities, we have

$$\begin{aligned}\mathbf{p}_r(\pi) &= \mathbb{P}_T[c(E(r)) = \pi]\\ &= \frac{Z_T(c(E(r)) = \pi)}{\sum_{\rho\in C_r} Z_T(c(E(r)) = \rho)}.\end{aligned}$$

Since T is a tree, we can further write

$$\begin{aligned}\mathbf{p}_r(\pi) &= \frac{\prod_i Z_{T_i}(\pi(e_i) \notin c(E_{T_i}(v_i)))}{\sum_{\rho\in C_r}\prod_i Z_{T_i}(\rho(e_i)\notin c(E(v_i)))} \\ &= \frac{\prod_i \sum_{\tau\in C_{v_i}:\pi(e_i)\notin\tau} Z_{T_i}(c(E_{T_i}(v_i)) = \tau)}{\sum_{\rho\in C_r}\prod_i \sum_{\tau\in C_{v_i}:\rho(e_i)\notin\tau} Z_{T_i}(c(E_{T_i}(v_i)) = \tau)} \\ &= \frac{\prod_i \sum_{\tau\in C_{v_i}:\pi(e_i)\notin\tau} \mathbb{P}_{T_i}[c(E_{T_i}(v_i)) = \tau]}{\sum_{\rho\in C_r}\prod_i \sum_{\tau\in C_{v_i}:\rho(e_i)\notin\tau} \mathbb{P}_{T_i}[c(E_{T_i}(v_i)) = \tau]}.\end{aligned}$$

□

We will regard $f = (f_\pi)_{\pi\in C_r} : \mathbb{R}_{\geq 0}^{C_{v_1}} \times \mathbb{R}_{\geq 0}^{C_{v_2}} \times \cdots \times \mathbb{R}_{\geq 0}^{C_{v_d}} \to \mathcal{D}(C_r)$ as a function taking inputs $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_d)$ where $\mathbf{p}_i \in \mathbb{R}_{\geq 0}^{C_{v_i}}$ for $i \in [d]$. Note that in some cases, $\mathbf{p}$ might not encode a distribution.

3.3 Marginal bounds propagated by recursion

Now we can give marginal bounds of probabilities that propagated by the recursion. For brevity, we define some notations of marginals for any non-zero vector $\mathbf{p}_v \in \mathbb{R}_{\geq 0}^{C_v}$ with $v \in V$ and $c \in [q]$: Let $\mathbf{p}_v(c) = \sum_{\tau\in C_v: c\in\tau} \mathbf{p}_v(\tau)$, $\mathbf{p}_v(\bar{c}) = \sum_{\tau\in C_v: c\notin\tau} \mathbf{p}_v(\tau)$, and $\mathbf{p}_v(\bar{c_1}, \bar{c_2}) = \sum_{\tau\in C_v: c_1, c_2\notin\tau} \mathbf{p}_v(\tau)$.

Especially, for $\mathbf{p}_r \in \mathcal{D}(C_r)$, we define $\mathbf{p}_r(i, c) = \sum_{\pi\in C_r:\pi(e_i)=c} \mathbf{p}_r(\pi)$, and $\mathbf{p}_r(i, c_1, j, c_2) = \sum_{\pi\in C_r:\pi(e_i)=c_1,\pi(e_j)=c_2} \mathbf{p}_r(\pi)$.

By Lemma 11, we have

$$\mathbf{p}_r(\pi) = f_\pi(\{\mathbf{p}_i\}_{i\in[d]}) = \frac{\prod_i \mathbf{p}_{v_i}(\overline{\pi(e_i)})}{\sum_{\rho\in C_r}\prod_i \mathbf{p}_{v_i}(\overline{\rho(e_i)})}. \tag{2}$$

We have the following lemma similar to Lemma 9 on trees. Note that we do not require $\mathbf{p}_i$'s to be distributions in Lemma 12.

Lemma 12. *Given non-zero vector $\mathbf{p}_i \in \mathbb{R}_{\geq 0}^{C_{v_i}}$ for each $i \in [d]$ and $\mathbf{p}_r = f(\{\mathbf{p}_i\}_{i\in[d]})$, we have that for any color $a \in [q]$,*

$$\mathbf{p}_r(a) \leq \frac{d}{\beta - 1 + d} \quad \text{and} \quad \frac{\mathbf{p}_r(a)}{\mathbf{p}_r(\bar{a})} \leq \frac{d}{\beta - 1}.$$

Proof of Lemma 12. The proof is similar to that of Lemma 9. Define

$$A := \{\omega \in \Omega \mid \exists e \in E(r) : \omega(e) = a\}.$$

For any $\omega' \in A$, define

$$B(\omega') := \left\{\omega \in C_r \setminus A : d(\omega, \omega') = 1\right\}.$$

where d is the Hamming distance, i.e., the number of edges colored differently between the two colorings. We define

$$\iota(\omega, \omega') := \begin{cases} \frac{w(\omega)}{\sum_{\omega\in B(\omega')} w(\omega)} w(\omega'), & \omega \in B(\omega') \\ 0, & \text{otherwise} \end{cases}$$

where $w(\omega) := \prod_{i=1}^{d} \mathbf{p}_i(\overline{\omega(e_i)})$. Note that $\sum_{\omega \in C_r \setminus A} \iota(\omega, \omega') = w(\omega')$. For any $\omega' \in A$, assuming $\omega'(e_i) = a$, by (2),

$$\begin{aligned}\frac{w(\omega')}{\sum_{\omega \in B(\omega')} w(\omega)} &= \frac{\mathbf{p}_i(\overline{a})}{\sum_{a' \neq a} \mathbf{p}_i(\overline{a'})} \\ &= \frac{\sum_{\tau \in C_{v_i}, a \notin \tau} \mathbf{p}_i(\tau)}{\sum_{\tau \in C_{v_i}} \left(\sum_{a' \neq a} \mathbb{1}[a \notin \tau]\right) \mathbf{p}_i(\tau)} \\ &\leq \frac{\sum_{\tau \in C_{v_i}, a \notin \tau} \mathbf{p}_i(\tau)}{\sum_{\tau \in C_{v_i}, a \notin \tau} \left(\sum_{a' \neq a} \mathbb{1}[a \notin \tau]\right) \mathbf{p}_i(\tau)} \\ &\leq \frac{1}{\beta - 1}.\end{aligned}$$

The last inequality is because $\sum_{a' \neq a} \mathbb{1}[a \notin \tau] \geq \beta - 1$ for all $\tau \in C_{v_i}$. Then we have $\sum_{\omega' \in A} \iota(\omega, \omega') \leq \frac{d}{\beta - 1} w(\omega)$. Therefore by (2),

$$\begin{aligned}\mathbf{p}_r(a) &= \frac{\sum_{\omega' \in A} w(\omega')}{\sum_{\omega \in C_r \setminus A} w(\omega) + \sum_{\omega' \in A} w(\omega')} \\ &= \frac{\sum_{\omega' \in A} \sum_{\omega \in C_r \setminus A} \iota(\omega, \omega')}{\sum_{\omega \in \Omega \setminus A} w(\omega) + \sum_{\omega' \in A} \sum_{\omega \in C_r \setminus A} \iota(\omega, \omega')} \\ &= \frac{\sum_{\omega \in C_r \setminus A} \sum_{\omega' \in A} \iota(\omega, \omega')}{\sum_{\omega \in C_r \setminus A} \left(w(\omega) + \sum_{\omega' \in A} \iota(\omega, \omega')\right)} \\ &\leq \frac{d/(\beta - 1)}{1 + d/(\beta - 1)} = \frac{d}{\beta - 1 + d}.\end{aligned}$$

Similarly,

$$\frac{\mathbf{p}_r(a)}{\mathbf{p}_r(\overline{a})} = \frac{\sum_{\omega' \in A} w(\omega')}{\sum_{\omega \in C_r \setminus A} w(\omega)} = \frac{\sum_{\omega' \in A} \sum_{\omega \in C_r \setminus A} \iota(\omega, \omega')}{\sum_{\omega \in \Omega \setminus A} w(\omega)} \leq \frac{d}{\beta - 1}.$$

□

Lemma 13. *Given non-zero vector $\mathbf{p}_i \in \mathbb{R}_{\geq 0}^{C_{v_i}}$ for each $i \in [d]$ and $\mathbf{p}_r = f(\{\mathbf{p}_i\}_{i \in [d]})$, we have that for any color $a \in [q]$ and $j \in [d]$,*

$$\mathbf{p}_r(j, a) \leq \frac{\mathbf{p}_j(\overline{a})}{(\beta - 1) \sum_{\tau \in C_{v_j}} \mathbf{p}_j(\tau)}.$$

Proof. We have that

$$\begin{aligned}\mathbf{p}_r(j, a) &= \frac{\sum_{\substack{\sigma:\text{ coloring on } E(v) \setminus e_j \\ a \notin \sigma}} \prod_{i \neq j} \mathbf{p}_i(\overline{\sigma(v_i)}) \mathbf{p}_j(\overline{a})}{\sum_{\sigma:\text{ coloring on } E(r) \setminus e_j} \prod_{i \neq j} \mathbf{p}_i(\overline{\sigma(v_i)}) \left(\sum_{a' \notin \sigma} \mathbf{p}_j(\overline{a'})\right)} \\ &\leq \sup_{\sigma} \frac{\mathbf{p}_j(\overline{a})}{\sum_{a' \notin \sigma} \mathbf{p}_j(\overline{a'})} \\ &= \sup_{\sigma} \frac{\mathbf{p}_j(\overline{a})}{(\mathcal{L}(e_j) - \deg(v_j) + 1) \sum_{\tau \in C_{v_j}} \mathbf{p}_j(\tau) - \sum_{a' \in \sigma} \mathbf{p}_j(\overline{a'})} \\ &\leq \frac{\mathbf{p}_j(\overline{a}) / (\sum_{\tau \in C_{v_j}} \mathbf{p}_j(\tau))}{\mathcal{L}(e_j) - \deg(v_j) - \deg(r) + 1}.\end{aligned}$$

□

4 FPTAS for counting proper edge colorings on general graphs $q \geq 3\Delta$

In this section, we prove the following main algorithmic result, which is a formal version of Theorem 1.

Theorem 14. *Assume $\Delta \geq 4$. There exists a deterministic algorithm that outputs $\hat{Z}$ satisfying $(1-\delta)Z_{G,\mathcal{L}} \leq \hat{Z} \leq (1+\delta)Z_{G,\mathcal{L}}$ for any $(\Delta+2)$-extra edge coloring instance $(G,\mathcal{L})$ with maximum degree Δ and given error bound $0 < \delta < 1$ in time $\left(\frac{n}{\delta}\right)^{C(\Delta)}$, where n is the number of edges in G and $C(\Delta) = \mathcal{O}\left(\Delta^{\Delta \log \Delta} \log \Delta\right)$ is a universal constant only depends on Δ.*

Our key contribution is the following coupling independence result.

Theorem 15. *Let $(G = (V,E), \mathcal{L})$ be a $((1+\varepsilon)\Delta+1)$-extra list-edge-coloring instance. Then μ_E is $\left(1+\frac{2}{\varepsilon}\right)$-coupling independent. That is, for any $i \in E$, $a,b \in \mathcal{L}(i)$,*

$$\mathcal{W}_1\left(\mu_E^{i\leftarrow a}, \mu_E^{i\leftarrow b}\right) \leq 1 + \frac{2}{\varepsilon}.$$

We first set up our terminologies to argue about the Wasserstein distance. We define some upper bounds for the $\mathcal{W}_1$ distance between $\lambda\Delta$-extra list-colorings on Δ-degree graphs with s edges with respect to one different pinning. We will construct recursion on these upper bounds. An edge is *pendant* if one of its endpoints has degree exactly 1.

Definition 16 (Universal upper bounds for coupling independence)**.** *Define*

$$\kappa_{s,\Delta,\lambda} := \sup_{\substack{(G=(V,E),\mathcal{L}) \\ i\in E: i \text{ is pendant} \\ a,b\in\mathcal{L}(i), a\neq b}} \mathcal{W}_1\left(\mu_{E-i}^{i\leftarrow a}, \mu_{E-i}^{i\leftarrow b}\right)$$

where $(G,\mathcal{L})$ is taken over

1. *all graph $G=(V,E)$ such that $\deg(G) \leq \Delta$, $|E| \leq s$;*
2. *all color lists $\mathcal{L}$ such that $\forall e \in E: |\mathcal{L}(e)| \geq \deg(e) + \lambda\Delta + 1$.*

Remark. *It is clear from the definition that $\kappa_{s+1,\Delta,\lambda} \geq \kappa_{s,\Delta,\lambda}$ and $\kappa_{1,\Delta,\lambda} = 0$.*

Note that we only pin color on the edge i in Definition 16. If we need other pinnings, we can simply consider the pinnings as deleting the pinned edges and remove the pinned color from the lists of their adjacent edges.

The main lemma of this section is a recursion for $\kappa_{s,\Delta,\lambda}$ and leads to Theorem 15 immediately.

Lemma 17. *Let $\lambda = 1+\varepsilon$ for some $\varepsilon > 0$ and $s \geq 2$. Then*

$$\kappa_{s,\Delta,\lambda} \leq \frac{2}{2+\varepsilon}\left(\kappa_{s-1,\Delta,\lambda} + \frac{1}{2}\right).$$

The proof of Lemma 17 is based on a greedy one-step coupling of two marginal distributions with different pinning on a single edge. We describe the coupling in Section 4.1. The proofs of Lemma 17, Theorem 15 and Theorem 14 are in Section 4.2.

4.1 Decomposition of Wasserstein distance

The following lemma shows how we can go from $\kappa_{s,\Delta,\lambda}$ to $\kappa_{s-1,\Delta,\lambda}$ by one extra pinning.

Lemma 18. *For the instance $(G = (V,E),\mathcal{L})$, the pendant edge $i = \{u,v\} \in E$, and the colors $a,b \in \mathcal{L}(i)$ that fit into the definition of $\kappa_{s,\Delta,\lambda}$. Suppose $\deg(u) = 1$ and $\deg(v) \geq 2$, and $j \in N(i)$, $a,b \in \mathcal{L}(j)$. Then*

1. $\mathcal{W}_1\left(\mu_{E-i}^{\substack{i\leftarrow a\\ j\leftarrow b}}, \mu_{E-i}^{\substack{i\leftarrow b\\ j\leftarrow a}}\right) \leq 1 + \kappa_{s-1,\Delta,\lambda}$,
2. $\mathcal{W}_1\left(\mu_{E-i}^{\substack{i\leftarrow a\\ j\leftarrow b}}, \mu_{E-i}^{i\leftarrow b}\right), \mathcal{W}_1\left(\mu_{E-i}^{\substack{i\leftarrow b\\ j\leftarrow a}}, \mu_{E-i}^{i\leftarrow a}\right) \leq 1 + 2\kappa_{s-1,\Delta,\lambda}$,
3. $\mathcal{W}_1\left(\mu_{E-i}^{\substack{i\leftarrow a\\ b\notin N}}, \mu_{E-i}^{\substack{i\leftarrow b\\ a\notin N}}\right) = 0$.

Proof. Assume $j = \{v,w\}$.

1. We define a new instance $(G' = (V', E' = E - i), \mathcal{L}')$ by removing i, disconnecting j from v and delete a,b in the color lists of edges in $N(v)$. Then j becomes to a pendant edge. We have that

$$\mathcal{W}_1\left(\mu_{E-i;(G,\mathcal{L})}^{\substack{i\leftarrow a\\ j\leftarrow b}}, \mu_{E-i;(G,\mathcal{L})}^{\substack{i\leftarrow b\\ j\leftarrow a}}\right) = 1 + \mathcal{W}_1\left(\mu_{E-i-j;(G,\mathcal{L})}^{j\leftarrow b}, \mu_{E-i-j;(G,\mathcal{L})}^{j\leftarrow a}\right)$$
$$= 1 + \mathcal{W}_1\left(\mu_{E'-j;(G',\mathcal{L}')}^{j\leftarrow b}, \mu_{E'-j;(G',\mathcal{L}')}^{j\leftarrow a}\right).$$

   After the deletion of i and the removal of a or b from the color lists of $N(v)$, the number of extra colors of each edge remains unchanged. So by Definition 16, the Wasserstein distance $\mathcal{W}_1\left(\mu_{E'-j;(G',\mathcal{L}')}^{j\leftarrow b}, \mu_{E'-j;(G',\mathcal{L}')}^{j\leftarrow a}\right)$ is bounded by $\kappa_{s-1,\Delta,\lambda}$, so we have

$$\mathcal{W}_1\left(\mu_{E-i;(G,\mathcal{L})}^{\substack{i\leftarrow a\\ j\leftarrow b}}, \mu_{E-i;(G,\mathcal{L})}^{\substack{i\leftarrow b\\ j\leftarrow a}}\right) \leq 1 + \kappa_{s-1,\Delta,\lambda}.$$

2. By the law of total probability,

$$\mu_{E-i;(G,\mathcal{L})}^{\substack{i\leftarrow a\\ j\leftarrow b}} - \mu_{E-i;(G,\mathcal{L})}^{i\leftarrow b} = \mu_{E-i;(G,\mathcal{L})}^{\substack{i\leftarrow a\\ j\leftarrow b}} - \sum_{c\in\mathcal{L}(j)-b} \mu_{E-i;(G,\mathcal{L})}^{i\leftarrow b}(j\leftarrow c)\mu_{E-i;(G,\mathcal{L})}^{\substack{i\leftarrow b\\ j\leftarrow c}}$$
$$= \sum_{c\in\mathcal{L}(j)-b} \mu_{E-i;(G,\mathcal{L})}^{i\leftarrow b}(j\leftarrow c)\left(\mu_{E-i;(G,\mathcal{L})}^{\substack{i\leftarrow a\\ j\leftarrow b}} - \mu_{E-i;(G,\mathcal{L})}^{\substack{i\leftarrow b\\ j\leftarrow c}}\right).$$

   By proposition 6 this means

$$\mathcal{W}_1\left(\mu_{E-i;(G,\mathcal{L})}^{\substack{i\leftarrow a\\ j\leftarrow b}}, \mu_{E-i;(G,\mathcal{L})}^{i\leftarrow b}\right) \leq \sum_{c\in\mathcal{L}(j)-b} \mu_{E-i;(G,\mathcal{L})}^{i\leftarrow b}(j\leftarrow c)\mathcal{W}_1\left(\mu_{E-i;(G,\mathcal{L})}^{\substack{i\leftarrow a\\ j\leftarrow b}}, \mu_{E-i;(G,\mathcal{L})}^{\substack{i\leftarrow b\\ j\leftarrow c}}\right)$$
$$\leq 1 + \sum_{c\in\mathcal{L}(j)-b} \mu_{E-i;(G,\mathcal{L})}^{i\leftarrow b}(j\leftarrow c)\mathcal{W}_1\left(\mu_{E-i-j;(G,\mathcal{L})}^{\substack{i\leftarrow a\\ j\leftarrow b}}, \mu_{E-i-j;(G,\mathcal{L})}^{\substack{i\leftarrow b\\ j\leftarrow c}}\right). \tag{3}$$

   For each $c \in \mathcal{L}(j) - b$, we construct a new list-edge-coloring instance $(G' = (V, E'), \mathcal{L}')$ By removing j, appending a new edge j' to w, and removing b from the color lists of edges in $N(v)$. Then we have the following identities since the color constraints of each pair of edges are the same.

$$\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i-j;(G,\mathcal{L})} = \mu^{\substack{i\leftarrow a\\ j'\leftarrow b}}_{E'-i-j';(G',\mathcal{L}')}, \quad \mu^{\substack{i\leftarrow b\\ j\leftarrow c}}_{E-i-j;(G,\mathcal{L})} = \mu^{\substack{i\leftarrow c\\ j'\leftarrow c}}_{E'-i-j';(G',\mathcal{L}')}.$$

Applying these identities and triangle inequality to (3), we have

$$\begin{aligned}
\mathcal{W}_1\left(\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i;(G,\mathcal{L})}, \mu^{i\leftarrow b}_{E-i;(G,\mathcal{L})}\right) &\le 1 + \sum_{c\in\mathcal{L}(j)-b} \mu^{i\leftarrow b}_{E-i;(G,\mathcal{L})}(j\leftarrow c)\mathcal{W}_1\left(\mu^{\substack{i\leftarrow a\\ j'\leftarrow b}}_{E'-i-j';(G',\mathcal{L}')}, \mu^{\substack{i\leftarrow c\\ j'\leftarrow c}}_{E'-i-j';(G',\mathcal{L}')}\right)\\
&\le 1 + \sum_{c\in\mathcal{L}(j)-b} \mu^{i\leftarrow b}_{E-i;(G,\mathcal{L})}(j\leftarrow c)\Bigg(\mathcal{W}_1\left(\mu^{\substack{i\leftarrow a\\ j'\leftarrow b}}_{E'-i-j';(G',\mathcal{L}')}, \mu^{\substack{i\leftarrow a\\ j'\leftarrow c}}_{E'-i-j';(G',\mathcal{L}')}\right)\\
&\quad + \mathcal{W}_1\left(\mu^{\substack{i\leftarrow a\\ j'\leftarrow c}}_{E'-i-j';(G',\mathcal{L}')}, \mu^{\substack{i\leftarrow c\\ j'\leftarrow c}}_{E'-i-j';(G',\mathcal{L}')}\right)\Bigg).
\end{aligned}$$

Since we may remove the edge with the same pinning from the graph and remain the distribution unchanged, the two $\mathcal{W}_1$ distances are both bounded by $\kappa_{s-1,\Delta,\lambda}$, proving the second part of the lemma.

3. For the third part, notice that the available colors of all edges in $E-i$ are exactly the same, so $\mu^{\substack{i\leftarrow a\\ b\notin N}}_{E-i} = \mu^{\substack{i\leftarrow b\\ a\notin N}}_{E-i}$ and
$$\mathcal{W}_1\left(\mu^{\substack{i\leftarrow a\\ b\notin N}}_{E}, \mu^{\substack{i\leftarrow b\\ a\notin N}}_{E}\right) = 0. \qquad \square$$

Let $(G=(V,E),\mathcal{L})$ be a list-edge-coloring instance that fits into the constraints of $\kappa_{s,\Delta,\lambda}$ in Definition 16, $i\in E$ be a pendant edge, and $a,b\in\mathcal{L}(i)$ be colors. We do a one-step coupling to reduce the W_1 distance between graphs with s edges to that between graphs with $s-1$ edges using Proposition 6.

Denoting the two endpoints of i by u,v, we may assume u is the pendant vertex, that is $\deg(u)=1$ without loss of generality.

We denote the non-empty set $E(v)-i$ by N, and use integers from 1 to denote the edges in N so that $N=\{1,\ldots,d\}$. For every edge $j\in N$, define

$$\gamma_j := \frac{\mu^{i\leftarrow a}_{E-i}(j\leftarrow b)}{\mu^{i\leftarrow a}_{E-i}(b\notin N)}, \quad \delta_j := \frac{\mu^{i\leftarrow b}_{E-i}(j\leftarrow a)}{\mu^{i\leftarrow b}_{E-i}(a\notin N)}.$$

The following lemma describes the greedy coupling we use.

Lemma 19.

$$\begin{aligned}
\mu^{i\leftarrow a}_{E-i} - \mu^{i\leftarrow b}_{E-i} &= \sum_{j\in N}\frac{\gamma_j\wedge\delta_j}{1+\sum_k(\gamma_k\vee\delta_k)}\left(\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i} - \mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i}\right)\\
&\quad + \sum_{j\in N}\frac{(\gamma_j-\delta_j)\vee 0}{1+\sum_k(\gamma_k\vee\delta_k)}\left(\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i} - \mu^{i\leftarrow b}_{E-i}\right) + \sum_{j\in N}\frac{(\delta_j-\gamma_j)\vee 0}{1+\sum_k(\gamma_k\vee\delta_k)}\left(\mu^{i\leftarrow a}_{E-i} - \mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i}\right)\\
&\quad + \frac{1}{1+\sum_k(\gamma_k\vee\delta_k)}\left(\mu^{\substack{i\leftarrow a\\ b\notin N}}_{E-i} - \mu^{\substack{i\leftarrow b\\ a\notin N}}_{E-i}\right).
\end{aligned}$$

Proof. The criterion of decomposing $\mu^{i\leftarrow a}_{E-i} - \mu^{i\leftarrow b}_{E-i}$ is whether a or b appears in N. Without loss of generality, we assume $\mu^{i\leftarrow a}_{E-i}(b\notin N)\le\mu^{i\leftarrow b}_{E-i}(a\notin N)$, and denote $\mu^{i\leftarrow a}_{E-i}(b\notin N)/\mu^{i\leftarrow b}_{E-i}(a\notin N)$ by α. By the law of total probability, we have

$$\mu^{i\leftarrow a}_{E-i} = \sum_{j\in N}\mu^{i\leftarrow a}_{E-i}(j\leftarrow b)\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i} + \mu^{i\leftarrow a}_{E-i}(b\notin N)\mu^{\substack{i\leftarrow a\\ b\notin N}}_{E-i}. \tag{4}$$

It is also clear that

$$
\begin{aligned}
\mu^{i\leftarrow b}_{E-i} &= \alpha\mu^{i\leftarrow b}_{E-i} + (1-\alpha)\mu^{i\leftarrow b}_{E-i} \\
&= \alpha\left(\sum_{j\in N}\mu^{i\leftarrow b}_{E-i}(j\leftarrow a)\mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i} + \mu^{i\leftarrow b}_{E-i}(a\notin N)\mu^{\substack{i\leftarrow b\\ a\notin N}}_{E-i}\right) + (1-\alpha)\mu^{i\leftarrow b}_{E-i} \\
&= \alpha\sum_{j\in N}\mu^{i\leftarrow b}_{E-i}(j\leftarrow a)\mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i} + \mu^{i\leftarrow a}_{E-i}(b\notin N)\mu^{\substack{i\leftarrow b\\ a\notin N}}_{E-i} + (1-\alpha)\mu^{i\leftarrow b}_{E-i}.
\end{aligned}
$$

The point of the multiplier α is to align the coefficients of $\mu^{\substack{i\leftarrow a\\ b\notin N}}_{E-i}$ and $\mu^{\substack{i\leftarrow b\\ a\notin N}}_{E-i}$, so that they don't pair with other distributions, and will not introduce the pinning $\cdot\notin N$, which reduces the number of extra colors, into the recursion. By the above two decompositions,

$$
\begin{aligned}
\mu^{i\leftarrow a}_{E-i} - \mu^{i\leftarrow b}_{E-i} = &\sum_{j\in N}\mu^{i\leftarrow a}_{E-i}(j\leftarrow b)\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i} - \alpha\sum_{j\in N}\mu^{i\leftarrow b}_{E-i}(j\leftarrow a)\mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i} \\
&- (1-\alpha)\mu^{i\leftarrow b}_{E-i} \\
&+ \mu^{i\leftarrow a}_{E-i}(b\notin N)\left(\mu^{\substack{i\leftarrow a\\ b\notin N}}_{E-i} - \mu^{\substack{i\leftarrow b\\ a\notin N}}_{E-i}\right).
\end{aligned}
\tag{5}
$$

In general $\mu^{i\leftarrow a}_{E-i}(j\leftarrow b)$ and $\mu^{i\leftarrow b}_{E-i}(j\leftarrow a)$ do not equal, and we need to analyze them carefully. Then we can express $\mu^{i\leftarrow a}_{E-i}(j\leftarrow b), \mu^{i\leftarrow a}_{E-i}(b\notin N), \mu^{i\leftarrow b}_{E-i}(j\leftarrow a), \mu^{i\leftarrow b}_{E-i}(a\notin N)$ by them.

$$
\mu^{i\leftarrow a}_{E-i}(j\leftarrow b) = \frac{\mu^{i\leftarrow a}_{E-i}(j\leftarrow b)}{\mu^{i\leftarrow a}_{E-i}(b\notin N) + \sum_{k\in N}\mu^{i\leftarrow a}_{E-i}(k\leftarrow b)} = \frac{\gamma_j}{1+\sum_k\gamma_k}.
$$

We omitted the range of k for simplicity. Similarly,

$$
\mu^{i\leftarrow b}_{E-i}(j\leftarrow a) = \frac{\delta_j}{1+\sum_k\delta_k},\quad \mu^{i\leftarrow a}_{E-i}(b\notin N) = \frac{1}{1+\sum_k\gamma_k},\quad \mu^{i\leftarrow b}_{E-i}(a\notin N) = \frac{1}{1+\sum_k\delta_k},\quad \alpha = \frac{1+\sum_k\delta_k}{1+\sum_k\gamma_k}.
$$

Then we plug them into eq. (5).

$$
\begin{aligned}
\mu^{i\leftarrow a}_{E-i} - \mu^{i\leftarrow b}_{E-i} &= \sum_j\frac{\gamma_j}{1+\sum_k\gamma_k}\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i} - \sum_j\frac{\delta_j}{1+\sum_k\gamma_k}\mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i} \\
&\quad + \frac{\sum_k(\delta_k-\gamma_k)}{1+\sum_k\gamma_k}\mu^{i\leftarrow b}_{E-i} + \frac{1}{1+\sum_k\gamma_k}\left(\mu^{\substack{i\leftarrow a\\ b\notin N}}_{E-i} - \mu^{\substack{i\leftarrow b\\ a\notin N}}_{E-i}\right) \\
&= \sum_j\frac{\gamma_j\wedge\delta_j}{1+\sum_k\gamma_k}\left(\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i} - \mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i}\right) \\
&\quad + \sum_j\frac{(\gamma_j-\delta_j)\vee 0}{1+\sum_k\gamma_k}\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i} - \sum_j\frac{(\delta_j-\gamma_j)\vee 0}{1+\sum_k\gamma_k}\mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i} \\
&\quad + \frac{\sum_k(\delta_k-\gamma_k)}{1+\sum_k\gamma_k}\mu^{i\leftarrow b}_{E-i} + \frac{1}{1+\sum_k\gamma_k}\left(\mu^{\substack{i\leftarrow a\\ b\notin N}}_{E-i} - \mu^{\substack{i\leftarrow b\\ a\notin N}}_{E-i}\right).
\end{aligned}
$$

By adding $\frac{\sum_k(\delta_k-\gamma_k)\vee 0}{1+\sum_k\gamma_k}\left(\mu^{i\leftarrow a}_{E-i}-\mu^{i\leftarrow b}_{E-i}\right)$ on both sides, we get

$$\begin{aligned}\frac{1+\sum_k(\gamma_k\vee\delta_k)}{1+\sum_k\gamma_k}\left(\mu^{i\leftarrow a}_{E-i}-\mu^{i\leftarrow b}_{E-i}\right)&=\sum_j\frac{\gamma_j\wedge\delta_j}{1+\sum_k\gamma_k}\left(\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i}-\mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i}\right)\\&+\sum_j\frac{(\gamma_j-\delta_j)\vee 0}{1+\sum_k\gamma_k}\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i}-\sum_j\frac{(\delta_j-\gamma_j)\vee 0}{1+\sum_k\gamma_k}\mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i}\\&+\sum_j\frac{(\delta_j-\gamma_j)\vee 0}{1+\sum_k\gamma_k}\mu^{i\leftarrow a}_{E-i}-\sum_j\frac{(\delta_j-\gamma_j)\vee 0}{1+\sum_k\gamma_k}\mu^{i\leftarrow b}_{E-i}\\&+\frac{\sum_k(\delta_k-\gamma_k)}{1+\sum_k\gamma_k}\mu^{i\leftarrow b}_{E-i}+\frac{1}{1+\sum_k\gamma_k}\left(\mu^{\substack{i\leftarrow a\\ b\notin N}}_{E-i}-\mu^{\substack{i\leftarrow b\\ a\notin N}}_{E-i}\right).\end{aligned}$$

The highlighted terms sums to $-\sum_j\frac{(\gamma_j-\delta_j)\vee 0}{1+\sum_k\gamma_k}\mu^{i\leftarrow b}_{E-i}$. So we can pair the terms to get

$$\begin{aligned}\frac{1+\sum_k(\gamma_k\vee\delta_k)}{1+\sum_k\gamma_k}\left(\mu^{i\leftarrow a}_{E-i}-\mu^{i\leftarrow b}_{E-i}\right)&=\sum_j\frac{\gamma_j\wedge\delta_j}{1+\sum_k\gamma_k}\left(\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i}-\mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i}\right)\\&+\sum_j\frac{(\gamma_j-\delta_j)\vee 0}{1+\sum_k\gamma_k}\left(\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i}-\mu^{i\leftarrow b}_{E-i}\right)+\sum_j\frac{(\delta_j-\gamma_j)\vee 0}{1+\sum_k\gamma_k}\left(\mu^{i\leftarrow a}_{E-i}-\mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i}\right)\\&+\frac{1}{1+\sum_k\gamma_k}\left(\mu^{\substack{i\leftarrow a\\ b\notin N}}_{E-i}-\mu^{\substack{i\leftarrow b\\ a\notin N}}_{E-i}\right).\end{aligned}$$

Finally, by dividing $\frac{1+\sum_k(\gamma_k\vee\delta_k)}{1+\sum_k\gamma_k}$, we get the decomposition

$$\begin{aligned}\mu^{i\leftarrow a}_{E-i}-\mu^{i\leftarrow b}_{E-i}&=\sum_j\frac{\gamma_j\wedge\delta_j}{1+\sum_k(\gamma_k\vee\delta_k)}\left(\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i}-\mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i}\right)\\&+\sum_j\frac{(\gamma_j-\delta_j)\vee 0}{1+\sum_k(\gamma_k\vee\delta_k)}\left(\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i}-\mu^{i\leftarrow b}_{E-i}\right)+\sum_j\frac{(\delta_j-\gamma_j)\vee 0}{1+\sum_k(\gamma_k\vee\delta_k)}\left(\mu^{i\leftarrow a}_{E-i}-\mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i}\right)\\&+\frac{1}{1+\sum_k(\gamma_k\vee\delta_k)}\left(\mu^{\substack{i\leftarrow a\\ b\notin N}}_{E-i}-\mu^{\substack{i\leftarrow b\\ a\notin N}}_{E-i}\right).\end{aligned}$$

□

4.2 Proof of main theorems

Now we prove Lemma 17, which provides a recursion for $\kappa_{s,\Delta,\lambda}$.

Proof of Lemma 17. We consider every $\lambda\Delta+1$-extra color instance $(G=(V,E),\mathcal{L})$ with a pendant edge $i=\{u,v\}$ such that $\deg(G)\le\Delta$, $|E|\le s$. Suppose $\deg(u)=1$. If $\deg(v)=1$, then $\mathcal{W}_1\left(\mu^{i\leftarrow a}_{E-i},\mu^{i\leftarrow b}_{E-i}\right)=0$. In the following we assume $\deg(v)\ge 2$.

An application of Proposition 6 gives

$$\begin{aligned}\mathcal{W}_1\left(\mu^{i\leftarrow a}_{E-i},\mu^{i\leftarrow b}_{E-i}\right)&\le\sum_j\frac{\gamma_j\wedge\delta_j}{1+\sum_k(\gamma_k\vee\delta_k)}\mathcal{W}_1\left(\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i},\mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i}\right)\\&+\sum_j\frac{(\gamma_j-\delta_j)\vee 0}{1+\sum_k(\gamma_k\vee\delta_k)}\mathcal{W}_1\left(\mu^{\substack{i\leftarrow a\\ j\leftarrow b}}_{E-i},\mu^{i\leftarrow b}_{E-i}\right)+\sum_j\frac{(\delta_j-\gamma_j)\vee 0}{1+\sum_k(\gamma_k\vee\delta_k)}\mathcal{W}_1\left(\mu^{i\leftarrow a}_{E-i},\mu^{\substack{i\leftarrow b\\ j\leftarrow a}}_{E-i}\right)\\&+\frac{1}{1+\sum_k(\gamma_k\vee\delta_k)}\mathcal{W}_1\left(\mu^{\substack{i\leftarrow a\\ b\notin N}}_{E-i},\mu^{\substack{i\leftarrow b\\ a\notin N}}_{E-i}\right).\end{aligned}$$

By Lemma 18,

$$\begin{aligned}\mathcal{W}_1\left(\mu_{E-i}^{i\leftarrow a},\mu_{E-i}^{i\leftarrow b}\right) &\le 1-\frac{1}{1+\sum_k(\gamma_k\vee\delta_k)}+\frac{\sum_j(\gamma_j\wedge\delta_j)+2|\gamma_j-\delta_j|}{1+\sum_k(\gamma_k\vee\delta_k)}\kappa_{s-1,\Delta,\lambda}\\ &\le 1-\frac{1}{1+\sum_k(\gamma_k\vee\delta_k)}+\frac{\sum_j 2(\gamma_j\vee\delta_j)}{1+\sum_k(\gamma_k\vee\delta_k)}\kappa_{s-1,\Delta,\lambda}.\end{aligned}$$

Finally, the bound $\forall j\in N:\gamma_j,\delta_j\le\frac{1}{\lambda\Delta}$ from Corollary 10 gives

$$\mathcal{W}_1\left(\mu_{E-i}^{i\leftarrow a},\mu_{E-i}^{i\leftarrow b}\right)\le\frac{1}{1+\lambda}+\frac{2/\lambda}{1+1/\lambda}\kappa_{s-1,\Delta,\lambda}\le\frac{2}{2+\varepsilon}\left(\frac{1}{2}+\kappa_{s-1,\Delta,\lambda}\right).$$

Taking the supremum as in Definition 16 proves the lemma. □

Proof of Theorem 15. Lemma 17 shows that $\sup_s\kappa_{s,\Delta,\lambda}\le\frac{1}{2+\varepsilon}(1+2/\varepsilon)=1/\varepsilon$. And

$$\mathcal{W}_1\left(\mu_E^{i\leftarrow a},\mu_E^{i\leftarrow b}\right)\le 1+\mathcal{W}_1\left(\mu_{E-i}^{i\leftarrow a},\mu_{E-i}^{i\leftarrow b}\right).$$

To reduce to the pendant edge case, we may break i into two pendant edges i_1,i_2, connected to the two endpoints of i respectively, the color lists of i_1,i_2 are the same as i. We denote the new coloring instance by $(G',\mathcal{L}')$. Then we have

$$\mu_{E-i;(G,\mathcal{L})}^{i\leftarrow a}=\mu_{E-i_1-i_2;(G',\mathcal{L}')}^{\substack{i_2\leftarrow a\\ i_1\leftarrow a}},\quad \mu_{E-i;(G,\mathcal{L})}^{i\leftarrow b}=\mu_{E-i_1-i_2;(G',\mathcal{L}')}^{\substack{i_2\leftarrow b\\ i_1\leftarrow b}}.$$

So by triangle inequality, we have

$$\begin{aligned}&\mathcal{W}_1\left(\mu_{E;(G,\mathcal{L})}^{i\leftarrow a},\mu_{E;(G,\mathcal{L})}^{i\leftarrow b}\right)\\ &\le 1+\mathcal{W}_1\left(\mu_{E-i;(G,\mathcal{L})}^{i\leftarrow a},\mu_{E-i;(G,\mathcal{L})}^{i\leftarrow b}\right)\\ &=1+\mathcal{W}_1\left(\mu_{E-i_1-i_2;(G',\mathcal{L}')}^{\substack{i_2\leftarrow a\\ i_1\leftarrow a}},\mu_{E-i_1-i_2;(G',\mathcal{L}')}^{\substack{i_2\leftarrow b\\ i_1\leftarrow b}}\right)\\ &\le 1+\mathcal{W}_1\left(\mu_{E-i_1-i_2;(G',\mathcal{L}')}^{\substack{i_2\leftarrow a\\ i_1\leftarrow a}},\mu_{E-i_1-i_2;(G',\mathcal{L}')}^{\substack{i_2\leftarrow b\\ i_1\leftarrow a}}\right)+\mathcal{W}_1\left(\mu_{E-i_1-i_2;(G',\mathcal{L}')}^{\substack{i_2\leftarrow b\\ i_1\leftarrow a}},\mu_{E-i_1-i_2;(G',\mathcal{L}')}^{\substack{i_2\leftarrow b\\ i_1\leftarrow b}}\right)\\ &\le 1+\frac{2}{\epsilon}.\end{aligned}$$

□

Theorem 15 shows that any $((1+\varepsilon)\Delta+1)$-extra list-edge-coloring instance is $\left(1+\frac{2}{\varepsilon}\right)$-coupling independent. This is because adding additional pinnings can be viewed as generating a new $((1+\varepsilon)\Delta+1)$-extra list-edge-coloring instance by deleting the pinned edges and removing the corresponding colors from the lists of their adjacent edges.

The work of [CFG+24] designs an **FPTAS** for counting the partition function of any Gibbs distribution of permissive spin systems that is marginally bounded and coupling independent. A spin system is specified by a 4-tuple $S=(G=(V,E),q,A_E,A_V)$ where the state space is $[q]^V$ and the weight of a configuration is characterized by the matrices $A_E\in\mathbb{R}_{\ge0}^{q\times q}$ and $A_V\in\mathbb{R}_{\ge0}^{q}$. The Gibbs distribution is defined by:

$$\mu(\sigma)\propto w(\sigma):=\prod_{u,v\in E}A_E(\sigma(u),\sigma(v))\prod_{v\in V}A_V(\sigma(v)).$$

The normalizing factor of μ is called the partition function $Z:=\sum_{\sigma\in[q]^V}w(\sigma)$.

We say S is permissive if for any partial configuration $\tau \in [q]^\Lambda$ with $\Lambda \subseteq V$, the conditional partition function $Z^\tau = \sum_{\sigma:\tau\subset\sigma} w(\sigma) > 0$. For $\tau \in [q]^\Lambda$ with $\Lambda \subset V$, let μ_v^τ be the marginal distribution on $v \in V \setminus \Lambda$ conditional on the partial configuration τ. We say μ is b-marginally bounded if for any partial configuration $\tau \in [q]^\Lambda$ with $\Lambda \subseteq V$, any vertex $v \in V \setminus \Lambda$ and $c \in [q]$ such that $\mu_v^\tau > 0$, we have $\mu_v^\tau \geq b$.

The main result of [CFG+24] that we will use is as follows.

Theorem 20 ([CFG+24])**.** *Let $q \geq 2, b > 0, C > 0, \Delta \geq 3$ be constants. There exists a deterministic algorithm such that given a permissive spin system $\mathcal{S} = (G, q, A_E, A_V)$ and error bound $0 < \varepsilon < 1$, if the Gibbs distribution of $\mathcal{S}$ is b-marginally bounded and satisfies C-coupling independence, and the maximum degree of G is at most Δ, then it returns $\hat{Z}$ satisfying $(1-\varepsilon)Z \leq \hat{Z} \leq (1+\varepsilon)Z$ in time $\left(\frac{n}{\varepsilon}\right)^{f(q,b,C,\Delta)}$, where $f(q,b,C,\Delta) = \Delta^{\mathcal{O}\left(C\left(\log b^{-1} + \log C + \log\log \Delta\right)\right)} \log q$ is a constant.*

In our setting, the spin system is list-edge-coloring instance. Be careful with the parameters since the spins are on edges of the graph. In order to prove Theorem 14, we need the marginal lower bound from [GKM15].

Lemma 21 (Corollary of Lemma 3 in [GKM15])**.** *Fix a β-extra edge coloring instances $(G, \mathcal{L})$ of maximum degree Δ with Gibbs distribution μ. For any partial coloring τ on $\Lambda \subseteq E$, any $e \in E \setminus \Lambda$, and $c \in \mathcal{L}^\tau(e)$, it holds that*

$$\mu_e^\tau(c) \geq \frac{\left(1 - \frac{1}{|\mathcal{L}^\tau(e)| - \deg^\tau(e)}\right)^{\deg^\tau(e)}}{|\mathcal{L}^\tau(e)|} \geq \frac{(1-\frac{1}{\beta})^{2\Delta-2}}{\beta + 2\Delta - 2}.$$

Now we give the proof of our main theorem in this section.

Proof of Theorem 14. It is easy to verify that any β-extra edge coloring instance with $\beta \geq 1$ is permissive. From Theorem 15 and Lemma 21, we know that the Gibbs distribution μ of a $(\Delta+2)$-extra edge coloring instance $(G, \mathcal{L})$ is $(1+2\Delta)$-coupling independent and b-marginally bounded where $b = \frac{(1-\frac{1}{\Delta+2})^{2\Delta-2}}{3\Delta}$ is $\Omega\left(\frac{1}{\Delta}\right)$. Then by Theorem 20, there exists a deterministic algorithm that outputs $\hat{Z}$ satisfying $(1-\delta)Z_{G,\mathcal{L}} \leq \hat{Z} \leq (1+\delta)Z_{G,\mathcal{L}}$ in time $\left(\frac{n}{\delta}\right)^{C(\Delta)}$ where n is the number of edges in G and $C(\Delta) = \mathcal{O}\left(\Delta^{\Delta \log \Delta} \log \Delta\right)$. □

5 Strong spatial mixing for edge colorings on trees when $q > (3+o(1))\Delta$

In this section, we prove our theorem for strong spatial mixing.

Theorem 22. *Given a β-extra edge coloring instance $(G, \mathcal{L})$ where G is a tree of maximum degree Δ, the uniform distribution on such instance exhibits strong spatial mixing with exponential decay rate $1-\delta$ and constant $C = \max\{32q^{\frac{\Delta+2}{2}}\Delta^2(1-\delta)^{-3}, (1-\delta)^{-4}\}$ if $\beta > \max\{\Delta+50, (1+\eta_\Delta)\Delta+1\}$, where*

$$\delta = \frac{(1+(\beta-1-\Delta)/\Delta)^2 - (1+\eta_\Delta)^2}{2(1+(\beta-1-\Delta)/\Delta)^2}, \eta_\Delta = O\left(\frac{\log^2 \Delta}{\Delta}\right).$$

Specifically, if $\beta = \left(1 + \frac{\log^3 \Delta}{\Delta}\right)\Delta + 50$, then $\delta \approx \frac{\log^3 \Delta}{\Delta}$.

We already introduced the recursion for marginal probabilities of edge colorings on trees and derived certain marginal bounds in Section 3. We will then analyze its Jacobian matrix in Section 5.1. Using the bounds on the norm of the Jacobian matrix, we prove Theorem 22 in Section 5.2. Finally, we discuss the limit of our approach and possible further improvement in Section 5.3. A key ingredient in our bounds for the norm of Jacobian matrix is a bound for certain covariance matrices, which is addressed in Section 6.

5.1 Upper bound the 2-norm of Jacobian

5.1.1 The Jacobian

Recall the recursion f for marginals introduced in Section 3.2. We regard $f = (f_\pi)_{\pi \in C_r}$: $\mathbb{R}_{\geq 0}^{C_{v_1}} \times \mathbb{R}_{\geq 0}^{C_{v_2}} \times \cdots \times \mathbb{R}_{\geq 0}^{C_{v_d}} \to \mathcal{D}(C_r)$ as a function taking inputs $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_d)$ where $\mathbf{p}_i \in \mathbb{R}^{C_{v_i}}$ for $i \in [d]$. The Jacobian of f is a matrix $(\mathcal{J} f)(\mathbf{p}) \in \mathbb{R}^{C_r \times \bigcup_{i \in [d]} C_{v_i}}$. Since C_{v_i}'s are disjoint, for every $\tau \in \bigcup_{i \in [d]} C_{v_i}$, we will denote it by (i, τ) if $\tau \in C_{v_i}$ for clarity. Therefore,

$$(\mathcal{J} f)_{\pi,(i,\tau)}(\mathbf{p}) = \frac{\partial f_\pi}{\partial \mathbf{p}_i(\tau)}.$$

For each $i \in [d]$, define the matrix $\mathcal{J}_i \in \mathbb{R}^{C_r \times C_{v_i}}$ with entries

$$(\mathcal{J}_i f)_{\pi,\tau}(\mathbf{p}) = (\mathcal{J} f)_{\pi,(i,\tau)}(\mathbf{p})$$

for every $\pi \in C_r$ and $\tau \in C_{v_i}$.

We can write $\mathcal{J}_i f$ in a compact way.

Proposition 23. *Let $\mathbf{p}_r = f(\mathbf{p})$. Then*

$$(\mathcal{J}_i f)(\mathbf{p}) = \sum_{c \in \mathcal{L}(e_i)} \mathbf{a}_{i,c} \mathbf{b}_{i,c}^\top,$$

where $\mathbf{a}_{i,c^*} = f(\mathbf{p}) \odot \left[\mathbb{1}\left[\pi(e_i) = c^*\right] - \sum_{\pi' \in C_r : \pi'(e_i) = c^*} \mathbf{p}_r(\pi') \right]_{\pi \in C_r}$ *and* $\mathbf{b}_{i,c^*} = \left[\frac{\mathbb{1}[c^* \notin \tau]}{\sum_{\tau' \in C_{v_i} : c^* \notin \tau'} \mathbf{p}_i(\tau')} \right]_{\tau \in C_{v_i}}$. [2]

Proof. Let $q_{i,\rho} = \sum_{\tau \in C_{v_i} : \rho(e_i) \notin \tau} \mathbf{p}_i(\tau)$. For any $\pi \in C_r$, we write f_π as a function of $q_{i,\rho}$

$$f_\pi = \frac{\prod_i q_{i,\pi}}{\sum_{\rho \in C_r} \prod_i q_{i,\rho}}.$$

Then we can compute

$$\frac{\partial f_\pi}{\partial q_{i,\rho}} = \frac{1}{q_{i,\rho}} \left(\mathbb{1}\left[\rho = \pi\right] - f_\rho \right) f_\pi.$$

Therefore,

$$(\mathcal{J} f)_{\pi,(i,\tau)}(\mathbf{p}) = \frac{\partial f_\pi}{\partial \mathbf{p}_i(\tau)} = \sum_{\rho \in C_r} \frac{\partial f_\pi}{\partial q_{i,\rho}} \frac{\partial q_{i,\rho}}{\partial \mathbf{p}_i(\tau)} = \sum_{\rho \in C_r} \frac{1}{q_{i,\rho}} (\mathbb{1}\left[\rho = \pi\right] - f_\rho) f_\pi \cdot \mathbb{1}\left[\rho(e_i) \notin \tau\right].$$

We write $\mathcal{J}_i f$ explicitly:

$$(\mathcal{J}_i f)(\mathbf{p}) = \mathrm{diag}(f(\mathbf{p})) \cdot \sum_{\rho \in C_r} \frac{1}{q_{i,\rho}} \cdot \left[\mathbb{1}\left[\pi = \rho\right] - f_\rho(\mathbf{p}) \right]_{\pi \in C_r} \left[\mathbb{1}\left[\rho(e_i) \notin \tau\right] \right]_{\tau \in C_{v_i}}^\top.$$

[2] We write $\mathbf{u} \odot \mathbf{v}$ for their Hadamard product (entry-wise product).

Noting that $q_{i,\rho} = \mathbb{P}_{T_i}[\rho(e_i) \notin c(E_{T_i}(v_i))]$ only relies on the color $\rho(e_i)$, we have

$$(\mathcal{J}_i f)(\mathbf{p}) = \mathrm{diag}(f(\mathbf{p})) \sum_{c^* \in L(e_i)} \Big[\mathbb{1}[\pi(e_i) = c^*] - \sum_{\pi' \in C_r : \pi'(e_i) = c^*} \mathbf{p}_r(\pi)\Big]_{\pi \in C_r} \Big[\frac{\mathbb{1}[c^* \notin \tau]}{\sum_{\tau' \in C_{v_i} : c^* \notin \tau'} \mathbf{p}_i(\tau')}\Big]^{\top}_{\tau \in C_{v_i}}.$$

□

A well-known trick in the analysis of decay of correlation is to apply a potential function on the marginal recursion to amortize the contraction rate. Given an increasing potential function $\phi : [0,1] \to \mathbb{R}$, we define f^{ϕ} such that for any $\pi \in C_r$ and $\mathbf{m} \in \mathbb{R}^{C_{v_1} \times C_{v_2} \times \cdots \times C_{v_d}}$,

$$f^{\phi}_{\pi}(\mathbf{m}) = \phi\left(f_{\pi}\left(\left(\phi^{-1}(\mathbf{m}_1), \phi^{-1}(\mathbf{m}_2), \ldots, \phi^{-1}(\mathbf{m}_d)\right)\right)\right).$$

As a result, the Jacobian of f^{ϕ} can be obtained by the chain rule and the inverse function theorem as follows.

Proposition 24. *Given a smooth increasing function $\phi : [0,1] \to \mathbb{R}$ with derivative $\Phi = \phi'$, let $\mathbf{p} = \phi^{-1}(\mathbf{m})$. Then we have*

$$(\mathcal{J}_i f^{\phi})(\mathbf{m}) = \sum_{c} (\Phi(f(\mathbf{p})) \odot \mathbf{a}_{i,c})(\mathbf{b}_{i,c} \odot \Phi^{-1}(\mathbf{p}_i))^{\top}.$$

Taking $\Phi(x) = \frac{1}{\sqrt{x}}$, we have

$$(\mathcal{J}_i f^{\phi})(\mathbf{m}) = \sum_{c} \mathbf{a}^{\phi}_{i,c} (\mathbf{b}^{\phi}_{i,c})^{\top},$$

where $\mathbf{a}^{\phi}_{i,c^*} = \sqrt{f(\mathbf{p})} \odot \Big[\mathbb{1}[\pi(e_i) = c^*] - \sum_{\pi' \in C_r : \pi'(e_i) = c^*} \mathbf{p}_r(\pi')\Big]_{\pi \in C_r}$ and $\mathbf{b}^{\phi}_{i,c^*} = \Big[\frac{\mathbb{1}[c^* \notin \tau]\sqrt{\mathbf{p}_i(\tau)}}{\sum_{\tau' \in C_{v_i} : c^* \notin \tau'} \mathbf{p}_i(\tau')}\Big]_{\tau \in C_{v_i}}$.

5.1.2 Bounding $\left\|(\mathcal{J} f^{\phi})(\mathbf{p})\right\|_2$

In this section, we aim to derive an upper bound for the 2-norm of the Jacobian of the tree recursion. For brevity, we follow some notations which is defined in Section 3.3 of marginal probabilities w.r.t $\mathbf{p}_i$ and $\mathbf{p}_r$.

Also we introduce some notations of matrices used in later proof.

Definition 25. *For a distribution $\mathbf{p}$ over proper colorings on a broom $E_{T_v}(v) = \{e_1, \ldots, e_m\}$, we define $X_v = \{(i,c) : i \in [m], c \in \mathcal{L}(e_i)\}$ and its (local) covariance matrix $\mathsf{Cov}(\mathbf{p}) \in \mathbb{R}^{X_v \times X_v}$ with entries:*

$$\mathsf{Cov}(\mathbf{p})((i,c_1),(j,c_2)) = \sum_{\tau \in C_v : \tau(e_i) = c_1 \& \tau(e_j) = c_2} \mathbf{p}(\tau) - \left(\sum_{\tau \in C_v : \tau(e_i) = c_1} \mathbf{p}(\tau)\right)\left(\sum_{\tau \in C_v : \tau(e_j) = c_2} \mathbf{p}(\tau)\right).$$

Definition 26. *For a distribution over colorings $\mathbf{p}$ on a broom $E_{T_v}(v) = \{e_1, \ldots, e_m\}$, we define $X_v = \{(i,c) : i \in [m], c \in \mathcal{L}(e_i)\}$ and its diagonal matrix of mean vector $\Pi(\mathbf{p}) \in \mathbb{R}^{X_v \times X_v}$ as follows,*

$$\Pi(\mathbf{p}) = \mathrm{diag}\left\{\sum_{\tau \in C_v : \tau(e_i) = c} \mathbf{p}(\tau)\right\}_{(i,c) \in X_v}.$$

Now we define the notion of spectral independence on a broom.

Definition 27. *For any distribution over colorings $\mathbf{p}$ on a broom $E(v) = \{e_1, ..., e_m\}$, we say $\mathbf{p}$ is C-spectrally independent if it holds that*

$$\mathsf{Cov}(\mathbf{p}) \preceq C \cdot \Pi(\mathbf{p}).$$

The following is the main result in this section.

Condition 1 (marginal bound)**.** *For any $i \in [d]$, $\mathbf{p}_i$ is a distribution on C_{v_i} such that for any color a*

$$\mathbf{p}_i(a) \le \frac{|E_{T_i}(v_i)|}{\beta - 1 + |E_{T_i}(v_i)|},$$

and for $\mathbf{p}_r = (f_\pi(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_d))_{\pi \in C_v}$,

$$\frac{\mathbf{p}_r(i,a)}{\mathbf{p}_i(\bar{a})} \le \frac{1}{\beta - 1},$$

where $\beta \ge (1 + o(1))\Delta$.

Theorem 28. *For any $i \in [d]$, $\mathbf{p}_i = \phi^{-1}(\mathbf{m}_i)$ and $\mathbf{p}_r = f((\mathbf{p}_i)_{i \in [d]})$ satisfy Condition 1 and $(1+\eta)$-spectrally independent. Then $\beta \ge 1 + \frac{(1+\eta)\Delta}{\sqrt{1-2\delta}}$ implies that*

$$\left\| \mathcal{J} f^{\phi}(\mathbf{m}) \right\|_2 \le \frac{1 - \delta}{\sqrt{\Delta}}$$

where $\mathbf{m} = \phi(\mathbf{p})$ and $\phi(x) = 2\sqrt{x}$.

We will prove the theorem in Section 5.1.4 after introducing our key reduction in Section 5.1.3.

5.1.3 Dimension reduction

By the definition of 2-norm, we have that $\left\|(\mathcal{J} f^{\phi})(\mathbf{p})\right\|_2 = \sqrt{\lambda_{\max}((\mathcal{J} f^{\phi})(\mathbf{p})(\mathcal{J} f^{\phi})(\mathbf{p})^\top)}$. Let $\mathbf{A} := (\mathcal{J} f^{\phi})(\mathbf{p})(\mathcal{J} f^{\phi})(\mathbf{p})^\top$. We have that

$$\mathbf{A} = \sum_{i=1}^{d} (\mathcal{J}_i f^{\phi})(\mathbf{p})(\mathcal{J}_i f^{\phi})(\mathbf{p})^\top = \sum_{i=1}^{d} \sum_{c_1, c_2 \in \mathcal{L}(e_i)} \langle \mathbf{b}^{\phi}_{i,c_1}, \mathbf{b}^{\phi}_{i,c_2} \rangle \mathbf{a}^{\phi}_{i,c_1} (\mathbf{a}^{\phi}_{i,c_2})^\top.$$

The last equation simply follows from Proposition 24. The above calculation suggests that although the dimension of $\mathbf{A}$ is exponential in d, its rank is polynomial in d. In the following, we will find a much smaller matrix which can be used to upper bound $\mathbf{A}$. The idea is to use the trace method, namely to study $\mathrm{Tr}\left(\mathbf{A}^k\right)$. We have the following lemma.

Lemma 29. *For any positive semi-definite matrix $\mathbf{M} \in \mathbb{R}^{n \times n}$, we have that*

$$\lambda_{\max}(\mathbf{M}) = \lim_{k \to \infty} \left(\mathrm{Tr}(\mathbf{M}^k)\right)^{\frac{1}{k}}.$$

Proof of Lemma 29. Assume that $\lambda_1, \dots, \lambda_n$ are eigenvalues of $\mathbf{M}$ and $0 \le \lambda_1 \le \lambda_2 \le \dots \le \lambda_n$. $\mathbf{M}$ can be factored as $\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}$ where $\mathbf{\Lambda}$ is a diagonal matrix satisfying $\mathbf{\Lambda}(i,i) = \lambda_i$. Therefore,

$$\lim_{k\to\infty} \mathrm{Tr}\left(\mathbf{M}^k\right)^{\frac{1}{k}} = \lim_{k\to\infty} \mathrm{Tr}\left(\mathbf{Q}\mathbf{\Lambda}^k\mathbf{Q}^{-1}\right)^{\frac{1}{k}} = \lim_{k\to\infty} \mathrm{Tr}\left(\mathbf{\Lambda}^k\mathbf{Q}^{-1}\mathbf{Q}\right)^{\frac{1}{k}} = \lim_{k\to\infty} \left(\sum_{i=1}^{n} \lambda_i^k\right)^{\frac{1}{k}} = \lambda_n. \qquad \square$$

To simplify notations, we let

$$V(i, z_1, c_1, z_2, c_2) := \langle \mathbf{b}^{\phi}_{i,c_1}, \mathbf{b}^{\phi}_{i,c_2} \rangle \left(\mathbb{1}[z_1 = c_1] - \mathbf{p}_r(i, c_1)\right) \left(\mathbb{1}[z_2 = c_2] - \mathbf{p}_r(i, c_2)\right).$$

Then we can write $\mathbf{A}$ explicitly.

$$\mathbf{A}(\pi, \tau) = \sqrt{f(\mathbf{p})(\pi) f(\mathbf{p})(\tau)} \sum_{i=1}^{d} \sum_{c_1, c_2 \in \mathcal{L}(e_i)} V(i, \pi(e_i), c_1, \tau(e_i), c_2). \tag{6}$$

Let $g_k^{\pi}(i,c)$ denote

$$\sum_{\substack{\tau_1,\dots,\tau_{k-1}\in C_r\\ \tau_0=\pi}} \prod_{j=1}^{k-1} f(\mathbf{p})(\tau_j) \sum_{i_1,\dots,i_{k-1}\in[d],i_k=i} \sum_{\substack{c_{1,1}\in\mathcal{L}(e_{i_1})\\ \cdots\\ c_{1,k}\in\mathcal{L}(e_{i_k})}} \sum_{\substack{c_{2,1}\in\mathcal{L}(e_{i_1})\\ \cdots\\ c_{2,k-1}\in\mathcal{L}(e_{i_{k-1}})\\ c_{2,k}=c}} \prod_{j=1}^{k-1} V(i_j,\tau_{j-1}(i_j),c_{1,j},\tau_j(i_j),c_{2,j})$$

$$\times \langle \mathbf{b}^{\phi}_{i,c_{1,k}}, \mathbf{b}^{\phi}_{i,c}\rangle (\mathbb{1}\left[\tau_{k-1}(i)=c_{1,k}\right] - \mathbf{p}_r(i,c_{1,k})).$$

We omit π in g_k^{π} for brevity. Then we have that for any $\pi\in C_r$

$$\mathbf{A}^k(\pi,\pi) = f(\mathbf{p})(\pi)\sum_{i=1}^{d}\sum_{c\in\mathcal{L}(e_i)} g_k(i,c)(\mathbb{1}\left[\pi(i)=c\right]-\mathbf{p}_r(i,c)). \tag{7}$$

Fix π, then we will show that $\{g_k\}_{k\geq 1}$ can be computed recursively, which gives a simple representation of $\mathbf{A}^k(\pi,\pi)$. Let $X=\{(i,c)|i\in[d],c\in\mathcal{L}(e_i)\}$ be the set of all feasible edge-color pairs.

Lemma 30. *If $\mathbf{B}(\mathbf{p})\in\mathbb{R}^{X\times X}$ satisfies that*

$$\mathbf{B}(\mathbf{p})((i,c_2),(j,c_4)) = \sum_{c_3\in\mathcal{L}(e_j)} \frac{\mathbf{p}_j(\bar{c_3},\bar{c_4})}{\mathbf{p}_j(\bar{c_3})\mathbf{p}_j(\bar{c_4})} \times (\mathbf{p}_r(j,c_3,i,c_2)-\mathbf{p}_r(j,c_3)\mathbf{p}_r(i,c_2)).$$

Then we have that $g_k^{\top}=\alpha_\pi^{\top}\mathbf{B}^{k-1}$ where

$$\alpha_\pi(i,c_2) = \sum_{c_1\in\mathcal{L}(e_i)} \frac{\mathbf{p}_i(\bar{c_1},\bar{c_2})}{\mathbf{p}_i(\bar{c_1})\mathbf{p}_i(\bar{c_2})} \times (\mathbb{1}\left[\pi(i)=c_1\right]-\mathbf{p}_r(i,c_1)).$$

Proof of Lemma 30. For any $k>1$, we expand one layer of summation and get

$$\begin{aligned} g_k(i,c) &= \sum_{\tau_{k-1}\in C_r} f(\mathbf{p})(\tau_{k-1}) \sum_{i_{k-1}\in[d]} \sum_{c_{1,k}\in\mathcal{L}(e_i)} \sum_{c_{2,k-1}\in\mathcal{L}(e_{i_{k-1}})} g_{k-1}(i_{k-1},c_{2,k-1}) \\ &\quad \times \langle \mathbf{b}^{\phi}_{i,c_{1,k}},\mathbf{b}^{\phi}_{i,c}\rangle(\mathbb{1}\left[\tau_{k-1}(i)=c_{1,k}\right]-\mathbf{p}_r(i,c_{1,k}))(\mathbb{1}\left[\tau_{k-1}(i_{k-1})=c_{2,k-1}\right]-\mathbf{p}_r(i_{k-1},c_{2,k-1})) \\ &= \sum_{c_1\in\mathcal{L}(e_i)} \langle \mathbf{b}^{\phi}_{i,c_1},\mathbf{b}^{\phi}_{i,c}\rangle \sum_{j\in[d]} \sum_{c_{2,k-1}\in\mathcal{L}(e_j)} (\mathbf{p}_r(i,c_1,j,c_{2,k-1})-\mathbf{p}_r(i,c_1)\mathbf{p}_r(j,c_{2,k-1}))g_{k-1}(j,c_{2,k-1}). \end{aligned}$$

Recall that

$$\langle \mathbf{b}^{\phi}_{i,c_1},\mathbf{b}^{\phi}_{i,c}\rangle = \frac{\mathbf{p}_i(\bar{c_1},\bar{c})}{\mathbf{p}_i(\bar{c_1})\mathbf{p}_i(\bar{c})},$$

which indicates that $g_k^{\top}=g_{k-1}^{\top}\mathbf{B}(\mathbf{p})$. Now It is sufficient to prove that $g_1=\alpha_\pi$. Straight calculation shows that

$$g_1(i,c) = \sum_{c_1\in\mathcal{L}(e_i)} \langle \mathbf{b}^{\phi}_{i,c_1},\mathbf{b}^{\phi}_{i,c}\rangle(\mathbb{1}\left[\pi(i)=c_1\right]-\mathbf{p}_r(i,c_1)) = \alpha_\pi(i,c_2).$$ ☐

In the following, we omit $(\mathbf{p})$ in $\mathbf{B}(\mathbf{p})$ for brevity if there is no ambiguity. Lemma 30 directly indicates that we can use the 2-norm of $\mathbf{B}$ to upper bound that of $\mathbf{A}$.

Lemma 31. *Let $\mathbf{B}\in\mathbb{R}^{X\times X}$ and α_π denote the matrix and the vector defined in Lemma 30. Then we have that for any $k\geq 1$,*

$$\sum_{\pi\in C_r}\mathbf{A}^k(\pi,\pi) = \sum_{\pi\in C_r} f(\mathbf{p})(\pi)\alpha_\pi^{\top}\mathbf{B}^{k-1}\beta_\pi \tag{8}$$

where $\beta_\pi(j,c_4)=\mathbb{1}\left[\pi(j)=c_4\right]-\mathbf{p}_r(j,c_4)$, implying that $\lambda_{\max}(\mathbf{A})\leq\|\mathbf{B}\|_2$.

Proof of Lemma 31. Equation (8) immediately follows from Equation (7) and Lemma 30. Therefore, the maximum eigenvalue of $\mathbf{A}$ can be expressed as follows.

$$\begin{aligned}\lambda_{\max}(\mathbf{A}) &= \lim_{k\to\infty}\left(\sum_{\pi\in C_r} f(\mathbf{p})(\pi)\alpha_\pi^\top \mathbf{B}^{k-1}\beta_\pi\right)^{\frac{1}{k}}\\ &\le \lim_{k\to\infty}\left(\sum_{\pi\in C_r} f(\mathbf{p})(\pi)\left\|\alpha_\pi\right\|_2\left\|\mathbf{B}^{k-1}\beta_\pi\right\|_2\right)^{\frac{1}{k}}\\ &\le \lim_{k\to\infty}\left(\sum_{\pi\in C_r} f(\mathbf{p})(\pi)\left\|\alpha_\pi\right\|_2\left\|\beta_\pi\right\|_2\right)^{\frac{1}{k}}\left\|\mathbf{B}\right\|_2^{\frac{k-1}{k}}\\ &= \left\|\mathbf{B}\right\|_2.\end{aligned}$$

□

5.1.4 Bound the transition matrix

Let $D_T \in \mathbb{R}^{X\times X}$ be the diagonal matrix where $D_T((i,c),(i,c)) = \mathbf{p}_i(\bar{c})$. In this section, we give an upper bound for $\|\mathbf{B}\|_2$ as $\mathbf{B}$ can be represented as the product of covariance matrices of $\mathbf{p}_r$, $\mathbf{p}_i$ and some auxiliary diagonal matrices.

Proposition 32. *Let $C_i \in \mathbb{R}^{X_{v_i}\times|\mathcal{L}(e_i)|}$ denote the matrix satisfying that $C_i((j,c_1),c_2) = \mathbb{1}[c_1 = c_2]$ for any $i \in [d]$. Let $\mathbf{R} \in \mathbb{R}^{X\times X}$ denote*

$$\operatorname{diag}\{C_i^\top \mathsf{Cov}(\mathbf{p}_i)C_i\}_{i\in[d]}.$$

Then we have that

$$\mathbf{B} = \mathsf{Cov}(\mathbf{p}_r)D_T^{-1}\mathbf{R}D_T^{-1}.$$

Proof of Proposition 32. Note that for any $i \in [d]$ and $c_1, c_2 \in \mathcal{L}(e_i)$,

$$\begin{aligned}\mathbf{R}((i,c_1),(i,c_2)) &= C_i^\top \mathsf{Cov}(\mathbf{p}_i)C_i(c_1,c_2)\\ &= \mathbf{p}_i(\bar{c_1},\bar{c_2}) - \mathbf{p}_i(\bar{c_1})\mathbf{p}_i(\bar{c_2}).\end{aligned}$$

Therefore, for any $c_2 \in \mathcal{L}(e_i)$ and $c_4 \in \mathcal{L}(e_j)$, the following always holds.

$$\begin{aligned}\mathbf{B}((i,c_2),(j,c_4)) &= \sum_{c_3\in\mathcal{L}(e_j)}(\mathbf{p}_r(j,c_3,i,c_2) - \mathbf{p}_r(j,c_3)\mathbf{p}_r(i,c_2))\times\frac{\mathbf{p}_i(\bar{c_3},\bar{c_4})}{\mathbf{p}_i(\bar{c_3})\mathbf{p}_i(\bar{c_4})}\\ &= \sum_{c_3\in\mathcal{L}(e_j)}\mathsf{Cov}(\mathbf{p}_r)((i,c_2),(j,c_3))\left(\frac{\mathbf{p}_i(\bar{c_3},\bar{c_4})}{\mathbf{p}_i(\bar{c_3})\mathbf{p}_i(\bar{c_4})} - 1\right)\\ &= \sum_{(k,c_3):c_3\in\mathcal{L}(e_k)}\mathsf{Cov}(\mathbf{p}_r)((i,c_2),(k,c_3))\frac{1}{\mathbf{p}_k(\bar{c_3})}\mathbf{R}((k,c_3),(j,c_4))\frac{1}{\mathbf{p}_k(\bar{c_4})}.\end{aligned}$$

where the second equality comes from $\sum_{c_3\in\mathcal{L}(e_j)}(\mathbf{p}_r(j,c_3,i,c_2) - \mathbf{p}_r(j,c_3)\mathbf{p}_r(i,c_2)) = 0$. □

Then we can establish Theorem 28 through the above conclusions.

Proof of Theorem 28. The Loewner Order still holds under the Congruent transformation. Therefore, by spectral independence of $\mathbf{p}_i$, we have that

$$\mathbf{R} \preceq (1+\eta)\operatorname{diag}\{C_i^\top \Pi(\mathbf{p}_i)C_i\}_{i\in[d]} = (1+\eta)(I - D_T). \tag{9}$$

Plugging in $\mathrm{Cov}(\mathbf{p}_r) \le (1+\eta)\Pi(\mathbf{p}_r)$ and Equation (9), we get

$$\begin{aligned}\|\mathbf{B}\|_2 &\le (1+\eta)^2 \lambda_{\max}(\Pi(\mathbf{p}_r)D_T^{-1}(I-D_T)D_T^{-1}) \\ &= (1+\eta)^2 \max_{(i,c):c\in\mathcal{L}(e_i)} \frac{\mathbf{p}_r(i,c)\mathbf{p}_i(c)}{\mathbf{p}_i(\bar{c})^2}. \end{aligned} \tag{10}$$

Applying marginal bounds for Condition 1, we have that

$$\begin{aligned}\|\mathbf{B}\|_2 &\le (1+\eta)^2 \max_{(i,c):c\in\mathcal{L}(e_i)} \frac{\mathbf{p}_i(c)}{\mathbf{p}_i(\bar{c})(\beta-1)} \\ &\le (1+\eta)^2 \max_{i\in[d]} \frac{\deg(v_i)-1}{(\beta-1)^2} \\ &\le \frac{1-2\delta}{\Delta}\end{aligned}$$

for some $\delta > 0$. The last inequality follows from $\beta \ge 1 + \frac{(1+\eta)\Delta}{\sqrt{1-2\delta}}$. By Lemma 31, $\left\|(\mathcal{J}f^{\phi})(\mathbf{p})\right\|^2 \le \|\mathbf{B}\|_2$. Therefore,

$$\left\|(\mathcal{J}f^{\phi})(\mathbf{p})\right\|_2 \le \sqrt{\|\mathbf{B}\|_2} \le \frac{1-\delta}{\sqrt{\Delta}}. \qquad \square$$

5.2 Strong spatial mixing via contraction

As Theorem 28 gives the upper bounds on the 2-norm of the Jacobian matrix, we now proceed to demonstrate how these bounds can be used to prove strong spatial mixing via contraction. Specifically, we will quantify the decay of correlations using the derived bounds. Let $B_G(u,d)$ denote $\{u' \in V \mid \mathrm{dist}_G(u,u') = d\}$.

Proof of Theorem 22. Fix $e_r = (u,r) \in E$ and r is the root of the tree. Let τ_1 and τ_2 be two different feasible pinnings on $\Lambda \subseteq E \setminus \{e_r\}$. We use $(T',\mathcal{L}')$ to denote the edge coloring instance which is obtained by removing every $e \in \Lambda \setminus \partial_{\tau_1,\tau_2}$ from T and removing $\tau_1(e)$ from the lists of the neighbours of e. It is easy to verify that $(T',\mathcal{L}')$ is still a β-extra edge coloring instance.

Let $\ell := \min_{e\in\partial_{\tau_1,\tau_2}} \mathrm{dist}_{T'}(e_r,e) - 1$. It is trivial if $\ell = \infty$. Without loss of generality, assume $\ell \ge 3$. Let $\mathbf{p}_\ell$ and $\mathbf{p}'_\ell$ denote the marginal distribution over $\bigcup_{u\in B_{T'}(r,\ell)} E_{T'_u}(u)$ under the pinning τ_1 and τ_2 respectively. Let $f^{\phi,i\to i-1}$ and $f^{i\to i-1}$ denote the concatenation of recursive function f^{ϕ} and f of subtrees which is rooted at $B_{T'}(r,i-1)$ respectively. For simplicity, let $f^{\phi,i\to j} := f^{\phi,j+1\to j} \circ f^{\phi,j+2\to j+1} \circ \cdots \circ f^{\phi,i\to i-1}$ for any $i > j$ and it is the same for $f^{i\to j}$. We use $\mathbf{m}_\ell(t)$ denote the linear combination of $\phi(\mathbf{p}_\ell)$ and $\phi(\mathbf{p}'_\ell)$, that is, $\mathbf{m}_\ell(t) = t\phi(\mathbf{p}_\ell) + (1-t)\phi(\mathbf{p}'_\ell)$. Then we have that

$$\begin{aligned}\left\|f^{\phi,\ell\to 0}(\mathbf{m}_\ell(0)) - f^{\phi,\ell\to 0}(\mathbf{m}_\ell(1))\right\|_2 &= \left\|\int_0^1 (\mathcal{J}f^{\phi,\ell\to 0}(\mathbf{m}_\ell(t)))\cdot(\phi(\mathbf{p}_\ell)-\phi(\mathbf{p}'_\ell))\,\mathrm{d}t\right\|_2 \\ &\le \int_0^1 \left\|(\mathcal{J}f^{\phi,\ell\to 0}(\mathbf{m}_\ell(t)))\cdot(\phi(\mathbf{p}_\ell)-\phi(\mathbf{p}'_\ell))\right\|_2 \mathrm{d}t \\ &\le \max_{t\in[0,1]} \left\|(\mathcal{J}f^{\phi,\ell\to 0}(\mathbf{m}_\ell(t)))\right\|_2 \left\|(\phi(\mathbf{p}_\ell)-\phi(\mathbf{p}'_\ell))\right\|_2 \\ &\le \max_{t\in[0,1]} \left\|(\mathcal{J}f^{\phi,\ell\to 0}(\mathbf{m}_\ell(t)))\right\|_2 \max_{u\in B_{T'}(r,\ell-1)} \Delta^{\frac{\ell}{2}} \left\|\phi(\mathbf{p}_u)-\phi(\mathbf{p}'_u)\right\|_2\end{aligned} \tag{11}$$

where $\mathbf{p}_u$ and $\mathbf{p}'_u$ are the marginal distributions over $E_{T'_u}(u)$ on T'_u under the pinning τ_1 and τ_2 respectively. As $\phi(x) = 2\sqrt{x}$, $\|\phi(\mathbf{p}_u)-\phi(\mathbf{p}'_u)\|_2 \le \|\phi(\mathbf{p}'_u)\|_2 + \|\phi(\mathbf{p}_u)\|_2 = 4$. And we have that for any $t \in [0,1]$,

$$\begin{aligned}\left\|(\mathcal{J} f^{\phi,\ell\to 0}(\mathbf{m}_\ell(t)))\right\|_2 &= \left\|(\mathcal{J} f^{\phi,1\to 0}(\mathbf{m}_1(t)))\dots(\mathcal{J} f^{\phi,\ell\to\ell-1}(\mathbf{m}_\ell(t)))\right\|_2\\ &\le \prod_{i=1}^{\ell}\left\|\mathcal{J} f^{\phi,i\to i-1}(\mathbf{m}_i(t))\right\|_2\\ &= \prod_{i=1}^{\ell}\sup_{u\in B_{T'}(r,i-1)}\left\|\mathcal{J} f^{\phi}(\mathbf{m}_u(t))\right\|_2\\ &\le 16\Delta q\left(\sup_{u\in B_{T'}(r,d),d\le\ell-3}\left\|\mathcal{J} f^{\phi}(\mathbf{m}_u(t))\right\|_2\right)^{\ell-2} \end{aligned} \tag{12}$$

where $\mathbf{m}_i(t) := f^{\phi,\ell\to i}(\mathbf{m}_\ell(t)) = \phi(f^{\ell\to i}(\phi^{-1}(\mathbf{m}_\ell(t))))$ and $\mathbf{m}_i(t) = (\mathbf{m}_u(t))_{u\in B_{T'}(r,i-1)}$. Here Equation (12) follows from the following claim.

Claim 33. *For any $t \in [0,1]$, we have that*

1. $\left\|\mathcal{J} f^{\phi}(\mathbf{m}_u(t))\right\|_2 \le 4\sqrt{2\Delta q}$ *holds for* $u \in B_{T'}(r,\ell-1)$.
2. $\left\|\mathcal{J} f^{\phi}(\mathbf{m}_u(t))\right\|_2 \le 2\sqrt{2\Delta q}$ *holds for* $u \in B_{T'}(r,\ell-2)$.

Proof of Claim 33. Fix $u \in B_{T'}(r,\ell-1)$. By the concavity of ϕ, we have that for any $\tau \in C_u$ and $t \in [0,1]$,

$$\phi^{-1}(\mathbf{m}_u(t))(\tau) = \phi^{-1}(t\phi(\mathbf{p}_u(\tau)) + (1-t)\phi(\mathbf{p}'_u(\tau))) \le t\mathbf{p}_u(\tau) + (1-t)\mathbf{p}'_u(\tau);$$

$$\phi^{-1}(\mathbf{m}_u(t))(\tau) \ge t^2\mathbf{p}_u(\tau) + (1-t)^2\mathbf{p}'_u(\tau).$$

Since $\ell = \min_{e\in\partial_{\tau_1,\tau_2}} \mathrm{dist}_{T'}(e_r, e) - 1$, by Lemma 12 and Lemma 13, $\mathbf{p}_\ell$ and $\mathbf{p}'_\ell$ satisfy Condition 1, thus we have that

$$\left\|\mathcal{J} f^{\phi}(\mathbf{m}_u(t))\right\|_2^2 \le \left\|\mathbf{B}(\phi^{-1}(\mathbf{m}_u(t)))\right\|_1 \le \Delta q \times \frac{1}{\left(\frac{1}{2}\frac{\beta-1}{\beta-1+\Delta}\right)^2} \times 2 = 32\Delta q.$$

For any $u \in B_{T'}(r,\ell-2)$, by Lemma 12 and Lemma 13, $\phi^{-1}(\mathbf{m}_u(t))$ satisfies Condition 1. Therefore,

$$\left\|\mathcal{J} f^{\phi}(\mathbf{m}_u(t))\right\|_2^2 \le \left\|\mathbf{B}(\phi^{-1}(\mathbf{m}_u(t)))\right\|_1 \le \Delta q \times \frac{1}{\left(\frac{\beta-1}{\beta-1+\Delta}\right)^2} \times 2 = 8\Delta q. \qquad \square$$

The second equation follows from $\mathcal{J} f^{\phi,i\to i-1}(\mathbf{m}_i(t)) = \mathrm{diag}\{\mathcal{J} f^{\phi}(\mathbf{m}_u(t))\}_{u\in B_{T'}(r,i-1)}$. By Lemma 12, Lemma 13 and Lemma 36, for any $u \in B_{T'}(r,d)$ and $d \le \ell-3$, $\mathbf{m}_u(t)$ satisfies Condition 1 and is $(1+\eta_\Delta)$-spectrally independent. Plugging in Theorem 28 and Equation (11), we have that

$$\begin{aligned}\left\|\mu_{e_r}^{\tau_1} - \mu_{e_r}^{\tau_2}\right\|_{\mathrm{TV}} &\le \left\|\mu_{E_T(r)}^{\tau_1} - \mu_{E_T(r)}^{\tau_2}\right\|_{\mathrm{TV}}\\ &= \frac{1}{2}\left\|f^{\phi,\ell\to 0}(\mathbf{m}_\ell(0)) - f^{\phi,\ell\to 0}(\mathbf{m}_\ell(1))\right\|_1\\ &\le \frac{\sqrt{q^\Delta}}{2}\left\|f^{\phi,\ell\to 0}(\mathbf{m}_\ell(0)) - f^{\phi,\ell\to 0}(\mathbf{m}_\ell(1))\right\|_2\\ &\le 32q^{\frac{\Delta+2}{2}}\Delta^2(1-\delta)^{\ell-2}.\end{aligned}$$

We pick $C = \max\{32q^{\frac{\Delta+2}{2}}\Delta^2(1-\delta)^{-3}, (1-\delta)^{-4}\}$ to finish the proof. $\square$

5.3 Worst-case scenario

Though Theorem 28 establishes an upper bound on the 2-norm of the Jacobian matrix under certain conditions, in this section, we will introduce the "worst" pinning of q-edge coloring, which is the bottleneck of our analysis. In fact, with potential function $\phi(x) = 2\sqrt{x}$ and applying 2-norm for the correlation decay step, the best bound we expected to prove is $q > \left(\frac{3+\sqrt{5}}{2} + o(1)\right)\Delta \approx 2.618\Delta$. However what we can only prove strong spatial mixing for instances of $(1+o(1))\Delta$-extra edge colorings, as we currently lack a better upper bound for $\mathbf{R}$. Note that the pinning that maximizes eq. (10) is the same as the pinning in Theorem 35 and eq. (10) can indeed achieve the upper bound $\frac{\Delta-1}{(q-2\Delta+2)^2}$ under this worst pinning.

Before showing the worst-case scenario, we introduce the following technical lemma to calculate the eigenvalues of some simple matrices.

Lemma 34. *Given constant $k_1, k_2 \neq 0$, the eigenvalues of $k_1\mathbf{1}\mathbf{1}^\top + k_2\mathsf{Id}_n$ are either k_2 or $nk_1 + k_2$ where Id_n is the identity matrix in $\mathbb{R}^{n\times n}$.*

Proof of Lemma 34. The eigenvalues λ and the corresponding eigenvectors $\mathbf{v}$ satisfies that

$$(k_1\mathbf{1}\mathbf{1}^\top + k_2\mathsf{Id})\mathbf{v} = \lambda\mathbf{v} \implies k_1\langle \mathbf{1}, \mathbf{v}\rangle \mathbf{1} = (\lambda - k_2)\mathbf{v}. \tag{13}$$

Therefore, there are only two cases for Equation (13).

$$\begin{cases} \lambda = k_2 \wedge \mathbf{1} \perp \mathbf{v} \\ \mathbf{v} \parallel \mathbf{1} \end{cases}.$$

Plugging in $\mathbf{v} = \mathbf{1}$ and Equation (13), we get

$$k_1\langle \mathbf{1}, \mathbf{1}\rangle = \lambda - k_2 \implies \lambda = k_2 + nk_1. \qquad \square$$

Now we show the worst pinning and corresponding norm value of $\mathbf{B}$ and $(\mathcal{J}f^\phi)(\mathbf{p})$.

Theorem 35. *Under some specific pinning, the norm of $\mathbf{B}$ satisfies lower bound that*

$$\|(\mathcal{J}f^\phi)(\mathbf{p})\|_2^2 = \|\mathbf{B}\|_2 = \frac{\Delta - 1}{(q-\Delta)(q-2\Delta+2)}.$$

Then $\|\mathbf{B}\|_2 < \frac{1}{\Delta}$ implies that $q > \left(\frac{3+\sqrt{5}}{2} + o(1)\right)\Delta$.

Proof of Theorem 35. Consider the following instance of edge coloring $(G, \mathcal{L})$ generated from a q-coloring instance by pinning:

1. $\deg(r) = 1$, $\mathcal{L}(e_1) = \{\Delta, \dots, q\}$, that is, in original instance r has Δ children and we pin $e_2, \dots, e_\Delta$ with $1, 2, \dots, \Delta - 1$.
2. $\deg(v_1) = \Delta$ and for any $u \in N(v_1)$ and $u \neq r$, $\mathcal{L}(\{u, v_1\}) = \{\Delta, \dots, q\}$, that is, we assume u has $\Delta - 1$ children and we pin the children of u with color $1, 2, \dots, \Delta - 1$.

Then by Proposition 32, $\mathbf{B}$ is equivalent to a matrix in $\mathbb{R}^{(q-\Delta+1)\times(q-\Delta+1)}$ and holds for the following equation

$$\mathbf{B} = -\frac{\Delta - 1}{(q-\Delta+1)(q-\Delta)(q-2\Delta+2)}\left(\mathbf{1}\mathbf{1}^\top - \mathsf{Id}\right) + \frac{\Delta - 1}{(q-\Delta+1)(q-2\Delta+1)}\mathsf{Id}.$$

$\mathbf{B}$ is a symmetrical matrix, thus the singular values equals to the eigenvalues. Applying Lemma 34, we have that

$$\|\mathbf{B}\|_2 = \lambda_{\max}(\mathbf{B}) = \frac{\Delta - 1}{(q-\Delta)(q-2\Delta+2)}.$$

For $(\mathcal{J}f^{\phi})(\mathbf{p})$, since $|C_r| = |\mathcal{L}(e_1)| = q - \Delta + 1$, $(\mathcal{J}f^{\phi})(\mathbf{p})(\mathcal{J}f^{\phi})^{\top}(\mathbf{p}) = \mathbf{A} \in \mathbb{R}^{(q-\Delta+1)\times(q-\Delta+1)}$. Now it is sufficient to show that $\mathbf{A} = \mathbf{B}$. By Equation (6), we have that

$$
\begin{aligned}
\mathbf{A}(c_1, c_2) &= \frac{1}{q-\Delta+1} \sum_{c_3, c_4 \in \mathcal{L}(e_1)} \langle \mathbf{b}^{\phi}_{1,c_3}, \mathbf{b}^{\phi}_{1,c_4} \rangle \left(\mathbb{1}[c_1 = c_3] - \frac{1}{q-\Delta+1} \right) \left(\mathbb{1}[c_2 = c_4] - \frac{1}{q-\Delta+1} \right) \\
&= \begin{cases} \frac{\Delta-1}{(q-\Delta+1)(q-2\Delta+2)} & , c_1 = c_2 \\ \frac{1-\Delta}{(q-\Delta+1)(q-\Delta)(q-2\Delta+2)} & , c_1 \neq c_2 \end{cases} \\
&= \mathbf{B}(c_1, c_2).
\end{aligned}
$$

□

Theorem 35 indicates the limitations of our current analysis by providing a lower bound on the norm of the Jacobian matrix. This indicates the best bound one can expect to use 2-norm and the potential function $\phi(x) = 2\sqrt{x}$ is $q \approx 2.618\Delta$.

6 Covariance matrix on brooms

We will give an upper bound for the covariance matrix defined in Definition 25 in this section.

Lemma 36. *Given a tree T rooted with r of depth $\ell + 1$ and weight functions $\mathbf{p}_v : E_{T_v}(v) \to \mathbb{R}_{\geq 0}$ for any $v \in B_T(r, \ell)$. If $\ell \geq 2$ and $\beta \geq \Delta + 50$, then $\mathbf{p}_r := f^{\ell \to 0}((\mathbf{p}_v)_{v \in B(r,\ell)})$ is $(1 + \eta_\Delta)$-spectrally independent, i.e.*

$$\mathsf{Cov}(\mathbf{p}_r) \preceq (1 + \eta_\Delta)\Pi(\mathbf{p}_r).$$

where η_Δ is $O\left(\frac{\log^2 \Delta}{\Delta}\right)$.

6.1 Weighted edge coloring instance

Given a tree T and associated β-extra color lists $\mathcal{L}$ for all edges in T, an weighted edge coloring instance is a distribution over proper list colorings $(T, \mathcal{L})$. There are non-intersecting connected subgraphs $\{K_i\}_{i \in [N]}$ in T called *boundaries* and associated *weight functions* $\{w_i\}_{i \in [N]}$ such that $w_i : C_{v_i} \to \mathbb{R}_{\geq 0}$.

Definition 37 (Distribution of weighted list edge coloring)**.** *Given a list edge coloring instance $(T, \mathcal{L})$, boundaries $\{K_i\}_{i \in [N]}$ and weighted functions $\{w_i\}_{i \in [N]}$ as above, define the associated weighted edge coloring instance be the distribution μ on Ω such that*

$$\mu(\sigma) := \frac{1}{Z} \prod_{i \in [N]} w_i(\sigma|_{K_i}),$$

where $Z := \sum_{\sigma \in \Omega} \prod_{i \in [N]} w_i(\sigma|_{K_i})$ is the partition function.

Notice that, if we pick an $\omega_i \in \Omega_{K_i}$ for each $i \in [N]$, then

$$\mu^{\omega_1 \cup \cdots \cup \omega_N}(\sigma) \propto \begin{cases} \frac{1}{Z} \prod_{i \in [N]} w_i(\sigma|_{K_i}), & \forall i \in [N], \sigma|_{K_i} = \omega_i, \\ 0, & \text{otherwise}, \end{cases}$$

which means μ is a distribution over uniform edge colorings conditioned on a pinning on all K_i. This leads to the following lemma.

Lemma 38. *We have*

$$\mu = \sum_{\omega_1 \in \Omega_{K_1}, \cdots, \omega_N \in \Omega_{K_N}} \mu_{K_1 \cup \cdots \cup K_N}(\omega_1 \cup \cdots \cup \omega_N) \mu^{\omega_1 \cup \cdots \cup \omega_N}.$$

That is, the distribution on weighted coloring instance in Definition 37 can be expressed as a mixture of uniform distributions over edge colorings.

Proof. This is just an application of the total probability formula. □

With Lemma 38, we can prove weighted edge colorings inherit the marginal bound Lemma 9 on non-boundary edges.

Lemma 39 (Marginal upper bound - weighted version). *Consider a weighted edge coloring instance on a tree $T = (V, E)$ as defined in Definition 37, whose distribution is denoted by μ, and a pinning $\xi \in \Omega_S$ for some subset $S \subset E$. Then for any $i \in V$, $F \subseteq E_i \setminus \left(\bigcup_{i \in [N]} K_i \cup S\right)$ and color a: denoting $\beta := \min_{e \in F} \left\{\left|\mathcal{L}^\xi(e)\right| - \deg(e)\right\}$, we have*

$$\mu^\xi(a \in c(F)) \leq \frac{|F|}{\beta - 1 + |F|}.$$

Proof. We write μ^ξ into a mixture of uniform distributions on edge colorings by Lemma 38:

$$\mu^\xi = \sum_{\omega_1 \in \Omega_{K_1}, \cdots, \omega_N \in \Omega_{K_N}} \mu^\xi_{K_1 \cup \cdots \cup K_N}(\omega_1 \cup \cdots \cup \omega_N) \mu^{\xi \cup \omega_1 \cup \cdots \cup \omega_N}.$$

It suffices to prove the bound for any $\mu^{\xi \cup \omega_1 \cup \cdots \cup \omega_N}$. Notice that $\mu^{\xi \cup \omega_1 \cup \cdots \cup \omega_N}$ are uniform distributions on proper colorings defined on $T \setminus \left(\bigcup_{i \in [N]} K_i \cup S\right)$, with β-extra colors on all edges in F, so Lemma 9 applies and proves the lemma. □

We also need the following lower bound in the matrix trickle-down.

Lemma 40. *Consider a weighted edge coloring instance on a tree T with β-extra color lists $\mathcal{L}$. If $K = E(v)$ for some vertex v in T and none of the neighbours of K is in the boundary, $\tau \in \Omega_S$ is a pinning on some subset $S \subset E$, and the number of unpinned edges in K is k. Then for any unpinned $e \in K$ and $c \in \mathcal{L}^\tau(e)$,*

$$\mu^\tau_e(c) \geq \frac{(\beta - 1)^2}{(\beta + k - 2)(\beta + \Delta - 2)} \frac{1}{\ell^\tau_e - k + 1}.$$

where $\ell^\tau_e := |\mathcal{L}^\tau(e)|$.

Proof. We assume $e = \{u, v\}$, where $K = E(v)$ and $L := E(v)$. By Lemma 38, it suffices to prove the bound for any $\mu^{\sigma \cup \omega_1 \cup \cdots \cup \omega_N}$. Denoting ξ as the shortcut for $\sigma \cup \omega_1 \cup \cdots \cup \omega_N$, we can do a further decomposition:

$$\mu^\xi_e(c) = \sum_{\substack{\sigma_1 \in \Omega^\xi_{K \setminus \{e\}} \\ c \notin \sigma_1}} \mu^\xi_{K \setminus \{e\}}(\sigma_1) \cdot \sum_{\substack{\sigma_2 \in \Omega^{\xi \cup \sigma_1}_{L \setminus \{e\}} \\ c \notin \sigma_2}} \mu^{\xi \cup \sigma_1}_{L \setminus \{e\}}(\sigma_2) \cdot \mu^{\xi \cup \sigma_1 \cup \sigma_2}_e(c).$$

By assumption, K and L are not in the boundary, then Lemma 39 gives

$$\sum_{\substack{\sigma_1 \in \Omega^\xi_{K \setminus \{e\}} \\ c \notin \sigma_1}} \mu^\xi_{K \setminus \{e\}}(\sigma_1) \geq \frac{\beta - 1}{\beta + k - 2}$$

$$\sum_{\substack{\sigma_2 \in \Omega^{\xi \cup \sigma_1}_{L \setminus \{e\}} \\ c \notin \sigma_2}} \mu^{\xi \cup \sigma_1}_{L \setminus \{e\}}(\sigma_2) \geq \frac{\beta - 1}{\beta + \Delta - 2}.$$

Moreover, since $\mu_e^{\xi\cup\sigma_1\cup\sigma_2}$ is the uniform distribution over $\mathcal{L}^{\xi\cup\sigma_1\cup\sigma_2}(e)$, whose size is at most $\ell_e^\tau - k + 1$, it is lower bounded by $1/(\ell_e^\tau - k + 1)$.

Combining the three bounds, we have

$$\mu_e^\tau(c) \geq \frac{(\beta-1)^2}{(\beta+k-2)(\beta+\Delta-2)} \frac{1}{\ell_e^\tau - k + 1}. \qquad \square$$

After introducing the concept of weighted edge coloring instances, we now turn our attention to the marginal distribution on a broom. For simplicity we run the matrix trickle down theorem on one broom K in T, that does not intersect with the boundaries (actually on the quotient simplicial complex on K). Then it is necessary to look at the marginal distribution of μ on K under some pinning $\xi \in \Omega_F$ on a subset of edges F. We have

$$\mu_K^\xi(\tau) \propto \sum_{\substack{\sigma:\sigma|_K=\tau \\ \sigma|_F=\xi}} \prod_{i\in[N]} w_i(\sigma|_{K_i}). \tag{14}$$

The following lemma demonstrates the relation between the weighted coloring instance and the tree recursion.

Lemma 41. *Assume the tree T with root r is of depth $\ell+1$. Let the boundaries be all brooms on $\ell+1$-th level. For any $v \in B(r,\ell)$, the weight function on $E_{T_v}(v)$ is just $\mathbf{p}_v$, and $\mathbf{p} = (\mathbf{p}_v)_{v\in B(r,\ell)}$. Choose $K = E(r)$. Then for the simplicial complex defined as above, we have*

$$\mathbf{p}_r := f^{\ell\to 0}(\mathbf{p}) = \mu_K.$$

Proof. For simplicity, we assume that all leaves in T are at the same level. It can be generalized to any tree of depth $\ell+1$. We write the entries of $\mathbf{p}_r$ explicitly:

$$\begin{aligned}
\mathbf{p}_r(\pi) &\propto \prod_{v^1\in N(r)} \sum_{\substack{\sigma^1\in C_{v^1} \\ \pi((r,v^1))\notin\sigma^1}} \mathbf{p}_{v^1}(\sigma^1) \\
&\propto \prod_{v^1\in N(r)} \sum_{\substack{\sigma^1\in C_{v^1} \\ \pi((r,v^1))\notin\sigma^1}} \cdots \prod_{v^\ell\in N(v^{\ell-1})} \sum_{\substack{\sigma^\ell\in C_{v^{\ell-1}} \\ \sigma^{\ell-1}((v^{\ell-1},v^\ell))\notin\sigma^\ell}} \mathbf{p}_{v^\ell}(\sigma^\ell) \\
&= \sum_{\sigma|_K=\pi} \prod_{v\in B(v,\ell)} \mathbf{p}_v(\sigma|_{E_{T_v}(v)}).
\end{aligned}$$

Therefore, $\mathbf{p}_r = \mu_K$. $\square$

Suppose $K = \{e_i\}_{i\in[d]}$. Then $T \setminus K$ is composed of d disconnected subgraphs denoted by T_i such that T_i is adjacent to K in T (we define $T_i = \varnothing$ if e_i is pendant). Moreover, we define $\mathcal{K}_i := \{K_j \mid K_j \subseteq T_i\}$. Since both $\{T_i\}$ and $\{\mathcal{K}_i\}$ contain disjoint subgraphs and the σ in the summation in eq. (14) is determined by partial colorings $\sigma|_{T_i}$, we can define $\mathcal{S}_{i,c}^\xi \subseteq \Omega_{T_i}^\xi$ by the proper colorings on T_i such that is compatible with e_i being colored c.

$$\begin{aligned}
\mu_K^\xi(\tau) &\propto \sum_{\sigma_1\in\mathcal{S}_{1,\tau(e_1)}^\xi} \sum_{\sigma_2\in\mathcal{S}_{2,\tau(e_2)}^\xi} \cdots \sum_{\sigma_N\in\mathcal{S}_{N,\tau(e_N)}^\xi} \prod_{i\in[d]} \prod_{j:K_j\in\mathcal{K}_i} w_j(\sigma_i|_{K_j}) \\
&= \prod_{i\in[d]} \sum_{\sigma_i\in\mathcal{S}_{i,\tau(e_i)}^\xi} \prod_{j:K_j\in\mathcal{K}_i} w_j(\sigma_i|_{K_i}).
\end{aligned}$$

Define

$$p_{e_i,c}^\xi = \sum_{\sigma_i\in\mathcal{S}_{i,c}^\xi} \prod_{j:K_j\in\mathcal{K}_i} w_j(\sigma_i|_{K_i}).$$

Then we have $\mu_K^\xi(\tau) \propto \prod_{i\in[d]} p_{e_i,\tau(e_i)}^\xi$. This proves the following lemma.

Lemma 42 (Marginal distribution on a broom)**.** *Consider a list edge coloring instance $(T, \mathcal{L})$, boundaries $\{K_i\}_{i\in[N]}$ weighted functions $\{w_i\}_{i\in[N]}$, and a pinning $\xi \in \Omega_F$ for some subset $F \subset E$. Then for a broom K disjoint from boundaries there exists constants $p^{\xi}_{e,c}$ for $e \in K$, $c \in \mathcal{L}^{\xi}(e)$ such that*

$$\mu^{\xi}_K(\tau) \propto \prod_{i\in[d]} p^{\xi}_{e_i,\tau(e_i)}.$$

Let $q^{\xi}_e = \sum_{c\in\mathcal{L}^{\xi}(e)} p^{\xi}_{e,c}$ and $q^{\xi}_{e,f} = \sum_{c\in\mathcal{L}^{\xi}(e)\cap\mathcal{L}^{\xi}(f)} p^{\xi}_{e,c} p^{\xi}_{f,c}$.

Next we present some bounds on the quantities $p^{\xi}_{e,c}$ and $q^{\xi}_{\cdot}$

Lemma 43. *For any $e_i \in K, \xi \in \Omega_F$ for some subset $F \subset E$,*

$$\frac{p^{\xi}_{e_i,c}}{q^{\xi}_{e_i}} \le \frac{1}{\ell^{\xi}_{e_i}}.$$

Proof. Denote the subtree induced by T_i and e_i by $\tilde{T}_i$. We consider the weighted coloring instance on $(\tilde{T}_i, \mathcal{L})$ with boundaries $\mathcal{K}_i$ and weighted functions $\{w_j\}_{K_j\in\mathcal{K}_i}$ and denote the associated distribution by ν. Then by Definition 37,

$$\nu^{\xi}_e(c) = \frac{p^{\xi}_{e_i,c}}{q^{\xi}_{e_i,c}}.$$

Then

$$\frac{p^{\xi}_{e_i,c}}{q^{\xi}_{e_i}} = \sum_{\sigma\in C_{T_i,\mathcal{L}}} \nu^{\sigma\cup\xi}_e(c)\nu^{\xi}_{T_i}(\sigma).$$

By Lemma 39, $\nu^{\sigma\cup\xi}_e(c) \le \frac{1}{\beta}$. On the other hand, $\sum_{\sigma\in C_{T_i,\mathcal{L}}} \nu^{\xi}_{T_i}(\sigma) = 1$. So $\frac{p_{e,c}}{q_e} \le \frac{1}{\beta}$. ☐

Lemma 44. *For $e, f \in K$ and $\xi \in \Omega_F$ for some subgraph F,*

$$q^{\xi}_{e,f} \le \frac{q^{\xi}_e q^{\xi}_f}{\beta}.$$

Proof. Direct calculation yields

$$\begin{aligned}
q^{\xi}_{e,f} &= \sum_{c\in\mathcal{L}(e)\cap\mathcal{L}(f)} p^{\xi}_{e,c} p^{\xi}_{f,c} \\
&\le \left(\sum_{c\in\mathcal{L}(e)} \left(p^{\xi}_{e,c}\right)^2\right)^{1/2} \left(\sum_{c\in\mathcal{L}(f)} (p^{\xi}_{f,c})^2\right)^{1/2} \\
&\le \frac{\sqrt{q^{\xi}_e q^{\xi}_f}}{\beta} \left(\sum_{c\in\mathcal{L}^{\xi}(e)} p^{\xi}_{e,c}\right)^{1/2} \left(\sum_{c\in\mathcal{L}^{\xi}(f)} p^{\xi}_{f,c}\right)^{1/2} \\
&= \frac{q^{\xi}_e q^{\xi}_f}{\beta}.
\end{aligned}$$

☐

6.2 Simplicial complexes

First we introduce simplicial complexes to encode the edge coloring instance. Given a universe U, a *simplicial complex* $\mathcal{C} \subseteq 2^U$ is a collection of subsets of U that is downward close, which means that if $\sigma \in \mathcal{C}$ and $\sigma' \subseteq \sigma$, then $\sigma' \in \mathcal{C}$. Every element $\sigma \in \mathcal{C}$ is called a *face*, and a face that is not a proper subset of any other face is called a *maximal face* or a *facet*. The dimension of a face σ is $\mathsf{dim}(\sigma) := |\sigma|$, namely the size of σ. For every $k \geq 0$, let $\mathcal{C}_k := \{\sigma \in \mathcal{C} : |\sigma| = k\}$ be the set of faces of dimension k. Specifically, $\mathcal{C}_0 = \{\varnothing\}$. The dimension of $\mathcal{C}$ is the maximum dimension of faces in $\mathcal{C}$.

Besides, we say $\mathcal{C}$ is a *pure* n-dimensional simplicial complex if all maximal faces in $\mathcal{C}$ are of dimension n. In this work we only deal with pure simplicial complexes. In a pure simplicial complexe, we define the co-dimension of a face σ as $\mathsf{codim}(\sigma) := n - \mathsf{dim}(\sigma)$.

Let π_n be a distribution over the maximal faces $\mathcal{C}_n$. We use the pair $(\mathcal{C}, \pi_n)$ to denote a *weighted simplicial complex* where for each $1 \leq k < n$, the distribution π_n induces a distribution π_k over $\mathcal{C}_k$. Formally, for every $1 \leq k < n$ and every $\sigma' \in \mathcal{C}_k$, $\pi_k(\sigma')$ is proportional to the sum of weights of maximal faces containing σ. Formally,

$$\pi_k(\sigma') := \frac{1}{\binom{n}{k}} \sum_{\sigma \in \mathcal{C}_n : \sigma' \subset \sigma} \pi_n(\sigma).$$

It can be easily verified that π_k is a distribution on $\mathcal{C}_k$. Sometimes, we omit the subscript when $k = 1$, i.e., we write π for π_1.

Also we define the simplicial complexes generated by pinning a face in $\mathcal{C}$. For a face $\tau \in \mathcal{C}$ of dimension k, we define its *link* as

$$\mathcal{C}_\tau = \{\sigma \setminus \tau : \sigma \in \mathcal{C} \wedge \tau \subseteq \sigma\}.$$

Obviously, $\mathcal{C}_\tau$ is a pure $(n-k)$-dimensional simplicial complex. Similarly, for every $1 \leq j \leq n-k$, $\mathcal{C}_{\tau,j}$ is the set of faces in $\mathcal{C}_\tau$ of dimension j. We also use $\pi_{\tau,j}$ to denote the *marginal distribution* on $\mathcal{C}_{\tau,j}$. Formally, for every $\sigma \in \mathcal{C}_{\tau,j}$,

$$\pi_{\tau,j}(\sigma) := \mathbb{P}_{\alpha \sim \pi_{k+j}}[\alpha = \tau \cup \sigma : \tau \subseteq \alpha] = \frac{\pi_{k+j}(\tau \cup \sigma)}{\binom{k+j}{k} \cdot \pi_k(\tau)}.$$

We also drop the subscript when $j = 1$, i.e., we write π_τ for $\pi_{\tau,1}$. We define a random walk P_τ with stationary distribution π_τ as

$$P_\tau(x, y) = \frac{\pi_{\tau,2}(\{x, y\})}{2\pi_\tau(x)}.$$

6.3 Matrix trickle down on a broom

Now we define the corresponding simplicial complex to the weighted edge coloring instance defined in Section 6.1. Recall that we are dealing with a weighted edge coloring instance on a tree $T = (V, E)$ with β-extra color lists $\mathcal{L}$. $K = E_i$ for some $i \in V$ and $|K| = d$. We are going to show that the distribution of weighted colorings on K can be represented by a weighted simplicial complex.

Since any proper edge coloring $\sigma : E \to [q]$ could be regarded as a set of pairs of edge and color, namely $\{(e, c) \in E \times [q] : \sigma(e) = c\}$, the weighted edge-coloring instance restricted on K can be naturally represented as a pure weighted simplicial complex $(\mathcal{C}, \pi_d)$ where $\mathcal{C}$ consists of all proper partial colorings in $(K, \mathcal{L}|_K)$ and $\pi_d = \mu_K$.

Before introducing the matrix trickle-down theorem, we define notations for matrices related to π_τ. Define $\Pi_\tau \in \mathbb{R}^{\mathcal{C}_1 \times \mathcal{C}_1}$ as $\Pi_\tau := \mathrm{diag}(\pi_\tau)$ supported on $\mathcal{C}_{\tau,1} \times \mathcal{C}_{\tau,1}$, and $\pi_\tau := [\pi_\tau(x)]_{x \in \mathcal{C}_1}$ be a vector supported on $C_{\tau,1}$

For convenience, define the pseudo inverse $\Pi_\tau^{-1} \in \mathbb{R}^{\mathcal{C}_1 \times \mathcal{C}_1}$ of Π_τ as $\Pi_\tau^{-1}(x,x) = \pi_\tau(x)^{-1}$ for $x \in \mathcal{C}_{\tau,1}$ and 0 otherwise. Similarly, the pseudo inverse square root $\Pi_\tau^{-1/2} \in \mathbb{R}^{\mathcal{C}_1 \times \mathcal{C}_1}$is defined as $\Pi_\tau^{-1/2}(x,x) = \pi_\tau(x)^{-1/2}$ for $x \in \mathcal{C}_{\tau,1}$ and 0 otherwise. When $\tau = \varnothing$, we omit the subscript.

Recalling Definition 25 and Definition 26, the following lemma relates the covariance of μ_K and the matrices defined in this section.

Proposition 45.

$$\Pi(\mu_K) = d\Pi$$

and

$$\mathsf{Cov}(\mu_K) = d\left((d-1)\left(\Pi P - \frac{d}{d-1}\pi\pi^\top\right) + \Pi\right)$$

in the sense of padding with zeros.

The proof is by direct calculation. We apply the matrix trickle-down theorem on $(\mathcal{C}, \pi_d)$ to prove the following lemma. The main idea of the proof is almost the same as that in [WZZ24] except for substituting the number of feasible colors to the weights of feasible colors w.r.t $\{p_{e,c}\}$ after pinning. And the construction of matrix upper bound is simpler since the line graph of K is a clique. We include the details in Appendix A.

Lemma 46. *If K is not adjacent to boundaries, i.e. $\{K_i\}$ and $\beta \geq \Delta + 50$, then the simplicial complex $(\mathcal{C}, \pi_d)$ defined as above satisfying*

$$\Pi P - \frac{d}{d-1}\pi\pi^\top \preceq \frac{\eta_\Delta}{d-1}\Pi,$$

where η_Δ is $\mathcal{O}\left(\frac{\log^2 \Delta}{\Delta}\right)$.

Then Lemma 36 is derived directly from Lemma 46 and Proposition 45.

7 Tight bound for weak spatial mixing

In this section, we prove the tight bound for weak spatial mixing of q-edge coloring instance on trees. Note that the strong spatial mixing of q-edge coloring instance is a special case of spatial mixing of list instance (for list coloring instance, weak spatial mixing is equivalent to strong spatial mixing). Therefore the theorem stated in this section is a weak version of spatial mixing conclusions and thus we can give a tight bound for trees.

The main theorem for weak spatial mixing on trees is as follows.

Theorem 47. *Given a tree $T = (V, E)$ with n vertices, m edges and maximum degree Δ. Suppose that instance $(T, \mathcal{L})$ is a q-edge coloring instance (i.e. for any $e \in E$, $\mathcal{L}(e) = [q]$). Then we have that*

1. *If $q \geq 2\Delta - 1$, the edge coloring instance satisfies weak spatial mixing with rate $1 - \frac{1-\varepsilon}{2\Delta-(1+\varepsilon)}$, where $\varepsilon = \max\{\frac{\Delta-1}{\Delta}, \frac{e-1}{e}\}$.*
2. *If $q \leq 2\Delta - 2$, there exists an instance that does not satisfy weak spatial mixing.*

Consider the following example which simply shows that hardness of weak spatial mixing:

Example 1. *Consider a $(\Delta-1)$-regular tree with depth d (the depth of the root r is 0) and d is an even number. Let Λ be the edges between vertices of depth d and $d-1$. Let τ_1 be the pinning over Λ which only uses $1,2,\dots,\Delta-1$ and τ_2 be the pinning over Λ which only uses $\Delta,\dots,2\Delta-2$. It is easy to verify that for any $\{u,v\}\in E\setminus\Lambda$ and $\mathrm{dep}(u)=\mathrm{dep}(v)+1$,*

$$\forall\sigma\in\Omega^{\tau_1},,\sigma(\{u,v\})\in\begin{cases}\{1,\dots,\Delta-1\} & ,2\mid\mathrm{dep}(u)\\ \{\Delta,\dots,2\Delta-2\} & ,2\nmid\mathrm{dep}(u)\end{cases}$$

$$\forall\sigma\in\Omega^{\tau_2},\sigma(\{u,v\})\in\begin{cases}\{\Delta,\dots,2\Delta-2\} & ,2\mid\mathrm{dep}(u)\\ \{1,\dots,\Delta-1\} & ,2\nmid\mathrm{dep}(u)\end{cases}$$

where $\mathrm{dep}(u)$ is the depth of u. Therefore, for any $e\in E\setminus\Lambda$, we have that

$$\|\mu_e^\sigma-\mu_e^\tau\|_{\mathrm{TV}}=1$$

which demonstrates that the weak spatial mixing does not hold.

The proof scheme is also using the idea of correlation decay and we use the uniform distribution as bridge to prove the weak spatial mixing property. And we use another recursion which is different from that in strong spatial mixing. Instead of considering a broom of edges, we specify the marginal probability of a pendant edge and then generalize to every edge. Since the lists of feasible colors are clear, we use $\mathbb{P}_T[\cdot]$ to denote $\mathbb{P}_{T,\mathcal{L}}[\cdot]$ for simplicity.

Lemma 48 (One-step contraction)**.** *Suppose $(T=(V,E),\mathcal{L})$ is a q-edge coloring instance, where T is a tree with a pendant edge $e=\{r',r\}$ on its root r (that is, $\deg(r')=1$) and τ is the pinning over a set of edges Λ whose edges are incident to leaf vertices. If for any $e_i=\{v_i,r\}\in E$, the subtree T_i with pendant edge e_i satisfies that*

$$\forall c\in[q],\left|\mathbb{P}_{T_i\cup\{e_i\}}[c(e_i)=c\mid c(\Lambda)=\tau]-\frac{1}{q}\right|\le\delta$$

where $\delta<\frac{1}{q}$ is a universal constant, then we have that

$$\forall c\in[q],\left|\mathbb{P}_T[c(e)=c\mid c(\Lambda)=\tau]-\frac{1}{q}\right|\le\frac{2\Delta-2}{q(1-\delta\,|q-2\Delta+2|)}\delta.$$

Proof of Lemma 48. Assume that there are d children of r. Let $P_{r,c}$ denote $\mathbb{P}_T[c(e)=c\mid c(\Lambda)=\tau]$ and $P_{i,c}$ denote $\mathbb{P}_{T_i\cup\{e_i\}}[c(e_i)=c\mid c(\Lambda)=\tau]$. Then we have the following recursion for any $c\in[q]$,

$$P_{r,c}=\frac{\sum_{A\in C_r,c\notin A}\prod_{i=1}^{d}P_{i,A_i}}{(q-d)\sum_{A\in C_r}\prod_{i=1}^{d}P_{i,A_i}}.$$

$\delta < \frac{1}{q}$ implies that $|C_r| = q^{\underline{d}}$. Therefore, the following should be true for any $c \in [q]$,

$$\begin{aligned}
\frac{1}{P_{r,c}} &= q - d + \sum_{i=1}^{d} P_{i,c} \frac{(q-d) \sum_{A \in C_r, A_i = c} \prod_{j \neq i} P_{j,A_j}}{\sum_{A \in C_r, c \notin A} \prod_{j=1}^{d} P_{j,A_j}} \\
&= q - d + \sum_{i=1}^{d} P_{i,c} \frac{\sum_{A \in C_r, c \notin A} \prod_{j \neq i} P_{j,A_j}}{\sum_{A \in C_r, c \notin A} P_{i,A_i} \prod_{j \neq i} P_{j,A_j}} \\
&\leq q - d + \sum_{i=1}^{d} (\frac{1}{q} + \delta) \frac{\sum_{A \in C_r, c \notin A} \prod_{j \neq i} P_{j,A_j}}{(\frac{1}{q} - \delta) \sum_{A \in C_r, c \notin A} \prod_{j \neq i} P_{j,A_j}} \\
&= q + \frac{2dq\delta}{1 - q\delta}. \qquad (15)
\end{aligned}$$

In the same way, we can show that $\frac{1}{P_{r,c}} \geq q - \frac{2dq\delta}{1+q\delta}$. Combine this inequality and Equation (15), we get

$$\frac{-2d\delta}{q - \delta q(q-2d)} \leq P_{r,c} - \frac{1}{q} \leq \frac{2d\delta}{q + \delta q(q-2d)} \implies |P_{r,c} - \frac{1}{q}| \leq \frac{2d}{q(1 - \delta\,|q-2d|)} \delta.$$

When $q > 2d$, $\frac{2d}{q(1-\delta(q-2d))}\delta$ is monotone increasing with respect to d. Then $d \leq \Delta - 1$ implies that

$$\left| P_{r,c} - \frac{1}{q} \right| \leq \frac{2\Delta - 2}{q(1 - \delta\,|q - 2\Delta + 2|)} \delta. \qquad \square$$

Lemma 21 shows that the q-edge coloring instance admits the marginal lower bound, which is a start point of recursive contraction. Now we can prove Theorem 47.

Proof of Theorem 47. We prove weak spatial mixing for pendant edges first. Let d denote $\min_{e' \in \Lambda} \mathrm{dist}_T(e', e)$. Lemma 21 implies that if $d = 2$,

$$\frac{1}{eq} \leq \mathbb{P}_T\left[c(e) = c \mid c(\Lambda) = \tau\right] \leq \frac{1}{q - \Delta + 1}.$$

The right hand side inequality trivially follows from the recursion in the proof of Lemma 48. Therefore, we have that

$$\forall c \in [q], \left| \mathbb{P}_T\left[c(e) = c \mid c(\Lambda) = \tau\right] - \frac{1}{q} \right| \leq \max\left\{ \frac{\Delta - 1}{\Delta}, \frac{e-1}{e} \right\} \frac{1}{q} < \frac{1}{q} \qquad (16)$$

which serves as the base case of recursive contraction of the marginal probability. Let $\varepsilon = \max\left\{\frac{\Delta-1}{\Delta}, \frac{e-1}{e}\right\}$. Therefore, Plugging Equation (16) and Lemma 48, we get that for $d \geq 2$,

$$\forall c \in [q], \left| \mathbb{P}_T\left[c(e) = c \mid c(\Lambda) = \tau\right] - \frac{1}{q} \right| \leq \frac{\varepsilon}{q} \left(1 - \frac{1 - \varepsilon}{2\Delta - (1 + \varepsilon)} \right)^{d-2}. \qquad (17)$$

For general edge $e = \{u, v\}$, we split e into two pendant edges $e_1 = \{u, w\}$ and $e_2 = \{w, v\}$ by adding a new vertex w to V and $\mathcal{L}(e_1) = \mathcal{L}(e_2) = \mathcal{L}(e)$. Let T' denote the new graph after splitting e. Then we have that for any $c \in [q]$,

$$\begin{aligned}
\mathbb{P}_T\left[c(e) = c \mid c(\Lambda) = \tau\right] &= \mathbb{P}_{T'}\left[c(e_1) = c \mid c(e_1) = c(e_2), c(\Lambda) = \tau\right] \\
&= \frac{\mathbb{P}_{T'}\left[c(e_1) = c(e_2) = c \mid c(\Lambda) = \tau\right]}{\mathbb{P}_{T'}\left[c(e_1) = c(e_2) \mid c(\Lambda) = \tau\right]} \\
&= \frac{\mathbb{P}_{T'}\left[c(e_1) = c \mid c(\Lambda) = \tau\right] \mathbb{P}_{T'}\left[c(e_2) = c \mid c(\Lambda) = \tau\right]}{\sum_{c' \in [q]} \mathbb{P}_{T'}\left[c(e_1) = c' \mid c(\Lambda) = \tau\right] \mathbb{P}_{T'}\left[c(e_2) = c' \mid c(\Lambda) = \tau\right]}.
\end{aligned}$$

The last equation follows from the disconnection between e_1 and e_2. For $d \geq 2$, let $\eta = \frac{\varepsilon}{q}\left(1 - \frac{1-\varepsilon}{2\Delta-(1+\varepsilon)}\right)^{d-2}$ for simplicity. By Equation (17), we have that

$$\frac{(\frac{1}{q}-\eta)^2}{q(\frac{1}{q}+\eta)^2} \leq \mathbb{P}_T\left[c(e) = c \mid c(\Lambda) = \tau\right] \leq \frac{(\frac{1}{q}+\eta)^2}{q(\frac{1}{q}-\eta)^2}.$$

Therefore, for $d \geq 2$, plugging $\eta \leq \frac{\varepsilon}{q}$ and the above inequality implies that

$$|\mathbb{P}_T\left[c(e) = c \mid c(\Lambda) = \tau\right] - \frac{1}{q}| \leq \frac{4\eta}{(1-q\eta)^2} \leq \frac{4\varepsilon}{q(1-\varepsilon)^2}\left(1 - \frac{1-\varepsilon}{2\Delta-(1+\varepsilon)}\right)^{d-2}.$$

We pick $C = \max\left\{\frac{8\varepsilon}{(1-\varepsilon)^2}(1 - \frac{1-\varepsilon}{2\Delta-(1+\varepsilon)})^{-2}, \left(1 - \frac{1-\varepsilon}{2\Delta-(1+\varepsilon)}\right)^{-1}\right\}$ to finish the first part of the proof.

For the second part, it is trivial after applying Example 1. □

References

[ALG21a] Dorna Abdolazimi, Kuikui Liu, and Shayan Oveis Gharan. A matrix trickle-down theorem on simplicial complexes and applications to sampling colorings. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 161–172. IEEE, IEEE, 2021. 2, 3, 35

[ALG21b] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. *SIAM Journal on Computing*, (0):FOCS20–1, 2021. 2

[CCFV25] Charlie Carlson, Xiaoyu Chen, Weiming Feng, and Eric Vigoda. Optimal mixing for randomly sampling edge colorings on trees down to the max degree. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 5418–5433. SIAM, 2025. 2

[CFG+24] Xiaoyu Chen, Weiming Feng, Heng Guo, Xinyuan Zhang, and Zongrui Zou. Deterministic counting from coupling independence. *arXiv preprint arXiv:2410.23225*, 2024. 2, 16, 17

[CG24] Zongchen Chen and Yuzhou Gu. Fast sampling of b-matchings and b-edge covers. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA 2024)*, pages 4972–4987, 2024. 3

[CLMM23] Zongchen Chen, Kuikui Liu, Nitya Mani, and Ankur Moitra. Strong spatial mixing for colorings on trees and its algorithmic applications. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 810–845. IEEE, 2023. 2, 3, 4

[CV25] Charlie Carlson and Eric Vigoda. Flip dynamics for sampling colorings: Improving (11/6—ε) using a simple metric. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2194–2212. SIAM, 2025. 2

[CZ23] Xiaoyu Chen and Xinyuan Zhang. A near-linear time sampler for the Ising model with external field. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 4478–4503. SIAM, 2023. 2

[DHP20] Michelle Delcourt, Marc Heinrich, and Guillem Perarnau. The Glauber dynamics for edge-colorings of trees. *Random Structures & Algorithms*, 57(4):1050–1076, 2020. 2

[DMKLP25] Lucas De Meyer, František Kardoš, Aurélie Lagoutte, and Guillem Perarnau. An algorithmic Vizing's theorem: toward efficient edge-coloring sampling with an optimal number of colors. *arXiv preprint arXiv:2501.11541*, 2025. 2

[EGH+19] Charilaos Efthymiou, Andreas Galanis, Thomas P Hayes, Daniel Štefankovič, and Eric Vigoda. Improved strong spatial mixing for colorings on trees. *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, 2019. 2

[GKM15] David Gamarnik, Dmitriy Katz, and Sidhant Misra. Strong spatial mixing of list coloring of graphs. *Random Structures & Algorithms*, 46(4):599–613, 2015. 17

[HJNP19] Marc Heinrich, Alice Joffard, Jonathan Noel, and Aline Parreau. Unpublished manuscript. https://hoanganhduc.github.io/events/CoRe2019/CoRe_2019_Open_Problems.pdf, 2019. 2

[JVV86] Mark R Jerrum, Leslie G Valiant, and Vijay V Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical computer science*, 43:169–188, 1986. 2

[Moi19] Ankur Moitra. Approximate counting, the Lovász local lemma, and inference in graphical models. *Journal of the ACM*, 66(2):10:1–10:25, 2019. 2

[Vil21] Cédric Villani. *Topics in optimal transportation*, volume 58. American Mathematical Soc., 2021. 6

[WZZ24] Yulin Wang, Chihao Zhang, and Zihan Zhang. Sampling proper colorings on line graphs using (1+ o (1)) Δ colors. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1688–1699, 2024. 2, 3, 31, 37, 43

A Proofs for matrix trickle-down process

In this section, we use the language of simplicial complexes as in Section 6. That is, for a weighted edge coloring instance on a tree T with β-extra color losts $\mathcal{L}$ and distribution μ, we consider the colorings on a broom K as a weighted simplicial complex $(\mathcal{C}, \pi_{|K|})$ such that $\pi_{|K|} = \mu_K$.

Throughout this section, we consider K as an edge set. For a pinning $\tau \in \mathcal{C}_i$ with $0 \le i \le |K| - 2$, we define $K_\tau = \{e \in K \mid e \notin \tau\}$ and $\mathsf{col}(\tau) = \{c \mid \exists e \in K, \tau(e) = c\}$. Let $\mathsf{Id}_\tau \in \mathbb{R}^{\mathcal{C}_1 \times \mathcal{C}_1}$ be the identity matrix restricted on $\mathcal{C}_{\tau,1}$. Define $\mathsf{Adj}_\tau \in \mathbb{R}^{\mathcal{C}_1 \times \mathcal{C}_1}$ such that $\mathsf{Adj}_\tau(ec, fc) = 1$ if $ec, fc \in \mathcal{C}_{\tau,1}$ and all other entries are 0.

A.1 Matrix trickle-down theorem

The following proposition is the main tool we use in this section.

Proposition 49 (Theorem 1.3 in [ALG21a]). *Given a pure d-dimensional weighted simplicial complex $(\mathcal{C}, \pi_d)$, if there exists a family of matrices $\left\{M_\tau \in \mathbb{R}^{\mathcal{C}_1 \times \mathcal{C}_1}\right\}$ satisfying*

- *For every $\tau \in \mathcal{C}_{d-2}$,*

$$\Pi_\tau P_\tau - 2\pi_\tau \pi_\tau^\top \preceq M_\tau \preceq \frac{1}{5}\Pi_\tau;$$

- *For every face $\tau \in \mathcal{C}_{d-k}$ with $k \geq 3$, one of the following two conditions hold:*

  1. $$M_\tau \preceq \frac{k-1}{3k-1}\Pi_\tau \quad \text{and} \quad \mathbf{E}_{x\sim\pi_\tau}\left[M_{\tau\cup\{x\}}\right] \preceq M_\tau - \frac{k-1}{k-2}M_\tau \Pi_\tau^{-1} M_\tau.$$

  2. *$(\mathcal{C}_\tau, \pi_{\tau,k})$ is the product of M pure weighted simplicial complexes $(\mathcal{C}^{(1)}, \pi^{(1)}), \ldots (\mathcal{C}^{(M)}, \pi^{(M)})$ of dimension $n_1, \ldots, n_M$ respectively and*

$$M_\tau = \sum_{i\in[M]:n_i\geq 2} \frac{n_i(n_i-1)}{k(k-1)} \cdot M_{\tau\cup\eta_{-i}}$$

  where $\eta_{-i} = \eta \setminus \mathcal{C}_1^{(i)}$ for an arbitrary $\eta \in \mathcal{C}_{\tau,k}$.

Then for every face $\tau \in \mathcal{C}_{d-k}$ with $k \geq 2$, it holds that

$$\Pi_\tau P_\tau - \frac{k}{k-1}\pi_\tau\pi_\tau^\top \preceq M_\tau \preceq \frac{k-1}{3k-1}\Pi_\tau.$$

In particular, $\lambda_2(P_\tau) \leq \lambda_1(\Pi_\tau^{-1}M_\tau)$.

A.2 Base case

We do matrix trickle down on $\mu_K(\tau)$. Consider the base case: Assume the free edges in K are e, f and other edges are pinned with assignment τ. Let $\mathbf{p}_{ef} = (p_{e,c}, p_{f,c})_{c\in\mathcal{L}^\tau(e)\cup\mathcal{L}^\tau(f)}$ and $\mathbf{p}_{fe} = (p_{f,c}, p_{e,c})_{c\in\mathcal{L}^\tau(e)\cup\mathcal{L}^\tau(f)}$. We omit the superscript τ in q_e^τ and $q_{e,f}^\tau$ in base case part since it is clear in the context. Moreover, let $\tilde{\mathbf{p}}_e = (0, p_{e,c})_{c\in\mathcal{L}^\tau(e)\cup\mathcal{L}^\tau(f)}, \tilde{\mathbf{p}}_f = (p_{f,c}, 0)_{c\in\mathcal{L}^\tau(e)\cup\mathcal{L}^\tau(f)}$, and define $\mathbb{1}_e, \mathbb{1}_f \in \mathbb{R}^{\mathcal{C}_{\tau,1}}$ such that $\mathbb{1}_e(i,c) = \mathbb{1}[i = e], \mathbb{1}_f(i,c) = \mathbb{1}[i = f]$. Finally, we define $\mathbf{1}_{ef} \in \mathbb{R}^{\mathcal{C}_{\tau,1}\times\mathcal{C}_{\tau,1}}$ such that $\mathbf{1}_{ef}((i,c_1),(j,c_2) = \mathbb{1}[c_1 = c_2 \wedge i \neq j]$.

$$\begin{aligned}
&\Pi_\tau P_\tau - 2\pi_\tau\pi_\tau^\top \\
&= \frac{1}{2(q_e q_f - q_{e,f})^2}\mathrm{diag}(\mathbf{p}_{ef})\left((q_e q_f - q_{e,f})(\mathbb{1}_e\mathbb{1}_f^\top + \mathbb{1}_f\mathbb{1}_e^\top - \mathbf{1}_{ef}) - \frac{(q_e\mathbb{1}_f + q_f\mathbb{1}_e)(q_e\mathbb{1}_f + q_f\mathbb{1}_e)^\top}{2} + \mathbf{p}_{fe}\mathbf{p}_{fe}^\top\right)\mathrm{diag}(\mathbf{p}_{ef}) \\
&\preceq \frac{1}{2(q_e q_f - q_{e,f})^2}\mathrm{diag}(\mathbf{p}_{ef})\left(\mathbf{p}_{fe}\mathbf{p}_{fe}^\top - (q_e q_f - q_{e,f})\mathbf{1}_{ef})\right)\mathrm{diag}(\mathbf{p}_{ef}) \\
&\preceq \frac{1}{2(q_e q_f - q_{e,f})^2}\mathrm{diag}(\mathbf{p}_{ef})\left(2(\tilde{\mathbf{p}}_e\tilde{\mathbf{p}}_e^\top + \tilde{\mathbf{p}}_f\tilde{\mathbf{p}}_f^\top) - (q_e q_f - q_{e,f})\mathbf{1}_{ef})\right)\mathrm{diag}(\mathbf{p}_{ef}).
\end{aligned}$$

We do a row summation to bound the first term. Firstly,

$$\tilde{\mathbf{p}}_e\tilde{\mathbf{p}}_e^\top + \tilde{\mathbf{p}}_f\tilde{\mathbf{p}}_f^\top \preceq q_e\mathrm{diag}(\tilde{\mathbf{p}}_e) + q_f\mathrm{diag}(\tilde{\mathbf{p}}_f).$$

Then,

$$\frac{1}{(q_e q_f - q_{e,f})^2}\mathrm{diag}(\mathbf{p}_{ef})(\tilde{\mathbf{p}}_e\tilde{\mathbf{p}}_e^\top + \tilde{\mathbf{p}}_f\tilde{\mathbf{p}}_f^\top)\mathrm{diag}(\mathbf{p}_{ef}) \preceq \frac{1}{(q_e q_f - q_{e,f})^2}\mathrm{diag}(\mathbf{p}_{ef})(q_e\mathrm{diag}(\tilde{\mathbf{p}}_e) + q_f\mathrm{diag}(\tilde{\mathbf{p}}_f))\mathrm{diag}(\mathbf{p}_{ef}),$$

which is a diagonal matrix. On the entry (ec, ec), it equals

$$\begin{aligned}
\frac{q_f p_{ec}}{(q_e q_f - q_{e,f})^2}p_{fc}p_{ec} &= \frac{2q_f q_e}{q_f q_e} \cdot \frac{p_{ec}}{q_e} \cdot \frac{q_f q_e}{q_f q_e - q_{fe}} \cdot \frac{(q_f - p_{fc})p_{ec}}{2(q_f q_e - q_{fe})} \cdot \frac{p_{fc}}{q_f - p_{fc}} \\
&\leq 2 \cdot \frac{1}{\beta} \cdot \frac{\beta}{\beta - 1} \cdot \pi_\tau(ec) \cdot \frac{1}{\beta - 1} && \text{(Lemmas 43 and 44)} \\
&= \frac{2}{(\beta-1)^2}\pi_\tau(ec).
\end{aligned}$$

Applying the same argument to (fc, fc), we have

$$\Pi_\tau P_\tau - 2\pi_\tau\pi_\tau^\top \preceq -\frac{1}{2(q_eq_f - q_{e,f})^2}\mathrm{diag}(\mathbf{p}_{ef})\left((q_eq_f - q_{e,f})\mathbf{1}_{ef}\right)\mathrm{diag}(\mathbf{p}_{ef}) + \frac{2}{(\beta-1)^2}\Pi_\tau.$$

Let M_τ be a block diagonal matrix with blocks M_τ^c:

$$M_\tau^c = \frac{p_{e,c}p_{f,c}}{2(q_eq_f - q_{e,f})}\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} + \frac{2}{(\beta-1)^2}\Pi_\tau^c. \tag{18}$$

Then we have the base case inequality

$$\Pi_\tau P_\tau - 2\pi_\tau\pi_\tau^\top \preceq \mathrm{diag}(M_\tau^c).$$

A.3 Induction

The induction step in Proposition 49 is to show that for every τ with codim$(\tau) = k > 2$ and connected G_τ,

$$\mathbf{E}_{x\sim\pi_\tau}\left[M_{\tau\cup\{x\}}\right] \preceq M_\tau - \frac{k-1}{k-2}M_\tau\Pi_\tau^{-1}M_\tau. \tag{19}$$

For every τ and $c \in [q]$, we will define a matrix $M_\tau^c \in \mathbb{R}^{K^c\times K^c}$ and let M_τ be the block diagonal matrix with block M_τ^c for every $c \in [q]$. It is not hard to see that we only require

$$\mathbf{E}_{x\sim\pi_\tau}\left[M_{\tau\cup\{x\}}^c\right] \preceq M_\tau^c - \frac{k-1}{k-2}M_\tau^c(\Pi_\tau^c)^{-1}M_\tau^c \tag{20}$$

to hold for every c and τ, where $(\Pi_\tau^c)^{-1}$ is Π_τ^{-1} restricted on $K^c \times K^c$. We now describe our construction of M_τ^c for a fixed color c. We write M_τ^c into the sum of a diagonal matrix and an off-diagonal matrix, i.e.,

$$M_\tau^c = \frac{1}{k-1}(A_\tau^c + \Pi_\tau^c B_\tau^c), \tag{21}$$

where A_τ^c is an off-diagonal matrix and B_τ^c is a diagonal matrix.

From now on, when c is clear from the context, we will omit the superscript c for matrices. For example, we will write M_τ, Π_τ, Adj_τ, Id_τ, A_τ, B_τ, ... instead of M_τ^c, Π_τ^c, Adj_τ^c, Id_τ^c, A_τ^c, B_τ^c, ... respectively. Plugging the above construction of M_τ into (20) and remembering that the superscript c has been omitted, we obtain

$$(k-1)\cdot\mathbf{E}_{x\sim\pi_\tau}\left[A_{\tau\cup\{x\}} + \Pi_{\tau\cup\{x\}}B_{\tau\cup\{x\}}\right] \preceq (k-2)\cdot(A_\tau + \Pi_\tau B_\tau) - (A_\tau + \Pi_\tau B_\tau)\,\Pi_\tau^{-1}\,(A_\tau + \Pi_\tau B_\tau).$$

Here we need the following inequality of matrices.

Lemma 50 (Corollary 12 in [WZZ24]). *Let $A_1, \ldots, A_n$ be a collection of symmetric matrices and $\Pi \succeq 0$. Then*

$$\left(\sum_{i=1}^n A_i\right)\Pi\left(\sum_{i=1}^n A_i\right) \preceq n\sum_{i=1}^n A_i\Pi A_i.$$

It follows from Lemma 50 that

$$(A_\tau + \Pi_\tau B_\tau)\,\Pi_\tau^{-1}\,(A_\tau + \Pi_\tau B_\tau) \preceq 2A_\tau\Pi_\tau^{-1}A_\tau + 2\Pi_\tau B_\tau^2.$$

As a result, in order for (19) to hold, we only need to design A_τ and B_τ satisfying

$$(k-1)\cdot\mathbf{E}_{x\sim\pi_\tau}\left[A_{\tau\cup\{x\}}\right] - (k-2)\cdot A_\tau + 2A_\tau\Pi_\tau^{-1}A_\tau \preceq (k-2)\Pi_\tau B_\tau - (k-1)\cdot\mathbf{E}_{x\sim\pi_\tau}\left[\Pi_{\tau\cup\{x\}}B_{\tau\cup\{x\}}\right] - 2\Pi_\tau\left(B_\tau\right)^2. \tag{22}$$

A.4 Construction of A^i_τ

We define

$$A_\tau = a_k \cdot (k-1) \cdot \mathbf{E}_{\sigma \sim \pi_{\tau,k-2}} \left[A_{\tau \cup \sigma} \right]. \tag{23}$$

Then we can deduce the following relation between A_τ's whose co-dimensions differ by one.

Lemma 51.

$$\mathbf{E}_{x \sim \pi_\tau} \left[A_{\tau \cup \{x\}} \right] = \frac{k-2}{k-1} \cdot \frac{a_{k-1}}{a_k} A_\tau(ec, fc).$$

where $h = h_\tau \geq 1$.

Proof. For any $ec, fc \in K_\tau$,

$$\begin{aligned}
\mathbf{E}_{x \sim \pi_\tau} \left[A_{\tau \cup \{x\}} \right](ec, fc) &= (k-2) \sum_{x \in \mathcal{C}_{\tau,1}} \frac{1}{k} \mu^\tau_{K_x}(x) a_{k-1} \sum_{\sigma \in \mathcal{C}_{\tau \cup \{x\}, k-3}} \frac{2}{(k-1)(k-2)} \mu^{\tau \cup \{x\}}_{K_\sigma}(\sigma) A_{\tau \cup \{x\} \cup \sigma}(ec, fc) \\
&= \frac{2}{k(k-1)} \sum_{x \in \mathcal{C}_{\tau,1}} \sum_{\sigma \in \mathcal{C}_{\tau \cup \{x\}, k-3}} \mu^\tau_{K_x}(x) \mu^{\tau \cup \{x\}}_{K_\sigma}(\sigma) a_{k-1} A_{\tau \cup \{x\} \cup \sigma}(ec, fc) \\
&= \frac{2}{k(k-1)} \sum_{\sigma' \in \mathcal{C}_{\tau, k-2}} \mu^\tau_{K_{\sigma'}}(\sigma') A_{\tau \cup \sigma'}(uc, vc)(k-2) a_{k-1} \\
&= (k-2) \mathbf{E}_{\sigma' \sim \pi_{\tau,k-2}} \left[A_{\tau \cup \sigma'}(ec, fc) \right] \\
&= \frac{k-2}{k-1} \cdot \frac{a_{k-1}}{a_k} A_\tau(ec, fc).
\end{aligned}$$

□

It follows from the definition that A_τ is proportional to the expectation of the base cases when the boundary is drawn from $\pi_{\tau,k-2}$. For some technical reasons, we would like to isolate those boundaries containing the color c. This leads us to the following lemma.

Lemma 52.

$$A_\tau = \frac{2a_k}{k} \sum_{\substack{\omega \in \mathcal{C}_{\tau,k} \\ c \notin \mathrm{col}(\omega)}} \pi_{\tau,k}(\omega) A^{i,\omega}_\tau,$$

where $A^{i,\omega}_\tau$ *is the matrix supported on* $K^c_\tau \times K^c_\tau$ *such that*

$$A^\omega_\tau(uc, vc) = -\frac{p^{\tau \cup \omega|_{K_\tau \setminus \{e,f\}}}_{e,c} p^{\tau \cup \omega|_{K_\tau \setminus \{e,f\}}}_{f,c}}{\overline{q}^{\tau \cup \omega|_{K_\tau \setminus \{e,f\}}}_e \overline{q}^{\tau \cup \omega|_{K_\tau \setminus \{e,f\}}}_f - \overline{q}^{\tau \cup \omega|_{K_\tau \setminus \{e,f\}}}_{ef}},$$

and $\overline{q}^\cdot_e := q^\cdot_e - p^\cdot_{ec}$, $\overline{q}^\cdot_f := q^\cdot_f - p^\cdot_{fc}$, *and* $\overline{q}^\cdot_{ef} := q^\cdot_{ef} - p^\cdot_{ec} p^\cdot_{fc}$.

Proof. In the proof we use $c \notin \cdot$ as a shortcut for $c \notin \mathrm{col}(\cdot)$. By the definition of A_τ,

$$\begin{aligned}
A_\tau(ec, fc) &= (k-1) a_k \frac{2}{k(k-1)} \sum_{\substack{\sigma \in \mathcal{C}_{\tau,k-2} \\ c \notin \sigma}} \mu^\tau_{K_\tau \setminus \{e,f\}}(\sigma) \frac{p^{\tau \cup \sigma}_{ec} p^{\tau \cup \sigma}_{fc}}{q^{\tau \cup \sigma}_e q^{\tau \cup \sigma}_f - q^{\tau \cup \sigma}_{ef}} \\
&= \frac{2a_k}{k} \sum_{\substack{\sigma \in \mathcal{C}_{\tau,k-2} \\ c \notin \sigma}} \Big(\sum_{\xi \in \mathcal{C}_{\tau \cup \sigma, 2}} \mu^{\tau \cup \sigma}_{\{e,f\}}(\xi) \Big) \mu^\tau_{K_\tau \setminus \{e,f\}}(\sigma) \frac{p^{\tau \cup \sigma}_{ec} p^{\tau \cup \sigma}_{fc}}{q^{\tau \cup \sigma}_e q^{\tau \cup \sigma}_f - q^{\tau \cup \sigma}_{ef}}.
\end{aligned}$$

The equality is becase if $c \in \sigma$, the $p_{\cdot,c}^{\tau\cup\sigma}$ terms vanish. Notice that for any $\sigma \in \mathcal{C}_{\tau,k-2}$ and $\xi \in \mathcal{C}_{\tau\cup\sigma,2}$,

$$\begin{aligned}1 &= \sum_{\xi\in\mathcal{C}_{\tau\cup\sigma,2}} \mu_{\{e,f\}}^{\tau\cup\sigma}(\xi) \\ &= \frac{\sum_{\xi\in\mathcal{C}_{\tau\cup\sigma,2}} \mu_{\{e,f\}}^{\tau\cup\sigma}(\xi)}{\sum_{\substack{\xi\in\mathcal{C}_{\tau\cup\sigma,2}\\ c\notin\xi}} \mu_{\{e,f\}}^{\tau\cup\sigma}(\xi)} \sum_{\substack{\xi\in\mathcal{C}_{\tau\cup\sigma,2}\\ c\notin\xi}} \mu_{\{e,f\}}^{\tau\cup\sigma}(\xi) \\ &= \frac{q_e^{\tau\cup\sigma} q_f^{\tau\cup\sigma} - q_{ef}^{\tau\cup\sigma}}{\overline{q}_e^{\tau\cup\sigma}\overline{q}_f^{\tau\cup\sigma} - \overline{q}_{ef}^{\tau\cup\sigma}} \sum_{\substack{\xi\in\mathcal{C}_{\tau\cup\sigma,2}\\ c\notin\xi}} \mu_{\{e,f\}}^{\tau\cup\sigma}(\xi).\end{aligned}$$

Multiply this expression in the former equation, we have

$$\begin{aligned}A_\tau(ec,fc) &= \frac{2a_k}{k} \sum_{\substack{\sigma\in\mathcal{C}_{\tau,k-2}\\ c\notin\sigma}} \mu_{K_\tau\setminus\{e,f\}}^{\tau}(\sigma) \frac{q_e^{\tau\cup\sigma} q_f^{\tau\cup\sigma} - q_{ef}^{\tau\cup\sigma}}{\overline{q}_e^{\tau\cup\sigma}\overline{q}_f^{\tau\cup\sigma} - \overline{q}_{ef}^{\tau\cup\sigma}} \sum_{\substack{\xi\in\mathcal{C}_{\tau\cup\sigma,2}\\ c\notin\xi}} \mu_{\{e,f\}}^{\tau\cup\sigma}(\xi) \frac{p_{ec}^{\tau\cup\sigma} p_{fc}^{\tau\cup\sigma}}{q_e^{\tau\cup\sigma} q_f^{\tau\cup\sigma} - q_{ef}^{\tau\cup\sigma}} \\ &= \frac{2a_k}{k} \sum_{\substack{\sigma\in\mathcal{C}_{\tau,k-2}\\ c\notin\sigma}} \mu_{K_\tau\setminus\{e,f\}}^{\tau}(\sigma) \sum_{\substack{\xi\in\mathcal{C}_{\tau\cup\sigma,2}\\ c\notin\xi}} \mu_{\{e,f\}}^{\tau\cup\sigma}(\xi) \frac{p_{ec}^{\tau\cup\sigma} p_{fc}^{\tau\cup\sigma}}{\overline{q}_e^{\tau\cup\sigma}\overline{q}_f^{\tau\cup\sigma} - \overline{q}_{ef}^{\tau\cup\sigma}} \\ &= \frac{2a_k}{k} \sum_{\substack{\omega\in\mathcal{C}_{\tau,k}\\ c\notin\sigma}} \mu_{K_\tau}^{\tau}(\omega) \frac{p_{ec}^{\tau\cup\omega|_{K_\tau\setminus\{e,f\}}} p_{fc}^{\tau\cup\omega|_{K_\tau\setminus\{e,f\}}}}{\overline{q}_e^{\tau\cup\omega|_{K_\tau\setminus\{e,f\}}}\overline{q}_f^{\tau\cup\omega|_{K_\tau\setminus\{e,f\}}} - \overline{q}_{ef}^{\tau\cup\omega|_{K_\tau\setminus\{e,f\}}}},\end{aligned}$$

and the lemma follows. □

A.5 Spectral analysis of A_τ

For every $\omega \in \mathcal{C}_{\tau,k}$ such that $c \notin \mathsf{col}((\omega \cup \tau)|_{K_\tau})$, define $\Xi_\tau^\omega \in \mathbb{R}^{K^c \times K^c}$ as the diagonal matrix such that for every $ec \in K_\tau^c$:

$$\Xi_\tau^\omega(ec,ec) = \frac{p_{ec}^{\tau\cup\omega|_{K\setminus\{e\}}}}{\overline{q}_e^{\tau\cup\omega|_{K\setminus\{e\}}}}.$$

Lemma 53.

$$A_\tau^\omega = \Xi_\tau^\omega(-\mathsf{Adj}_\tau + \mathcal{R}_\tau^\omega)\Xi_\tau^\omega,$$

where $\rho(\mathcal{R}_\tau^{i,\omega}) \le \frac{2(k-1)}{\beta-1}$.

Proof. Let $ec, fc \in K_\tau^c$. To ease the notation, when τ and ω are clear from the context, we use $\Gamma(e)$ and $\Gamma(e,f)$ to denote the partial coloring $(\tau\cup\omega)|_{V\setminus\{e\}}$ and $(\tau\cup\omega)|_{V\setminus\{e,f\}}$ respectively. Using our new notations, we have

$$\Xi_\tau(ec,ec) = \frac{p_{ec}^{\Gamma(e)}}{\overline{q}_e^{\Gamma(e,f)}}.$$

Observing that since $c \notin \mathsf{col}\left((\tau\cup\omega)|_{K_\tau}\right)$, we have

$$\begin{aligned}p_{e,c}^{\Gamma(e)} &= p_{e,c}^{\Gamma(e,f)}, \\ p_{f,c}^{\Gamma(f)} &= p_{e,c}^{\Gamma(e,f)}.\end{aligned}$$

So we can write $A_\tau^{i,\omega}$ as

$$\begin{aligned} A_\tau^{i,\omega}(uc,vc) &= -\frac{p_{e,c}^{\Gamma(e)} p_{f,c}^{\Gamma(f)}}{\overline{q}_e^{\Gamma(e,f)}\overline{q}_f^{\Gamma(e,f)} - \overline{q}_{ef}^{\Gamma(e,f)}} \\ &= \frac{p_{e,c}^{\Gamma(e)} p_{f,c}^{\Gamma(f)}}{\overline{q}_e^{\Gamma(e)}\overline{q}_f^{\Gamma(f)}}\Big(-1 + \mathcal{R}_\tau^{\omega}(ec,fc)\Big) \end{aligned}$$

where

$$\mathcal{R}_\tau^{i,\omega}(ec,fc) = \frac{-\overline{q}_e^{\Gamma(e)}\overline{q}_f^{\Gamma(f)} + \overline{q}_e^{\Gamma(e,f)}\overline{q}_f^{\Gamma(e,f)} - \overline{q}_{ef}^{\Gamma(e,f)}}{\overline{q}_e^{\Gamma(e,f)}\overline{q}_f^{\Gamma(e,f)} - \overline{q}_{ef}^{\Gamma(e,f)}}. \tag{24}$$

Notice that

$$\begin{aligned} \left|\overline{q}_e^{\Gamma(e,f)}\overline{q}_f^{\Gamma(e,f)} - \overline{q}_e^{\Gamma(e)}\overline{q}_f^{\Gamma(f)}\right| &= \left|\overline{q}_e^{\Gamma(e,f)} p_{f,\omega(e)}^{\Gamma(e,f)} + \overline{q}_f^{\Gamma(e,f)} p_{e,\omega(f)}^{\Gamma(e,f)} - p_{f,\omega(e)}^{\Gamma(e,f)} p_{e,\omega(f)}^{\Gamma(e,f)}\right| \\ &\le \overline{q}_e^{\Gamma(e,f)} p_{f,\omega(e)}^{\Gamma(e,f)} + \overline{q}_f^{\Gamma(e,f)} p_{e,\omega(f)}^{\Gamma(e,f)}. \end{aligned}$$

Therefore,

$$\left|\mathcal{R}_\tau^{i,\omega}(ec,fc)\right| \le \frac{\overline{q}_e^{\Gamma(e,f)} p_{f,\omega(e)}^{\Gamma(e,f)} + \overline{q}_f^{\Gamma(e,f)} p_{e,\omega(f)}^{\Gamma(e,f)}}{\overline{q}_e^{\Gamma(e,f)}\overline{q}_f^{\Gamma(e,f)} - \overline{q}_{ef}^{\Gamma(e,f)}},$$

which equals $\mu_{\{e\};\mathcal{L}'}^{\Gamma(e,f)}(\omega(e)) + \mu_{\{f\};\mathcal{L}'}^{\Gamma(e,f)}(\omega(f))$, defining $\mathcal{L}'$ be the color lists obtained by removing c from the color lists of e, f. By Lemma 43, this is bounded by $\frac{2}{\beta-1}$ since the modified coloring istance is $(\beta-1)$-extra. The lemma then follows by doing a row summation to $\mathcal{R}_\tau^{\omega}$. □

Proposition 54. *Consider non-zero coefficients λ_i, $i \in [N]$, matrix $A = \sum_{i\in[N]} \lambda_i A_i$ and $\Sigma \preceq 0$, and all matrices are square, of the same size and symemtric. Then*

$$A\Sigma A \preceq \Big(\sum_{i\in[N]} \lambda_i\Big)\Big(\sum_{i\in[N]} \lambda_i A_i \Sigma A_i\Big).$$

Proof.

$$\begin{aligned} A\Sigma A &= \sum_{i,j\in[N], i<j} \lambda_i\lambda_j\big(A_i\Sigma A_j + A_i\Sigma A_j\big) + \sum_{i\in[N]} \lambda_i^2 A_i\Sigma A_i \\ &\preceq \sum_{i,j\in[N]} \lambda_i\lambda_j\big(A_i\Sigma A_i + A_j\Sigma A_j\big) + \sum_{i\in[N]} \lambda_i^2 A_i\Sigma A_i \\ &= \Big(\sum_{i\in[N]} \lambda_i\Big)\Big(\sum_{i\in[N]} \lambda_i A_i\Sigma A_i\Big). \end{aligned}$$

□

Define $C_\tau := \sum_{\omega\in\mathcal{C}_\tau c\notin\omega} \pi_\tau(\omega)$, and $\mathcal{C}'_{\tau,k} := \left\{\omega \in \mathcal{C}_{\tau,k} \;\middle|\; c \notin \omega\right\}$, then the following lemma holds.

Lemma 55. $A_\tau \Pi_\tau^{-1} A_\tau \preceq \frac{4a_k^2 C_\tau}{k^2} \sum_{\omega\in\mathcal{C}'_{\tau,k}} \pi_\tau(\omega) A_\tau^\omega \Pi_\tau^{-1} A_\tau^\omega \preceq \frac{4a_k^2}{k^2} \sum_{\omega\in\mathcal{C}'_{\tau,k}} \pi_\tau(\omega) A_\tau^\omega \Pi_\tau^{-1} A_\tau^\omega$.

Proof. By Lemma 53 and Proposition 54. □

In the following discussion, let $\gamma = \left(1 + \frac{\Delta-1}{\beta-1}\right)^3 \frac{1}{\beta-1}$.

Lemma 56. $A_\tau^\omega \Pi_\tau^{-1} A_\tau^\omega \preceq \gamma k \cdot \Xi_\tau^\omega \left((\mathsf{Adj}_\tau)^2 + \frac{4(k-1)^2}{(\beta-1)^2} \mathsf{Id}_\tau \right) \Xi_\tau^{i,\omega}$.

Proof. By Lemmas 39, 40 and 43 for any $ec \in K_\tau^c$,

$$\Xi_\tau^\omega \Pi_\tau^{-1}(uc, uc) \le k \frac{(\beta+k-2)(\beta+\Delta-2)}{(\beta-1)^2} \frac{\ell_u - k + 1}{\ell_u - k - \Delta + 1} \le k \frac{(\beta+\Delta-2)^2}{(\beta-1)^2} \frac{\beta+\Delta-2}{\beta-1} \le k \left(1 + \frac{\Delta-1}{\beta-1}\right)^3.$$

Then it follows from Lemma 53 and $\Xi_\tau^{i,\omega} \preceq \frac{1}{\beta-1} \cdot \mathsf{Id}_\tau$ that

$$A_\tau^\omega \Pi_\tau^{-1} A_\tau^\omega \preceq 2k\gamma \Xi_\tau^{i,\omega} \left((\mathsf{Adj}_\tau)^2 + \frac{4(k-1)^2}{(\beta-1)^2} \mathsf{Id}_\tau \right) \Xi_\tau^\omega. \qquad \square$$

Lemma 57. $\frac{1}{k} \sum_{\omega \in \mathcal{C}_{\tau,k}} \pi_{\tau,k}(\omega) \Xi_\tau^\omega \Pi_\tau^{-1} = \mathsf{Id}_\tau$.

Proof. At entry (ec, ec) such that $\pi_\tau(ec) \neq 0$, the LHS is

$$\begin{aligned}
\frac{1}{k} \sum_{\omega \in \mathcal{C}_{\tau,k}} \mu_{K_\tau}^\tau(\omega) \pi_\tau^{-1}(ec) \frac{p_{e,c}^{\tau \cup \omega|_{K \setminus \{e\}}}}{\bar{q}_e^{\tau \cup \omega|_{K \setminus \{e\}}}} &= \frac{1}{k} \sum_{\omega' \in \mathcal{C}_{\tau, K_\tau}} \mu_{K_\tau}^\tau(\omega') \pi_\tau^{-1}(ec) \frac{p_{e,c}^{\tau \cup \omega'}}{\bar{q}_e^{\tau \cup \omega'}} \\
&= \frac{1}{k} \sum_{\omega' \in \mathcal{C}_{\tau, K_\tau}} \mu_{K_\tau \setminus \{e\}}^\tau(\omega') \mu_{\{e\}}^{\tau \cup \omega'}(c) \pi_\tau^{-1}(ec) \\
&= \frac{1}{k} \sum_{\omega' \in \mathcal{C}_{\tau, K_\tau}} \mu_{K_\tau}^\tau(\omega \cup \{ec\}) \pi_\tau^{-1}(ec) = \pi_\tau^{-1}(ec) \pi_\tau(ec) = 1.
\end{aligned}$$

$\square$

We are now ready to bound $\left((k-1) \cdot \mathbf{E}_{x \sim \pi_\tau} \left[A_{\tau \cup \{x\}} \right] - (k-2) \cdot A_\tau + 4 A_\tau \Pi_\tau^{-1} A_\tau \right)$ in the LHS of Equation (22).

Lemma 58. *There exists a sequence of non-negative numbers* $\{a_h\}_{0 \le h \le \Delta}$ *such that*

$$\Pi_\tau^{-\frac{1}{2}} \left((k-1) \cdot \mathbf{E}_{x \sim \pi_\tau} \left[A_{\tau \cup \{x\}} \right] - (k-2) \cdot A_\tau + 2 A_\tau \Pi_\tau^{-1} A_\tau \right) \Pi_\tau^{-\frac{1}{2}} \preceq \frac{8\gamma(k-1)}{\beta-1} \Big(1 + \frac{\Delta}{\beta-1} \big(1 + \frac{2}{\beta-1} \big) \Big).$$

Proof. Applying Lemma 52 and Lemma 55, we obtain

$$\mathsf{LHS} \preceq \Pi_\tau^{-\frac{1}{2}} \left(\frac{2(k-2)}{k} (a_{k-1} - a_k) \sum_{\omega \in \mathcal{C}'_{\tau,k}} A_\tau^\omega + \frac{8 a_k^2}{k^2} \cdot \sum_{\omega \in \mathcal{C}'_{\tau,k}} A_\tau^\omega \Pi_\tau^{-1} A_\tau^\omega \right) \Pi_\tau^{-\frac{1}{2}}.$$

Then by Lemma 53 and Lemma 56, we can bound above by

$$\mathsf{LHS} \preceq \frac{2}{k} \Pi_\tau^{-\frac{1}{2}} \Xi_\tau^\omega \Bigg((k-2)(a_{k-1} - a_k) \sum_{\omega \in \mathcal{C}'_{\tau,k}} \pi_\tau(\omega) (-\mathsf{Adj}_\tau + \frac{2(k-1)}{\beta-1} \mathsf{Id}_\tau) \tag{25}$$

$$+4\gamma a_k^2 \cdot \sum_{\omega \in \mathcal{C}'_{\tau,k}} \pi_\tau(\omega) \Big(\mathsf{Adj}_\tau^2 + \frac{4(k-1)^2}{(\beta-1)^2} \mathsf{Id}_\tau \Big) \Bigg) \Xi_\tau^\omega \Pi_\tau^{-\frac{1}{2}}, \tag{26}$$

where $\gamma = \left(1 + \frac{\Delta-1}{\beta-1}\right)^3 \frac{1}{\beta-1}$.

We want to find a sequence of $\{a_k\}$ so that the spectral radius of the following matrices $\tilde{A}_k$ appearing in the non-remainder terms in Equation (25) is small:

$$\tilde{A}_h := -(k-2)(a_k - a_{k-1}) \mathsf{Adj}_\tau + 4 a_k^2 \gamma \left(\mathsf{Adj}_\tau \right)^2.$$

Since the spectrum of Adj_τ is $\{-1,(k-1)\}$, the spectrum of $\tilde{A}_k$ is

$$\left\{(k-2)(a_{k-1}-a_k)+4\gamma a_k^2,\ -(k-1)(k-2)(a_{k-1}-a_k)+4\gamma(k-1)^2a_k^2\right\}.$$

Define

$$a_k=\frac{1}{1+4\gamma(k-2)}(2\le k\le\Delta).$$

Then we have

$$\rho(\tilde{A}_k)\le\frac{4\gamma(1+4\gamma(k-2))h}{(4\gamma(k-3)+1)(4\gamma(k-2)+1)^2}\le 4\gamma(k-1) \tag{27}$$

for $k\ge 3$. In particular, when $k=2$, $\rho(\tilde{A}_h)\le 4a_k^2\gamma(k-1)^2=4\gamma$, which is consistent with the above bound. So we have $\rho(\tilde{A}_k)\le 4\gamma(k-1)$ for $h\ge 1$. Note that $\Xi_\tau^\omega\le\frac{1}{\beta-1}\cdot\mathsf{Id}_\tau$, it then follows from Equation (27) and Lemma 57 that

$$\begin{aligned}&\frac{2}{k}\Pi_\tau^{-\frac12}\Xi_\tau^\omega\left(-(k-2)(a_{k-1}-a_k)\sum_{\omega\in\mathcal{C}'_{\tau,k}}\pi_\tau(\omega)\mathsf{Adj}_\tau+4\gamma a_k^2\sum_{\omega\in\mathcal{C}'_{\tau,k}}\mathsf{Adj}_\tau^2\right)\Xi_\tau^\omega\Pi_\tau^{-\frac12}\\&\le\frac{8\gamma(k-1)}{\beta-1}\mathsf{Id}_\tau.\end{aligned}\tag{28}$$

A direct calculation yields that

$$\begin{aligned}&\frac{2}{k}\Pi_\tau^{-\frac12}\Xi_\tau^\omega\left((k-2)(a_{k-1}-a_k)\sum_{\omega\in\mathcal{C}'_{\tau,k}}\pi_\tau(\omega)\frac{2(k-1)}{\beta-1}\mathsf{Id}_\tau+4\gamma a_k^2\cdot\sum_{\omega\in\mathcal{C}'_{\tau,k}}\pi_\tau(\omega)\frac{4(k-1)^2}{(\beta-1)^2}\mathsf{Id}_\tau\right)\Xi_\tau^\omega\Pi_\tau^{-\frac12}\\&\le\frac{8\gamma(k-1)^2}{(\beta-1)^2}\left(\frac{2}{\beta-1}+1\right)\mathsf{Id}_\tau\le\frac{8\gamma(k-1)\Delta}{(\beta-1)^2}\left(\frac{2}{\beta-1}+1\right)\mathsf{Id}_\tau.\end{aligned}\tag{29}$$

Combining Equation (28) and Equation (29) finishes the proof. □

A.6 Construction of B_τ

For τ of co-dimension $k>2$, we introduce coefficients $\{b_k\}_{3\le h\le\Delta}$ whose values will be determined later, and define B_τ as follows:

$$B_\tau(ec,ec)=b_k \tag{30}$$

for any $e\in K_\tau$ and all other entries are 0 When $k=2$, this is exactly the base case considered in Appendix A.2. According to (18), we have $b_1=\frac{1}{(\beta-1)^2}$.

Notice that if $\mathrm{codim}(\tau)=k\ge 3$, then

$$\mathbf{E}_{x\sim\pi_\tau}\left[\Pi_{\tau\cup\{x\}}b_k\right]=\pi_\tau(ec)^{-1}\sum_{x\in\mathcal{C}_\tau}\pi_\tau(x)\pi_{\tau\cup\{x\}}(ec)b_{k-1}=\sum_{x\in\mathcal{C}_\tau}\pi_{\tau\cup\{ec\}}(x)b_{k-1}=b_{k-1}. \tag{31}$$

where the second equality follows from the fact that $\pi_\tau(x)\pi_{\tau\cup\{x\}}(vc)=\pi_{\tau\cup\{vc\}}(x)\pi_\tau(vc)$.

By the discussion in the last section and the above definition, now eq. (22) becomes

$$(k-2)b_k-(k-1)b_{k-1}-2b_k^2\ge\frac{8\gamma(k-1)\Delta}{(\beta-1)^2}\left(\frac{2}{\beta-1}+1\right)$$

for all $3\le k\le\Delta$.

Assume $\beta \geq 11$. Since

$$
\begin{aligned}
\Pi_\tau^{-1/2} A_\tau \Pi_\tau^{-1/2} &= \frac{2a_k}{k} \cdot \Pi_\tau^{-1/2} \sum_{\omega \in \mathcal{C}'_{\tau,k}} \pi_\tau(\omega) A_\tau^\omega \Pi_\tau^{-1/2} \\
&= \frac{2a_k}{k} \cdot \Pi_\tau^{-1/2} \sum_{\omega \in \mathcal{C}'_{\tau,k}} \pi_\tau(\omega) \Xi_\tau^\omega (-\mathsf{Adj}_\tau + \mathcal{R}_\tau^\omega) \Xi_\tau^\omega \Pi_\tau^{-1/2} \\
&\preceq \frac{2a_k}{(\beta-1)} \left(1 + \frac{2(k-1)}{\beta - 1}\right) \mathsf{Id}_\tau \\
&\preceq \frac{1}{(\beta-1)} \frac{1 + \frac{2(k-1)}{\beta-1}}{1 + \frac{4(k-2)\left(1+\frac{\Delta-1}{\beta-1}\right)^3}{\beta-1}} \mathsf{Id}_\tau \\
&\preceq \frac{1}{\beta-1} \mathsf{Id}_\tau \\
&\preceq \frac{1}{10} \mathsf{Id}_\tau,
\end{aligned}
\tag{32}
$$

we have $M_\tau = \frac{\sum_i A_\tau + \Pi_\tau B_\tau}{k-1} \preceq \frac{k-1}{3k-1} \Pi_\tau$ as long as $B_\tau \preceq \left(\frac{(k-1)^2}{3k-1} - \frac{1}{10}\right) \mathsf{Id}_\tau$. We strengthen this constraint to $B_\tau \preceq \left(\frac{1}{5} - \frac{1}{10}\right) \mathsf{Id}_\tau = \frac{1}{10} \mathsf{Id}_\tau$.

For brevity, we denote $8\gamma\Delta\left(\frac{2}{\beta-1} + 1\right)$ by $C(\Delta)$ in the following calculation. Therefore, our constraints for $\{b_k\}_{1 \leq k \leq \Delta}$ are

$$
\begin{cases}
(k-2)b_k - (k-1)b_{k-2} \geq 2b_k^2 + \frac{C(\Delta)(k-1)}{(\beta-1)^2}, & 3 \leq k \leq \Delta; \\
b_k \leq \frac{1}{10}, & 2 \leq k \leq \Delta.
\end{cases}
\tag{$\blacktriangle$}
$$

It follows from Lemma 26 in [WZZ24] that there is a feasible solution of eq. ($\blacktriangle$) as long as $\beta \geq c\sqrt{\Delta} \log^2 \Delta + 2c$, where $c = \sqrt{20(1 + 2C(\Delta))}$. And the solution $b_k \leq \frac{1}{(\beta-1)^2}\left(1 + (6 + 16C(\Delta))\Delta \log^2 \Delta\right)$. Notice that if $\beta - 1 \geq \max\{\Delta, 10\}$, then $C(\Delta) \leq \frac{10}{\Delta - 1}$, and the solution exists if $\beta - 1 \geq 20 \log^2 \Delta + 2\sqrt{\frac{200}{\Delta-1}}$.

Putting all constraints to β together, we have

$$
\beta - 1 \geq \max\left\{\Delta, 10, 20 \log^2 \Delta + 2\sqrt{\frac{200}{\Delta - 1}}\right\},
$$

which can be unified to a single bound that $\beta \geq \Delta + 50$.

A.7 Proof of Lemma 46

Proof of Lemma 46. For any $\tau \in \mathcal{C}_{d-k}$ with $k \geq 2$, we construct the matrices A_τ and B_τ as Appendix A.4 and Appendix A.6. Then we have

$$
\rho(\Pi_\tau^{-1/2} A_\tau \Pi_\tau^{-1/2}) + \rho(B_\tau) \leq \frac{1}{\beta - 1} + \frac{1}{(\beta-1)^2} + \frac{(6 + 16C(\Delta))\Delta \log^2 \Delta}{(\beta - 1)^2}.
$$

When $\beta \geq \Delta + 50$, the above term is upper bounded by $\eta_\Delta := \frac{1 + (6 + \frac{160}{\Delta-1}) \log^2 \Delta}{\Delta} + \frac{1}{\Delta^2}$. Applying Proposition 49, we have

$$
\rho(P_\tau - \frac{k}{k-1} \mathbf{1}\pi_\tau^\top) \leq \rho(\Pi_\tau^{-1} M_\tau) \leq \frac{\eta_\Delta}{k-1}.
$$

Taking $\tau = \varnothing$, we obtain that

$$
\Pi P - \frac{d}{d-1} \pi_\tau \pi_\tau^\top \preceq \frac{\eta_\Delta}{d-1} \Pi. \qquad \square
$$