Decay of Correlation for Edge Colorings When $q > 3\Delta$

Zejia Chen
Shanghai Jiao Tong University, China

Yulin Wang
Shanghai Jiao Tong University, China

Chihao Zhang
Shanghai Jiao Tong University, China

Zihan Zhang
Graduate Institute for Advanced Studies, SOKENDAI, Tokyo, Japan

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 $q \geq 2\Delta - 1$.

2012 ACM Subject Classification Theory of computation → Generating random combinatorial structures

Keywords and phrases Strong Spatial Mixing, Edge Coloring, Approximate Counting

Digital Object Identifier 10.4230/LIPIcs.ICALP.2025.54

Category Track A: Algorithms, Complexity and Games

Related Version *Full Version*: https://arxiv.org/abs/2502.06586

Funding *Chihao Zhang*: The author is supported by National Natural Science Foundation of China No. 62372289.

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 [10, 1, 16, 3]. 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 [16] established the spectral independence property, a notion to measure the correlation in high-dimensional distributions [2], whenever $q > (2 + o(1))\Delta$ for general graphs and the work of [3] establish approximate tensorization of variance on trees when $q \geq \Delta + 1$. Note that Glauber dynamics is known to be reducible

© Zejia Chen, Yulin Wang, Chihao Zhang, and Zihan Zhang;
licensed under Creative Commons License CC-BY 4.0
52nd International Colloquium on Automata, Languages, and Programming (ICALP 2025).
Editors: Keren Censor-Hillel, Fabrizio Grandoni, Joël Ouaknine, and Gabriele Puppis
Article No. 54; pp. 54:1–54:18
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

when $q < 2\Delta$ on general graphs [13] 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 [9]. 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 [14].

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 [6] is a notion stronger than spectral independence, and thanks to recent work of [5], building on the machinery of Moitra [15], 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 $q > 3.618\Delta$ [4, 5] 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. [11, 8]). It is conjectured that SSM holds whenever $q \geq \Delta + 1$ and in [8], 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 [7] 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 [8]. 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 [1] in the way of [16] 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

[1] Theses quantities are referred to as "subdistributions" in [8]

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 6. Verification for technical lemmas are deferred to the full version of this paper.

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 $f\colon \Omega \to \mathbb{R}$ defined on a finite domain Ω, we use $\Big[f(x)\Big]_{x\in\Omega}$ 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.

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 $\lambda_i, 1 \leq i \leq k$ and distributions $\{\mu_i\}_{1\leq i\leq k}$, $\{\nu_i\}_{1\leq i\leq k}$ 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

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

The proof of Proposition 6 can be found in the full version.

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 $\partial_{\tau_1,\tau_2} = \{e \in \Lambda \mid \tau_1(e) \neq \tau_2(e)\}$, we have that*

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

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*

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

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

▶ Remark 9. 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 10.** *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:*

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

The proof of Lemma 10 is deferred to the full version of this paper. Especially, when $|F| = 1$, we have the following corollary.

▶ **Corollary 11.** *Let $(G = (V, E), \mathcal{L})$ be a β-extra list-edge-coloring instance. Then for any $e \in E, v \in V, a \in \mathcal{L}(e)$,*

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

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 $v_1, v_2, \ldots, v_d$. 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.

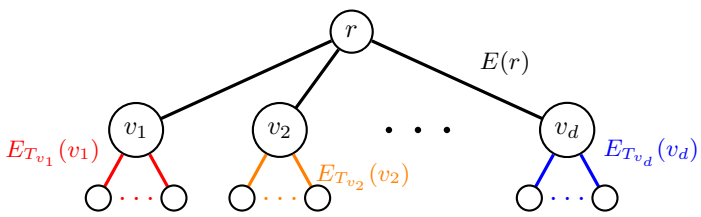

Figure 1 Brooms on a tree.

▶ **Lemma 12.** *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*:*

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

We provide the verification of Lemma 12 in the full version. 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 12, we have

$$\mathbf{p}_r(\pi) = f_\pi(\{\mathbf{p}_i\}_{i\in[d]}) = \frac{\prod_i \mathbf{p}_{v_i}(\overline{\pi(e_i)})}{\sum_{\rho \in C_r} \prod_i \mathbf{p}_{v_i}(\overline{\rho(e_i)})}. \tag{1}$$

We have the following lemmas similar to Lemma 10 on trees whose proofs are also in the full version. Note that we do not require $\mathbf{p}_i$'s to be distributions in Lemma 13 and Lemma 14.

▶ **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 $a \in [q]$,*

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

▶ **Lemma 14.** *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 $j \in [d]$,*

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

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 15.** *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 16.** *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 17** (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 18. 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 17. 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 16 immediately.

▶ **Lemma 19.** *Let $\lambda = 1+\varepsilon$ for some $\varepsilon > 0$ and $s \geq 2$. Then*

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

Proof of Theorem 16. Lemma 19 shows that $\sup_s \kappa_{s,\Delta,\lambda} \leq \frac{1}{2+\varepsilon}(1+2/\varepsilon) = 1/\varepsilon$. And

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

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

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

◀

The proof of Lemma 19 is based on a greedy one-step coupling of two marginal distributions with different pinning on a single edge.

Let $(G = (V, E), \mathcal{L})$ be a list-edge-coloring instance that fits into the constraints of $\kappa_{s,\Delta,\lambda}$ in Definition 17, $i \in E$ 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

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

The following lemma describes the greedy coupling we use.

▶ **Lemma 20.**

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

The first, second and third lines in the RHS correspond to the couplings with 1, 2 and 0 discrepancies respectively. With the marginal bounds in Section 3, the chance of 2 discrepancies is controlled by the chance of 0 discrepancy when the color lists are Δ-extra, and we have a constant rate of discrepancy decay and proves Lemma 20.

We leave the proofs of Lemmas 19 and 20 to the full version of this paper.

4.1 Proof of Theorem 15

Theorem 16 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 [5] 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 $A_E \in \mathbb{R}_{\geq 0}^{q\times q}$ and $A_V \in \mathbb{R}_{\geq 0}^{q}$. 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 [5] that we will use is as follows.

▶ **Theorem 21** ([5]). *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 15, we need the marginal lower bound from [12].

▶ **Lemma 22** (Corollary of Lemma 3 in [12]). *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 15. It is easy to verify that any β-extra edge coloring instance with $\beta \geq 1$ is permissive. From Theorem 16 and Lemma 22, 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 21, 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 23.** *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 23 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. Due to space limitations, proofs of technical lemmas and bounding the covariance matrices is addressed in the full version.

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. It derives from direct computation.

▶ **Proposition 24.** *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]

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 $\mathbf{m} \in \mathbb{R}^{C_{v_1} \times C_{v_2} \times \cdots \times C_{v_d}}$,

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

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

▶ **Proposition 25.** *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

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

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

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.

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

▶ **Definition 26.** *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:*

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

▶ **Definition 27.** *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,*

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

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

▶ **Definition 28.** *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

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

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

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

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

▶ **Theorem 29.** *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* $\beta \geq 1 + \frac{(1+\eta)\Delta}{\sqrt{1-2\delta}}$ *implies that*

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

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

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

To simplify notations, we let

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

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

$$\mathbf{A}(\pi, \tau) = \sqrt{f(\mathbf{p})(\pi)f(\mathbf{p})(\tau)} \sum_{i=1}^{d} \sum_{c_1,c_2\in\mathcal{L}(e_i)} V(i, \pi(e_i), c_1, \tau(e_i), c_2). \tag{2}$$

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

$$\sum_{\substack{\tau_1,\ldots,\tau_{k-1}\in C_r\\ \tau_0=\pi}} \prod_{j=1}^{k-1} f(\mathbf{p})(\tau_j)$$
$$\times \sum_{\substack{i_1,\ldots,i_{k-1}\in[d]\\ i_k=i}} \sum_{\substack{c_{1,1}\in\mathcal{L}(e_{i_1})\\ \vdots\\ c_{1,k}\in\mathcal{L}(e_{i_k})}} \sum_{\substack{c_{2,1}\in\mathcal{L}(e_{i_1})\\ \vdots\\ c_{2,k-1}\in\mathcal{L}(e_{i_{k-1}})\\ c_{2,k}=c}} \prod_{j=1}^{k-1} V(i_j, \tau_{j-1}(i_j), c_{1,j}, \tau_j(i_j), c_{2,j})$$
$$\times \langle \mathbf{b}^{\phi}_{i,c_{1,k}}, \mathbf{b}^{\phi}_{i,c}\rangle(\mathbb{1}\left[\tau_{k-1}(i) = c_{1,k}\right] - \mathbf{p}_r(i, c_{1,k})).$$

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

$$\mathbf{A}^k(\pi, \pi) = f(\mathbf{p})(\pi) \sum_{i=1}^{d} \sum_{c\in\mathcal{L}(e_i)} g_k(i, c)(\mathbb{1}\left[\pi(i) = c\right] - \mathbf{p}_r(i, c)). \tag{3}$$

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

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

Then we have that $g^{\top}_k = \alpha^{\top}_{\pi}\mathbf{B}^{k-1}$ where

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

It directly indicates that we can upper bound the maximum degree of $\mathbf{A}$ by the 2-norm of $\mathbf{B}$. In the following, we omit $(\mathbf{p})$ in $\mathbf{B}(\mathbf{p})$ for brevity if there is no ambiguity. Lemma 31 directly indicates that we can use the 2-norm of $\mathbf{B}$ to upper bound that of $\mathbf{A}$.

▶ **Lemma 32.** *Let $\mathbf{B} \in \mathbb{R}^{X \times X}$ and α_π denote the matrix and the vector defined in Lemma 31. Then we have that for any $k \geq 1$,*

$$\sum_{\pi \in C_r} \mathbf{A}^k(\pi, \pi) = \sum_{\pi \in C_r} f(\mathbf{p})(\pi) \alpha_\pi^\top \mathbf{B}^{k-1} \beta_\pi \tag{4}$$

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

The details in dimension reduction from $\mathbf{A}$ to $\mathbf{B}$ including the proofs of Lemma 30, Lemma 31 and Lemma 32 are in the full version.

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 33.** *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*

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

Then we have that

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

It can be verified via direct computation. Then we can establish Theorem 29 through the above proposition.The detailed proof of Proposition 33 and Theorem 29 are in the full version.

5.2 Strong spatial mixing via contraction

As Theorem 29 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. We need an upper bound for the covariance matrix defined in Definition 26, i.e. the spectral independence property, which can be found in the full version of this paper. Let $B_G(u, d)$ denote $\{u' \in V \mid \mathrm{dist}_G(u, u') = d\}$.

▶ **Lemma 34.** *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 $O\left(\frac{\log^2 \Delta}{\Delta}\right)$.

Now we are able to prove Theorem 23. Suppose there are two pinnings τ_1, τ_2 that differ on edges below $B_T(r, \ell)$. The key idea for proving Theorem 23 is to leverage the bound of Jacobian matrix from Theorem 29 on brooms $E_{T_u}(u)$ for any vertices $u \in B_T(r, k)$ with $k < \ell - 2$, while for the remaining two levels of brooms, specifically those corresponding to vertices in $B_T(r, \ell - 2)$ and $B_T(r, \ell - 1)$, we apply trivial bounds of Jacobian. Note that we "abandon" the lowest two levels of brooms due to the constraint for spectral independence from Lemma 34. The complete proof of Theorem 23 is included in the full version.

5.3 Worst-case scenario

Though Theorem 29 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) in the full version is the same as the pinning in Theorem 36 and it 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 35.** *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 $\mathbb{R}^{n\times n}$.*

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

▶ **Theorem 36.** *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 $q > \left(\frac{3+\sqrt{5}}{2} + o(1)\right)\Delta$.

For the worst case scenario, consider the following instance of edge coloring $(G, \mathcal{L})$ generated from a q-coloring instance by pinning:

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

Theorem 36 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 $q \approx 2.618\Delta$.

The formal process for computing $\mathbf{B}$ are in the full version. Theorem 36 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 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 37.** *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 38.** 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 $1, 2, \ldots, \Delta - 1$ and τ_2 be the pinning over Λ which only uses $\Delta, \ldots, 2\Delta - 2$. It is easy to verify that for any $\{u, v\} \in E \setminus \Lambda$ and $\mathrm{dep}(u) = \mathrm{dep}(v) + 1$,

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

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

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 39** (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*

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

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

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

The formal process for computing $\mathbf{B}$ are in the full version of this paper. Theorem 36 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 $q \approx 2.618\Delta$.

References

1 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. doi:10.1109/FOCS52979.2021.00024.

2 Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. *SIAM Journal on Computing*, 53(6):FOCS20–1, 2021.

3 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. doi:10.1137/1.9781611978322.184.

4 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. doi:10.1137/1.9781611978322.71.

5 Xiaoyu Chen, Weiming Feng, Heng Guo, Xinyuan Zhang, and Zongrui Zou. Deterministic counting from coupling independence. *arXiv preprint arXiv:2410.23225*, 2024. doi:10.48550/arXiv.2410.23225.

6 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. doi:10.1137/1.9781611977554.CH170.

7 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. doi:10.1137/1.9781611977912.178.

8 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. doi:10.1109/FOCS57990.2023.00053.

9 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.

10 Michelle Delcourt, Marc Heinrich, and Guillem Perarnau. The Glauber dynamics for edge-colorings of trees. *Random Structures & Algorithms*, 57(4):1050–1076, 2020. doi:10.1002/RSA.20960.

11 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.

12 David Gamarnik, Dmitriy Katz, and Sidhant Misra. Strong spatial mixing of list coloring of graphs. *Random Structures & Algorithms*, 46(4):599–613, 2015. doi:10.1002/RSA.20518.

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

14 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. doi:10.1016/0304-3975(86)90174-X.

15 Ankur Moitra. Approximate counting, the Lovász local lemma, and inference in graphical models. *Journal of the ACM*, 66(2):10:1–10:25, 2019. doi:10.1145/3268930.

16 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. doi:10.1145/3618260.3649724.