RAPID MIXING FROM SPECTRAL INDEPENDENCE BEYOND THE BOOLEAN DOMAIN

WEIMING FENG, HENG GUO, YITONG YIN, AND CHIHAO ZHANG

ABSTRACT. We extend the notion of spectral independence (introduced by Anari, Liu, and Oveis Gharan [ALO20]) from the Boolean domain to general discrete domains. This property characterises distributions with limited correlations, and implies that the corresponding Glauber dynamics is rapidly mixing.

As a concrete application, we show that Glauber dynamics for sampling proper q-colourings mixes in polynomial-time for the family of triangle-free graphs with maximum degree Δ provided $q \geq (\alpha^* + \delta)\Delta$ where $\alpha^* \approx 1.763$ is the unique solution to $\alpha^* = \exp(1/\alpha^*)$ and $\delta > 0$ is any constant. This is the first efficient algorithm for sampling proper q-colourings in this regime with possibly unbounded Δ. Our main tool of establishing spectral independence is the recursive coupling by Goldberg, Martin, and Paterson [GMP05].

1. Introduction

Let V be a set of variables, each of which takes values from a discrete domain of size $q \geq 2$. Sampling from a complicated joint distribution μ over the state space $[q]^V = \{0, 1, \dots, q-1\}^V$ is an important yet intricate computational task. The *Markov chain Monte Carlo (MCMC)* method is the most powerful and flexible technique to design efficient samplers. We will focus on *Glauber dynamics* in this paper, which is one of the simplest and most widely used Markov chains. In each step, it does the following:

(1) choose a variable uniformly at random;
(2) resample the value of the variable according to its marginal distribution conditioned on the values of all other variables.

Denote by μ_t the distribution of the state after t steps. It is usually straightforward to show that μ_t converges to the desired distribution μ as t tends to ∞. However, the more challenging task is to understand *how fast* the distance between μ_t and μ converges to 0. This rate of convergence is known as the *mixing time*. Many tools have been invented towards proving fast convergence, or the so-called rapid mixing property of Glauber dynamics. We refer the reader to [LP17] for a recent monograph on this topic.

Distributions of interest often have rich and complicated landscapes, which makes analysing the convergence rate of Glauber dynamics a long-standing challenge in theoretical computer science. To tackle this challenge, various techniques were introduced, such as canonical paths [JS89] and path coupling [BD97]. In a more recent line of work [DK17, Opp18, KO20, AL20], an interesting new method of analysing the mixing time emerged via the so-called "local-to-global" argument for high-dimensional expanders. This technique has played a central role in a few recent breakthrough results, such as uniform sampling of matroid bases [ALOV19, CGM19, ALOV20],[1] and a tight analysis for the hardcore model [ALO20] and more generally for anti-ferromagnetic 2-spin systems [CLV20].

(Weiming Feng, Yitong Yin) State Key Laboratory for Novel Software Technology, Nanjing University, 163 Xianlin Avenue, Nanjing, Jiangsu Province, China. E-mail: fengwm@smail.nju.edu.cn, yinyt@nju.edu.cn

(Heng Guo) School of Informatics, University of Edinburgh, Informatics Forum, Edinburgh, EH8 9AB, United Kingdom. E-mail: hguo@inf.ed.ac.uk

(Chihao Zhang) John Hopcroft Center for Computer Science, Shanghai Jiao Tong University, 800 Dongchuan Road, Minhang District, Shanghai, China. E-mail: chihao@sjtu.edu.cn

This research was supported by the National Key R&D Program of China 2018YFB1003202 and the NSFC Nos. 61722207, 61672275 and 61902241.

[1] The bases exchange walk for matroids can be viewed as Glauber dynamics as follows. Consider r variables, where r is the rank of the matroid. Each variable can take values from the ground set subject to the matroid constraint. Each bases exchange move is exactly resampling the value of a randomly chosen variable conditioned on the assignment of all other variables.

Of particular interest to us is the work of Anari, Liu, and Oveis Gharan [ALO20]. In order to apply the result of Alev and Lau [AL20], they introduced *spectral independence*, which is the property that the correlation matrix of μ and all of its conditional distributions have bounded maximum eigenvalues. They focused on the $q = 2$ case. Formally, for each feasible[2] $\sigma_\Lambda \in \{0,1\}^\Lambda$ where $\Lambda \subseteq V$, Anari, Liu and Oveis Gharan defined a *signed pairwise influence matrix* $I_\mu^{\sigma_\Lambda}$ by $I_\mu^{\sigma_\Lambda}(u,v) \triangleq (\mu_v^{\sigma_\Lambda,u\leftarrow 1}(1) - \mu_v^{\sigma_\Lambda,u\leftarrow 0}(1)) \cdot \mathbf{1}[u \neq v]$ for all $u, v \in V \setminus \Lambda$, where $\mu_v^{\sigma_\Lambda,u\leftarrow i}(i = 0,1)$ is the marginal distribution on v induced from μ conditional on the configuration on Λ fixed as σ_Λ and that u is fixed to i. In [ALO20], a distribution μ over $\{0,1\}^V$ is said to be spectrally independent if for any $\Lambda \subseteq V$, any feasible $\sigma_\Lambda \in \{0,1\}^\Lambda$, the maximum eigenvalue $\lambda_{\max}(I_\mu^{\sigma_\Lambda})$ can be upper bounded appropriately. They proved that the spectral independence property implies rapid mixing of Glauber dynamics. Using this tool, they confirmed a long-standing conjecture: Glauber dynamics for the Gibbs distribution of the hardcore model is rapidly mixing up to the uniqueness threshold. Later on, Chen, Liu and Vigoda [CLV20] further extended the mixing results to general antiferromagnetic 2-spin systems.

Despite the success in the Boolean domain, the machinery developed by Anari, Liu and Oveis Gharan does not handle many important distributions, such as the Gibbs distribution of Potts models where $q > 2$ can be any positive integer. Therefore a natural question is whether the approach developed in [ALO20, CLV20], or more specifically the notion of spectral independence, can be generalised beyond the Boolean domain. We note two interconnected difficulties for this task: (1) when $q > 2$, there are many non-equivalent choices for the definition of influence between two variables $u, v \in V$; and (2) it is not clear whether the elegant connection [ALO20, Theorem 3.1] between the "local" random walks of [AL20] and the spectrum of the influence matrix still holds beyond the Boolean domain.

Our first contribution is to introduce the following generalised influence matrix. This definition allows us to recover the part relevant to rapid mixing in the aforementioned result [ALO20, Theorem 3.1] for the more general setting.

Definition 1.1 (Influence Matrix). Let μ be a distribution over $[q]^V$. Fix any $\Lambda \subseteq V$ and any feasible $\sigma_\Lambda \in [q]^\Lambda$. For any distinct $u, v \in V \setminus \Lambda$, we define the *(pairwise) influence* of u on v by

$$\Psi_\mu^{\sigma_\Lambda}(u,v) \triangleq \max_{i,j \in \Omega_u^{\sigma_\Lambda}} d_{\mathrm{TV}}\left(\mu_v^{\sigma_\Lambda,u\leftarrow i}, \mu_v^{\sigma_\Lambda,u\leftarrow j}\right), \tag{1}$$

where $\Omega_u^{\sigma_\Lambda} \triangleq \left\{i \in [q] \mid \mu_u^{\sigma_\Lambda}(i) > 0\right\}$ denotes the set of possible values of u given condition σ_Λ, $d_{\mathrm{TV}}(\cdot,\cdot)$ denotes the total variation distance between two distributions, and for $c = i$ or j, $\mu_v^{\sigma_\Lambda,u\leftarrow c}$ is the marginal distribution on v induced from μ conditional on the configuration on Λ fixed as σ_Λ and that u is fixed to c.

Furthermore, let $\Psi_\mu^{\sigma_\Lambda}(u,v) \triangleq 0$ for $u = v$ and write $\Psi_\mu^{\sigma_\Lambda}$ for the *(pairwise) influence matrix* whose entries are given by $\Psi_\mu^{\sigma_\Lambda}(u,v)$.

In our definition, $\Psi_\mu^{\sigma_\Lambda}(u,v)$ is the maximum influence on v caused by a single disagreement on u conditional on σ_Λ. The entries of $\Psi_\mu^{\sigma_\Lambda}$ are total variation distances and are therefore non-negative. We remark that our definition is not identical to the original influence matrix $I_\mu^{\sigma_\Lambda}$ in [ALO20] even in the Boolean domain since the latter is signed. Nevertheless, if $q = 2$, it holds that $\Psi_\mu^{\sigma_\Lambda}(u,v) = \left|I_\mu^{\sigma_\Lambda}(u,v)\right|$.

With the definition of the influence matrix, we define spectral independence for general $q \geq 1$ as follows.

Definition 1.2 (Spectral Independence). We say a distribution μ over $[q]^V$, where $n = |V|$, is (C, η)-*spectrally independent*, if every $0 \leq k \leq n-2$, $\Lambda \subseteq V$ of size k and any feasible $\sigma_\Lambda \in [q]^\Lambda$, the spectral radius $\rho\left(\Psi_\mu^{\sigma_\Lambda}\right)$ of the influence matrix $\Psi_\mu^{\sigma_\Lambda}$ satisfies

$$\rho\left(\Psi_\mu^{\sigma_\Lambda}\right) \leq C \quad \text{and} \quad \frac{\rho\left(\Psi_\mu^{\sigma_\Lambda}\right)}{n-k-1} \leq \eta.$$

[2]A configuration $\sigma \in \{0,1\}^V$ is *feasible* if $\mu(\sigma) > 0$. A partial configuration $\sigma_\Lambda \subseteq \{0,1\}^\Lambda$ for $\Lambda \subseteq V$ is *feasible* if it can be extended to a feasible configuration.

Consider the Glauber dynamics for a general distribution μ and let $P_{\text{Glauber}} \in \mathbb{R}_{\geq 0}^{\Omega\times\Omega}$ be its transition matrix. It is well-known that the Glauber dynamics converges to stationary distribution μ when P_{Glauber} is irreducible, see e.g. [LP17].

The rate of convergence of Glauber dynamics is captured by the *mixing time*, defined as:

$$\forall\, 0 < \varepsilon < 1, \quad T_{\text{mix}}(\varepsilon) = \max_{\boldsymbol{x}_0 \in \Omega} \min \left\{ t \mid d_{\text{TV}} \left(P^t_{\text{Glauber}}(\boldsymbol{x}_0, \cdot), \mu \right) \leq \varepsilon \right\}.$$

Our main theorem states that the Glauber dynamics for μ is rapidly mixing if μ is spectrally independent.

Theorem 1.3. *Let μ be a distribution over $[q]^V$. If μ is (C,η)-spectrally independent for $C \geq 0$ and $0 \leq \eta < 1$, then the Glauber dynamics for μ has mixing time*

$$T_{mix}(\varepsilon) \leq \frac{n^{1+2C}}{(1-\eta)^{2+2C}} \left(\log \frac{1}{\varepsilon \mu_{\min}} \right),$$

where $n = |V|$ and $\mu_{\min} \triangleq \min\{\mu(\sigma) \mid \sigma \in [q]^V \wedge \mu(\sigma) > 0\}$.

This generalises a similar result by Anari, Liu and Oveis Gharan [ALO20] for $q = 2$. Their proof is based on a linear algebra argument which completely characterises the spectrum of their influence matrix in terms of the spectrum of the local random walks, so that the result of Alev and Lau [AL20] applies. However, it is not clear whether a similar argument exists for general q. Instead our main contribution is a new coupling based argument to connect spectral independence to rapid mixing of Glauber dynamics, which holds for any $q \in \mathbb{N}$. To be more specific, we also utilises the result of Alev and Lau [AL20]. We show that the second largest eigenvalue of the local random walk can be bounded in terms of the spectral radius of our influence matrix (see Lemma 3.6). In order to relate these two quantities, we employed a coupling analysis reminiscent of the work of Hayes [Hay06]. See Section 3 for an overview of our proof.

To apply our result, one needs to verify the spectral independence property, which is equivalent to bound the spectral radius of an influence matrix. This is not an easy task in general. A more tractable way is to bound the induced 1-norm or the induced ∞-norm of the influence matrix, which are upper bounds of its spectral radius.

Corollary 1.4. *Let μ be a distribution over $[q]^V$, where $n = |V|$. If there exist two constants $C \geq 0$ and $0 \leq \eta < 1$ such that for every $0 \leq k \leq n-2$, $\Lambda \subseteq V$ of size k and any feasible $\sigma_\Lambda \in [q]^\Lambda$, the influence matrix $\Psi_\mu^{\sigma_\Lambda}$ satisfies one of following two conditions:*

- ***bounded all-to-one influence**:*

$$\left\| \Psi_\mu^{\sigma_\Lambda} \right\|_1 \triangleq \max_{v \in V\setminus\Lambda} \sum_{u \in V\setminus\Lambda} \Psi_\mu^{\sigma_\Lambda}(u,v) \leq \min\left\{C, \eta(n-k-1)\right\}$$

- ***bounded one-to-all influence**:*

$$\left\| \Psi_\mu^{\sigma_\Lambda} \right\|_\infty \triangleq \max_{u \in V\setminus\Lambda} \sum_{v \in V\setminus\Lambda} \Psi_\mu^{\sigma_\Lambda}(u,v) \leq \min\left\{C, \eta(n-k-1)\right\}$$

then the Glauber dynamics for μ has mixing time

$$T_{mix}(\varepsilon) \leq \frac{n^{1+2C}}{(1-\eta)^{2+2C}} \left(\log \frac{1}{\varepsilon \mu_{\min}} \right),$$

where $\mu_{\min} \triangleq \min\{\mu(\sigma) \mid \sigma \in [q]^V \wedge \mu(\sigma) > 0\}$.

The conditions in Corollary 1.4 have been previously established for the hardcore model [ALO20] (all-to-one influence) and more generally for anti-ferromagnetic 2-spin systems [CLV20] (one-to-all influence).[3] Such conditions are quite natural for *Gibbs distributions* induced by *q-spin systems*. Roughly speaking, a

[3] Although in [ALO20] and [CLV20], the corresponding conditions were established for the signed influence matrix $I_\mu^{\sigma_\Lambda}$, they are still applicable to our Corollary 1.4 since $\|\Psi_\mu^{\sigma_\Lambda}\|_1 = \|I_\mu^{\sigma_\Lambda}\|_1$ and $\|\Psi_\mu^{\sigma_\Lambda}\|_\infty = \|I_\mu^{\sigma_\Lambda}\|_\infty$ when $q = 2$.

q-spin system is defined on a graph $G = (V, E)$, where vertices represent random variables that take values in $[q]$, and edges model pairwise interactions. Both "bounded all-to-one influence" and "bounded one-to-all influence" can be viewed as some forms of the *spatial mixing* or *correlation decay* property of the q-spin systems. This property roughly says that the influence between two vertices decays rapidly with respect to their distance in the graph G and has been widely exploited to design efficient samplers for the Gibbs distribution. For antiferromagnetic 2-spin systems, the rapid mixing regimes obtained by [ALO20, CLV20] match the best known correlation decay results [Wei06, LLY13, GL18, SS20]. We show that our notion of spectral independence can also be used to obtain efficient sampling algorithms up to known correlation decay regime for multi-spin systems [GMP05, GKM15].

1.1. **Application to spin systems.** As a concrete application, we consider an important multi-spin system, i.e. proper graph q-colourings, or equivalently the anti-ferromagnetic Potts model with the temperature going to negative infinity. A graph q-colouring instance is specified by $(G, [q])$, where $[q] = \{0, 1, \dots, q-1\}$ is a set of colours and $G = (V, E)$ is a simple undirected graph. A proper colouring $X \in [q]^V$ assigns each vertex $v \in V$ a colour $X_v \in [q]$ such that $X_u \neq X_v$ for all $\{u, v\} \in E$. Let Ω denote the set of all proper colourings and μ denote the uniform distribution over Ω. In this concrete setting, the Glauber dynamics works as follows. The chain starts from an arbitrary proper colouring $X \in \Omega$, and in each step, it does:

(1) pick a vertex $v \in V$ uniformly at random;
(2) update X_v by choosing a colour from $[q] \setminus \{X_u \mid \{v, u\} \in E\}$ uniformly at random.

When $q \geq \Delta + 2$, the chain converges to μ for any initial colouring X. However, it is a notorious open problem that whether the condition $q \geq \Delta + 2$ also guarantees rapid mixing. We make some progress towards this problem by proving the following result.

Let $\alpha^* \approx 1.763\dots$ be the positive root of the equation $x^x = \mathrm{e}$. Using Theorem 1.3, we obtain the following.

Theorem 1.5. *Let $\delta > 0$ be a constant. For any graph colouring instance $(G, [q])$ where G is triangle-free and $q \geq (\alpha^* + \delta)\Delta$, the Glauber dynamics on $(G, [q])$ has mixing time*

$$T_{mix}(\varepsilon) \leq \left(9\mathrm{e}^5 n\right)^{2+9/\delta} \log\left(\frac{q}{\varepsilon}\right),$$

where n is the number of vertices in G and $\Delta \geq 3$ is the maximum degree of G.

While Theorem 1.5 is stated for graph q-colouring instances, the mixing time upper bound holds for the more general list colouring problem (see Theorem 6.1). In fact, the same rapid mixing bound holds as long as the marginal probabilities are always appropriately upper bounded. This is formally stated by Condition 6.2 and Theorem 6.3.

It is instructive to compare Theorem 1.5 with the vast body of literature on this problem. The study was initiated by the pioneering work of Jerrum [Jer95] and of Salas and Sokal [SS97], who showed $O(n \log n)$ mixing time if $q \geq (2 + \delta)\Delta$. So far, in general graphs, the best result is the $O(n^2)$ mixing time when $q \geq (\frac{11}{6} - \varepsilon_0)\Delta$ for some absolute small constant $\varepsilon_0 > 0$ [Vig00, CDM+19]. For restricted families of graphs, there is a long line of work that studied the mixing time of Glauber dynamics under various conditions [DF01, Hay03, HV03, GMP05, HV06, Mol04, Hay13, DFHV13]. A few results most relevant to Theorem 1.5 are listed in Table 1. The triangle-free condition, or more generally the requirement on the girth of the graph, has played an important role to improve the dependency of q and Δ. For a more complete picture, we refer the reader to the survey [FV07].

In addition to algorithms based on Glauber dynamics mentioned above, using the reduction from sampling to counting [JVV86], one can obtain sampling algorithms from approximate counting algorithms [GK12, LY13, LSS19]. The current best FPTAS for counting q-colourings is given by Liu, Sinclair and Srivastava [LSS19]. The algorithm has running time $n^{f(\Delta)}$ where $f(\Delta) = \exp(\mathrm{poly}(\Delta))$ in (1) general graphs with $q \geq 2\Delta$; (2) triangle-free graphs with $q \geq (\alpha^* + \delta)\Delta + \beta(\delta)$. Therefore, their algorithm does not run in polynomial-time if $\Delta = \omega(1)$.

	Regime	Girth	Other requirement	Mixing time $T_{\mathsf{mix}}(\frac{1}{4\mathrm{e}})$
[GMP05]	$q > \alpha^*\Delta$	≥ 4	$\Delta = O(1)$ and neighbourhood amenable	$O(n^2)$
[HV06]	$q \geq (\alpha^* + \delta)\Delta$	≥ 4	$\Delta = \Omega(\log n)$	$O(\frac{n}{\delta}\log n)$
[DFHV13]	$q \geq (\alpha^* + \delta)\Delta$	≥ 5	$\Delta \geq \Delta_0(\delta)$	$O(\frac{n}{\delta}\log n)$
This work	$q \geq (\alpha^* + \delta)\Delta$	≥ 4	–	$O((9\mathrm{e}^5 n)^{2+9/\delta}\log q)$

Table 1. Mixing time results for sampling proper graph q-colourings.

Compared with previous results, we achieved a $q \geq (\alpha^* + \delta)\Delta$ bound in triangle-free graphs without any additional requirements. The condition in Theorem 1.5 matches the best known *strong spatial mixing* regime for graph proper q-colourings [GMP05, GKM15].

Theorem 1.5 is proved via verifying the sufficient condition in Corollary 1.4. In fact, we apply the recursive coupling technique introduced by Goldberg, Martin and Paterson [GMP05] to bound the total influence caused by one vertex (the "one-to-all infleunce"), namely to bound the induced ∞-norm of the influence matrix $\Psi_{\mu}^{\sigma_\Lambda}$, while the original approach in [GMP05] only provides bounds for "one-to-one influence". Comparing to the traditional path coupling analysis, the power of the spectral independence approach lies in the fact that we can avoid considering the worst case scenario for the influence matrix in Definition 1.1. For path coupling, to avoid the worst case analysis one needs to establish so-called local uniformity [Hay13], which is difficult and causes various technical conditions in the results listed above. In contrast, the method based on the spectral independence bypasses this obstacle.

The downside of our result, similar to those of [ALO20, CLV20], is that the running time has a high exponent depending on how close the parameters are to the threshold. Nonetheless, unlike the algorithm of [LSS19], our exponent remains a constant even if $\Delta = \omega(1)$, as long as we are below the threshold.

Finally, we remark that our refinement of recursive coupling argument might find applications in other problems. Armed with our notion of spectral independence, we essentially proved that the success of recursive coupling implies rapid mixing of Glauber dynamics for *any* graph. This form of algorithmic implication was only known for special families of graphs like amenable graphs [GMP05] and planar graphs [YZ13] before.

2. Preliminaries

2.1. **Linear algebra**. Let $v \in \mathbb{C}^n$ be an n-dimensional vector. For any integer $p \geq 1$, the ℓ_p-norm of v is defined by $\|v\|_p = (\sum_{i=1}^n |v_i|^p)^{1/p}$. Let $A \in \mathbb{C}^{n\times n}$ be a matrix. For any integer $p \geq 1$, the induced ℓ_p-norm of A is defined by $\|A\|_p = \sup_{v\in\mathbb{C}^n : \|v\|_p = 1} \|Av\|_p$. Let $\lambda_1, \lambda_2, \ldots \lambda_n \in \mathbb{C}$ be the eigenvalues of A. The *spectral radius* of A is defined by $\rho(A) \triangleq \max_{1\leq i\leq n} |\lambda_i|$. The following relation is well-known.

Proposition 2.1 ([HJ12, Theorem 5.6.9. & Corollary 5.6.14]). *Let $A \in \mathbb{C}^{n\times n}$ be a matrix. For any integer $p \geq 1$, it holds that $\rho(A) \leq \|A\|_p$ and $\lim_{k\to\infty} \|A^k\|_p^{1/k} = \rho(A)$.*

2.2. **Total variation distance and coupling**. Let μ and ν be two distributions over state space Ω. The *total variation distance* between μ and ν is defined by

$$d_{\mathrm{TV}}(\mu, \nu) \triangleq \frac{1}{2}\sum_{\boldsymbol{x}\in\Omega} |\mu(\boldsymbol{x}) - \nu(\boldsymbol{x})| .$$

A coupling of μ and ν is a joint distribution $(X, Y) \in \Omega \times \Omega$ such that the marginal distribution of X is μ and the marginal distribution of Y is ν. The following result is the well-known coupling inequality.

Proposition 2.2 ([LP17, Proposition 4.7]). *Let μ and ν be two distributions over state space Ω. For any coupling (X, Y) of μ and ν, it holds that*

$$d_{\mathrm{TV}}(\mu, \nu) \leq \mathbf{Pr}[X \neq Y].$$

Furthermore, there exists an optimal coupling (X, Y) *such that* $d_{\mathrm{TV}}(\mu, \nu) = \mathbf{Pr}[X \neq Y]$.

2.3. **Markov chain and mixing time.** Let Ω be a finite set which is the state space. A Markov chain $(X_t)_{t\geq 0}$ on Ω is specified by transition matrix $P \in \mathbb{R}_{\geq 0}^{\Omega\times\Omega}$. We often identify the transition matrix with the corresponding Markov chain. The Markov chain is *irreducible* if for any $x, y \in \Omega$, there is a $t \geq 0$ such that $P^t(x, y) > 0$. The Markov chain is *aperiodic* if for any $x \in \Omega$, $\gcd\{t > 0 \mid P^t(x, x) > 0\} = 1$. A distribution π (viewed as a row vector) on Ω is *stationary* with respect to a Markov chain P if $\pi P = \pi$. If a Markov chain P is irreducible and aperiodic, then P has a unique stationary distribution. A Markov chain is *reversible* with respect to a distribution π if the following *detailed balance* condition holds

$$\forall x, y \in \Omega, \quad \pi(x)P(x, y) = \pi(y)P(y, x), \tag{2}$$

which implies that π is a stationary distribution of P. All Markov chains considered in this paper are reversible. In the following we state a few well-known spectral properties of reversible Markov chains.

Proposition 2.3 ([LP17, Lemma 12.2]). *Let Ω be a finite set with $|\Omega| = n$. Let π be a distribution with support Ω. Let $P \in \mathbb{R}_{\geq 0}^{\Omega\times\Omega}$ be the transition matrix of a Markov chain that is reversible with respect to π. Then*

- *P has n real eigenvalues $1 = \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots \lambda_n \geq -1$;*
- *there exist real eigenvectors $f_1, f_2, \ldots, f_n \in \mathbb{R}^\Omega$ such that $Pf_i = \lambda_i f_i$ for all $1 \leq i \leq n$, $f_1 = \vec{1}$ is a one-vector, and for any $1 \leq i, j \leq n$,*

$$\sum_{x\in\Omega} f_i(x) f_j(x) \pi(x) = \mathbf{1}[i = j].$$

We remark that Proposition 2.3 holds if P is reversible to π and the support of π is Ω. It does not require P to be irreducible. The following proposition bounds the mixing time of Markov chain.

Proposition 2.4 ([LP17, Theorem 12.4]). *Let Ω be a state space with $|\Omega| = n \geq 2$. Let π be a distribution with support Ω. Let $P \in \mathbb{R}_{\geq 0}^{\Omega\times\Omega}$ be the transition matrix of a Markov chain that is reversible with respect to π. Let $1 = \lambda_1 \leq \lambda_2 \leq \ldots \lambda_n \leq -1$ be the real eigenvalues of P. Define the the* absolute *spectral gap*

$$\gamma_\star \triangleq 1 - \lambda_\star = 1 - \max\{|\lambda_i| \mid 2 \leq i \leq n\}.$$

Let $\pi_{\min} \triangleq \min_{x\in\Omega} \pi(x)$. If $\gamma_\star > 0$, then it holds that

$$\forall\, 0 < \varepsilon < 1, \quad T_{mix}(\varepsilon) \leq \frac{1}{\gamma_\star}\left(\log \frac{1}{\varepsilon \pi_{\min}}\right),$$

where $T_{mix}(\varepsilon) \triangleq \max_{x\in\Omega} \min\{t \mid d_{\mathrm{TV}}(P^t(x, \cdot), \pi) \leq \varepsilon\}$ denotes the mixing time of Markov chain.

Note that the reversible chain P is irreducible and aperiodic if the absolute spectral gap $\gamma_\star > 0$. Proposition 2.4 says that P converges to the unique stationary distribution π rapidly if $\gamma_\star$ is bounded away from 0. See [LP17, Theorem 12.4] for a formal proof of Proposition 2.4.

We will use the following proposition to bound the absolute value of the second largest eigenvalue of P. Similar results appeared in [LP17, Theorem 13.1] and [Che98].

Proposition 2.5. *Let Ω be a state space with $n = |\Omega| \geq 2$. Let π be a distribution with support Ω. Let $P \in \mathbb{R}_{\geq 0}^{\Omega\times\Omega}$ be the transition matrix of a Markov chain that is reversible with respect to π. Then the second largest eigenvalue of P satisfies*

$$\forall t \geq 1, \quad |\lambda_2|^t \leq d(t) \triangleq \max_{x,y\in\Omega} d_{\mathrm{TV}}\left(P^t(x, \cdot), P^t(y, \cdot)\right).$$

Proof. Define a distance function δ on Ω as:

$$\forall x, y \in \Omega: \quad \delta(x,y) \triangleq \mathbf{1}[x \neq y].$$

For every function $f : \Omega \to \mathbb{R}$, define its Lipschitz constant with respect to δ as

$$\mathrm{Lip}(f) \triangleq \max_{x,y\in\Omega: x\neq y} \frac{|f(x)-f(y)|}{\delta(x,y)}.$$

Fix a pair $x, y \in \Omega$, we use $C(x,y)$ to denote the optimal coupling between $P^t(x,\cdot)$ and $P^t(y,\cdot)$. Note that

$$P^t f(x) = \mathbf{E}_{X\sim P^t(x,\cdot)}\left[f(X)\right].$$

Then for any $t \geq 1$, any function $f : \Omega \to \mathbb{R}$ and any $x, y \in \Omega$,

$$\left|P^t f(x) - P^t f(y)\right| = \left|\mathbf{E}_{(X,Y)\sim C(x,y)}\left[f(X)-f(Y)\right]\right| \leq \mathbf{E}_{(X,Y)\sim C(x,y)}\left[|f(X)-f(Y)|\right],$$

where the equality holds due to linearity of expectation. Then for any $t \geq 1$, any f and any x, y,

$$\left|P^t f(x) - P^t f(y)\right| \leq \mathrm{Lip}(f)\mathbf{Pr}_{(X,Y)\sim C(x,y)}\left[X \neq Y\right] = \mathrm{Lip}(f) d_{\mathrm{TV}}\left(P^t(x,\cdot), P^t(y,\cdot)\right) \leq \mathrm{Lip}(f) d(t).$$

Note that the inequality above holds for all $x, y \in \Omega$. It implies that $\mathrm{Lip}(P^t f) \leq \mathrm{Lip}(f) d(t)$.

Recall $|\Omega| = n$. Let $f_1, f_2, \ldots, f_n \in \mathbb{R}^\Omega$ be the eigenvectors in Proposition 2.3, where $f_1 = \vec{1}$. Let $f = f_2$ be the eigenvector of λ_2, we have

$$|\lambda_2|^t \cdot \mathrm{Lip}(f_2) = \mathrm{Lip}(\lambda_2^t f_2) = \mathrm{Lip}(P^t f_2) \leq \mathrm{Lip}(f_2) d(t).$$

Note that $f_2 \neq \vec{0}$. Since $f_1 = \vec{1}$ is a constant vector and $\sum_{x\in\Omega} f_1(x)\pi(x)f_2(x) = \sum_{x\in\Omega}\pi(x)f_2(x) = 0$, vector f_2 can not be a constant vector. Thus, $\mathrm{Lip}(f_2) > 0$, we have $|\lambda_2|^t \leq d(t)$ for all $t \geq 1$. □

One powerful technique to bound $d_{\mathrm{TV}}\left(P^t(x,\cdot), P^t(y,\cdot)\right)$ is the *coupling of Markov chain*. A coupling of P is a joint random process $(X_t, Y_t)_{t\geq 0}$ such that $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ individually follow the transition rule of P, and if $X_k = Y_k$, then $X_t = Y_t$ for all $t \geq k$. The following result follows from Proposition 2.2.

Proposition 2.6. *Let P be a Markov chain on state space Ω with a stationary distribution π. Let $X \in \Omega$ be a state. Let $(X_t, Y_t)_{t\geq 0}$ be a coupling of Markov chain such that $X_0 = x_0$ and $Y_0 = y_0$. Then*

$$\forall t \geq 1, \quad d_{\mathrm{TV}}\left(P^t(x_0,\cdot), P^t(y_0,\cdot)\right) \leq \mathbf{Pr}\left[X_t \neq Y_t\right].$$

3. Proof overview

In this section, we overview our proof of the main theorem (Theorem 1.3). We actually prove a slightly more general result. We first introduce the following definition of $(\eta_0, \eta_1, \ldots, \eta_{n-2})$-spectral independence, which is analogous to a similar notion in [ALO20].

Definition 3.1 ($(\eta_0, \eta_1, \ldots, \eta_{n-2})$-Spectral Independence)**.** We say a distribution μ over $[q]^V$ with $|V| = n$ is $(\eta_0, \eta_1, \ldots, \eta_{n-2})$-*spectrally independent*, if for every $0 \leq k \leq n-2$, $\Lambda \subseteq V$ of size k and any feasible $\sigma_\Lambda \in [q]^\Lambda$, the spectral radius $\rho\left(\Psi_\mu^{\sigma_\Lambda}\right)$ of influence matrix $\Psi_\mu^{\sigma_\Lambda}$ satisfies

$$\rho\left(\Psi_\mu^{\sigma_\Lambda}\right) \leq \eta_k.$$

Since Glauber dynamics is reversible with respect to μ, its transition matrix has real eigenvalues. The following theorem gives a lower bound on its spectral gap when μ is spectrally independent.

Theorem 3.2. *Let μ be a distribution over $[q]^V$, where $n = |V|$. Let $\eta_0, \eta_1, \ldots, \eta_{n-2}$ be a sequence where $0 \leq \eta_k < n-k-1$ for all $0 \leq k \leq n-2$. If μ is $(\eta_0, \eta_1, \ldots, \eta_{n-2})$-spectrally independent, then the Glauber dynamics for μ has spectral gap*

$$1 - \lambda_2(P_{\mathrm{Glauber}}) \geq \frac{1}{n}\prod_{k=0}^{n-2}\left(1 - \frac{\eta_k}{n-k-1}\right),$$

where $\lambda_2(P_{\mathrm{Glauber}})$ is the second largest eigenvalue of transition matrix P_{Glauber}.

We prove this theorem on general domain of size $q \ge 2$. The main theorem (Theorem 1.3) is a corollary of Theorem 3.2, because if μ is (C,η)-spectrally independent, then μ is $(\eta_0,\eta_1,\dots,\eta_{n-2})$-spectrally independent with $\eta_k = \min\{C, \eta(n-k-1)\}$. We first give a proof overview of Theorem 3.2, then prove main theorem (Theorem 1.3) via Theorem 3.2 in Section 3.3.

3.1. **Glauber dynamics and local random walks.** To prove Theorem 3.2, we first interpret the Glauber dynamics on $[q]^V$ as a down-up random walk on simplicial complexes. Then we apply the local-to-global theorem due to Alev and Lau [AL20] to reduce the task of analysing Glauber dynamics (a global random walk) to the task of analysing local random walks. Similar routines have been applied in several previous works [ALOV19, CGM19, ALO20, CLV20].

Definition 3.3 (Local Random Walk). For any subset $\Lambda \subseteq V$, any feasible partial configuration $\sigma_\Lambda \in [q]^\Lambda$, define local random walk P_{σ_Λ} on $U_{\sigma_\Lambda} = \{(u,c) \in \overline{\Lambda} \times [q] \mid \mu_u^{\sigma_\Lambda}(c) > 0\}$ as

$$\forall (u,i),(v,j) \in U_{\sigma_\Lambda}, \quad P_{\sigma_\Lambda}((u,i),(v,j)) \triangleq \frac{\mathbf{1}[u \neq v]}{|V|-|\Lambda|-1}\mu_v^{\sigma_\Lambda, u\leftarrow i}(j), \tag{3}$$

where $\overline{\Lambda} = V \setminus \Lambda$, and $\mu_v^{\sigma_\Lambda, u\leftarrow i}$ is the marginal distribution on v induced from μ conditional on the configuration on Λ fixed as σ_Λ and that u is fixed to i.

Lemma 3.5 below shows that the second largest eigenvalue $\lambda_2(P_{\text{Glauber}})$ of Glauber dynamics is small as long as the second largest eigenvalues $\lambda_2(P_{\sigma_\Lambda})$[4] of local random walks are all small.

Condition 3.4. *Let μ be a distribution over $[q]^V$, where $n = |V|$. There exists a sequence $\alpha_0,\alpha_1,\dots,\alpha_{n-2}$ such that for every $0 \le k \le n-2$, $\Lambda \subseteq V$ of size k and any feasible $\sigma_\Lambda \in [q]^\Lambda$, the transition matrix P_{σ_Λ} satisfies*

$$\lambda_2(P_{\sigma_\Lambda}) \le \alpha_k,$$

where $\lambda_2(P_{\sigma_\Lambda})$ is the second largest eigenvalue of the matrix P_{σ_Λ}.

Lemma 3.5 ([AL20]). *Let μ be a distribution over $[q]^V$, where $n = |V|$. Let $\alpha_0,\alpha_1,\dots,\alpha_{n-2}$ be a sequence where $0 \le \alpha_i < 1$ for all $0 \le i \le n-2$. If μ satisfies Condition 3.4 with $\alpha_0,\alpha_1,\dots,\alpha_{n-2}$, then the Glauber dynamics for μ has spectral gap*

$$1 - \lambda_2(P_{\text{Glauber}}) \ge \frac{1}{n}\prod_{k=0}^{n-2}(1-\alpha_k),$$

where $\lambda_2(P_{\text{Glauber}})$ is the second largest eigenvalue of transition matrix P_{Glauber}.

Lemma 3.5 relates Glauber dynamics to local random walks, which provides a powerful tool to analyse Glauber dynamics, because the state space of local random walks are exponentially smaller compared to that of Glauber dynamics. Lemma 3.5 (proved in Section 4) is an easy corollary of the main result in [AL20].

3.2. **Analysis of local random walks.** Our remaining task is to bound the second largest eigenvalues of local random walks.

Our main technical contribution is the following lemma (proved in Section 5), which states that for distribution μ over $[q]^V$ with general domain size $q \ge 2$, these second largest eigenvalues of local random walks are always small if μ is spectrally independent.

Lemma 3.6. *Let μ be a distribution over $[q]^V$, where $n = |V|$. If μ is $(\eta_0,\eta_1,\dots,\eta_{n-2})$-spectrally independent, then μ satisfies Condition 3.4 with a sequence $\alpha_0,\alpha_1,\dots,\alpha_{n-2}$ such that*

$$\forall 0 \le k \le n-2: \quad \alpha_k = \frac{\eta_k}{n-k-1}.$$

[4]Local random walk P_{σ_Λ} has real eigenvalues because P_{σ_Λ} is reversible. See Section 4 for more details.

For the special case $q = 2$, Anari, Liu and Oveis Gharan [ALO20] proved a similar version of Lemma 3.6. They used a linear algebra argument to identify the second largest eigenvalue of the local random walk with the largest eigenvalue of the signed influence matrix. In such analysis, some key identities crucially rely on that $q = 2$, which makes it hard to extend it to general domains of size $q > 2$.

Alternatively, we propose a new *coupling* based argument to show the rapid mixing of the local random walk P_{σ_Λ}, assuming the spectral independence, which implies an upper bound of $\lambda_2(P_{\sigma_\Lambda})$. Specifically, we construct a coupling $(X_t, Y_t)_{t\geq 0}$ for each local random walk, and show that the two chains coalesce (namely $X_t = Y_t$) quickly if μ is spectrally independent. Then we use Proposition 2.5 and Proposition 2.6 to bound the second largest eigenvalue. Our coupling argument is simple and combinatorial, reminiscent of an analysis by Hayes [Hay06]. And it has the advantage of being applicable to joint distributions with general domain sizes. Note that here we are only giving an upper bound for $\lambda_2(P_{\sigma_\Lambda})$, which is sufficient for our purpose, rather than establishing the equality as in [ALO20] for the case with $q = 2$. Detailed analysis is given in Section 5.

3.3. **Proof of main theorem.** It is straightforward to verify that Theorem 3.2 is a corollary of Lemma 3.5 and Lemma 3.6. We now use Theorem 3.2 to prove the main theorem (Theorem 1.3).

Proof of Theorem 1.3. Since μ is (C, η)-spectrally independent (Definition 1.2) for $C \geq 0$ and $1 \leq \eta < 1$, by Definition 3.1, μ is $(\eta_0, \eta_1, \ldots, \eta_{n-2})$-spectrally independent for

$$\eta_k = \min\left\{C, \eta(n-k-1)\right\}.$$

By Theorem 3.2, we have

$$1 - \lambda_2(P_{\mathrm{Glauber}}) \geq \frac{1}{n}\prod_{k=0}^{n-2}\left(1 - \frac{\eta_k}{n-k-1}\right) \geq \frac{1}{n}\prod_{k=0}^{n-2}\left(1 - \min\left\{\frac{C}{n-k-1}, \eta\right\}\right) = \frac{1}{n}\prod_{k=1}^{n-1}\left(1 - \min\left\{\frac{C}{k}, \eta\right\}\right).$$

Thus, the spectral gap has the following lower bound

$$\begin{aligned}
1 - \lambda_2(P_{\mathrm{Glauber}}) &\geq \frac{1}{n}\left(\prod_{k=1}^{2+2C-1}(1-\eta)\right)\left(\prod_{k=2+2C}^{n-1}\left(1 - \frac{C}{k}\right)\right) \geq \frac{(1-\eta)^{2+2C}}{n}\prod_{k=2+2C}^{n}\left(1 - \frac{C}{k}\right) \\
&\geq \frac{(1-\eta)^{2+2C}}{n}\exp\left(-\sum_{k=2+2C}^{n}\frac{2C}{k}\right) \geq \frac{(1-\eta)^{2+2C}}{n}\exp\left(-2C\sum_{k=2}^{n}\frac{1}{k}\right) \\
(\star) \qquad &\geq \frac{(1-\eta)^{2+2C}}{n}\exp\left(-2C\ln n\right) = \frac{(1-\eta)^{2+2C}}{n^{1+2C}},
\end{aligned}$$

where $(\star)$ holds because $\sum_{k=2}^{n}\frac{1}{k} \leq \ln n$.

Since the transition matrix of Glauber dynamics is positive semi-definite, all of its eigenvalues are real [DGU14, AL20]. Let the eigenvalues be $1 = \lambda_1(P_{\mathrm{Glauber}}) \geq \lambda_2(P_{\mathrm{Glauber}}) \geq \ldots \geq \lambda_{|\Omega|}(P_{\mathrm{Glauber}}) \geq 0$, where $\Omega \subseteq [q]^V$ is the support of μ. The absolute spectral gap of Glauber dynamics has the following lower bound

$$\gamma_\star = 1 - \lambda_\star = 1 - \max_{2\leq i\leq|\Omega|}|\lambda_i(P_{\mathrm{Glauber}})| = 1 - \lambda_2(P_{\mathrm{Glauber}}) \geq \frac{(1-\eta)^{2+2C}}{n^{1+2C}}.$$

By Proposition 2.4, we have

$$T_{\mathrm{mix}}(\varepsilon) \leq \frac{1}{\gamma_\star}\left(\log\frac{1}{\varepsilon\mu_{\min}}\right) \leq \frac{n^{1+2C}}{(1-\eta)^{2+2C}}\left(\log\frac{1}{\varepsilon\mu_{\min}}\right). \qquad \square$$

4. Simplicial complexes and Glauber dynamics

In this section, we relate the Glauber dynamics to the random walk on simplicial complexes. As explained in the proof overview, we reduce the task of giving an upper bound for the second largest eigenvalue of Glauber dynamics to the task of giving an upper bound for the second largest eigenvalues of local random walks (Lemma 3.5).

4.1. **Simplicial complexes and random walks.** Let U be a ground set. A *simplicial complex* $X \subseteq 2^U$ is a family of subset that is downward closed, i.e. if $\alpha \in X$, then $\beta \in X$ for all $\beta \subseteq \alpha$. Each subset $\alpha \in X$ is called a *face*. The *dimension* of a face α is its size $|\alpha|$.[5] We use $X(j)$ to denote the set of faces with dimension j. The dimension of a simplicial complex X is the maximum dimension of all its faces. We call X a *pure* d-dimensional simplicial complex if every maximal face of X is of dimension d. We only consider pure simplicial complexes in this paper.

We consider the weighted simplicial complexes. Let X be a pure d-dimensional simplicial complex. Given a weight function $\Pi : X(d) \to \mathbb{R}_{\geq 0}$, define the induced weights for all faces in X by

$$\forall \alpha \in X, \quad \Pi(\alpha) = \sum_{\beta \in X(d) : \beta \supseteq \alpha} \Pi(\beta). \tag{4}$$

For each face $\alpha \in X$, the *link* X_α is simplicial complexes defined by

$$X_\alpha \triangleq \{\beta \setminus \alpha \mid \beta \in X \land \alpha \subseteq \beta\}.$$

Let Π_α be the weight of X_α induced from Π, i.e. for each face $\beta \in X_\alpha$,

$$\Pi_\alpha(\beta) \triangleq \Pi(\alpha \uplus \beta).$$

The *one-skeleton* of X_α is a weighted graph $G_\alpha = (V_\alpha, E_\alpha, \Phi_\alpha)$, where $V_\alpha = X_\alpha(1)$ is the set of singletons, $E_\alpha = X_\alpha(2)$ is the set of 2-dimensional faces, and $\Phi_\alpha(u, v) = \Pi_\alpha(\{u, v\})$ for all $\{u, v\} \in E_\alpha$. We use P_α to denote the simple (non-lazy) random walk on one-skeleton G_α. The transition probability is defined by

$$\forall u, v \in V_\alpha, \quad P_\alpha(u, v) \triangleq \begin{cases} \frac{\Phi_\alpha(u,v)}{\sum_{w:\{u,w\} \in E_\alpha} \Phi_\alpha(u,w)} & \text{if } \{u, v\} \in E_\alpha; \\ 0 & \text{if } \{u, v\} \notin E_\alpha. \end{cases} \tag{5}$$

Given a pure d-dimensional weighted simplicial complexes (X, Π), define the following down-up random walk $P_d^\vee$ on $X(d)$. Suppose the current state is $\sigma_t \in X(d)$, the next state $\sigma_{t+1} \in X(d)$ is generated as follows

- (down walk) pick $x \in \sigma_t$ uniformly at random, and drop x to obtain $\sigma' = \sigma_t \setminus \{x\} \in X(d-1)$;
- (up walk) sample $\sigma_{t+1} \in X(d)$ satisfying $\sigma' \subseteq \sigma_{t+1}$ with probability proportional to $\Pi(\sigma_{t+1})$.

Therefore, the transition matrix of down-up random walk is defined by

$$\forall \alpha, \beta \in X(d), \quad P_d^\vee(\alpha, \beta) \triangleq \begin{cases} \sum_{\tau \in X(d-1) : \tau \subset \alpha} \frac{\Pi(\alpha)}{d \cdot \Pi(\tau)} & \text{if } \alpha = \beta; \\ \frac{\Pi(\beta)}{d \cdot \Pi(\alpha \cap \beta)} & \text{if } \alpha \cap \beta \in X(d-1); \\ 0 & \text{otherwise.} \end{cases}$$

The relation between down-up random walk $P_d^\vee$ and random walks one-skeletons P_α was studied in many works [Opp18, KO20, AL20]. Note that both random walks $P_d^\vee$ and P_α are reversible. By Proposition 2.3, both of them have real eigenvalues.

Definition 4.1 (local-spectral expander [Opp18, KO20, AL20]). Let (X, Π) be a pure d-dimensional weighted simplicial complexes. We say that (X, Π) is a $(\gamma_0, \gamma_1, \ldots, \gamma_{d-2})$-local-spectral expander if for any $0 \leq k \leq d-2$, it holds that

$$\max\{\lambda_2(P_\alpha) \mid \alpha \in X(k)\} \leq \gamma_k,$$

[5] In some papers, such as [KO20, AL20], the dimension is defined to be $|\alpha| - 1$.

where $\lambda_2(P_\alpha)$ stands for the second largest eigenvalue of P_α, and P_α, as defined in (5), is the transition matrix for the simple (non-lazy) random walk on one-skeleton of link X_α.

Theorem 4.2 ([AL20]). *Let (X,Π) be a pure d-dimensional weighted simplicial complexes. If (X,Π) is a $(\gamma_0,\gamma_1,\ldots,\gamma_{d-2})$-local-spectral expander, then*

$$\lambda_2(P_d^\vee) \le 1 - \frac{1}{d}\prod_{k=0}^{d-2}(1-\gamma_k),$$

where $\lambda_2(P_d^\vee)$ is the second largest eigenvalue of down-up random walk $P_d^\vee$.

We remark that the chain $P_d^\vee$ is denoted as $P_{d-1}^\triangledown$ in [AL20].

4.2. **Connections to Glauber dynamics.** Let μ be a distribution over $[q]^V$, where $|V| = n$. Let $\Omega \subseteq [q]^V$ be the support of μ. We define a ground set of nq elements

$$U \triangleq \{(u,c) \mid u \in V \wedge c \in [q]\}.$$

For each (possibly partial) configuration $\sigma \in [q]^\Lambda$ where $\Lambda \subseteq V$, we associate with it a face $f_\sigma \subseteq U$ as

$$f_\sigma \triangleq \{(u,\sigma_u) \mid u \in \Lambda\}.$$

Let X be the downward closure of the family of faces $\{f_\sigma \mid \sigma \in \Omega\}$. Then X is a pure n-dimensional simplicial complex. For each maximal face $f_\sigma \in X$ where $\sigma \in \Omega$, we assign a weight according to μ

$$\Pi(f_\sigma) = \mu(\sigma),$$

and each face in X obtains an induced weight from (4). Hence, (X,Π) is a weighted pure n-dimensional simplicial complex. The following observation is straightforward to verify. One way to understand it is to view the state space $[q]^V$ as the set of bases of a partition matroid.

Observation 4.3. *The Glauber dynamics on μ is precisely the down-up random walk on $X(n)$.*

Let $\Lambda \subseteq V$ be a subset of variables. For every feasible partial configuration $\sigma = \sigma_\Lambda \in [q]^\Lambda$, there exists a face $f_\sigma = \{(u,\sigma_u) \mid u \in \Lambda\}$ in X, and vice versa.

To simplify the notation, we use P_σ to denote the simple (non-lazy) random walk P_{f_σ} (defined in (5)) on one-skeleton of link X_{f_σ}. By definition, P_σ is a random walk on $U_\sigma = \{(u,c) \in \overline{\Lambda} \times [q] \mid \mu_u^\sigma(c) > 0\}$, where $\overline{\Lambda} = V \setminus \Lambda$. Fix $\boldsymbol{x} = (u,i) \in U_\sigma$ and $\boldsymbol{y} = (v,j) \in U_\sigma$. The weight of edge $\{\boldsymbol{x},\boldsymbol{y}\}$ in one-skeleton of link X_{f_σ} is given by

$$\Phi_{f_\sigma}(\boldsymbol{x},\boldsymbol{y}) = \sum_{\substack{\tau\in\Omega \\ \tau_\Lambda=\sigma,\tau_u=i,\tau_v=j}} \mu(\tau) = \mathbf{Pr}_{X\sim\mu}\left[X_\Lambda = \sigma \wedge X_u = i \wedge X_v = j\right].$$

Thus,

$$\forall (u,i),(v,j) \in U_\sigma, \quad P_\sigma((u,i),(v,j)) = \frac{\mathbf{1}\left[u \ne v\right]}{|V| - |\Lambda| - 1}\mu_v^{\sigma,u\leftarrow i}(j),$$

where $\mu_v^{\sigma,u\leftarrow i}$ is the marginal distribution on v induced from μ conditional on the configurations on Λ fixed as σ and that u is fixed to i. This is precisely the local random walk in Definition 3.3. Note that P_σ is reversible because the random walk on one-skeleton is reversible.

Note that if μ satisfies Condition 3.4 with $\alpha_0,\alpha_1,\ldots,\alpha_{n-2}$, then the n-dimensional weighted simplicial complex (X,Π) defined above is a $(\gamma_0,\gamma_1,\ldots,\gamma_{n-2})$-local-spectral expander with $\gamma_k = \alpha_k$ for all $0 \le k \le n-2$. Hence, Lemma 3.5 is a corollary of Theorem 4.2.

5. Analysis of local random walks

In this section, we prove Lemma 3.6, which states that the second largest eigenvalues of local random walks are small if the distribution μ is spectrally independent. The new ingredient of this part is a coupling based argument for the above implication for distributions with general domain size q.

5.1. **Proof of Lemma 3.6**. Fix a subset $\Lambda \subseteq V$ with $0 \le |\Lambda| \le n-2$, and a feasible partial configuration $\sigma_\Lambda \in [q]^\Lambda$. To simplify the notation, we use σ to denote σ_Λ. Consider the random walk P_σ defined in (3). Recall the state space of P_σ is defined by

$$U_\sigma \triangleq \left\{(u,i) \in \overline{\Lambda} \times [q] \mid \mu_u^\sigma(i) > 0\right\}. \tag{6}$$

Thus $|U_\sigma| \ge 2$ because $|\overline{\Lambda}| = |V \setminus \Lambda| \ge 2$ and σ is feasible. Consider a random walk Q_σ:

$$Q_\sigma \triangleq \frac{n-|\Lambda|-1}{n-|\Lambda|} P_\sigma + \frac{1}{n-|\Lambda|} I_\sigma, \tag{7}$$

where $I_\sigma \in \mathbb{R}_{\ge 0}^{U_\sigma \times U_\sigma}$ is the identity matrix. In other words, in each step, with probability $\frac{1}{n-|\Lambda|}$, the random walk Q_σ stays at the current state; otherwise, Q_σ evolves in the same way as P_σ.

Define a distribution π over U_σ as

$$\forall (u,i) \in U_\sigma, \quad \pi(u,i) \triangleq \frac{1}{n-|\Lambda|} \mu_u^\sigma(i). \tag{8}$$

Note that $\sum_{i \in \Omega_u^\sigma} \mu_u^\sigma(i) = 1$, where $\Omega_u^\sigma \triangleq \{i \in [q] \mid \mu_u^\sigma(i) > 0\}$. Thus $\sum_{(u,i)\in U_\sigma} \pi(u,i) = 1$ and π is well-defined. We claim that both P_σ and Q_σ are reversible with respect to π. For any $(u,i), (v,j) \in U_\sigma$, we verify the detailed balance equation. If $u = v$, then it is straightforward to verify

$$\pi(u,i) P_\sigma((u,i),(v,j)) = 0 = \pi(v,j) P_\sigma((v,j),(u,i));$$

otherwise $u \ne v$, then

$$\begin{aligned}\pi(u,i) P_\sigma((u,i),(v,j)) &= \frac{\mu_u^\sigma(i) \cdot \mu_v^{\sigma, u \leftarrow i}(j)}{(n-|\Lambda|)(n-|\Lambda|-1)} = \frac{\mathbf{Pr}_{X\sim\mu}\left[X_u = i \wedge X_v = j \mid X_\Lambda = \sigma\right]}{(n-|\Lambda|)(n-|\Lambda|-1)} \\ &= \frac{\mu_v^\sigma(j) \cdot \mu_u^{\sigma, v \leftarrow j}(i)}{(n-|\Lambda|)(n-|\Lambda|-1)} = \pi(v,j) P_\sigma((v,j),(u,i)).\end{aligned}$$

Since Q_σ is a lazy version of P_σ, Q_σ is also reversible to π. By (6) and (8), the support of π is U_σ. By Proposition 2.3, P_σ and Q_σ both have $|U_\sigma|$ real eigenvalues. Let $\lambda_2(P_\sigma)$ and $\lambda_2(Q_\sigma)$ denote the second largest eigenvalues of P_σ and Q_σ. By the definition of Q_σ in (7), we have the following proposition.

Proposition 5.1. $\lambda_2(Q_\sigma) = \frac{n-|\Lambda|-1}{n-|\Lambda|} \lambda_2(P_\sigma) + \frac{1}{n-|\Lambda|}$.

Proposition 5.1 is a basic result in linear algebra. We claim the following result about $\lambda_2(Q_\sigma)$.

Lemma 5.2. $\lambda_2(Q_\sigma) \le \frac{\rho(\Psi_\mu^\sigma)+1}{n-|\Lambda|}$.

The proof of Lemma 5.2 is deferred to the next subsection. We now use Proposition 5.1 and Lemma 5.2 to prove Lemma 3.6. Suppose μ is $(\eta_0, \eta_1, \ldots, \eta_{n-2})$-spectrally independent (Definition 3.1). By Proposition 5.1, it holds that

$$\lambda_2(P_\sigma) = \frac{n-|\Lambda|}{n-|\Lambda|-1}\left(\lambda_2(Q_\sigma) - \frac{1}{n-|\Lambda|}\right) \le \frac{\rho(\Psi_\mu^\sigma)}{n-|\Lambda|-1} \le \frac{\eta_k}{n-k-1}, \text{ where } k = |\Lambda|.$$

The above inequality holds for any $\Lambda \subseteq V$ with $0 \le |\Lambda| \le n-2$ and any feasible $\sigma \in [q]^\Lambda$. This implies μ satisfies Condition 3.4 with $\alpha_0, \alpha_1, \ldots, \alpha_{n-2}$ such that $\alpha_k = \frac{\eta_k}{n-k-1}$.

5.2. **A coupling based analysis.** We now prove Lemma 5.2. We will use coupling to give an upper bound of $\lambda_2(Q_\sigma)$. First we define a matrix A:

$$A \triangleq \frac{1}{n-|\Lambda|}\left(\left(\Psi_\mu^\sigma\right)^T + I\right), \tag{9}$$

where I is the identity matrix. For any $t \geq 1$, define

$$d(t) \triangleq \max_{x_0, y_0 \in U_\sigma} d_{\mathrm{TV}}\left(Q_\sigma^t(x_0,\cdot), Q_\sigma^t(y_0,\cdot)\right). \tag{10}$$

Lemma 5.3. *For any $t \geq 1$, $d(t) \leq \left\|A^{t-1}\right\|_1$.*

Proof. By the definitions of Q_σ in (7) and P_σ in (3), we have

$$\forall (u,i),(v,j) \in U_\sigma, \quad Q_\sigma((u,i),(v,j)) = \frac{\mu_v^{\sigma,u\leftarrow i}(j)}{n-|\Lambda|},$$

where if $u = v$, the distribution $\mu_u^{\sigma,u\leftarrow i}(j) = \mathbf{1}[i=j]$ for all $j \in [q]$. Let $X_0, X_1, X_2, \ldots \in U_\sigma$ be the sequence of random states generated by Q_σ, where $X_t = (X_t^{\mathsf{vtx}}, X_t^{\mathsf{val}})$, $X_t^{\mathsf{vtx}} \in V \setminus \Lambda$ and $X_t^{\mathsf{val}} \in [q]$. By definition of Q_σ in (7), given $X_{t-1} = (u,i)$, the random pair $X_t = (v,j)$ can be generated by the following procedure

- sample $v \in V \setminus \Lambda$ uniformly at random;
- sample $j \in [q]$ from the distribution $\mu_v^{\sigma,u\leftarrow i}(\cdot)$.

Next, we define a coupling procedure C. Let $(X_t)_{t\geq 0}$ be the random walk Q_σ starting from $X_0 = \boldsymbol{x}_0 \in U_\sigma$, and $(Y_t)_{t\geq 0}$ be the random walk Q_σ starting from $Y_0 = \boldsymbol{y}_0 \in U_\sigma$, where $\boldsymbol{x}_0$ and $\boldsymbol{y}_0$ achieve the maximum in (10). Consider each transition step $(X,Y) \to (X',Y')$. Suppose $X = (u_x, i_x)$ and $Y = (u_y, i_y)$. Then $X' = (u'_x, i'_x)$ and $Y' = (u'_y, i'_y)$ are generated as follows:

- sample $v \in V \setminus \Lambda$ uniformly at random, set $u'_x = u'_y = v$;
- sample (i'_x, i'_y) from the optimal coupling of $\mu_v^{\sigma,u_x\leftarrow i_x}$ and $\mu_v^{\sigma,u_y\leftarrow i_y}$, where $v = u'_x = u'_y$.

It is easy to verify that C is a coupling of Markov chain Q_σ. By Proposition 2.6, we have

$$\forall t \geq 1, \quad d(t) = \max_{\boldsymbol{x}_0,\boldsymbol{y}_0 \in U_\sigma} d_{\mathrm{TV}}\left(Q_\sigma^t(\boldsymbol{x}_0,\cdot), Q_\sigma^t(\boldsymbol{y}_0,\cdot)\right) \leq \mathbf{Pr}_C\left[X_t \neq Y_t\right]. \tag{11}$$

Hence, we only need to bound the right-hand-side of (11).

Denote $X_t = (X_t^{\mathsf{vtx}}, X_t^{\mathsf{val}})$ and $Y_t = (Y_t^{\mathsf{vtx}}, Y_t^{\mathsf{val}})$. By the definition of the coupling procedure C, it holds that $X_t^{\mathsf{vtx}} = Y_t^{\mathsf{vtx}}$ for all $t \geq 1$, and

$$\forall t \geq 1, u \in V \setminus \Lambda, \quad \mathbf{Pr}_C\left[X_t^{\mathsf{vtx}} = Y_t^{\mathsf{vtx}} = u\right] = \frac{1}{n-|\Lambda|}. \tag{12}$$

For any $t \geq 1$, we define a column vector $e_t \in \mathbb{R}_{\geq 0}^{V\setminus\Lambda}$ such that

$$\forall u \in V \setminus \Lambda, \quad e_t(u) \triangleq \mathbf{Pr}_C\left[X_t^{\mathsf{vtx}} = Y_t^{\mathsf{vtx}} = u \wedge X_t^{\mathsf{val}} \neq Y_t^{\mathsf{val}}\right].$$

Then $d(t) \leq \mathbf{Pr}_C\left[X_t \neq Y_t\right] = \sum_{u\in V\setminus\Lambda} e_t(u) = \|e_t\|_1$ for all $t \geq 1$. By (12), we have

$$\forall u \in V \setminus \Lambda, \quad e_1(u) \leq \mathbf{Pr}_C\left[X_t^{\mathsf{vtx}} = Y_t^{\mathsf{vtx}} = u\right] = \frac{1}{n-|\Lambda|}. \tag{13}$$

Recall $\Omega_u^\sigma \triangleq \{i \in [q] \mid \mu_u^\sigma(i) > 0\}$ for each $u \in V \setminus \Lambda$, and the state space of the random walk Q_σ is $U_\sigma = \left\{(u,i) \mid u \in \overline{\Lambda} \wedge i \in \Omega_u^\sigma\right\}$. For any $t \geq 2$, we have for all $u \in V \setminus \Lambda$,

$$\begin{aligned}
e_t(u) &= \mathbf{Pr}_C\left[X_t^{\mathsf{vtx}} = Y_t^{\mathsf{vtx}} = u \wedge X_t^{\mathsf{val}} \neq Y_t^{\mathsf{val}}\right] \\
&= \sum_{v \in V \setminus \Lambda} \sum_{\substack{i,j \in \Omega_v^\sigma \\ i \neq j}} \Big(\mathbf{Pr}_C\left[X_t^{\mathsf{vtx}} = Y_t^{\mathsf{vtx}} = u \wedge X_t^{\mathsf{val}} \neq Y_t^{\mathsf{val}} \mid X_{t-1}^{\mathsf{vtx}} = Y_{t-1}^{\mathsf{vtx}} = v \wedge X_{t-1}^{\mathsf{val}} = i \wedge Y_{t-1}^{\mathsf{val}} = j\right] \\
&\qquad\qquad \times \mathbf{Pr}_C\left[X_{t-1} = (v,i) \wedge Y_{t-1} = (v,j)\right]\Big) \\
&= \sum_{v \in V \setminus \Lambda} \sum_{\substack{i,j \in \Omega_v^\sigma \\ i \neq j}} \frac{1}{n - |\Lambda|} d_{\mathrm{TV}}\left(\mu_u^{\sigma, v \leftarrow i}, \mu_u^{\sigma, v \leftarrow j}\right) \mathbf{Pr}_C\left[X_{t-1} = (v,i) \wedge Y_{t-1} = (v,j)\right].
\end{aligned}$$

The first equality is obtained from the chain rule, together with the facts that for $t \geq 2$, (1) $\mathbf{Pr}_C\left[X_{t-1}^{\mathsf{vtx}} = Y_{t-1}^{\mathsf{vtx}}\right] = 1$; (2) $X_t^{\mathsf{val}} \neq Y_t^{\mathsf{val}}$ only if $X_{t-1}^{\mathsf{val}} \neq Y_{t-1}^{\mathsf{val}}$; (3) $X_{t-1}^{\mathsf{val}}, Y_{t-1}^{\mathsf{val}} \in \Omega_v^\sigma$ if $X_{t=1}^{\mathsf{vtx}} = Y_{t-1}^{\mathsf{vtx}} = v$ since Q_σ is a random walk over U_σ. The last equality is obtained using the definition of the coupling C. It holds because $X_t^{\mathsf{vtx}} = Y_t^{\mathsf{vtx}}$ are sampled from $V \setminus \Lambda$ uniformly at random and $X_t^{\mathsf{val}}, Y_t^{\mathsf{val}}$ are sampled from the optimal coupling between $\mu_u^{\sigma, v \leftarrow i}$ and $\mu_u^{\sigma, v \leftarrow j}$. By the definition of the influence matrix Ψ_μ^σ in (1) and the definition of the matrix A in (9), we have that for any $u, v \in V \setminus \Lambda$, any $i, j \in \Omega_u^\sigma$ such that $i \neq j$, it holds that

$$\frac{1}{n - |\Lambda|} d_{\mathrm{TV}}\left(\mu_u^{\sigma, v \leftarrow i}, \mu_u^{\sigma, v \leftarrow j}\right) \leq A(u,v).$$

Hence, for any $t \geq 2$, we have that for all $u \in V \setminus \Lambda$, $e_t(u)$ can be bounded by

$$e_t(u) \leq \sum_{v \in V \setminus \Lambda} \sum_{\substack{i,j \in \Omega_v^\sigma \\ i \neq j}} A(u,v) \mathbf{Pr}_C\left[X_{t-1} = (v,i) \wedge Y_{t-1} = (v,j)\right] = \sum_{v \in V \setminus \Lambda} A(u,v) e_{t-1}(v) = (A e_{t-1})(u). \tag{14}$$

All e_t are non-negative vectors and A is a non-negative matrix. Combining (11), (13) and (14), we have that for any $t \geq 1$,

$$d(t) \leq \|e_t\|_1 \leq \left\|A^{t-1} e_1\right\|_1 \leq \left\|A^{t-1}\right\|_1 \|e_1\|_1 \leq \left\|A^{t-1}\right\|_1. \qquad \square$$

Now, we are ready to prove Lemma 5.2.

Proof of Lemma 5.2. Recall Q_σ is a random walk over U_σ, π is defined in (8). Since Q_σ is reversible with respect to π and the support of π is U_σ, by Proposition 2.5 and Lemma 5.3, for any $t \geq 1$,

$$|\lambda_2(Q_\sigma)|^t \leq d(t) \leq \left\|A^{t-1}\right\|_1.$$

We may assume that $\lambda_2(Q_\sigma) > 0$, as otherwise Lemma 5.2 holds trivially. We have

$$\forall t \geq 1, \quad \lambda_2(Q_\sigma)^{\frac{t}{t-1}} \leq \left\|A^{t-1}\right\|_1^{\frac{1}{t-1}}.$$

Let $t \to \infty$ in both sides, we have

$$\lambda_2(Q_\sigma) = \lim_{t \to \infty} \lambda_2(Q_\sigma)^{\frac{t}{t-1}} \leq \lim_{t \to \infty} \left\|A^{t-1}\right\|_1^{\frac{1}{t-1}} = \rho(A),$$

where the last equality holds due to Proposition 2.1. Note that if $\lambda \in \mathbb{C}$ is an eigenvalue of $(\Psi_\mu^\sigma)^T$, then $\lambda + 1$ is an eigenvalue of $(\Psi_\mu^\sigma)^T + I$, and $|\lambda + 1| \leq |\lambda| + 1$. By the definition of A, we have

$$\lambda_2(Q_\sigma) \leq \rho(A) = \frac{1}{n - |\Lambda|} \rho\left(\left(\Psi_\mu^\sigma\right)^T + I\right) \leq \frac{\rho\left(\left(\Psi_\mu^\sigma\right)^T\right) + 1}{n - |\Lambda|} = \frac{\rho(\Psi_\mu^\sigma) + 1}{n - |\Lambda|}. \qquad \square$$

6. Rapid mixing for list colourings

An instance of the list colouring is a pair $(G, \boldsymbol{L})$ where $G = (V, E)$ is a simple undirected graph and $\boldsymbol{L} = \{L(v) \mid v \in V\}$ is a collection of *colour lists* associated to each vertex $v \in V$. A proper list colouring X assigns each vertex $v \in V$ a colour $X_v \in L(v)$ such that $X_u \neq X_v$ for all $\{u, v\} \in E$. Let $\Omega_{G,\boldsymbol{L}}$ denote the set of proper list colourings and $\mu_{G,\boldsymbol{L}}$ denote the uniform distribution over $\Omega_{G,\boldsymbol{L}}$.

The Glauber dynamics on $(G, \boldsymbol{L})$ is defined as follows. The chain starts from an arbitrary proper list colouring $X \in \Omega_{G,\boldsymbol{L}}$. In each step, the chain does the following:

- pick a vertex $v \in V$ uniformly at random;
- update X_v by a uniformly at random colour from $L(v) \setminus \{X_u \mid \{v, u\} \in E\}$.

We prove the following rapid mixing result for list colourings.

Theorem 6.1. *Let $(G = (V, E), \boldsymbol{L})$ be an instance of list colouring where $\boldsymbol{L} = \{L(v) \mid v \in V\}$. Let $\Delta \geq 3$ be the maximum degree of G and $\delta > 0$ be a constant. If G is triangle-free and for every $v \in V$, it holds that*

$$|L(v)| - \deg_G(v) \geq (\alpha^* + \delta - 1)\Delta, \tag{15}$$

then the Glauber dynamics on $(G, \boldsymbol{L})$ satisfies

$$T_{mix}(\varepsilon) \leq \left(9\mathrm{e}^5 n\right)^{1+9/\delta} \cdot \log\left(\frac{M}{\varepsilon}\right).$$

where $M \triangleq \prod_{v \in V} |L(v)|$.

Note that Theorem 1.5 is a corollary of Theorem 6.1, in which $M = q^n$.

In order to prove Theorem 6.1, we define a partial order $\preceq$ among list-colouring instances. Let $(G' = (V', E'), \boldsymbol{L}')$ and $(G = (V, E), \boldsymbol{L})$ be two list colouring instances where $\boldsymbol{L}' = \{L'(v) \mid v \in V'\}$ and $\boldsymbol{L} = \{L(v) \mid v \in V\}$. We say $(G', \boldsymbol{L}') \preceq (G, \boldsymbol{L})$ if there exists a vertex $v \in V$ satisfying

- $G' = G[V \setminus \{v\}]$;
- for every $u \in \Gamma_G(v)$, it holds that $L'(u) \subseteq L(u)$ and $|L(u) \setminus L'(u)| \leq 1$;
- for every $u \in V' \setminus \Gamma_G(v)$, it holds that $L'(u) = L(u)$.

Here, $\Gamma_G(v)$ denotes the neighbourhood of v in graph G. We remark that in the definition above, for each $u \in \Gamma_G(v)$, we can rewrite the requirement as $L'(u) = L(u) \setminus \{c\}$ for some colour c. This colour c is not necessarily in $L(u)$ (in which case $L'(u) = L(u)$ and can be distinct for different $u \in \Gamma_G(v)$).

Intuitively, $(G', \boldsymbol{L}') \preceq (G, \boldsymbol{L})$ means that one can obtain $(G', \boldsymbol{L}')$ from $(G, \boldsymbol{L})$ by removing one vertex v and change the colour lists of the neighbours of v by removing at most one color. We call a family of list-colouring instances $\mathcal{L}$ *downward closed* if for every $(G, \boldsymbol{L}) \in \mathcal{L}$ and every $(G', \boldsymbol{L}')$ such that $(G', \boldsymbol{L}') \preceq (G, \boldsymbol{L})$, we have $(G', \boldsymbol{L}') \in \mathcal{L}$.

The *downward closure* of an instance $(G, \boldsymbol{L})$ is the minimum downward closed family of instances containing $(G, \boldsymbol{L})$.

Consider the following condition for a family of list colouring instances $\mathcal{L}$.

Condition 6.2. *Let $\chi > 0$, $0 < \varepsilon_1 < 1$ and $\varepsilon_2 > 0$. It holds that*

- *the maximum degree of instances in $\mathcal{L}$ is at most χ;*
- *for any $(G = (V, E), \boldsymbol{L}) \in \mathcal{L}$, a proper list colouring exists, and for any vertex $v \in V$ satisfying $\deg_G(v) \leq \chi - 1$, it holds that*

$$\forall c \in L(v): \quad \mu_{v,(G,\boldsymbol{L})}(c) \leq \frac{\varepsilon_1}{\deg_G(v)}; \tag{16}$$

  for any vertex $v \in V$, it holds that

$$\forall c \in L(v): \quad \mu_{v,(G,\boldsymbol{L})}(c) \leq \frac{1}{\varepsilon_2 \chi + 1}. \tag{17}$$

We have the following theorem.

Theorem 6.3. *Let $0 < \varepsilon_1 < 1$ and $\varepsilon_2 > 0$ be two constants. The following result holds for any $\chi > 0$. Let $\mathscr{L}$ be a downward closed family of list-colouring instances satisfying Condition 6.2 with parameters χ, ε_1 and ε_2. For any $(G = (V, E), \boldsymbol{L}) \in \mathscr{L}$, the mixing time of Glauber dynamics satisfies*

$$T_{mix}(\varepsilon) \le \left(9e^{\frac{2}{\varepsilon_2}}\right)^{\left(1+\frac{1}{(1-\varepsilon_1)\varepsilon_2}\right)} n^{1+\frac{2}{(1-\varepsilon_1)\varepsilon_2}} \cdot \log\left(\frac{M}{\varepsilon}\right),$$

where $M = \prod_{v \in V} |L(v)|$.

Theorem 6.1 is actually a corollary of Theorem 6.3 via verifying Condition 6.2. We will prove Theorem 6.3 first. The proof of Theorem 6.1 is deferred to Section 6.4.

6.1. **Analysis of mixing time**.

In the following, we assume $\mathscr{L}$ is downward closed and satisfies Condition 6.2. Let $\chi > 0$, $0 < \varepsilon_1 < 1$ and $\varepsilon_2 > 0$ be the parameters promised by Condition 6.2.

For any list colouring instance $(G, \boldsymbol{L})$ where $G = (V, E)$, recall $\mu_{G,\boldsymbol{L}}$ is the uniform distribution over all proper list colourings. Define the matrix $R_{G,\boldsymbol{L}} \in \mathbb{R}_{\ge 0}^{V \times V}$ by

$$\forall u, v \in V, \quad R_{G,\boldsymbol{L}}(u, v) = \max_{c_1, c_2 \in L(u)} d_{\mathrm{TV}}\left(\mu_{v,(G,\boldsymbol{L})}^{u \leftarrow c_1}, \mu_{v,(G,\boldsymbol{L})}^{u \leftarrow c_2}\right), \tag{18}$$

where for $c = c_1$ or c_2, $\mu_{v,(G,\boldsymbol{L})}^{u \leftarrow c}$ denotes the marginal distribution on v projected from $\mu_{G,\boldsymbol{L}}$ conditional on the colour of u is fixed as c. The matrix R is essentially the same as the influence matrix $\Psi_\mu^{\sigma_\Lambda}$ in (1), except that in the case of $u = v$, $R_{G,\boldsymbol{L}}(v, v) = 0$ if and only if $|L(v)| = 1$ (thus $c_1 = c_2$). Namely,

$$R_{G,\boldsymbol{L}}(v, v) = \max_{c_1, c_2 \in L(v)} d_{\mathrm{TV}}\left(\mu_{v,(G,\boldsymbol{L})}^{v \leftarrow c_1}, \mu_{v,(G,\boldsymbol{L})}^{v \leftarrow c_2}\right) = \mathbf{1}\left[|L(v)| > 1\right].$$

Roughly speaking, each entry $R_{G,\boldsymbol{L}}(u, v)$ is the influence of u on v given two different colours of u. The key to apply Theorem 1.3 is to bound the total influence of u on all other vertices.

Lemma 6.4. *For any instance $(G = (V, E), \boldsymbol{L}) \in \mathscr{L}$,*

$$\forall u \in V, \quad \sum_{v \in V: v \ne u} R_{G,\boldsymbol{L}}(u, v) \le \min\left\{\left(1 - \frac{1}{3e^{1/\varepsilon_2}}\right)(|V| - 1), \frac{1}{(1 - \varepsilon_1)\varepsilon_2}\right\}.$$

We first use Lemma 6.4 to prove the main theorem for list colouring (Theorem 6.3). Then we prove Lemma 6.4 in Section 6.2 and Section 6.3.

To prove Theorem 6.3, we will also need the following notion of pinning.

Definition 6.5 (instance induced by pinning). Let $(G = (V, E), \boldsymbol{L})$ be a list colouring instance. Let $\Lambda \subseteq V$ be a subset of vertices and $\sigma \in \otimes_{v \in \Lambda} L(v)$ a partial colouring on Λ. Define $\mathsf{Pin}_{G,\boldsymbol{L}}(\Lambda, \sigma) = (\widetilde{G}, \widetilde{\boldsymbol{L}})$ as the induced list colouring instance after the pinning σ, where $\widetilde{G} = G[V \setminus \Lambda]$ is the subgraph of G induced by $V \setminus \Lambda$, and $\widetilde{\boldsymbol{L}} = \{\widetilde{L}(v) \mid v \in V \setminus \Lambda\}$ is defined by for all $v \in V \setminus \Lambda$,

$$\widetilde{L}(v) = L(v) \setminus \{\sigma_u \mid u \in \Lambda \wedge \{u, v\} \in E\}.$$

It is clear that for any Λ and σ, $\mathsf{Pin}_{G,\boldsymbol{L}}(\Lambda, \sigma)$ is in the downward closure of $(G, \boldsymbol{L})$.

Now, we are ready to prove Theorem 6.3.

Proof of Theorem 6.3. It suffices to verify that every $(G, \boldsymbol{L}) \in \mathscr{L}$ is (C, η)-spectrally independent, which implies the theorem by Theorem 1.3. Fix a list colouring instance $(G = (V, E), \boldsymbol{L}) \in \mathscr{L}$. Fix a subset $\Lambda \subseteq V$ with $|\Lambda| \le n - 2$ and a feasible partial colouring $\sigma_\Lambda \in \otimes_{v \in \Lambda} L(v)$. Let $(\widetilde{G}, \widetilde{\boldsymbol{L}}) = \mathsf{Pin}_{G,\boldsymbol{L}}(\Lambda, \sigma_\Lambda)$, where $\widetilde{G} = G[V \setminus \Lambda]$ and $\widetilde{\boldsymbol{L}} = \{\widetilde{L}(v) \mid v \in V \setminus \Lambda\}$. Note that for any $u \in V \setminus \Lambda$, $\widetilde{L}(u)$ contains precisely the feasible colours for u conditional on σ_Λ. Then, by the definition of $\Psi_\mu^{\sigma_\Lambda}$ in (1),

$$\forall u, v \in V \setminus \Lambda \text{ with } u \neq v, \quad \Psi_\mu^{\sigma_\Lambda}(u,v) = \max_{c_1,c_2 \in \widetilde{L}(u)} d_{\mathrm{TV}}\left(\mu_{v,(G,\boldsymbol{L})}^{\sigma_\Lambda, u \leftarrow c_1}, \mu_{v,(G,\boldsymbol{L})}^{\sigma_\Lambda, u \leftarrow c_2}\right)$$

(by Definition 6.5)

$$= \max_{c_1,c_2 \in \widetilde{L}(u)} d_{\mathrm{TV}}\left(\mu_{v,(\widetilde{G},\widetilde{\boldsymbol{L}})}^{u \leftarrow c_1}, \mu_{v,(\widetilde{G},\widetilde{\boldsymbol{L}})}^{u \leftarrow c_2}\right)$$

$$= R_{\widetilde{G},\widetilde{\boldsymbol{L}}}(u,v).$$

Also by the definition of $\Psi_\mu^{\sigma_\Lambda}$, for any $v \in V \setminus \Lambda$, it holds that $\Psi_\mu^{\sigma_\Lambda}(v,v) = 0$. Since $\mathscr{L}$ is downward closed, $(\widetilde{G}, \widetilde{\boldsymbol{L}}) \in \mathscr{L}$. By Lemma 6.4,

$$\left\|\Psi_\mu^{\sigma_\Lambda}\right\|_\infty = \max_{u \in V \setminus \Lambda} \sum_{v \in V \setminus \Lambda} \Psi_\mu^{\sigma_\Lambda}(u,v) = \max_{u \in V \setminus \Lambda} \sum_{v \in V \setminus \Lambda : v \neq u} R_{\widetilde{G},\widetilde{\boldsymbol{L}}}(u,v)$$

$$\leq \min\left\{\left(1 - \frac{1}{3\mathrm{e}^{1/\varepsilon_2}}\right)(n - |\Lambda| - 1), \frac{1}{(1-\varepsilon_1)\varepsilon_2}\right\}.$$

Hence, the list colouring instance $(G, \boldsymbol{L}) \in \mathscr{L}$ satisfies bound one-to-all influence condition in Corollary 1.4 with $C = \frac{1}{(1-\varepsilon_1)\varepsilon_2}$ and $\eta = 1 - \frac{1}{3\mathrm{e}^{1/\varepsilon_2}}$. By Corollary 1.4, Glauber dynamics on $(G, \boldsymbol{L})$ has mixing time

$$T_{\mathrm{mix}}(\varepsilon) \leq \frac{n^{1+\frac{2}{(1-\varepsilon_1)\varepsilon_2}}}{\left(\frac{1}{3\mathrm{e}^{1/\varepsilon_2}}\right)^{2+\frac{2}{(1-\varepsilon_1)\varepsilon_2}}} \cdot \log\left(\frac{1}{\varepsilon \mu_{\min}}\right) \leq \left(9\mathrm{e}^{\frac{2}{\varepsilon_2}}\right)^{\left(1+\frac{1}{(1-\varepsilon_1)\varepsilon_2}\right)} n^{1+\frac{2}{(1-\varepsilon_1)\varepsilon_2}} \cdot \log\left(\frac{M}{\varepsilon}\right),$$

where the last inequality holds because $\frac{1}{\mu_{\min}} \leq M = \prod_{v \in V} |L(v)|$. □

The two upper bounds in Lemma 6.4 are proved in Section 6.2 and Section 6.3 respectively.

6.2. **An easy coupling analysis.** We now prove the first part of Lemma 6.4, namely,

Lemma 6.6. *Let $\mathscr{L}$ be a downward closed family of list colouring instances satisfying Condition 6.2 with parameters $\chi > 0$, $0 < \varepsilon_1 < 1$ and $\varepsilon_2 > 0$. For any instance $(G = (V,E), \boldsymbol{L}) \in \mathscr{L}$, it holds that*

$$\forall u \in V, \quad \sum_{v \in V : v \neq u} R_{G,\boldsymbol{L}}(u,v) \leq \left(1 - \frac{1}{3\mathrm{e}^{1/\varepsilon_2}}\right)(|V| - 1).$$

To prove Lemma 6.6, we need the following well-known recursion of list colouring.

Proposition 6.7 ([GK12, LY13, GKM15]). *Let $(G = (V,E), \boldsymbol{L}) \in \mathscr{L}$ be a list colouring instance. Let $v_1, v_2, \ldots, v_m$ denote the neighbours of v in G. Let $c \in L(v)$ be a colour. Let G_v be the subgraph of G induced by $V \setminus \{v\}$. For each $1 \leq i \leq m$, define a colour list $\boldsymbol{L}_{i,c} = \{L_{i,c}(u) \mid u \in V \setminus \{v\}\}$, where $L_{i,c}(u) = L(u) \setminus \{c\}$ for all $u = v_j$ and $j < i$, and $L_{i,c}(u) = L(u)$ for other vertices. It holds that for any $c \in L(v)$,*

$$\mu_{v,(G,\boldsymbol{L})}(c) = \frac{\prod_{i=1}^m \left(1 - \mu_{v_i,(G_v,\boldsymbol{L}_{i,c})}(c)\right)}{\sum_{c' \in L(v)} \prod_{i=1}^m \left(1 - \mu_{v_i,(G_v,\boldsymbol{L}_{i,c'})}(c')\right)}.$$

We first derive upper and lower bounds for marginal probabilities from Condition 6.2.

Lemma 6.8. *Let $\mathscr{L}$ be a downward closed family of list colouring instances satisfying Condition 6.2 with parameters $\chi > 0$ and $0 < \varepsilon_1 < 1$ and $\varepsilon_2 > 0$. For any instance $(G = (V,E), \boldsymbol{L}) \in \mathscr{L}$, it holds that*

$$\forall c \in L(v), \quad \frac{1}{\mathrm{e}^{1/\varepsilon_2}|L(v)|} \leq \mu_{v,(G,\boldsymbol{L})}(c) \leq \frac{1}{\varepsilon_2 \chi + 1}.$$

Proof. The upper bound is directly from Condition 6.2. So we only need to prove the lower bound.

Fix an instance $(G, \boldsymbol{L}) \in \mathscr{L}$. Since $\mathscr{L}$ is downward closed, each instance $(G_v, \boldsymbol{L}_{i,c}) \in \mathscr{L}$, where $(G_v, \boldsymbol{L}_{i,c})$ is defined in Proposition 6.7. By the recursion in Proposition 6.7, we have

$$\mu_{v,(G,\boldsymbol{L})}(c) = \frac{\prod_{i=1}^{m}\left(1-\mu_{v_i,(G_v,\boldsymbol{L}_{i,c})}(c)\right)}{\sum_{c'\in L(v)}\prod_{i=1}^{m}\left(1-\mu_{v_i,(G_v,\boldsymbol{L}_{i,c'})}(c')\right)} \geq \frac{\left(1-\frac{1}{\varepsilon_2\chi+1}\right)^{\chi}}{|L(v)|} \geq \frac{1}{\mathrm{e}^{1/\varepsilon_2}\,|L(v)|}.$$

This proves the lower bound. □

Now, we are ready to prove Lemma 6.6.

Proof of Lemma 6.6. Consider the list colouring instance $(G = (V, E), \boldsymbol{L})$. Fix a vertex u and two colours $c_1, c_2 \in L(u)$. Define a list colouring instance $\mathcal{L}_1 = (G_u, \boldsymbol{L}_1) = \mathsf{Pin}_{G,\boldsymbol{L}}(\{u\}, c_1)$, where G_u is the subgraph of G induced by $V \setminus \{u\}$ and $\boldsymbol{L}_1 = \{L_1(w) \mid w \in V \setminus \{u\}\}$. Define a list colouring instance $\mathcal{L}_2 = (G_u, \boldsymbol{L}_2) = \mathsf{Pin}_{G,\boldsymbol{L}}(\{u\}, c_2)$, where $\boldsymbol{L}_2 = \{L_2(w) \mid w \in V \setminus \{u\}\}$. Then

$$\forall v \neq u, \quad \mu^{u\leftarrow c_1}_{v,(G,\boldsymbol{L})}(\cdot) = \mu_{v,\mathcal{L}_1}(\cdot), \quad \mu^{u\leftarrow c_2}_{v,(G,\boldsymbol{L})}(\cdot) = \mu_{v,\mathcal{L}_2}(\cdot).$$

Since $(G, \boldsymbol{L}) \in \mathscr{L}$ and $\mathscr{L}$ is downward closed, it holds that both $\mathcal{L}_1, \mathcal{L}_2 \in \mathscr{L}$. By Lemma 6.8, for any $v \neq u$,

$$\forall c \in L_1(v): \quad \mu_{v,\mathcal{L}_1}(c) \geq \frac{1}{\mathrm{e}^{1/\varepsilon_2}\,|L_1(v)|}$$
$$\forall c \in L_2(v): \quad \mu_{v,\mathcal{L}_2}(c) \geq \frac{1}{\mathrm{e}^{1/\varepsilon_2}|L_2(v)|}.$$

On the other hand, since $\mathcal{L}_1, \mathcal{L}_2 \in \mathscr{L}$, for any $v \in V$, it holds that $|L_1(v)| \geq 2$ and $|L_2(v)| \geq 2$ (otherwise, the upper bound for the marginals in Condition 6.2 cannot hold). By the definitions of $\mathcal{L}_1$ and $\mathcal{L}_2$, it holds that $|L_1(v) \cap L_2(v)| \geq \min\{|L_1(v), L_2(v)|\} - 1$ and $\big|\,|L_1(v)| - |L_2(v)|\,\big| \leq 1$. Thus, we can couple $\mu_{v,\mathcal{L}_1}(\cdot)$ and $\mu_{v,\mathcal{L}_2}(\cdot)$ with success probability at least

$$\sum_{c\in L_1(v)\cap L_2(v)} \min\left\{\frac{1}{\mathrm{e}^{1/\varepsilon_2}\,|L_1(v)|}, \frac{1}{\mathrm{e}^{1/\varepsilon_2}\,|L_2(v)|}\right\} \geq \frac{\min\{|L_1(v), L_2(v)|\} - 1}{\mathrm{e}^{1/\varepsilon_2}\max\{|L_1(v)|, |L_2(v)|\}}$$
$$\geq \frac{1}{e^{1/\varepsilon_2}} \cdot \frac{\min\{|L_1(v)|, |L_2(v)|\} - 1}{\min\{|L_1(v)|, |L_2(v)|\} + 1} \geq \frac{1}{3\mathrm{e}^{1/\varepsilon_2}}.$$

By the coupling inequality (Proposition 2.2), we have for any $c_1, c_2 \in L(u)$ and any $v \neq u$,

$$d_{\mathrm{TV}}\left(\mu^{u\leftarrow c_1}_{v,(G,\boldsymbol{L})}, \mu^{u\leftarrow c_2}_{v,(G,\boldsymbol{L})}\right) = d_{\mathrm{TV}}\left(\mu_{v,\mathcal{L}_1}, \mu_{v,\mathcal{L}_2}\right) \leq 1 - \frac{1}{3\mathrm{e}^{1/\varepsilon_2}}.$$

By the definition of $R_{G,\boldsymbol{L}}$, we have that

$$\sum_{v\in V: v\neq u} R_{G,\boldsymbol{L}}(u, v) \leq \left(1 - \frac{1}{3\mathrm{e}^{1/\varepsilon_2}}\right)(|V| - 1)\,. \qquad \square$$

6.3. **Recursive coupling**. We then prove the second part of Lemma 6.4.

Lemma 6.9. *Let $\mathscr{L}$ be a downward closed family of list colouring instances satisfying Condition 6.2 with parameters $\chi > 0$, $0 < \varepsilon_1 < 1$ and $\varepsilon_2 > 0$. For any instance $(G = (V, E), \boldsymbol{L}) \in \mathscr{L}$, it holds that*

$$\forall u \in V, \quad \sum_{v\in V: v\neq u} R_{G,\boldsymbol{L}}(u, v) \leq \frac{1}{(1-\varepsilon_1)\varepsilon_2}.$$

We use the following lemma to prove Lemma 6.9.

Definition 6.10 (self-avoiding walk (SAW)). A path $P = (v_1, v_2, \ldots, v_\ell)$ in a graph G is called a *self-avoiding walk* (SAW) if each v_i and v_{i+1} are adjacent and $v_i \neq v_j$ for all $i \neq j$.

Lemma 6.11. *Let $\mathscr{L}$ be a downward closed family of list colouring instances satisfying Condition 6.2 with parameters $\chi > 0$, $0 < \varepsilon_1 < 1$ and $\varepsilon_2 > 0$. For any instance $(G = (V, E), \boldsymbol{L}) \in \mathscr{L}$ and any two vertices $u, v \in V$ with $u \neq v$, it holds that*

$$R_{G,\boldsymbol{L}}(u, v) \leq \frac{1}{\varepsilon_1 \varepsilon_2} \sum_{\substack{\text{SAW } P = (v_1, v_2, \ldots, v_\ell) \\ u = v_1 \text{ and } v = v_\ell}} \prod_{k=1}^{\ell-1} \frac{\varepsilon_1}{|\Gamma_G(v_k) \setminus \{v_i \mid i < k\}|}, \tag{19}$$

where $\Gamma_G(v_k)$ is the neighbourhood of v_k in G.

We remark that the denominator of each ratio in the RHS of (19) is positive because $v_{k+1} \in \Gamma_G(v_k) \setminus \{v_i \mid i < k\}$ for all $1 \leq k \leq \ell - 1$. Lemma 6.11 is proved in Section 6.3.1 via a recursive coupling argument.

Now, we are ready to prove Lemma 6.9.

Proof of Lemma 6.9. Fix $(G = (V, E), \boldsymbol{L}) \in \mathscr{L}$. For any vertex $u \in V$ and any integer $\ell \geq 1$, we use P^u_ℓ to denote the set of all SAWs from u that contains ℓ vertices. Formally, $P^u_\ell \triangleq \{P = (v_1, v_2, \ldots, v_\ell) \mid P \text{ is a SAW}, v_1 = u\}$. We claim that

$$\forall u \in V, \ell \geq 1, \quad \sum_{\text{SAW } P = (v_1, v_2, \ldots, v_\ell) \in P^u_\ell} \prod_{k=1}^{\ell-1} \frac{\varepsilon_1}{|\Gamma_G(v_k) \setminus \{v_i \mid i < k\}|} \leq \varepsilon_1^{\ell-1}. \tag{20}$$

We now use (20) to prove Lemma 6.9. By Lemma 6.11, for any $u \in V$,

$$\begin{aligned}
\sum_{v \in V: v \neq u} R_{G,\boldsymbol{L}}(u, v) &\leq \frac{1}{\varepsilon_1 \varepsilon_2} \cdot \sum_{v \in V: v \neq u} \sum_{\substack{\text{SAW } P = (v_1, v_2, \ldots, v_\ell) \\ u = v_1 \text{ and } v = v_\ell}} \prod_{k=1}^{\ell-1} \frac{\varepsilon_1}{|\Gamma_G(v_k) \setminus \{v_i \mid i < k\}|} \\
(\star) \quad &\leq \frac{1}{\varepsilon_1 \varepsilon_2} \cdot \sum_{\ell=2}^{\infty} \sum_{\text{SAW } P = (v_1, v_2, \ldots, v_\ell) \in P^u_\ell} \prod_{k=1}^{\ell-1} \frac{\varepsilon_1}{|\Gamma_G(v_k) \setminus \{v_i \mid i < k\}|} \\
(\text{by } (20)) \quad &\leq \frac{1}{\varepsilon_1 \varepsilon_2} \cdot \sum_{\ell=2}^{\infty} \varepsilon_1^{\ell-1} = \frac{1}{(1 - \varepsilon_1)\varepsilon_2},
\end{aligned}$$

where $(\star)$ is due to the fact that $v \neq u$ implies all SAWs in consideration are of length at least 2. This proves Lemma 6.9.

We then prove (20) by an induction on ℓ. If $\ell = 1$, the LHS of (20) is 1, thus (20) holds trivially. Suppose (20) holds for all $\ell \leq t$, we prove it for $\ell = t + 1$. Let $P^{u \to v}_t$ denote the set of all SAWs from u to v that contains t vertices. Formally,

$$P^{u \to v}_t \triangleq \{P = v_1, v_2, \ldots, v_t \mid P \text{ is a SAW}, v_1 = u, v_t = v\}.$$

Hence, $P^u_t = \bigcup_{v \in V} P^{u \to v}_t$. If $P \in P^u_{t+1}$ is a SAW such that $P = v_1, v_2, \ldots, v_t, v_{t+1}$, then the prefix $P' = v_1, v_2, \ldots, v_t$ is in the set $P^{u \to v_t}_t$ and $v_{t+1} \in \Gamma_G(v_t) \setminus \{v_i \mid i < t\}$, and vice versa. This implies that

$$P^u_{t+1} = \bigcup_{v \in V} \left\{(P, w) \mid P = (v_1 = u, v_2, \ldots, v_t = v) \in P^{u \to v}_t, w \in \Gamma_G(v) \setminus \{v_i \mid i < t\}\right\}, \tag{21}$$

where (P, w) is the path obtained by appending w at the end of the path P. We have

$$
\begin{aligned}
&\forall u \in V, && \sum_{\text{SAW } P=(v_1,v_2,\ldots,v_{t+1})\in P^u_{t+1}} \prod_{k=1}^{t} \frac{\varepsilon_1}{|\Gamma_G(v_k)\setminus\{v_i \mid i<k\}|}\\
&\text{(by (21))} && = \sum_{v\in V}\sum_{\text{SAW } P=(v_1,v_2,\ldots,v_t)\in P^{u\to v}_{t}}\ \sum_{w\in\Gamma_G(v)\setminus\{v_i\mid i<t\}} \prod_{k=1}^{t} \frac{\varepsilon_1}{|\Gamma_G(v_k)\setminus\{v_i \mid i<k\}|}\\
&(\star) && \le \varepsilon_1\cdot\sum_{v\in V}\sum_{\text{SAW } P=(v_1,v_2,\ldots,v_t)\in P^{u\to v}_{t}} \prod_{k=1}^{t-1} \frac{\varepsilon_1}{|\Gamma_G(v_k)\setminus\{v_i \mid i<k\}|}\\
& && = \varepsilon_1\cdot\sum_{\text{SAW } P=(v_1,v_2,\ldots,v_t)\in P^{u}_{t}} \prod_{k=1}^{t-1} \frac{\varepsilon_1}{|\Gamma_G(v_k)\setminus\{v_i \mid i<k\}|}\\
&\text{(by I.H.)} && \le \varepsilon_1^t.
\end{aligned}
$$

The inequality $(\star)$ holds because $\Gamma_G(v)\setminus\{v_i \mid i<t\} = \Gamma_G(v_t)\setminus\{v_i \mid i<t\}$ (due to $v_t = v$). We remark the $(\star)$ is an inequality rather than an equality because $\Gamma_G(v)\setminus\{v_i \mid i<t\}$ can be empty. This proves (20). □

6.3.1. *Influence bounds via recursion.* Now we prove Lemma 6.11. The proof technique is based on the "recursive coupling" introduced by Goldberg, Martin and Paterson [GMP05].

We introduce some definitions. Let $(G, \boldsymbol{L})$ be a list colouring instance, where $G = (V, E)$. Fix a vertex $u \in V$ and two colours $c_1, c_2 \in L(u)$. Let $w_1, w_2, \ldots, w_m$ denote the neighbours of u in graph G, where $m = \deg_G(u)$. For any $0 \le k \le m$, we define a list colouring instance $(G_u, \boldsymbol{L}^{c_1,c_2}_{u,k})$: The graph $G_u = G[V\setminus\{u\}]$ is obtained by removing vertex u from G. The colour list $\boldsymbol{L}^{c_1,c_2}_{u,k}$ is obtained by removing the colour c_1 from the lists $L(w_\ell)$ for $\ell < k$, and removing the colour c_2 for the lists $L(w_\ell)$ for $\ell > k$. Formally,

$$
\forall v \in V\setminus\{u\}: \quad L^{c_1,c_2}_{u,k}(v) = \begin{cases} L(v)\setminus\{c_1\} & \text{if } v \in \{w_1, w_2, \ldots, w_{k-1}\}\\ L(v)\setminus\{c_2\} & \text{if } v \in \{w_{k+1}, w_{k+2}, \ldots, w_m\}\\ L(v) & \text{if } v \notin \Gamma_G(u) \text{ or } v = w_k. \end{cases} \tag{22}
$$

Lemma 6.12. *Let $\mathscr{L}$ be a downward closed family of list colouring instances satisfying Condition 6.2 with parameters $\chi > 0$, $0 < \varepsilon_1 < 1$ and $\varepsilon_2 > 0$. For any $(G = (V, E), \boldsymbol{L}) \in \mathscr{L}$, the following result holds. Fix a pair of vertices $u, v \in V$. Let $w_1, w_2, \ldots, w_{\deg_G(u)}$ denote the neighbours of u in G. Let $c_1, c_2 \in L(u)$ be the colours achieving the maximum in (18) (breaking ties arbitrarily). Then,*

$$
R_{G,\boldsymbol{L}}(u, v) \le \begin{cases} 1 & \textit{if } u = v;\\ 0 & \textit{if } u \textit{ and } v \textit{ are disconnected in } G;,\\ \sum_{k=1}^{\deg_G(u)} \alpha_k \cdot R_{G_u,\boldsymbol{L}^{c_1,c_2}_{u,k}}(w_k, v) & \textit{otherwise.} \end{cases}
$$

where for all $1 \le k \le \deg_G(u)$,

$$
\alpha_k = \min\left(\frac{\varepsilon_1}{\deg_{G_u}(w_k)}, \frac{1}{\varepsilon_2\chi + 1}\right).
$$

We remark that if $\deg_{G_u}(w_k) = 0$, then by convention we have $\frac{\varepsilon_1}{\deg_{G_u}(w_k)} = \infty$ and thus $\alpha_k = \frac{1}{\varepsilon_2\chi+1}$.

Now we use Lemma 6.12 to derive Lemma 6.11 and defer the proof of Lemma 6.12 to Section 6.3.2.

Proof of Lemma 6.11. Suppose $(G = (V, E), \boldsymbol{L}) \in \mathscr{L}$. It is clear that the instance $(G_u, \boldsymbol{L}^{c_1,c_2}_{u,k})$ obtained from $(G, \boldsymbol{L})$ is also in $\mathscr{L}$. Hence, we can use Lemma 6.12 recursively. This implies that for any $(G, \boldsymbol{L}) \in \mathscr{L}$, any

$u, v \in V$,

$$R_{G,\boldsymbol{L}}(u,v) \le \sum_{\substack{\text{SAW } P=(v_1,v_2,\ldots,v_\ell)\\ u=v_1 \text{ and } v=v_\ell}} \prod_{k=2}^{\ell} \min\left(\frac{\varepsilon_1}{|\Gamma_G(v_k)\setminus\{v_i \mid i<k\}|}, \frac{1}{\varepsilon_2\chi+1}\right). \tag{23}$$

If u and v are disconnected, then $R_{G,\boldsymbol{L}}(u,v)=0$, and in this case, the RHS of (19) is 0 because there is no SAW from u to v, thus (19) holds. So in the following we assume that u and v are connected.

Our goal is to prove (19). Comparing (23) with (19), the main difference is the range of k in the product. We will trade the last factor of $\frac{1}{\varepsilon_2\chi+1}$ for $k=\ell$ by a factor of $\frac{\varepsilon_1}{\chi}$ for $k=1$, with a loss of $\frac{1}{\varepsilon_1\varepsilon_2}$.

More precisely, by (23), we have

$$\begin{aligned}
R_{G,\boldsymbol{L}}(u,v) &\overset{(\star)}{\le} \frac{1}{\varepsilon_2\chi+1} \sum_{\substack{\text{SAW } P=(v_1,v_2,\ldots,v_\ell)\\ u=v_1 \text{ and } v=v_\ell}} \left(\prod_{k=2}^{\ell-1} \frac{\varepsilon_1}{|\Gamma_G(v_k)\setminus\{v_i \mid i<k\}|}\right)\\
(\text{as } 0<\deg_G(v_1)\le\chi) \quad &\le \frac{\chi}{\varepsilon_1(\varepsilon_2\chi+1)} \sum_{\substack{\text{SAW } P=(v_1,v_2,\ldots,v_\ell)\\ u=v_1 \text{ and } v=v_\ell}} \left(\prod_{k=1}^{\ell-1} \frac{\varepsilon_1}{|\Gamma_G(v_k)\setminus\{v_i \mid i<k\}|}\right)\\
&\le \frac{1}{\varepsilon_1\varepsilon_2} \sum_{\substack{\text{SAW } P=(v_1,v_2,\ldots,v_\ell)\\ u=v_1 \text{ and } v=v_\ell}} \left(\prod_{k=1}^{\ell-1} \frac{\varepsilon_1}{|\Gamma_G(v_k)\setminus\{v_i \mid i<k\}|}\right),
\end{aligned}$$

where inequality $(\star)$ holds due to (23) and $\ell \ge 2$ (since $u \ne v$). Note that in the formula above, it holds that $|\Gamma_G(v_k)\setminus\{v_i \mid i<k\}| > 0$ for all $1 \le k \le \ell-1$ because $v_{k+1} \in \Gamma_G(v_k)\setminus\{v_i \mid i<k\}$. This proves Lemma 6.11. ◻

6.3.2. *Establish recursion via coupling.* Next, we prove Lemma 6.12. Fix an instance $(G=(V,E),\boldsymbol{L}) \in \mathscr{L}$. Fix a vertex $u \in V$. Let $c_1, c_2 \in L(u)$ be the colours achieving the maximum in (18) (breaking ties arbitrarily). Our goal is to bound

$$R_{G,\boldsymbol{L}}(u,v) = \max_{c_1,c_2\in L(u)} d_{\mathrm{TV}}\left(\mu^{u\leftarrow c_1}_{v,(G,\boldsymbol{L})}, \mu^{u\leftarrow c_2}_{v,(G,\boldsymbol{L})}\right).$$

If $u=v$, then $R_{G,\boldsymbol{L}}(u,v) \le 1$. If u and v are disconnected in G, then $R_{G,\boldsymbol{L}}(u,v)=0$. In the rest of this section, we assume $u \ne v$ and u, v are connected in graph G.

Let $w_1, w_2, \ldots, w_m$ denote the neighbours of u in G, where $m=\deg_G(u)$. We construct a graph G' from G as follows. We remove the vertex u from the graph G, add m new vertices $u_1, u_2, \ldots, u_m$, and then add m new edges $\{u_i, w_i\}$. Finally, we define a set of colour lists $\boldsymbol{L}' = \{L'(v) \mid v \in V\setminus\{u\} \cup \{u_1,u_2,\ldots,u_m\}\}$ as

$$L'(v) \triangleq \begin{cases} L(u) & \text{if } v \in \{u_1,u_2,\ldots,u_m\}\\ L(v) & \text{if } v \in V\setminus\{u\}.\end{cases}$$

This defines a new list colouring instance $(G',\boldsymbol{L}')$. Figure 1 gives a small example.

For each $0 \le k \le m$, we define a set of partial colourings σ_k on $\{u_1,u_2,\ldots,u_m\}$ by

$$\sigma_k(u_j) \triangleq \begin{cases} c_1 & \text{if } 1 \le j \le k\\ c_2 & \text{if } k+1 \le j \le m.\end{cases}$$

Then, it holds that $\mu^{u\leftarrow c_1}_{v,(G,\boldsymbol{L})} = \mu^{\sigma_m}_{v,(G',\boldsymbol{L}')}$ and $\mu^{u\leftarrow c_2}_{v,(G,\boldsymbol{L})} = \mu^{\sigma_0}_{v,(G',\boldsymbol{L}')}$. By the triangle inequality, we have

$$d_{\mathrm{TV}}\left(\mu^{u\leftarrow c_1}_{v,(G,\boldsymbol{L})}, \mu^{u\leftarrow c_2}_{v,(G,\boldsymbol{L})}\right) = d_{\mathrm{TV}}\left(\mu^{\sigma_0}_{v,(G',\boldsymbol{L}')}, \mu^{\sigma_m}_{v,(G',\boldsymbol{L}')}\right) \le \sum_{k=1}^{m} d_{\mathrm{TV}}\left(\mu^{\sigma_{k-1}}_{v,(G',\boldsymbol{L}')}, \mu^{\sigma_k}_{v,(G',\boldsymbol{L}')}\right). \tag{24}$$

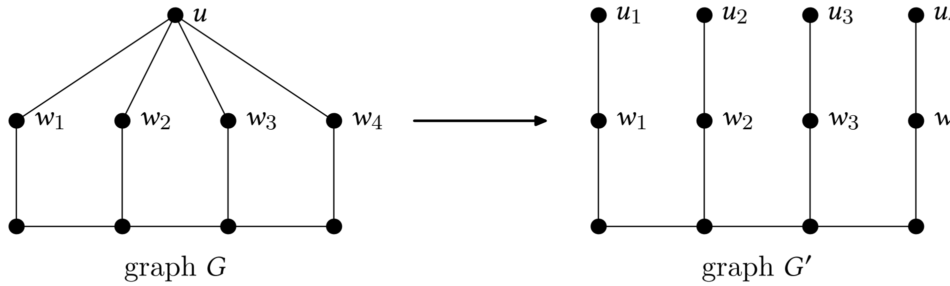

FIGURE 1. Split vertex u to modify the graph G to G'

We now bound $d_{\mathrm{TV}}\left(\mu^{\sigma_{k-1}}_{v,(G',\boldsymbol{L}')},\mu^{\sigma_k}_{v,(G',\boldsymbol{L}')}\right)$ for each $1 \le k \le m$. Consider the following coupling procedure C.

- sample $c, c' \in L'(w_k) = L(w_k)$ from the optimal coupling of $\mu^{\sigma_{k-1}}_{w_k,(G',\boldsymbol{L}')}$ and $\mu^{\sigma_k}_{w_k,(G',\boldsymbol{L}')}$.
- sample c_v, c'_v from the optimal coupling of $\mu^{\sigma_{k-1},w_k\leftarrow c}_{v,(G',\boldsymbol{L}')}$ and $\mu^{\sigma_k,w_k\leftarrow c'}_{v,(G',\boldsymbol{L}')}$.

By the definition of σ_k and σ_{k-1}, they differ only at one vertex u_k. By the construction of the graph G', u_k is only adjacent to w_k. Then conditional on the colour of w_k, the colour of u_k is independent from the colour of v. Hence, in this coupling, we know that $c_v \ne c'_v$ can happen only if $c \ne c'$. Since c, c' are sampled from the optimal coupling, we have $\mathbf{Pr}_C\left[c \ne c'\right] = d_{\mathrm{TV}}\left(\mu^{\sigma_{k-1}}_{w_k,(G',\boldsymbol{L}')},\mu^{\sigma_k}_{w_k,(G',\boldsymbol{L}')}\right)$. Therefore,

$$\begin{aligned} d_{\mathrm{TV}}\left(\mu^{\sigma_{k-1}}_{v,(G',\boldsymbol{L}')},\mu^{\sigma_k}_{v,(G',\boldsymbol{L}')}\right) &\le \mathbf{Pr}_C\left[c_v \ne c'_v\right] \\ &\le d_{\mathrm{TV}}\left(\mu^{\sigma_{k-1}}_{w_k,(G',\boldsymbol{L}')},\mu^{\sigma_k}_{w_k,(G',\boldsymbol{L}')}\right)\cdot \max_{c,c'\in L'(w_k)} d_{\mathrm{TV}}\left(\mu^{\sigma_{k-1},w_k\leftarrow c}_{v,(G',\boldsymbol{L}')},\mu^{\sigma_k,w_k\leftarrow c'}_{v,(G',\boldsymbol{L}')}\right). \end{aligned}$$

Recall that the graph G_u is obtained by removing u from G, and the colour lists $\boldsymbol{L}^{c_1,c_2}_{u,k}$ is defined in (22). We can further derive

$$\begin{aligned} d_{\mathrm{TV}}\left(\mu^{\sigma_{k-1}}_{v,(G',\boldsymbol{L}')},\mu^{\sigma_k}_{v,(G',\boldsymbol{L}')}\right) &\le d_{\mathrm{TV}}\left(\mu^{\sigma_{k-1}}_{w_k,(G',\boldsymbol{L}')},\mu^{\sigma_k}_{w_k,(G',\boldsymbol{L}')}\right)\cdot \max_{c,c'\in L'(w_k)} d_{\mathrm{TV}}\left(\mu^{\sigma_{k-1},w_k\leftarrow c}_{v,(G',\boldsymbol{L}')},\mu^{\sigma_k,w_k\leftarrow c'}_{v,(G',\boldsymbol{L}')}\right) \\ (\star) \qquad &= d_{\mathrm{TV}}\left(\mu^{\sigma_{k-1}}_{w_k,(G',\boldsymbol{L}')},\mu^{\sigma_k}_{w_k,(G',\boldsymbol{L}')}\right)\cdot \max_{c,c'\in L^{c_1,c_2}_{u,k}(w_k)} d_{\mathrm{TV}}\left(\mu^{w_k\leftarrow c}_{v,(G_u,\boldsymbol{L}^{c_1,c_2}_{u,k})},\mu^{w_k\leftarrow c'}_{v,(G_u,\boldsymbol{L}^{c_1,c_2}_{u,k})}\right) \\ (25) \qquad &= d_{\mathrm{TV}}\left(\mu^{\sigma_{k-1}}_{w_k,(G',\boldsymbol{L}')},\mu^{\sigma_k}_{w_k,(G',\boldsymbol{L}')}\right)\cdot R_{G_u,\boldsymbol{L}^{c_1,c_2}_{u,k}}(w_k,v). \end{aligned}$$

Equation $(\star)$ holds due to $L'(w_k) = L(w_k) = L^{c_1,c_2}_{u,k}(w_k)$ and the definitions of instances $(G',\boldsymbol{L}')$ and $(G_u,\boldsymbol{L}^{c_1,c_2}_{u,k})$.

Now, our task is reduced to bound $d_{\mathrm{TV}}\left(\mu^{\sigma_{k-1}}_{w_k,(G',\boldsymbol{L}')},\mu^{\sigma_k}_{w_k,(G',\boldsymbol{L}')}\right)$. Let S denote $\{u_1,\dots,u_m\}$. We define two list colouring instances $(G^*_1,\boldsymbol{L}^*_1) = \mathsf{Pin}_{G',\boldsymbol{L}'}(S,\sigma_{k-1})$ and $(G^*_2,\boldsymbol{L}^*_2) = \mathsf{Pin}_{G',\boldsymbol{L}'}(S,\sigma_k)$. Then we have $\mu^{\sigma_{k-1}}_{w_k,(G',\boldsymbol{L}')} = \mu_{w_k,(G^*_1,\boldsymbol{L}^*_1)}$ and $\mu^{\sigma_k}_{w_k,(G',\boldsymbol{L}')} = \mu_{w_k,(G^*_2,\boldsymbol{L}^*_2)}$. Thus

$$d_{\mathrm{TV}}\left(\mu^{\sigma_{k-1}}_{w_k,(G',\boldsymbol{L}')},\mu^{\sigma_k}_{w_k,(G',\boldsymbol{L}')}\right) = d_{\mathrm{TV}}\left(\mu_{w_k,(G^*_1,\boldsymbol{L}^*_1)},\mu_{w_k,(G^*_2,\boldsymbol{L}^*_2)}\right). \tag{26}$$

Besides, $G^*_1 = G^*_2 = G_u$ and both $(G^*_1,\boldsymbol{L}^*_1),(G^*_2,\boldsymbol{L}^*_2)$ can be obtained from $(G,\boldsymbol{L})$ by removing u and removing certain colours from $L(u_k)$ for $k = 1,\dots,m$. So we have that $(G^*_1,\boldsymbol{L}^*_1),(G^*_2,\boldsymbol{L}^*_2) \in \mathscr{L}$ since $\mathscr{L}$ is downward closed. Moreover, the two collections of colour lists $\boldsymbol{L}^*_1 = \left\{L^*_1(v) \mid v \in V\setminus\{u\}\right\}$ and $\boldsymbol{L}^*_2 = \left\{L^*_2(v) \mid v \in V\setminus\{u\}\right\}$ can only differ at w_k where $L^*_1(w_k) = L(w_k)\setminus\{c_2\}$ and $L^*_2(w_k) = L(w_k)\setminus\{c_1\}$.

We prove an auxiliary lemma.

Lemma 6.13. *Let* $\deg_G(w_k)$ *denote the degree of* w_k *in* G.

$$d_{\mathrm{TV}}\left(\mu_{w_k,(G_1^*,\boldsymbol{L}_1^*)},\mu_{w_k,(G_2^*,\boldsymbol{L}_2^*)}\right)\le\min\left(\frac{\varepsilon_1}{\deg_G(w_k)-1},\frac{1}{\varepsilon_2\chi+1}\right).$$

Proof. It suffices to prove that

$$d_{\mathrm{TV}}\left(\mu_{w_k,(G_1^*,\boldsymbol{L}_1^*)},\mu_{w_k,(G_2^*,\boldsymbol{L}_2^*)}\right)=\max\left\{\mu_{w_k,(G_1^*,\boldsymbol{L}_1^*)}(c_1),\mu_{w_k,(G_2^*,\boldsymbol{L}_2^*)}(c_2)\right\}. \tag{27}$$

To see that (27) implies the lemma, note that $(G,\boldsymbol{L})\in\mathscr{L}$, thus $\deg_{G_u}(w_k)=\deg_G(w_k)-1\le\chi-1$. Since $(G_1^*,\boldsymbol{L}_1^*),(G_2^*,\boldsymbol{L}_2^*)\in\mathscr{L}$, Condition 6.2 gives

$$\max\left\{\mu_{w_k,(G_1^*,\boldsymbol{L}_1^*)}(c_1),\mu_{w_k,(G_2^*,\boldsymbol{L}_2^*)}(c_2)\right\}\le\min\left(\frac{\varepsilon_1}{\deg_{G_u}(w_k)},\frac{1}{\varepsilon_2\chi+1}\right)=\min\left(\frac{\varepsilon_1}{\deg_G(w_k)-1},\frac{1}{\varepsilon_2\chi+1}\right).$$

It remains to verify (27). Note that assuming Condition 6.2, the distributions in (27) are well-defined.

Let $(G_u,\widetilde{\boldsymbol{L}})$ be a list colouring instance where $\widetilde{\boldsymbol{L}}=\left\{\widetilde{L}(v)\mid v\in V\setminus\{u\}\right\}$ differs from $\boldsymbol{L}_1^*$ and $\boldsymbol{L}_2^*$ only on w_k, and $\widetilde{L}(w_k)=L(w_k)$. For each colour c, define $n(c)$ as the number of proper list colourings of $(G_u,\widetilde{\boldsymbol{L}})$ such that the colour of w_k is c. Note that $n(c)=0$ if $c\notin\widetilde{L}(w_k)$. Define

$$N\triangleq\sum_{c\in\widetilde{L}(w_k)\setminus\{c_1,c_2\}}n(c).$$

We claim that

$$d_{\mathrm{TV}}\left(\mu_{w_k,(G_1^*,\boldsymbol{L}_1^*)},\mu_{w_k,(G_2^*,\boldsymbol{L}_2^*)}\right)=\frac{\max\{n(c_1),n(c_2)\}}{N+\max\{n(c_1),n(c_2)\}}. \tag{28}$$

This implies (27) as the RHS of (28) equals $\max\left\{\mu_{w_k,(G_1^*,\boldsymbol{L}_1^*)}(c_1),\mu_{w_k,(G_2^*,\boldsymbol{L}_2^*)}(c_2)\right\}$. To show (28), we may assume $n(c_1)\ge n(c_2)$ first. Then,

$$\begin{aligned}
d_{\mathrm{TV}}\left(\mu_{w_k,(G_1^*,\boldsymbol{L}_1^*)},\mu_{w_k,(G_2^*,\boldsymbol{L}_2^*)}\right)&=\frac{1}{2}\left(\sum_{c\in L(w_k)\setminus\{c_1,c_2\}}\left|\frac{n(c)}{N+n(c_1)}-\frac{n(c)}{N+n(c_2)}\right|+\frac{n(c_1)}{N+n(c_1)}+\frac{n(c_2)}{N+n(c_2)}\right)\\
(\text{as } n(c_1)\ge n(c_2))\qquad&=\frac{1}{2}\left(\frac{N(n(c_1)-n(c_2))}{(N+n(c_1))(N+n(c_2))}+\frac{n(c_1)N+n(c_2)N+2n(c_1)n(c_2)}{(N+n(c_1))(N+n(c_2))}\right)\\
&=\frac{n(c_1)}{N+n(c_1)}.
\end{aligned}$$

Similarly, $d_{\mathrm{TV}}\left(\mu_{w_k,(G_1^*,\boldsymbol{L}_1^*)},\mu_{w_k,(G_2^*,\boldsymbol{L}_2^*)}\right)=\frac{n(c_2)}{N+n(c_2)}$ if $n(c_2)>n(c_1)$. This shows (28). □

Combining (26) and Lemma 6.13, we have

$$\begin{aligned}
d_{\mathrm{TV}}\left(\mu^{\sigma_{k-1}}_{u_k,(G',\boldsymbol{L}')},\mu^{\sigma_k}_{u_k,(G',\boldsymbol{L}')}\right)&=d_{\mathrm{TV}}\left(\mu_{w_k,(G_1^*,\boldsymbol{L}_1^*)},\mu_{w_k,(G_2^*,\boldsymbol{L}_2^*)}\right)\\
(\text{by Lemma 6.13})\qquad&\le\min\left(\frac{\varepsilon_1}{\deg_G(w_k)-1},\frac{1}{\varepsilon_2\chi+1}\right)\\
&=\min\left(\frac{\varepsilon_1}{\deg_{G_u}(w_k)},\frac{1}{\varepsilon_2\chi+1}\right), \qquad (29)
\end{aligned}$$

where G_u is the subgraph of G induced by $V\setminus\{u\}$. By (24), (25) and (29), we have

$$R_{G,\boldsymbol{L}}(u,v)\le\sum_{k=1}^{\deg_G(u)}\min\left(\frac{\varepsilon_1}{\deg_{G_u}(w_k)},\frac{1}{\varepsilon_2\chi+1}\right)\cdot R_{G_u,\boldsymbol{L}_{u,k}^{c_1,c_2}}(w_k,v).$$

This proves Lemma 6.12.

6.4. **Verify Condition 6.2 (Proof of Theorem 6.1).** We first introduce the following lemma.

Lemma 6.14. *Let $(G = (V, E), \mathbf{L})$ be an instance of list colouring where G is triangle-free and $\mathbf{L} = \{L(v) \mid v \in V\}$. Let $\Delta \geq 3$ be the maximum degree of G and $\delta > 0$ be a constant. Assume for every $v \in V$, it holds that*

$$|L(v)| - \deg_G(v) \geq (\alpha^* + \delta - 1)\Delta.$$

Let $\mathscr{L}$ be the downward closure of $(G, \mathbf{L})$. Then $\mathscr{L}$ satisfies Condition 6.2 with parameters $\chi = \Delta$, $\varepsilon_1 = 1 - \frac{\delta}{\alpha^*+\delta}$ and $\varepsilon_2 = 0.4 + \delta$.

It is clear that Theorem 6.1 is a consequence of Lemma 6.14 and Theorem 6.3.

Proof of Theorem 6.1. Suppose the instance $(G, \mathbf{L})$ satisfies the condition in (15). By Lemma 6.14, the downward closure $\mathscr{L}$ of $(G, \mathbf{L})$ satisfies Condition 6.2 with parameters $\chi = \Delta$, $\varepsilon_1 = 1 - \frac{\delta}{\alpha^*+\delta}$ and $\varepsilon_2 = 0.4 + \delta$. By Theorem 6.3, we have

$$T_{\mathsf{mix}}(\varepsilon) \leq \left(9\mathrm{e}^{\frac{2}{\varepsilon_2}}\right)^{\left(1+\frac{1}{(1-\varepsilon_1)\varepsilon_2}\right)} n^{1+\frac{2}{(1-\varepsilon_1)\varepsilon_2}} \cdot \log\left(\frac{M}{\varepsilon}\right).$$

Note that $\frac{2}{\varepsilon_2} = \frac{2}{0.4+\delta} \leq 5$ and $\frac{1}{(1-\varepsilon_1)\varepsilon_2} = \frac{\alpha^*+\delta}{\delta(0.4+\delta)} \leq \frac{1}{\delta} \cdot \frac{\alpha^*}{0.4} \leq \frac{9}{2\delta}$. Thus, we have

$$T_{\mathsf{mix}}(\varepsilon) \leq \left(9\mathrm{e}^5\right)^{\left(1+\frac{9}{2\delta}\right)} n^{1+\frac{9}{\delta}} \cdot \log\left(\frac{M}{\varepsilon}\right) \leq \left(9\mathrm{e}^5 n\right)^{1+9/\delta} \cdot \log\left(\frac{M}{\varepsilon}\right).$$ □

6.4.1. *Proof of Lemma 6.14.* We first remark that $\chi \geq 3$. We then claim that every instance $(G = (V, E), \mathbf{L} = \{L(v) \mid v \in V\}) \in \mathscr{L}$ satisfies

$$\forall v \in V : |L(v)| - \deg_G(v) \geq (\alpha^* + \delta - 1)\chi \tag{30}$$

and G is triangle-free. To see this, we only need to notice that (30) is preserved by the $\preceq$ relation, namely if $(G', \mathbf{L}')$ satisfies (30) and $(G, \mathbf{L}) \preceq (G', \mathbf{L}')$, then $(G, \mathbf{L})$ satisfies (30) as well. This holds since by the definition of $\preceq$, $(G, \mathbf{L})$ can be obtained from $(G', \mathbf{L}')$ by removing some vertex v and removing at most one colour from the colour lists of v's neighbours. Therefore, once the size of the colour list of certain vertex u decreases by one, its degree must decrease by one as well. So the LHS of (30) never decreases. Besides, it is easy to see all graphs in $\mathscr{L}$ are triangle-free.

By (30) and $\chi \geq 3$, for any $(G, \mathbf{L}) \in \mathscr{L}$, $|L(v)| \geq \deg_G(v) + 3(\alpha^* + \delta - 1) \geq \deg_G(v) + 2$ for any vertex v. One can construct a proper list colouring using a simple greedy procedure. Hence, a proper list colouring exists for any instance in $\mathscr{L}$.

We fix a list colouring instance $(G = (V, E), \mathbf{L}) \in \mathscr{L}$. We first prove

$$\mu_{v,(G,\mathbf{L})}(c) \leq \frac{1}{(\alpha^* + \delta - 1)\chi} \overset{(\star)}{\leq} \frac{1}{(0.4 + \delta)\chi + 1}, \tag{31}$$

where $(\star)$ holds due to $\chi \geq 3$, so we can pick $\varepsilon_2 = 0.4 + \delta$. Conditional on any colouring of $\Gamma_G(v)$, vertex v has at least $(\alpha^* + \delta - 1)\chi$ available colours and therefore the marginal probability is at most $\frac{1}{(\alpha^*+\delta-1)\chi} \leq \frac{1}{(0.4+\delta)\chi+1}$. Since $\mu_{v,(G,\mathbf{L})}(c)$ is a convex combination of these conditional probabilities, the upper bound follows.

Next, fix a vertex $v \in V$ with $\deg_G(v) \leq \chi - 1$. We prove $\mu_{v,(G,\mathbf{L})}(c) \leq \frac{1-\delta/(\alpha^*+\delta)}{\deg_G(v)}$, so we can pick $\varepsilon_1 = 1 - \frac{\delta}{\alpha^*+\delta}$. Let $\Gamma_G^+(v) = \Gamma_G(v) \cup \{v\}$ denote inclusive neighbourhood of v. We show that, conditional on any colouring σ of $V \setminus \Gamma_G^+(v)$, the marginal probability $\mu^\sigma_{v,(G,\mathbf{L})}(c) \leq \frac{1-\delta/(\alpha^*+\delta)}{\deg_G(v)}$. Define a new instance $(\widetilde{G}, \widetilde{\mathbf{L}}) = \mathsf{Pin}_{G,\mathbf{L}}(V \setminus \Gamma_G^+(v), \sigma)$, where $\mathsf{Pin}_\cdot(\cdot)$ is in Definition 6.5. Since $\mathscr{L}$ is downward closed, $(\widetilde{G}, \widetilde{\mathbf{L}}) \in \mathscr{L}$. Let $m = \deg_G(v) = \deg_{\widetilde{G}}(v)$. It suffices to prove that

$$\forall c \in L(v) = \widetilde{L}(v), \quad \mu_{v,(\widetilde{G},\widetilde{\mathbf{L}})}(c) \leq \frac{1 - \delta/(\alpha^* + \delta)}{m}. \tag{32}$$

Note that if $m = 0$, (32) holds trivially. If $m = 1$ or $m = 2$, by $(\widetilde{G}, \widetilde{\boldsymbol{L}}) \in \mathscr{L}$, $\chi \geq 3$ and (31), we have

$$\mu_{v,(\widetilde{G},\widetilde{L})}(c) \leq \frac{1}{(\alpha^* + \delta - 1)\chi} \leq \frac{1}{3(\alpha^* + \delta - 1)} \overset{(\star)}{\leq} \frac{1 - \delta/(\alpha^* + \delta)}{2} \leq \frac{1 - \delta/(\alpha^* + \delta)}{m},$$

where $(\star)$ holds because $\frac{1}{3(\alpha^*+\delta-1)} \leq \frac{\alpha^*}{2(\alpha^*+\delta)}$ for all $\delta > 0$.

Now, we assume $m \geq 3$. Let $v_1, v_2, \dots, v_m$ denote the neighbours of v in $\widetilde{G}$. For each $1 \leq i \leq m$, define $s_i = |\widetilde{L}(v_i)|$, and for any colour b, let $\delta_{i,b} = 1$ if $b \in \widetilde{L}(v_i)$; and $\delta_{i,b} = 0$ if $b \notin \widetilde{L}(v_i)$. Since $\widetilde{G}$ is a triangle-free graph, we have for any $\forall c \in L(v) = \widetilde{L}(v)$,

$$\mu_{v,(\widetilde{G},\widetilde{L})}(c) = \frac{\prod_{i=1}^m (s_i - \delta_{i,c})}{\sum_{b \in L(v)} \prod_{i=1}^m (s_i - \delta_{i,b})} = \frac{\prod_{i=1}^m \left(1 - \frac{\delta_{i,c}}{s_i}\right)}{\sum_{b \in L(v)} \prod_{i=1}^m \left(1 - \frac{\delta_{i,b}}{s_i}\right)} \leq \frac{1}{\sum_{b \in L(v)} \prod_{i=1}^m \left(1 - \frac{\delta_{i,b}}{s_i}\right)}. \tag{33}$$

We give a lower bound for denominator. Let $s_v = |L(v)|$. By the AM-GM inequality, we have

$$\sum_{b \in L(v)} \prod_{i=1}^m \left(1 - \frac{\delta_{i,b}}{s_i}\right) \geq s_v \left(\prod_{b \in L(v)} \prod_{i=1}^m \left(1 - \frac{\delta_{i,b}}{s_i}\right)\right)^{1/s_v} = s_v \left(\prod_{i=1}^m \prod_{b \in L(v) \cap \widetilde{L}(v_i)} \left(1 - \frac{1}{s_i}\right)\right)^{1/s_v}, \tag{34}$$

where the last equality holds because $\delta_{i,b} = 1$ if and only if $b \in \widetilde{L}(v_i)$. Note that $(L(v) \cap \widetilde{L}(v_i)) \subseteq \widetilde{L}(v_i)$ and $s_i = |\widetilde{L}(v_i)|$, which implies $|L(v) \cap \widetilde{L}(v_i)| \leq s_i$. We have that

$$\prod_{i=1}^m \prod_{b \in L(v) \cap \widetilde{L}(v_i)} \left(1 - \frac{1}{s_i}\right) \geq \prod_{i=1}^m \left(1 - \frac{1}{s_i}\right)^{s_i}.$$

Let $p = (\alpha^* + \delta - 1)m + 0.5$. Since $(\widetilde{G}, \widetilde{\boldsymbol{L}}) \in \mathscr{L}$ and $m = \deg_G(v) \leq \chi - 1$, for all $1 \leq i \leq m$,

$$s_i \geq (\alpha^* + \delta - 1)\chi \geq (\alpha^* + \delta - 1)(m + 1) \geq (\alpha^* + \delta - 1)m + 0.5 = p.$$

Note that $p > 1$ because $m \geq 3$. Also note that $f(x) = (1 - 1/x)^x$ is increasing when $x \geq 1$. Then we have $\prod_{i=1}^m \left(1 - \frac{1}{s_i}\right)^{s_i} \geq \left(1 - \frac{1}{p}\right)^{mp}$. By (34), we have

$$\sum_{b \in L(v)} \prod_{i=1}^m \left(1 - \frac{\delta_{i,b}}{s_i}\right) \geq s_v \left(1 - \frac{1}{p}\right)^{\frac{mp}{s_v}}.$$

Since $(\widetilde{G}, \widetilde{\boldsymbol{L}}) \in \mathscr{L}$ and $m = \deg_G(v) \leq \chi - 1$, $s_v \geq m + (\alpha^* + \delta - 1)\chi \geq m + (\alpha^* + \delta - 1)(m + 1) \geq m + p$. By the fact that $p > 1$, we have $\frac{1}{s_v} \leq \frac{1}{m+p}$ and $\left(1 - \frac{1}{p}\right)^{-\frac{mp}{s_v}} \leq \left(1 - \frac{1}{p}\right)^{-\frac{mp}{m+p}}$. By (33), we have

$$\mu_{v,(\widetilde{G},\widetilde{L})}(c) \leq \frac{1}{s_v} \left(1 - \frac{1}{p}\right)^{-\frac{mp}{s_v}} \leq \frac{1}{m + p} \left(1 - \frac{1}{p}\right)^{-\frac{mp}{m+p}}. \tag{35}$$

To proof (32), we define the following function

$$f(m) \triangleq \frac{m + p}{m} \left(1 - \frac{1}{p}\right)^{\frac{mp}{m+p}} = \frac{(\alpha^* + \delta)m + 0.5}{m} \left(1 - \frac{1}{(\alpha^* + \delta - 1)m + 0.5}\right)^{\frac{m((\alpha^*+\delta-1)m+0.5)}{(\alpha^*+\delta)m+0.5}}$$

By definition, $\mu_{v,(\widetilde{G},\widetilde{L})}(c) \leq \frac{1}{mf(m)}$. In Lemma A.1, we show that $f(m)$ is a decreasing function for $m \geq 3$. Thus, we have

$$f(m) \geq \lim_{x \to \infty} f(x) = (\alpha^* + \delta) \exp\left(-\frac{1}{\alpha^* + \delta}\right).$$

Thus, we have

$$\mu_{v,(\widetilde{G},\widetilde{L})}(c) \leq \frac{1}{mf(m)} \leq \frac{1}{m} \cdot \frac{1}{\alpha^* + \delta} \exp\left(\frac{1}{\alpha^* + \delta}\right) \overset{(\star)}{\leq} \frac{1}{m} \cdot \frac{\alpha^*}{\alpha^* + \delta} = \frac{1 - \delta/(\alpha^* + \delta)}{m}.$$

where $(\star)$ is due to the fact that $\exp\left(\frac{1}{\alpha^*}\right) = \alpha^*$. This proves (32) for all $m \geq 3$.

6.5. **Tightness of the marginal upper bound**. Our whole analysis relies on the upper bound of the marginal probabilities, which states that the marginal probability of v taking a specific colour is less than the reciprocal value of v's degree. Similar properties were also used in analysing strong spatial mixing [GMP05, GKM15] or zero-free regions [LSS19] for graph colourings. A natural question is whether the upper bound can be further improved.

In this section, we show that the bound $|L(v)| > \alpha^*\Delta \pm O(1)$ is tight for our technique based on the upper bound of the marginal probabilities. This means that the bound in Theorem 6.1 is the best we can achieve using current techniques. However, we also remark that the construction below only applies to list colouring instances.

We show that there exists a list colouring instance $(G, \boldsymbol{L})$ with triangle-free G such that if $|L(v)| < \alpha^*\Delta - 3$ for some vertex v, then $(G, \boldsymbol{L})$ does not have the desired marginal upper bound. Consider a star $G = (V, E)$ with $(\Delta + 1)$ vertices, where $V = \{v, v_1, v_2, \ldots, v_\Delta\}$ and $E = \{\{v, v_i\} \mid 1 \leq i \leq \Delta\}$. Define colour lists $\boldsymbol{L}$ by $L(v) = [q] = \{0, 1, \ldots, q-1\}$ and $L(v_i) = [q-1] = \{0, 1, \ldots, q-2\}$ for all $1 \leq i \leq \Delta$.

Proposition 6.15. *If $q < \alpha^*\Delta - 3$, then $\mu_{v,(G,\boldsymbol{L})}(c) > \frac{1}{\deg_G(v)} = \frac{1}{\Delta}$, where c is the colour $q - 1$.*

Proof. We can calculate the probability that v takes the colour $c = q - 1$ as follows

$$\mu_{v,(G_1,\boldsymbol{L}_1)}(c) = \frac{(q-1)^\Delta}{(q-1)(q-2)^\Delta + (q-1)^\Delta} = \frac{1}{(q-2)\left(1 - \frac{1}{q-1}\right)^{(\Delta-1)} + 1} \geq \frac{1}{(q-2)\exp\left(-\frac{\Delta-1}{q-1}\right) + 1}.$$

If $q < \alpha^*\Delta - 3$, we can verify that $(q-2)\exp\left(-\frac{\Delta-1}{q-1}\right) + 1 < \Delta$. This proves the proposition. □

Note that the graph G is a tree, which means that no matter how large we assume the girth of the graph to be, such barrier of marginal upper bounds still exists.

Indeed, the upper bound (16) in Condition 6.2 is only required for vertices v with $\deg_G(v) \leq \chi - 1$, but a simple modification of the instance above can provide a counter example to Condition 6.2. Similar barriers of the marginal upper bound also appear in [GMP05, GKM15, LSS19]. Finally, we remark that the barrier discussed in this section only applies for our current technique, which is solely based on marginal upper bounds. It may still be possible to improve the dependence between the number of colours and the degree of the graph by exploiting spectral independence (Definition 1.2) through other means.

References

[AL20] Vedat Levi Alev and Lap Chi Lau. Improved analysis of higher order random walks and applications. In *STOC*, pages 1198–1211, 2020.

[ALO20] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. *arXiv preprint arXiv:2001.00303*, 2020. FOCS 2020, to appear.

[ALOV19] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. Log-concave polynomials II: high-dimensional walks and an FPRAS for counting bases of a matroid. In *STOC*, pages 1–12, 2019.

[ALOV20] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. Log-concave polynomials IV: Exchange properties, tight mixing times, and faster sampling of spanning trees. *CoRR*, abs/2004.07220, 2020.

[BD97] Russ Bubley and Martin Dyer. Path coupling: A technique for proving rapid mixing in Markov chains. In *Proceedings of the 38th IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 223–231, 1997.

[CDM+19] Sitan Chen, Michelle Delcourt, Ankur Moitra, Guillem Perarnau, and Luke Postle. Improved bounds for randomly sampling colorings via linear programming. In *SODA*, pages 2216–2234, 2019.

[CGM19] Mary Cryan, Heng Guo, and Giorgos Mousa. Modified log-sobolev inequalities for strongly log-concave distributions. In *FOCS*, pages 1358–1370, 2019.

[Che98] Mu-Fa Chen. Trilogy of couplings and general formulas for lower bound of spectral gap. In *Probability towards 2000 (New York, 1995)*, volume 128 of *Lect. Notes Stat.*, pages 123–136. Springer, New York, 1998.

[CLV20] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Rapid mixing of glauber dynamics up to uniqueness via contraction. *arXiv preprint arXiv:2004.09083*, 2020. FOCS 2020, to appear.

[DF01] Martin E. Dyer and Alan M. Frieze. Randomly colouring graphs with lower bounds on girth and maximum degree. In *FOCS*, pages 579–587, 2001.

[DFHV13] Martin E. Dyer, Alan M. Frieze, Thomas P. Hayes, and Eric Vigoda. Randomly coloring constant degree graphs. *Random Struct. Algorithms*, 43(2):181–200, 2013.

[DGU14] Martin E. Dyer, Catherine Greenhill, and Mario Ullrich. Structure and eigenvalues of heat-bath Markov chains. *Linear Algebra Appl.*, 454:57–71, 2014.

[DK17] Irit Dinur and Tali Kaufman. High dimensional expanders imply agreement expanders. In *FOCS*, pages 974–985. IEEE Computer Society, 2017.

[FV07] Alan M. Frieze and Eric Vigoda. A survey on the use of markov chains to randomly sample colourings. *Oxford Lecture Series in Mathematics and its Applications*, 34:53, 2007.

[GK12] David Gamarnik and Dmitriy Katz. Correlation decay and deterministic FPTAS for counting colorings of a graph. *J. Discrete Algorithms*, 12:29–47, 2012.

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

[GL18] Heng Guo and Pinyan Lu. Uniqueness, spatial mixing, and approximation for ferromagnetic 2-spin systems. *ACM Trans. Comput. Theory*, 10(4):Art. 17, 25, 2018.

[GMP05] Leslie Ann Goldberg, Russell A. Martin, and Mike Paterson. Strong spatial mixing with fewer colors for lattice graphs. *SIAM J. Comput.*, 35(2):486–517, 2005.

[Hay03] Thomas P. Hayes. Randomly coloring graphs of girth at least five. In *STOC*, pages 269–278, 2003.

[Hay06] Thomas P. Hayes. A simple condition implying rapid mixing of single-site dynamics on spin systems. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 39–46, 2006.

[Hay13] Thomas P. Hayes. Local uniformity properties for glauber dynamics on graph colorings. *Random Struct. Algorithms*, 43(2):139–180, 2013.

[HJ12] Roger A. Horn and Charles R. Johnson. *Matrix analysis.* Cambridge university press, 2012.

[HV03] Thomas P. Hayes and Eric Vigoda. A non-markovian coupling for randomly sampling colorings. In *FOCS*, pages 618–627, 2003.

[HV06] Thomas P. Hayes and Eric Vigoda. Coupling with the stationary distribution and improved sampling for colorings and independent sets. *Ann. Appl. Probab.*, 16, 2006.

[Jer95] Mark Jerrum. A very simple algorithm for estimating the number of k-colorings of a low-degree graph. *Random Struct. Algorithms*, 7(2):157–165, 1995.

[JS89] Mark Jerrum and Alistair Sinclair. Approximating the permanent. *SIAM J. Comput.*, 18(6):1149–1178, 1989.

[JVV86] Mark Jerrum, Leslie G. Valiant, and Vijay V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoret. Comput. Sci.*, 43:169–188, 1986.

[KO20] Tali Kaufman and Izhar Oppenheim. High order random walks: Beyond spectral gap. *Combinatorica*, 40(1):245–281, 2020.
[LLY13] Liang Li, Pinyan Lu, and Yitong Yin. Correlation decay up to uniqueness in spin systems. In *SODA*, pages 67–84. SIAM, 2013.
[LP17] David A Levin and Yuval Peres. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.
[LSS19] Jingcheng Liu, Alistair Sinclair, and Piyush Srivastava. A deterministic algorithm for counting colorings with 2Δ colors. In *FOCS*, pages 1380–1404, 2019.
[LY13] Pinyan Lu and Yitong Yin. Improved FPTAS for multi-spin systems. In *RANDOM*, pages 639–654, 2013.
[Mol04] Michael Molloy. The glauber dynamics on colorings of a graph with high girth and maximum degree. *SIAM J. Comput.*, 33(3):721–737, 2004.
[Opp18] Izhar Oppenheim. Local spectral expansion approach to high dimensional expanders part I: Descent of spectral gaps. *Discret. Comput. Geom.*, 59(2):293–330, 2018.
[SS97] Jesús Salas and Alan D Sokal. Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem. *J. Stat. Phys.*, 86(3):551–579, 1997.
[SS20] Shuai Shao and Yuxin Sun. Contraction: A unified perspective of correlation decay and zero-freeness of 2-spin systems. In *ICALP*, volume 168 of *LIPIcs*, pages 96:1–96:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.
[Str06] Adam W. Strzebonski. Cylindrical algebraic decomposition using validated numerics. *J. Symb. Comput.*, 41(9):1021–1038, 2006.
[Vig00] Eric Vigoda. Improved bounds for sampling colorings. *J. Math. Phys.*, 41(3):1555–1569, 2000.
[Wei06] Dror Weitz. Counting independent sets up to the tree threshold. In *STOC*, pages 140–149, 2006.
[YZ13] Yitong Yin and Chihao Zhang. Approximate counting via correlation decay on planar graphs. In *SODA*, pages 47–66. SIAM, 2013.

Appendix A. Computer assisted proof

We give a computer-assisted proof for the following technical lemma used in the analysis for colouring.

Lemma A.1. *Let $\alpha^* \approx 1.763\ldots$ be the solution of $\alpha^* = \exp\left(\frac{1}{\alpha^*}\right)$ and $\delta > 0$ a real number. Define*

$$f(x) = \frac{(\alpha^* + \delta)x + 0.5}{x}\left(1 - \frac{1}{(\alpha^* + \delta - 1)x + 0.5}\right)^{\frac{x((\alpha^*+\delta-1)x+0.5)}{(\alpha^*+\delta)x+0.5}}.$$

The function f is decreasing for $x \in [3, \infty)$.

Proof. Let $a = \alpha^* + \delta - 1 > 0.763$. Direct calculation yields $f'(x) = A \cdot B$ with

$$A = \left(2x^2(2ax - 1)(2(a+1)x + 1)\right)^{-1}\left(1 - \frac{2}{1 + 2ax}\right)^{\frac{x(1+2ax)}{1+2(a+1)x}};$$

$$B = 1 + 2x - 4a^2x^2 + 8a(1+a)x^3 + x\left(-1 + 8a^3x^3 - 2ax(1+2x) + 4a^2x^2(1+2x)\right)\ln\left(1 - \frac{2}{1+2ax}\right).$$

It is easy to see that $A > 0$, so we only need to verify that $B < 0$. To see this, note that the term

$$-1 + 8a^3x^3 - 2ax(1+2x) + 4a^2x^2(1+2x) = 2ax(1+2x)(2ax-1) + \left(8a^3x^3 - 1\right) > 0$$

for any $x \geq 3$ and $a \geq \alpha^* - 1$. It follows from the Taylor series of $\ln(1-z)$ that

$$\ln\left(1 - \frac{2}{1+2ax}\right) \leq -\frac{2}{1+2ax} - \frac{2}{(1+2ax)^2} - \frac{8/3}{(1+2ax)^3} - \frac{4}{(1+2ax)^4}.$$

Therefore we have

$$\begin{aligned} B &\leq 1 + 2x - 4a^2x^2 + 8a(1+a)x^3 \\ &\quad + x\left(-1 + 8a^3x^3 - 2ax(1+2x) + 4a^2x^2(1+2x)\right) \\ &\quad \cdot \left(-\frac{2}{1+2ax} - \frac{2}{(1+2ax)^2} - \frac{8/3}{(1+2ax)^3} - \frac{4}{(1+2ax)^4}\right) \overset{(\star)}{<} 0, \end{aligned}$$

where $(\star)$ is verified by the following Mathematica code:

```
Resolve[Exists[x,1+2x-4a^2x^2+8a(1+a)x^3+x(-1+8a^3x^3-2a x(1+2x)+4
    a^2x^2(1+2x))*(-(2/(1+2a*x))-2/(1+2a*x)^2-(8/3)/(1+2a*x)^3-4/(1+2a*x)^4)>=0 && a>763/1000 && x>=3]]
```

□

Here we used the Resolve command in Mathematica. This is a rigorous implementation of a quantifier elimination algorithm, which determines the feasibility of a collection of polynomial inequalities. For more details, see [Str06].