Rapid Mixing via Coupling Independence for Spin Systems with Unbounded Degree

Xiaoyu Chen
State Key Laboratory for Novel Software Technology, New Cornerstone Science Laboratory, Nanjing University, China

Weiming Feng
School of Computing and Data Science, The University of Hong Kong, China

Abstract

We develop a new framework to prove the mixing or relaxation time for the Glauber dynamics on spin systems with unbounded degree. It works for general spin systems including both 2-spin and multi-spin systems. As applications for this approach:

- We prove the optimal $O(n)$ relaxation time for the Glauber dynamics of random q-list-coloring on an n-vertices triangle-tree graph with maximum degree Δ such that $q/\Delta > \alpha^\star$, where $\alpha^\star \approx 1.763$ is the unique positive solution of the equation $\alpha = \exp(1/\alpha)$. This improves the $n^{1+o(1)}$ relaxation time for Glauber dynamics obtained by the previous work of Jain, Pham, and Vuong (2022). Besides, our framework can also give a near-linear time sampling algorithm under the same condition.
- We prove the optimal $O(n)$ relaxation time and near-optimal $\widetilde{O}(n)$ mixing time for the Glauber dynamics on hardcore models with parameter λ in *balanced* bipartite graphs such that $\lambda < \lambda_c(\Delta_L)$ for the max degree Δ_L in left part and the max degree Δ_R of right part satisfies $\Delta_R = O(\Delta_L)$. This improves the previous result by Chen, Liu, and Yin (2023).

At the heart of our proof is the notion of *coupling independence* which allows us to consider multiple vertices as a huge single vertex with exponentially large domain and do a "coarse-grained" local-to-global argument on spin systems. The technique works for general (multi) spin systems and helps us obtain some new comparison results for Glauber dynamics.

2012 ACM Subject Classification Theory of computation → Random walks and Markov chains

Keywords and phrases coupling independence, Glauber dynamics, mixing times, relaxation times, spin systems

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2025.68

Category RANDOM

Related Version *Full Version*: `https://arxiv.org/abs/2407.04672` [10]

Funding Weiming Feng acknowledges the support of Dr. Max Rössler, the Walter Haefner Foundation, and the ETH Zürich Foundation during his affiliation with ETH Zürich, where this work was done.

Acknowledgements We would like to thank Heng Guo for bringing the paper [30] to our attention and for the helpful discussions at the beginning of this project. We thank Eric Vigoda for suggesting a better statement of technical results. We also thank Charlie Carlson, Yitong Yin, and Xinyuan Zhang for their helpful discussions.

1 Introduction

The *spin system* is a fundamental probabilistic graphical model. It is defined on a graph $G = (V, E)$, where every vertex is a random variable and every edge models the local interactions. Each variable takes a value from a discrete domain $[q] = \{1, 2, \ldots, q\}$. Each

vertex has a vector $b \in \mathbb{R}^q_{\geq 0}$ called the *external field* and each edge has a symmetric matrix $A \in \mathbb{R}^{q\times q}_{\geq 0}$ called the *interaction matrix*. The spin system defines a Gibbs distribution over $[q]^V$ such that for any configuration $\sigma \in [q]^V$,

$$\mu(\sigma) \propto \prod_{v\in V} b(\sigma_v) \prod_{e=\{u,v\}\in E} A(\sigma_u, \sigma_v).$$

The spin system covers many important distributions including the uniform distribution of graph colorings, the Ising model, the hardcore gas model in Physics, and a broad class of undirected graphical models in machine learning [42].

Sampling from the Gibbs distribution is a central algorithmic task for spin systems. The *Glauber dynamics* is a fundamental Markov chain Monte Carlo (MCMC) method for sampling from high-dimensional distributions. Given a distribution μ over $[q]^V$, it starts from an arbitrary $X \in \Omega(\mu)$, where $\Omega(\mu) \subseteq [q]^V$ is the support of μ. In each step, it updates the current state X as follows:

- pick a variable $v \in V$ uniformly at random;
- resample the value of X_v from the conditional distribution $\mu_v(\cdot \mid X_{V\setminus v})$.

It is well-known that if the state space $\Omega(\mu)$ is connected through the moves of Glauber dynamics, then the distribution μ is the unique stationary distribution for the Glauber dynamics.

In this paper, we study the convergence rate of the Glauber dynamics. Let $(X_t)_{t\geq 0}$ denote the random sequence generated by the Glauber dynamics. Let $P : \Omega(\mu) \times \Omega(\mu) \to [0, 1]$ denote the transition matrix of the Glauber dynamics. Many notions capture the convergence rate. The most standard one is the *mixing time*, which is defined by

$$T^{\mathrm{GD}}_{\mathrm{mix}}(\mu, \varepsilon) := \max_{X_0\in\Omega(\mu)} \min\left\{t > 0 \mid \mathcal{D}_{\mathrm{TV}}\left(P^t(X_0, \cdot) \parallel \mu\right) \leq \varepsilon\right\}, \tag{1}$$

where $\mathcal{D}_{\mathrm{TV}}\left(P^t(X_0, \cdot) \parallel \mu\right)$ denotes the standard *total variation distance* between μ and the distribution of X_t. In words, if the Glauber dynamics starts from the worst initial state X_0, the mixing time is the minimum number t such that the total variation distance between $P^t(X_0, \cdot)$ and μ is below a sufficiently small constant. Another widely used notion is the *relaxation time*. A standard fact says that P only has non-negative real eigenvalues $1 = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{|\Omega|} \geq 0$ [22]. The gap $\lambda_1 - \lambda_2 = 1 - \lambda_2$ is called the *spectral gap* of Glauber dynamics. The relaxation time is defined by

$$T^{\mathrm{GD}}_{\mathrm{rel}}(\mu) := \frac{1}{1-\lambda_2}.$$

It is well known that $T^{\mathrm{GD}}_{\mathrm{mix}}(\mu, \varepsilon) = O(T^{\mathrm{GD}}_{\mathrm{rel}}(\mu) \log \frac{1}{\varepsilon\mu_{\min}})$, where $\mu_{\min} = \min_{\sigma\in\Omega(\mu)} \mu(\sigma)$.

Recently, a series of works [4, 1, 3] studied Glauber dynamics using high dimensional expanders. An important notion called *spectral independence* was developed during this process. Anari, Liu, and Oveis Gharan [3] first introduced spectral independence for Boolean distributions. The follow-up works [16, 24] then generalized it to non-Boolean distributions. For example, for a Boolean distribution μ over $\{-1, +1\}^{[n]}$, the *influence matrix* $\Psi \in \mathbb{R}^{n\times n}_{\geq 0}$ is defined by $\Psi(u, v) := \mathbf{Pr}_{X\sim\mu}[X_v = + \mid X_u = +] - \mathbf{Pr}_{X\sim\mu}[X_v = + \mid X_u = -]$. A distribution μ is C-spectrally independent if the maximum eigenvalue of Ψ is at most C. If every conditional distribution of μ is C-spectrally independent, then by the local-to-global argument [1], both relaxation and mixing time of Glauber dynamics are bounded by $n^{O(C)}$, where n is the number of variables. Given this polynomial bound $n^{O(C)}$, many works tried to obtain an improved or even the optimal mixing/relaxation time for Glauber dynamics, especially when μ is a Gibbs distribution defined by spin systems. Chen, Liu and Vigoda [21] proved that for spin systems on bounded degree graphs, the spectral independence implies both $O(n \log n)$ optimal mixing time and $O(n)$ optimal relaxation time.

The next question is how to deal with spin systems on *unbounded degree* graphs. Many works [29, 11, 2, 15, 12] focused on this question. Significant progress was made, especially for 2-spin systems ($q = 2$). [29] first studied coloring and weighted independent sets (hardcore model) in high-girth graphs and proved the near-optimal $n^{1+o(1)}$ relaxation time. [11] introduced a stronger variant of spectral independence called *complete spectral independence*, and proved the optimal $O(n)$ relaxation time for anti-ferromagnetic 2-spin systems in the uniqueness regime. To obtain the optimal mixing time, [2] made the first step and defined a new notion called *entropic independence*. After a line of work [2, 15, 12], the optimal $O(n \log n)$ mixing time was established for a broad class of 2-spin systems.

Most of the previous techniques [11, 2, 15, 12] for unbounded degree graphs are restricted to the 2-spin systems. We consider the following question in this paper.

How to prove the optimal mixing/relaxation time
for Glauber dynamics on (multi) spin systems with unbounded degree?

To the best of our knowledge, the only previous result beyond 2-spin systems is the $n^{1+o(1)}$ relaxation time for graph coloring [29]. However, [29] relies on the coupling analysis for colorings in [28], which makes it difficult to be generalized to other spin systems.

In this work, we develop a new framework for proving mixing/relaxation time for the Glauber dynamics on general spin systems including both 2-spin and multi-spin systems. Our new framework is based on a stronger variant of the spectral independence known as the *coupling independence*, which is already used implicitly or explicitly in many previous works [40, 5, 14, 18, 19, 34]. A spin system μ on $[q]^V$ is C-coupling independent if for any $v \in V$ and $a, b \in [q]$, there is a coupling (X, Y) where $X \sim \mu^{v \leftarrow a}$ and $Y \sim \mu^{v \leftarrow b}$ such that

$$\mathbb{E}\left[d_H(X, Y)\right] \leq C.$$

Here, $d_H(X, Y) = |\{v \in V \mid X_v \neq Y_v\}|$ denotes the hamming distance between X and Y and $\mu^{v \leftarrow a}$ the distribution induced by μ conditional on v taking the value a.

Given a spin system on a graph G with Gibbs distribution μ, we show that if μ and all the conditional distributions induced by μ satisfy the coupling independence and the maximum degree of G is greater than a large constant, then the following comparison results hold for Glauber dynamics.

- **Relaxation time comparison.** The relaxation time satisfies $T_{\text{rel}}^{\text{GD}}(\mu) = O(T_{\text{rel}}^{\text{GD}}(\mu^\star))$, where $\mu^\star$ is a conditional distribution obtained from μ by fixing the values on a subset $\Lambda \subseteq V$ of variables. The set Λ is chosen intentionally such that the induced subgraph $G[V \setminus \Lambda]$ on other vertices has smaller maximum degree. For many spin systems, the distribution $\mu^\star$ is in an "easy regime" so that the mixing/relaxation time for $\mu^\star$ is easy to analyze. We can bound the relaxation time for μ via this comparison result (see Theorem 9).
- **Mixing time comparison.** If μ is a monotone spin system (Definition 14) and the Glauber dynamics starts from a specific initial configuration, then the mixing time satisfies $T_{\text{mix}}^{\text{GD}}(\mu, \varepsilon) = \widetilde{O}(T_{\text{mix}}^{\text{GD}}(\mu^\star, \frac{1}{4e}))$, where $\widetilde{O}$ hides a polylog(n/ε) factor (See Theorem 15).

We obtain the relaxation/mixing time bounds via the above comparison results. In the relaxation time comparison result, the constant factor in $O(\cdot)$ is independent of the degree of the graph. In applications, the distribution $\mu^\star$ is in an "easy regime", hence, we can use some standard technique to show $T_{\text{rel}}^{\text{GD}}(\mu^\star) = O(n)$. The comparison result gives the optimal $T_{\text{rel}}^{\text{GD}}(\mu) = O(n)$ relaxation time. Similarly, in the applications of monotone systems, we can obtain the near-optimal $\widetilde{O}(n)$ mixing time for general graphs. Our comparison results only

hold for graphs with large maximum degrees. It does not cause any issue in applications, because coupling independence implies spectral independence, and for graphs with bounded maximum degree, [21] already established the optimal relaxation/mixing time.

Our proof techniques can also give a near-linear time (in input size) sampling algorithm (see Theorem 13). Furthermore, we introduce a general technique to establish coupling independence for 2-spin systems (Theorem 16). Specifically, many spectral independence results for 2-spin systems are proved by analyzing the decay of correlation in the self-avoiding walk tree [20, 13]. We show that all of such proofs can be translated to a proof of coupling independence.

Organization of the paper

In Section 1.1, we first exhibit some concrete applications. In Section 2, we give our technical results. In Section 3, we give an overview of proof techniques.

All the detailed analysis are given in the full version [10].

1.1 Applications

Let $G = (V, E)$ be a graph with maximum degree Δ and $[q] = \{1, 2, \ldots, q\}$ a set of colors. Given a set of color lists $L_v \subseteq [q], v \in V$, a proper list-coloring $X \in [q]^V$ assigns a color $X_v \in L_v$ to each vertex $v \in V$ such that adjacent vertices receive different colors. In a special case when $L_v = [q]$ for all $V \in V$, the list coloring becomes the standard graph q-coloring. We use μ to denote the uniform distribution over all proper list-colorings in G. For the list coloring, in each step, the Glauber dynamics picks a random vertex v and update its color to a random available color. There is a long line of works studying the mixing and relaxation time of Glauber dynamics e.g. [33, 45, 9].

In the era of spectral independence, the proper list-coloring has been re-studied by a series of works [16, 24, 19]. Though the technique varies, all these works established some coupling independence results for the proper list-coloring. For list colorings on triangle-free graphs, let $\alpha^\star \approx 1.763$ denote the unique positive solution to the equation $\alpha = \exp(1/\alpha)$. When $|L_v| > (\alpha^\star + \delta)\Delta$, the $O_\delta(1)$-coupling independence can be established by techniques in [16, 24]. Our framework gives the optimal relaxation time of Glauber dynamics even if the maximum degree of G is unbounded.

▶ **Theorem 1** (Coloring: Relaxation Time). *Let $\delta > 0$ be a constant. For any triangle-free graph $G = (V, E)$ and color lists $(L_v)_{v \in V}$, if $|L_v| \geq (\alpha^\star + \delta)\Delta$ for all $v \in V$, where $\Delta \geq 3$ is the maximum degree of G, then relaxation time of Glauber dynamics is $O_\delta(n)$, where n is the number of vertices in G.*

Under the condition of Theorem 1, the relaxation time of the Glauber dynamics has been studied by many previous works. Combining the spectral independence technique [1, 3] with the correlation decay analysis [27, 26], two independent works [16, 24] proved the polynomial relaxation time $n^{O(1/\delta)}$ of Glauber dynamics. For graphs with bounded maximum degree $\Delta = O(1)$, Chen, Liu and Vigoda [21] established the $O_{\Delta,\delta}(n)$ relaxation time, where $O_{\Delta,\delta}(\cdot)$ hides a constant factor like $\Delta^{O(\Delta^2/\delta)}$. For general graphs with possibly unbounded maximum degree, Jain, Pham and Vuong [29] proved the first almost linear relaxation time $O_\delta(ne^{(\log\log n)^2}) = O_\delta(n^{1+o(1)})$. Their proof combined the techniques in [21] with the coupling analysis in [28]. Compared to previous results, Theorem 1 gives the optimal linear relaxation time for general graphs.

We prove Theorem 1 by first verifying the coupling independence condition (Definition 7) and then applying our comparison result (Theorem 9). Theorem 1 requires $|L_v| \geq (\alpha^\star + \delta)\Delta$ because the current best coupling independence result requires this number of colors but our comparison result does not require such a strong condition. It is conjectured that $O_\delta(1)$-coupling independence should hold for proper list-coloring in general graphs when $|L_v| \geq (1+\delta)\Delta + O(1)$.

▶ **Conjecture 2** (Folklore). *Let $\delta > 0$ be a constant. For any graph $G = (V, E)$ with maximum degree Δ and color lists $(L_v)_{v \in V}$ such that $|L_v| \geq (1+\delta)\Delta + O(1)$ for all $v \in V$, the uniform distribution μ over all the proper list-colorings of G is $O_\delta(1)$-coupling independent.*

Our comparison framework can prove optimal relaxation time for Glauber dynamics on proper list-colorings of graphs (with potentially unbounded degree) as long as Conjecture 2 holds.

▶ **Proposition 3.** *If Conjecture 2 holds with $\delta > 0$, then for any list coloring instance in Conjecture 2, the relaxation time of Glauber dynamics is $O_\delta(n)$.*

The standard relation between relaxation time and mixing time implies that the Glauber dynamics mixes in time $O_\delta(n^2 \log q)$, which yields a sampling algorithm for the uniform distribution μ of graph colorings in time $O_\delta(\Delta n^2 \log q)$ because each step of Glauber dynamics can be simulated in time $O(\Delta)$. However, in terms of sampling algorithm, our technique would directly give an algorithm (not the Glauber dynamics) in time $\widetilde{O}_\delta(\Delta n)$. Since the input graph G contains Δn edges, the running time is near-linear in the input size.

▶ **Theorem 4** (Coloring: Algorithm). *Let $\delta > 0$ be a constant. There exists an algorithm such that given any $\varepsilon > 0$, any triangle-free graph $G = (V, E)$ and color lists $(L_v)_{v \in V}$, if $|L_v| \geq (\alpha^\star + \delta)\Delta$ for all $v \in V$, where $\Delta \geq 3$ is the maximum degree of G, it returns a random sample X satisfying $\mathcal{D}_{\mathrm{TV}}(X \parallel \mu) \leq \varepsilon$ in time $\Delta n (\log \frac{n}{\varepsilon})^{C(\delta)}$, where $C(\delta)$ is a constant depending only on δ.*

The next example is the hardcore model. Let $G = (V, E)$ be a graph. Let $\lambda > 0$ be the fugacity. The hardcore model defines a distribution μ over all independent sets $S \subseteq V$ in G such that $\mu(S) \propto \lambda^{|S|}$. Let $\Delta \geq 3$ denote the maximum degree of graph G. There is a critical threshold (a.k.a. uniqueness threshold) for the tree uniqueness phase transition [36]

$$\lambda_c(\Delta) := \frac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^\Delta}.$$

such that if $\lambda \leq \lambda_c(\Delta)$ the correlation between two vertices decays in their distance; if $\lambda > \lambda_c(\Delta)$, the long-range correlation exists. A computational phase transition occurs at the same threshold. If $\lambda < \lambda_c(\Delta)$, polynomial time sampling algorithm exists [46]; if $\lambda > \lambda_c(\Delta)$, the sampling problem is hard unless $\mathbf{NP} = \mathbf{RP}$ [44, 25]. The mixing and relaxation time of the Glauber dynamics for hardcore model were also extensively studied [41, 28, 23]. Recent works analyzed Glauber dynamics via spectral independence [3]. The optimal $O_\delta(n \log n)$ mixing time and the optimal $O_\delta(n)$ relaxation time were established when $\lambda \leq (1-\delta)\lambda_c(\Delta)$ for general graphs [21, 15, 12].

However, for the hardcore model on bipartite graphs, the picture is not very clear. Consider the hardcore model in a bipartite graph $G = (V = V_L \uplus V_R, E)$. Let Δ_L denote the maximum degree in the left part. Assume $3 \leq \Delta_L$. It is recently known that the uniqueness threshold for the hardcore model on the bipartite graph can be refined to $\lambda_c(\Delta_L) \geq \lambda_c(\Delta)$ where $\Delta \geq \Delta_L$ is the maximum degree of the bipartite graph [39, 13]. The Glauber dynamics is also proved to have polynomial mixing time when $\lambda < \lambda_c(\Delta_L)$ [13].

On the other side, when $\lambda > \lambda_c(\Delta_L)$, the lower bound in [44] does not hold for bipartite graphs and the problem is #BIS-hard [7], where #BIS is the problem of counting the independent sets in bipartite graphs. A line of works (e.g. [31, 8, 38, 17, 32]) studied various sampling algorithms in the low-temperature (large λ) regime.

Within the critical threshold $\lambda < \lambda_c(\Delta_L)$, we consider "balanced" bipartite graphs. Let Δ_R be the maximum degree in V_R. We say a bipartite graph is θ-balanced if $\Delta_R \le \theta\Delta_L$.

▶ **Theorem 5** (Bipartite Hardcore: Relaxation Time). *Let $\delta \in (0,1)$ and $\theta > 1$ be two constants. For any hardcore model on a θ-balanced bipartite graph G with fugacity λ, if $\lambda \le (1-\delta)\lambda_c(\Delta_L)$, then the relaxation time of Glauber dynamics is $O_{\delta,\theta}(n)$, where n is the number of vertices in G.*

For the mixing time, again, the standard relation gives $O_{\delta,\theta}(n^2)$ mixing time of the Glauber dynamics. However, since the bipartite graph hardcore is a monotone system, our technique also implies the $\widetilde{O}_{\delta,\theta}(n)$ *mixing time* of Glauber dynamics starting from the independent set containing all vertices in the left part: $X_0 = V_L$. Formally, for any $S \in \Omega(\mu)$,

$$T^{\mathrm{GD}}_{\mathrm{mix}}(\mu, \varepsilon \mid S) = \min\left\{t > 0 \mid \mathcal{D}_{\mathrm{TV}}\left(X_t \parallel \mu\right) \le \varepsilon \wedge X_0 = S\right\}.$$

▶ **Theorem 6** (Bipartite Hardcore: Mixing Time). *Let $\delta \in (0,1)$ and $\theta > 1$ be two constants. For any hardcore model on a θ-balanced bipartite graph G with fugacity λ, if $\lambda \le (1-\delta)\lambda_c(\Delta_L)$, then the mixing time of Glauber dynamic starting from V_L satisfies $T^{\mathrm{GD}}_{\mathrm{mix}}(\mu, \varepsilon \mid V_L) = n(\log \frac{n}{\varepsilon})^{C(\delta,\theta)}$, where $C(\delta,\theta)$ is a constant depending only on δ and θ.*

The previous work [13] established the $(\frac{\Delta_L \log n}{\lambda})^{O(1/\delta)} n^2$ relaxation time for the Glauber dynamics, which, by the standard relation, implies the $(\frac{\Delta_L \log n}{\lambda})^{O(1/\delta)} n^3 \log \frac{1+\lambda}{\lambda}$ mixing time. The previous result holds for general bipartite graphs as long as $\lambda \le (1-\delta)\lambda_c(\Delta_L)$. For balanced bipartite graphs, we obtained the optimal relaxation time $O_{\delta,\theta}(n)$ and the near-optimal mixing time $\widetilde{O}_{\delta,\theta}(n)$, which significantly improved the dependency to n and Δ_L compared to the previous result. For example, in the near critical case when $\lambda = (1-\delta)\lambda_c(\Delta_L) = \Theta(1/\Delta_L)$, previous result gives $\Delta_L^{O(1/\delta)} n^2 \cdot \mathrm{polylog}(n)$ relaxation time and $\Delta_L^{O(1/\delta)} n^3 \cdot \mathrm{polylog}(n)$ mixing time but our result gives $O(n)$ relaxation time and $n \cdot \mathrm{polylog}(n)$ mixing time. However, our result works only on balanced bipartite graphs. The result in [13] is still state-of-the-art for general bipartite graphs.

Finally, we point out that our technique could also recover many previous $O(n)$ relaxation time results for anti-ferromagnetic 2-spin systems in [11]. See Remark 10 for one example.

2 Technical Results

2.1 Coupling Independence

In this section, we give our general results for spin systems. Let $G = (V, E)$ be a graph. Let $[q] = \{1, 2, \ldots, q\}$ be a set of $q \ge 2$ spins. For each vertex $v \in V$, let vector $b_v \in \mathbb{R}^q_{\ge 0}$ be the *external field* at vertex v. For each edge $e \in E$, let symmetric matrix $A_e \in \mathbb{R}^{q\times q}_{\ge 0}$ be the *interaction matrix* at edge e. A spin system defines a *Gibbs distribution* μ over $[q]^V$ such that,

$$\forall \sigma \in [q]^V \quad \mu(\sigma) \propto w(\sigma) := \prod_{u\in V} b_u(\sigma_u) \prod_{e=\{v,w\}\in E} A_e(\sigma_v, \sigma_w).$$

We use $\Omega(\mu) \subseteq [q]^V$ to denote the support of the Gibbs distribution μ.

Let $\Lambda \subseteq V$ be a subset of vertices. Given any *pinning* $\tau \in [q]^{V\setminus\Lambda}$, we define a conditional distribution μ^τ by for any configuration $\sigma \in [q]^V$,

$$\mu^\tau(\sigma) \propto w^\tau(\sigma) := \mathbf{1}[\sigma_{V\setminus\Lambda} = \tau] \cdot \prod_{u\in\Lambda} b_u(\sigma_u) \prod_{\substack{e=\{v,w\}\in E\\ v,w\in\Lambda}} A_e(\sigma_v,\sigma_w) \prod_{\substack{e=\{v,w\}\in E\\ v\in\Lambda\wedge w\notin\Lambda}} A_e(\sigma_v,\tau_w). \tag{2}$$

In words, μ^τ is a Gibbs distribution obtained for μ by removing all edges $e \subseteq V\setminus\Lambda$ and putting a constraint that every vertex in $v \in V\setminus\Lambda$ must take the value τ_v. In particular, if τ is feasible (e.g. τ belongs to the support of the marginal distribution $\mu_{V\setminus\Lambda}$), then μ^τ is exactly the conditional distribution induced by μ given the condition τ. For all spin systems considered in this paper, it holds that $\sum_\sigma w^\tau(\sigma) > 0$ for all τ. The distribution in (2) is well-defined. Furthermore, for any subset S, we use μ^τ_S to denote the marginal distribution on S projected from μ^τ.

The following condition plays a key role in the proof of our main results. Let $\rho : V \to \mathbb{N}_{>0}$ be a function that maps every vertex $v \in V$ to a positive integer. We call the function ρ the Hamming weight function. For any two (possibly partial) configurations $\sigma, \tau \in [q]^\Lambda$, where $\Lambda \subseteq V$, define their weighted Hamming distance with respect to ρ by

$$H_\rho(\sigma,\tau) := \sum_{v\in\Lambda:\sigma(v)\neq\tau(v)} \rho(v). \tag{3}$$

▶ **Definition 7** (Coupling Independence). *Let $C \geq 1$ be a constant. A distribution μ over $[q]^V$ is said to be C-coupling independent (C-CI) if there exists Hamming weight function $\rho : V \to \mathbb{N}_{>0}$ such that the following holds. For any pinning $\sigma_1, \sigma_2 \in [q]^S$, where $S \subseteq V$ and σ_1, σ_2 disagree only at one vertex $v_0 \in S$, there exists a coupling (X, Y), where $X \sim \mu^{\sigma_1}$ and $Y \sim \mu^{\sigma_2}$, such that*

$$\frac{\mathbb{E}\left[H_\rho(X,Y)\right]}{\rho(v_0)} \leq C.$$

The notion of coupling independence was introduced explicitly in [14] to study the spectral independence property. For example, for Boolean distributions[1] ($q = 2$), given any pinning $\tau \in \{-,+\}^{V\setminus\Lambda}$, the $|\Lambda| \times |\Lambda|$ influence matrix [3] is defined by

$$\Psi^\tau_\mu(v,u) := \mu_u^{\tau\wedge v^+}(+) - \mu_u^{\tau\wedge v^-}(+), \tag{4}$$

where $u, v \in \Lambda$ and $\tau \wedge v^\pm$ denotes the pinning τ together with v taking the value $\pm$. A distribution μ is C-spectrally independent if the maximum eigenvalue of Ψ^τ_μ is at most C for any pinning τ. It is not hard to show that C-coupling independence implies C-spectral independence. Hence, recent works [14, 19, 18] utilized coupling independence to establish the spectral independence for various spin systems.

2.2 Compare Markov Chains via Coupling Independence

In this work, we find more applications for coupling independence beyond establishing spectral independence. We build some comparison results of Markov chains via coupling independence. As a by-product result, we also show that the coupling independence gives fast sampling algorithms.

Let μ be a Gibbs distribution over $[q]^V$ on graph $G = (V, E)$. For any $\Lambda \subseteq V$, we use $G[\Lambda]$ to denote the induced subgraph of G on vertex set Λ.

[1] The influence matrix and spectral independence are also defined for general distributions with $q \geq 2$. See [16] for the detailed definition.

▶ **Definition 8** (Relaxation Time with Pinning). *Let μ be a Gibbs distribution on graph $G = (V, E)$ with maximum degree Δ. Let $\eta \in [0, 1]$. Let $D(\eta)$ denote all subsets $\Lambda \subseteq V$ such that the maximum degree of $G[\Lambda]$ is at most $\eta\Delta$. Define*

$$T_{\mathrm{rel}}^{(\eta)}(\mu) := \max\left\{T_{\mathrm{rel}}^{\mathrm{GD}}(\mu^\tau) \mid \Lambda \in D(\eta) \wedge \tau \in [q]^{V\setminus\Lambda}\right\}.$$

In the above definition, μ^τ is a distribution on $[q]^V$. In every step, the Glauber dynamics picks $v \in V$ uniformly at random then resamples the value on v. If $v \notin \Lambda$, the value of v after resampling is always τ_v. Indeed, μ^τ is essentially the same as μ_Λ^τ. But, considering μ^τ would help us simplify some results and proofs. The following is our main comparison result.

▶ **Theorem 9** (Relaxation Time Comparison). *Let $M \geq 1$ and $0 < \eta \leq \frac{1}{2\lceil M\rceil}$ be two constants. There exists $\Delta_0 = \Omega(\frac{M^2}{\eta^2}\log\frac{M}{\eta})$ such that for any Gibbs distribution μ on graph G with the maximum degree $\Delta \geq \Delta_0$, if μ satisfies M-coupling independence, then the relaxation time of Glauber dynamics on μ satisfies*

$$T_{\mathrm{rel}}^{\mathrm{GD}}(\mu) \leq 2^{O(M/\eta)} \cdot T_{\mathrm{rel}}^{(\eta)}(\mu).$$

The theorem is proved in the full version ([10, Section 4]). See Section 3 for a proof overview.

The above theorem is a comparison result between two kinds of relaxation times. Consider the case when parameters of μ are close to the critical threshold so that the relaxation time $T_{\mathrm{rel}}^{\mathrm{GD}}(\mu)$ is hard to analyze. By choosing a sufficiently small η, suppose for any $\Lambda \in D(\eta)$ and any $\tau \in [q]^{V\setminus\Lambda}$, the conditional distribution μ^τ falls into an easy regime. The relaxation time $T_{\mathrm{rel}}^{(\eta)}(\mu)$ is easy to analyze. Theorem 9 boosts the relaxation time from an easy regime to the hard regime if μ satisfies the coupling independence and the maximum degree Δ is greater than a constant Δ_0.

When applying Theorem 9 to a specific spin system with Gibbs distribution μ, we first need to show that the μ satisfies the coupling independence property. Next, we choose a small constant η to guarantee that $T_{\mathrm{rel}}^{(\eta)}(\mu)$ is easy to analyze. Now, the constant parameter Δ_0 in Theorem 9 is fixed. If the maximum degree $\Delta \leq \Delta_0 = O(1)$ is bounded, then since the coupling independence implies the spectral independence, the previous work [21] already established the optimal relaxation time for μ. If the maximum degree $\Delta \geq \Delta_0$, we can apply our boosting result to bound the relaxation time. We show how to prove Theorem 1 and Proposition 3 via Theorem 9.

Proof Sketch of Theorem 1

Given a triangle-free graph $G = (V, E)$ and color lists $L_v \subseteq [q]$ with $|L_v| \geq (\alpha^\star + \delta)\Delta$ for all $v \in V$, let μ denote the uniform distribution over all proper list-colorings. By going through the analysis in [24], we can prove that μ satisfies $O(1/\delta)$-coupling independence. Let η be a parameter to be fixed later. For any $\Lambda \subseteq D(\eta)$, any pinning $\tau \in [q]^{V\setminus\Lambda}$, the distribution μ^τ is essentially the same as the distribution μ_Λ^τ because the coloring outside Λ is fixed by τ. By self-reducibility, μ_Λ^τ is a list coloring on $G' = G[\Lambda]$ with color list $L'_v = L_v \setminus \{\tau_u \mid u \notin \Lambda \wedge \{u, v\} \in E\}$. Let $\deg'(v)$ and $\deg(v)$ denote the degree of v in G' and G respectively. The new instance satisfies

$$\forall v \in \Lambda, \quad |L'_v| \geq |L_v| - (\deg(v) - \deg'(v)) \implies \frac{|L'_v|}{\Delta'} \geq \frac{|L_v| - \deg(v)}{\Delta'},$$

where Δ' denotes the maximum degree of G'. By the definition of $D(\eta)$, $\deg'(v) \leq \Delta' \leq \eta\Delta$. We have $|L_v| - \deg(v) > (\alpha^\star - 1)\Delta \geq \frac{\alpha^\star - 1}{\eta}\Delta'$. If we set the parameter $\eta \leq \frac{1}{10}$, then

$$\forall v \in \Lambda, \quad |L'_v| \geq 5\Delta'. \tag{5}$$

In this easy regime, one can use path coupling [6] to show $T_{\mathrm{rel}}^{(\eta)}(\mu) = O(n)$. To apply Theorem 9, we pick a small $\eta = O(\delta)$ and $\eta < \frac{1}{10}$. If $\Delta \geq \Delta_0 = \Theta(\frac{1}{\delta^4}\log\frac{1}{\delta})$, then

$$T_{\mathrm{rel}}^{(\eta)}(\mu) = 2^{O(1/\delta^2)}n = O_\delta(n).$$

On the other hand, if $\Delta \leq \Delta_0 = \Theta(\frac{1}{\delta^4}\log\frac{1}{\delta})$, then the maximum degree is bounded, we can use the result in [21] to obtain the relaxation time $T_{\mathrm{rel}}^{(\eta)}(\mu) = O_\delta(n)$ in the same order. This gives the proof sketch of Theorem 1. The only missing component is how to establish the coupling independence, which can be found in the full version ([10, Section 9]).

Proof of Proposition 3

To obtain (5), we only need to use the fact that $\alpha^\star > 1$. If we replace $\alpha^\star$ with $(1+\delta)$, then we can set $\eta \leq \frac{\delta}{5}$ and (5) still holds. The same analysis proves Proposition 3.

▶ Remark 10 (Hardcore Model in Uniqueness Regime). Theorem 9 could also rediscover some previous results. For example, for the hardcore model on a graph $G = (V, E)$ with fugacity $\lambda \leq (1-\delta)\Delta$, [11] proved the optimal $O_\delta(n)$ relaxation time. For a fixed λ, the hardcore model falls into an easy regime if we can reduce the maximum degree of the graph by a constant factor. The hardcore model in the uniqueness regime satisfies $O(1/\delta)$-coupling independence (which can be proved by Theorem 16 in this paper). Using a similar argument as that for list coloring, one can rediscover the optimal $O_\delta(n)$ relaxation time using Theorem 9.

We remark that the relaxation time result for the bipartite graph hardcore model (Theorem 5) is not a direct consequence from Theorem 9. We need to tweak the proof of Theorem 9 to prove Theorem 5. The proof of Theorem 5 is in the full version ([10, Section 8]).See Section 3 for a proof overview.

▶ Remark 11 (Compare Theorem 9 to the Technique in [11]). Another comparison result about relaxation time was given in [11]. The previous result considers general Boolean distribution (not necessarily Gibbs distribution) μ over $\{-,+\}^V$. Given a vector $\lambda = (\lambda_v)_{v\in V}$, $(\lambda * \mu)$ denotes the distribution such that for any $\sigma \in \{-,+\}^V$, $(\lambda * \mu) \propto \mu(\sigma)\prod_{v\in V:\sigma(v)=+}\lambda_v$. The result says if $(\lambda * \mu)$ is spectrally independent for all $\lambda \in (0,1]^V$, then one can compare $T_{\mathrm{rel}}^{\mathrm{GD}}(\mu)$ to $T_{\mathrm{rel}}^{\mathrm{GD}}(\lambda_\theta * \mu)$, where λ_θ is the vector with constant value $0 < \theta < 1$. When applying results to Gibbs distributions, here are some differences between Theorem 9 and the previous result in [11].

- Theorem 9 works for general domain $[q]$ but previous result works only for Boolean domain;
- The condition is incomparable. Theorem 9 requires coupling independence for μ and a degree lower bound for the underlying graph but the previous result requires spectral independence for a family of distributions;
- The easy regime is incomparable. The easy regime in Theorem 9 is the conditional distributions on a small degree subgraph but the easy regime in the previous result is $\lambda_\theta * \mu$;

For many spin systems, one can use Theorem 9 to establish the optimal $O(n)$ relaxation for Glauber dynamics, where n is the number of variables in the spin system. By the standard relation between mixing and relaxation time, the mixing time of Glauber dynamics can usually be bound by $O(n^2)$. Each transition of Glauber dynamics can be simulated in time $O(\Delta)$. Hence, one can obtain a sampling algorithm in time $O(\Delta n^2)$. Alternatively, we can give a faster sampling algorithm in time $\widetilde{O}(\Delta n)$ if the easy regime has near-linear mixing time.

▶ **Definition 12** (Mixing Time with Pinning). *Let μ be a Gibbs distribution on graph $G = (V, E)$ with maximum degree Δ. Let $\eta \in [0, 1]$. Let $D(\eta)$ denote all subsets $\Lambda \subseteq V$ such that the maximum degree of $G[\Lambda]$ is at most $\eta\Delta$. Define*

$$T^{(\eta)}_{\mathrm{mix}}(\mu) := \max\left\{ T^{\mathrm{GD}}_{\mathrm{mix}}\left(\mu^{\tau}, \frac{1}{4e}\right) \mid \Lambda \in D(\eta) \wedge \tau \in [q]^{V\setminus\Lambda} \right\}.$$

In words, for any pinning τ on $V \setminus \Lambda$ with $\Lambda \in D(\eta)$, $T^{(\eta)}_{\mathrm{mix}}(\mu)$ is an upper bound for the mixing time T of Glauber dynamics for μ^{τ} such that starting from the worst initial X_0, the total variation distance between X_T and μ^{τ} is at most $\frac{1}{4e}$.

▶ **Theorem 13** (Fast Sampling Algorithm). *Let $M \geq 1$ and $0 < \eta \leq \frac{1}{2\lceil M\rceil}$ be two constants. There exists an algorithm such that given any $\varepsilon \in (0, 1)$ and any Gibbs distribution μ on graph G with the maximum degree $\Delta \geq \Delta_0 = (\frac{10M}{\eta})^2 \log\frac{10M}{\eta}$, if μ satisfies M-coupling independence such that the weighted hamming distance ρ satisfies $\frac{\rho_{\max}}{\rho_{\min}} = \mathrm{poly}(n)$, then it returns a random sample X satisfying $\mathcal{D}_{\mathrm{TV}}(X \parallel \mu) \leq \varepsilon$ in time*

$$\Delta T^{(\eta)}_{\mathrm{mix}}(\mu)\left(\log\frac{n}{\varepsilon}\right)^{O(M/\eta)},$$

where n is the number of vertices in G and we use $\rho_{\max} = \max_{v\in V}\rho(v)$ and $\rho_{\min} = \min_{v\in V}\rho(v)$.

The theorem is proved in the full version ([10, Section 5]).See Section 3 for a proof overview.

In the above theorem, suppose $\frac{\rho_{\max}}{\rho_{\min}} = O(n^d)$ for some universal constant d, then the running time in above theorem should be $T^{\mathrm{GD}}_{\mathrm{mix}}(\mu, \eta)\Delta(\log\frac{n}{\varepsilon})^{C(M/\eta+d)}$ for some universal constant C. We then hide the constants C and d by $O(\cdot)$ in Theorem 13.

Theorem 4 can be obtained from Theorem 13. Consider the list coloring on a triangle free graphs $G = (V, E)$ with $|L_v| \geq (\alpha^\star + \delta)\Delta$. The uniform distribution μ satisfies $O(1/\delta)$-coupling-independence with standard Hamming weight $\rho(v) = 1$ for all $v \in V$. Take $\eta = O(1/\delta)$ be a small constant with $\eta < \frac{1}{10}$. By (5), a simple path coupling [6] shows that $T^{\mathrm{GD}}_{\mathrm{mix}}(\mu, \eta) = O(n \log n)$. Hence, if the maximum degree Δ is greater than Δ_0, we run the algorithm in Theorem 13 and the running time is $\Delta n \cdot \mathrm{polylog}(\frac{n}{\varepsilon})$. Otherwise, the maximum degree is bounded and the result in [21] gives the $O_\delta(n \log n)$ mixing time of the Glauber dynamics, then we can simulate Glauber dynamics to obtain a sampling algorithm. The proof of Theorem 4 is in the full version ([10, Section 9]).

The algorithm in Theorem 13 is *not* the Glauber dynamics. Roughly speaking, the algorithm uses some strategy to pick vertices and uses the Glauber update to resample the value of the picked vertices. However, for *monotone spin systems*, we can compare this algorithm to Glauber dynamics via censoring inequality [43] and then we can bound the mixing time of Glauber dynamics.

Let μ over $[q]^V$ be the Gibbs distribution. Define a partial order $\leq$ for $[q]^V$ as follows. For each $v \in V$, pick a *total order* $\leq_v$ on $[q]$. For any two $X, Y \in [q]^V$,

$$X \leq Y \quad \Longleftrightarrow \quad \forall v \in V, \quad X_v \leq_v Y_v. \tag{6}$$

For two distributions μ and ν over $[q]^V$, we say μ *is stochastic dominated by* ν (i.e., $\mu \preceq \nu$) if there is a coupling $\mathcal{C}$ between μ, ν such that $\mathbf{Pr}_{(X,Y)\sim\mathcal{C}}[X \leq Y] = 1$. Let P be the transition matrix of the Glauber dynamics on μ, which can be written as

$$P = \frac{1}{n}\sum_{v\in V} P_v, \tag{7}$$

where P_v performs *updates* at the $v \in V$ such that $P_v(X, Y) = \mu^{X_{V\setminus v}}(Y)$, for all $X, Y \in [q]^V$.

▶ **Definition 14** ([37, Chapter 22]). *We say μ is a* monotone spin system *if for every $v \in V$, P_v is ordering persistant, which means for any $X, Y \in [q]^V$ with $X \leq Y$, it holds that $P_v(X, \cdot) \preceq P_v(Y, \cdot)$.*

By the definition of the partial ordering $\leq$ in $[q]^V$, there is a unique maximum configuration for the ordering. Denote this state as X^+. Recall $T^{\mathrm{GD}}_{\mathrm{mix}}(\mu, \varepsilon \mid X^+)$ denotes the mixing time of Glauber dynamics starting from X^+.

▶ **Theorem 15** (Mixing Time Comparison). *Let $M \geq 1$ and $0 < \eta \leq \frac{1}{2\lceil M \rceil}$ be two constants. For any monotone spin system μ on graph G with the maximum degree $\Delta = \Omega(\frac{M^2}{\eta^2} \log \frac{M}{\eta})$, if μ satisfies M-coupling-independence such that the Hamming weight ρ satisfies $\frac{\rho_{\max}}{\rho_{\min}} = \mathrm{poly}(n)$, then the mixing time of Glauber dynamics starting from the maximum configuration satisfies*

$$T^{\mathrm{GD}}_{\mathrm{mix}}(\mu, \varepsilon \mid X^+) \leq \left(\log \frac{n}{\varepsilon}\right)^{O(M/\eta)} \cdot T^{(\eta)}_{\mathrm{mix}}(\mu),$$

where n is the number of vertices in graph G.

The theorem is proved in the full version ([10, Section 6]).See Section 3 for a proof overview.

The above theorem is of independent interest. Suppose μ is a monotone system with coupling independence property. The parameters of μ are in the critical regime and the underlying graph has an unbounded maximum degree. If we can choose a proper constant η such that $T^{(\eta)}_{\mathrm{mix}}(\mu) = O(n \log n)$, then the theorem gives a linear-optimal $\widetilde{O}(n)$ mixing time of Glauber dynamics for μ starting from the maximum configuration.

To obtain the near-linear mixing time for μ, some previous works [12, 15, 2] developed comparison techniques for the modified log-Sobolev (MLS) constants. Roughly speaking, if one can lower bound the MLS constant $\mathrm{mls}(\mu)$ of the Glauber dynamics for μ, then one can obtain the optimal $O(n \log n)$ mixing time. Previous works compared $\mathrm{mls}(\mu)$ to $\mathrm{mls}(\mu')$, where μ' is a distribution in the easy regime, and such comparison requires μ to satisfy certain entropic independence [2] condition. In general, it is not easy to verify the entropic independence condition and analyze $\mathrm{mls}(\mu')$ even if μ' is in an easy regime. Theorem 15 only requires the coupling independence condition and directly compares the mixing time. However, Theorem 15 requires monotone systems, and the final mixing result is restricted.

We remark that although the hardcore model in bipartite graphs is a monotone system, Theorem 6 is not a direct consequence from Theorem 15. We need to tweak the proof of Theorem 15 to prove Theorem 6. The proof of Theorem 6 is in the full version ([10, Section 8]). See Section 3 for a proof overview.

2.3 Establishing Coupling Independence

The next question is how to establish the coupling independence condition for spin systems. Previously, spectral independence was known for many spin systems. The coupling independence was often a by-product result when proving spectral independence. Hence, it is known for some specific spin systems such as subgraph world [14], b-matching [18] and coloring in high girth graphs [19].

In this paper, we give a tool to turn many existing spectral independence results into coupling independence results. A large family of spin systems is 2-spin systems. Let $G = (V, E)$ be a graph with maximum degree $\Delta \geq 3$. Let $0 \leq \beta \leq \gamma$ be the edge interactions such that $\gamma > 0$. Let $\lambda > 0$ be the external field. Let μ be the Gibbs distribution on G with parameters β, γ, λ such that for any $\sigma \in \{-, +\}$, $\mu(\sigma) \propto \lambda^{n_+(\sigma)} \beta^{m_+(\sigma)} \gamma^{m_-(\sigma)}$, where $n_+(\sigma)$ is the number of vertices v with $\sigma_v = +$ and $m_\pm(\sigma)$ is the number of edges $\{u, v\}$ with $\sigma_u = \sigma_v = \pm$. The 2-spin system is said to be ferromagnetic if $\beta\gamma > 1$ and anti-ferromagnetic if $\beta\gamma < 1$.

Anari, Liu, and Oveis Gharan [3] analyzed the spectral independence for the hardcore model. Chen, Liu, and Vigoda [20] extended the analysis to general 2-spin systems. Recall that the influence matrix Ψ_μ^τ is defined in (4). The maximum eigenvalue $\mathrm{Eig}_{\max}(\Psi_\mu^\tau)$ can be upper bounded by the total influence from one vertex

$$\mathrm{Eig}_{\max}(\Psi_\mu^\tau) \le \max_v \sum_{u \in V} |\Psi_\mu(v,u)|. \tag{8}$$

The RHS is called the *total influence bound*. For 2-spin systems, the analysis is performed on the Self-Avoiding-Walk (SAW) tree [46]. Roughly speaking, fix a vertex v, the SAW tree T_v enumerates all the SAWs in graph G starting from v. By defining a proper 2-spin system on T_v, one can use the total influence from the root in T_v to upper bound the total influence from v in G, and thus establish the spectral independence for Gibbs distribution μ. In [20], a weighted version of (8) is studied to deal with general 2-spin systems. We give the following result for coupling independence.

▶ **Theorem 16** (Informal version; see [10, Lemma 39] for the formal version)**.** *For 2-spin systems, the (weighted) total influence bound in the Self-Avoid-Walk tree implies coupling independence.*

As a consequence, all the spectral independence results for 2-spin systems in [20] can be turned into coupling independence results in black-box. For the hardcore model in bipartite graphs (Theorem 5 and Theorem 6), we can also use the above result to transform the total influence bound in [13] into coupling independence result.

Theorem 16 is proved by constructing a recursive coupling in the full version ([10, Section 7]). Fix a vertex v in G. We build a coupling (X, Y) between μ^{v^+} and μ^{v^-} and show the discrepancy between X and Y are bounded by the total influence in the SAW tree T_v. Suppose v has d neighbors $u_1, u_2, \ldots, u_d$. We split v into d copies $v_1, v_2, \ldots, v_d$ such that v_i only has one neighbor u_i. Define the pinning σ_i such that v_j for $j \ge i$ takes the value $+$ and v_j for $j < i$ takes the value $-$. Then $\mu^{v^+} = \mu^{\sigma_0}$ and $\mu^{v^-} = \mu^{\sigma_d}$. We couple each adjacent $\mu^{\sigma_{i-1}}$ and μ^{σ_i}, then merge them into a coupling between two endpoints μ^{σ_0} and μ^{σ_d}. For each adjacent pair, the only difference between σ_i and σ_{i-1} is the pinning at v_i. Hence, we first couple the only neighbor u_i of v_i then construct the coupling recursively if the coupling at u_i fails. This recursive processing essentially enumerates all SAWs from v. We can relate the coupling with the SAW tree to prove the theorem.

For multi-spin systems such as list-coloring, we can mimic the recursive coupling for 2-spin systems. Since the previous spectral independence results for list-coloring were also obtained via the SAW tree [24, 16], a similar proof gives the coupling independence.

3 Proof Overview

We give a proof overview for the relaxation time comparison result in Theorem 9. Let $G = (V, E)$ be a graph with maximum degree Δ. Let ℓ and k be two constant integers such that $\ell < k$. Their specific values will be fixed later. We first partition all the vertices in V into k parts $U_1, U_2, \ldots, U_k$ such that for any vertex $v \in V$, each U_i does not have more than $\frac{\eta}{\ell}\Delta$ neighbors of v, where $0 < \eta < 1$ is the parameter in Theorem 9. In other words, let $\Gamma_v = \{u \mid (u, v) \in E\}$ denote the set of neighbors of v in graph G. For any $i \in [k]$, $|\Gamma_v \cap U_i| \le \frac{\eta}{\ell}\Delta$. The existence of the partition is guaranteed by the Lovász local lemma. However, the local lemma requires the maximum degree Δ to be sufficiently large. That is why we require a lower bound for Δ in our technical results. We also remark that in our proof, the degree lower bound is used solely to ensure the existence of the partition. A similar partition appeared in the previous work [30].

The input Gibbs distribution μ over $[q]^V$ is a joint distribution of n variables $(X_v)_{v\in V}$, where each variable takes its value from $[q]$. Now, we can view μ as a joint distribution of k variables $Y = (Y_i)_{i\in[k]}$ such that each variable $Y_i = X_{U_i}$ takes its value from a huge domain $[q]^{U_i}$. We define the $k \leftrightarrow (k-\ell)$ down-up walk on Y. Given $Y = (Y_1, Y_2, \ldots, Y_k)$, the Markov chain does as follows

- Down-walk: Sample a set $S \in \binom{[k]}{\ell}$ of ℓ indices uniformly at random and then remove the configuration on the set S: $Y \to Y_{[k]\setminus S}$;
- Up-walk: Resample Y_S from μ conditional on $Y_{[k]\setminus S}$ and then go back to a full configuration $Y_{[k]\setminus S} \to Y_{[k]\setminus S} \cup Y_S$.

A full configuration $Y = (Y_1, Y_2, \ldots, Y_k)$ is on the level k. In the down-walk, we sample a random subset of indices $S \subseteq [k]$ with size ℓ. By dropping the configuration Y_S, we move from a full configuration at level k to a partial configuration at level $k-\ell$. In the up-walk, we resample Y_S and go back to the level k. The process can also be viewed as a kind of block dynamics for configuration $X \in [q]^V$. In every step, we pick a random subset $U_S = \cup_{i\in S} U_i \subseteq V$ of variables and resample $X(U_S)$ conditional on $X(V \setminus U_S)$.

We use local-to-global technique [35, 1, 3] to analyze the spectral gap of the $k \leftrightarrow (k-\ell)$ down-up walk for Y. The local-to-global technique suggests to analyze the relaxation time of $k \leftrightarrow 1$ down-up walk[2]. In the down walk, we pick a random S of size $|S| = k-1$ and drop Y_S. In the up-walk, we resample Y_S and go back to level k. We use coupling independence to analyze this $k \leftrightarrow 1$ down-up walk via path coupling. For simplicity, suppose μ satisfies C-coupling independence with standard Hamming distance ($\rho(v) = 1$ for all $v \in V$). We can view this $k \leftrightarrow 1$ down-up walk on Y as a block dynamics on $X \in [q]^V$, where it updates a block U_S in every step. Given two $X \in [q]^V$ and $X' \in [q]^V$ that disagree only at one vertex $v \in V$, say $v \in U_1$, we couple the transition of $k \leftrightarrow 1$ down-up walk. Let two $k \leftrightarrow 1$ down-up walks (starting from X and X', respectively) select the same random subset $S \subseteq [k]$ such that $|S| = k-1$.

- If $1 \in S$, which happens with probability $\frac{k-1}{k}$, then since $v \in U_1$ the value of v is removed in the down-walk, and thus X and X' can be coupled perfectly after the transition.
- If $1 \notin S$, which happens with probability $\frac{1}{k}$, then since $v \in U_1$, the disagreement at v may percolate to other blocks in the up-walk step. We use the coupling in the M-coupling independence to couple the up-walk so that the expected Hamming distance between X and X' after the transition is at most M.

Hence, the expected Hamming distance between X and X' after transition is at most $\frac{M}{k}$. If $k > M$, the path coupling gives the $O(\log n)$ mixing time and $O(1)$ relaxation time for this down-up walk. To apply the local-to-global technique, we also need to fix a configuration Y_Λ, where $\Lambda \subseteq [k]$ and $|\Lambda| = t \le \ell$, and consider the $(k-t) \leftrightarrow 1$ down-up walk for $Y_{[k]\setminus\Lambda}$. The same path coupling works if $k - t > M$. By choosing k and ℓ such that $k - \ell > M$ and using the local-to-global technique, we can show that the $k \leftrightarrow (k-\ell)$ down-up walk for Y has $O(1)$ relaxation time.

We then compare the $k \leftrightarrow (k-\ell)$ down-up for $Y = (Y_1, Y_2, \ldots, Y_k)$ to the Glauber dynamics for $X \in [q]^V$. Recall that $k \leftrightarrow (k-\ell)$ down-up walk is a kind of block dynamics for X. In every step, the block dynamics updates a subset $U_S = \cup_{i\in S} U_i$ with $|S| = \ell$. The update step is to resample $X(U_S)$ conditional on $X(V \setminus U_S)$. This step samples from the

[2] In [1, 3], the local walk is essentially defined as the $1 \leftrightarrow k$ up-down walk. Every state is Y_i for $i \in [k]$. In the up-walk, it extends Y_i to a full configuration Y. In the down-walk step, it samples a random index $j \in [k]$ and updates Y to Y_j. It is well-known that $1 \leftrightarrow k$ up-down walk and $k \leftrightarrow 1$ up-down walk has the same relaxation time.

conditional Gibbs distribution $\mu_{U_S}^{X(V\setminus U_S)}$ on subgraph $G[U_S]$. By the construction of the partition, the maximum degree of $G[U_S]$ is at most $\eta\Delta$ so that we have the relaxation time bound $T_{\mathrm{rel}}^{(\eta)}(\mu)$ for Glauber dynamics on $\mu_{U_S}^{X_{V\setminus U_S}}$. Let $T_{\mathrm{rel}}^{\text{down-up}}$ denote the relaxation time of $k \leftrightarrow (k-\ell)$ down-up walk. By some standard comparison argument between block dynamics and the Glauber dynamics, we can prove Theorem 9 by showing that

$$T_{\mathrm{rel}}^{\mathrm{GD}}(\mu) \le T_{\mathrm{rel}}^{\text{down-up}} \times T_{\mathrm{rel}}^{(\eta)}(\mu) = O(1) \times T_{\mathrm{rel}}^{(\eta)}(\mu).$$

Next, we explain how to get the near-linear time sampling algorithm in Theorem 13 and the mixing time in Theorem 15. Note that for $k \leftrightarrow 1$ down-up walk, the path coupling actually gives the $O(\log n)$ mixing time. For one update step, it selects a subset $S \subseteq [k]$ with $|S| = k-1$. Let i denote the missing index, i.e. $S \cup \{i\} = [k]$. The update step resamples Y_S conditional on Y_i. We can simulate this transition step using $(k-1) \leftrightarrow 1$ down-up walk for the conditional distribution on Y_S. This down-up walk also has the $O(\log n)$ mixing time. We do this recursively until we need to sample from a conditional distribution on $Y_{S'}$ with $|S'| = \ell$. Note that the maximum degree of the graph $G[U_{S'}]$ is at most $\eta\Delta$. Now, we simulate the Glauber dynamics for $O(T_{\mathrm{mix}}^{(\eta)}(\mu) \cdot \log n)$ steps to sample from the conditional distribution. The following informal algorithm generates an approximate random sample from μ within TV-distance error $\varepsilon = O(1)$. The formal algorithm is given in the full version ([10, Section 5]).

Algorithm $\mathsf{Sample}(X, \Lambda)$
Input: a subset $\Lambda \subseteq [k]$;
Output: the algorithm randomly updates the partial configuration X_{U_Λ} so that X_{U_Λ} becomes an independent approximate sample of $\mu_{U_\Lambda}^{\sigma}$, where $\sigma = X_{U_{[k]\setminus\Lambda}}$.

1. **if** $|\Lambda| = \ell$: update X_{U_Λ} for $O(T_{\mathrm{mix}}^{(\eta)}(\mu) \cdot \log n)$ steps, where every step uses the update rule of the Glauber dynamics on $\mu_{U_\Lambda}^{\sigma}$, where $\sigma = X_{U_{[k]\setminus\Lambda}}$.
2. **else**: run the update of $|\Lambda| \leftrightarrow 1$ down-up walk on $\mu_{U_\Lambda}^{\sigma}$ for $O(\log n)$ steps, where $\sigma = X_{U_{[k]\setminus\Lambda}}$ and every step does as follows:
   a. pick an index $i \in \Lambda$ uniformly at random and let $S = \Lambda \setminus \{i\}$;
   b. call $\mathsf{Sample}(X, S)$ to update X_{U_S}; `(recursion step)`

We can call $\mathsf{Sample}(X, [k])$ with an arbitrary feasible $X \in [q]^V$. When the algorithm terminates, X is updated to an approximate random sample from the distribution μ. Note that the parameter k is a constant. The total number of Glauber update steps is given by $N = (\log n)^{O(1)} T_{\mathrm{mix}}^{(\eta)}(\mu)$. Therefore, the overall running time of the algorithm is $O(\Delta N)$. Alternatively, we can think of the algorithm as follows: it first generates a random sequence of vertices $v_1, v_2, \ldots, v_N \in V$. In the t-th step, the algorithm randomly updates the value of v_t, conditioned on the values of other variables. This behavior is similar to standard Glauber dynamics for μ, except that the sequence $v_1, v_2, \ldots, v_N$ follows a non-trivial joint distribution. For monotone systems, we can compare this algorithm with Glauber dynamics using the censoring inequality.

The results for list-coloring are consequences of general technical results. However, we need to tweak the analysis to prove the results for hardcore model in bipartite graphs (Theorem 5 and Theorem 6). The reason is that for hardcore model on $G = (V_L \cup V_R, E)$, we only know $\lambda < \lambda_c(\Delta_L)$ but we cannot control the degree Δ_R in the right part V_R. Our technique can only prove the coupling independence for μ_L, which is the marginal distribution on V_L projected from μ. To prove the relaxation time and mixing time results, we first partition V_L into disjoint part $U_1, U_2, \ldots, U_k$ such that for any vertex $v \in V_R$, v has no more

than $\frac{\eta}{\ell}\Delta$ neighbors in each U_i. Again, the existence of the partition is guaranteed by the local lemma. Let $X \sim \mu_L$ be a partial configuration on V_L. We can define $Y = (Y_1, Y_2, \ldots, Y_k)$, where $Y_i = X_{U_i}$. By a similar proof, we show that the $k \leftrightarrow (k-\ell)$ down-up walk for Y is rapid mixing. We consider a global Markov chain over $\{-,+\}^{V_L \cup V_R}$ defined as follows. Let $\overline{X} \in \{-,+\}^{V_L \cup V_R}$ be a full configuration.

- Drop the right part to obtain $X \leftarrow \overline{X}(V_L)$;
- Update X using the $k \leftrightarrow (k-\ell)$ down-up walk for $Y = (Y_1, \ldots, Y_k)$, where $Y_i = X(U_i)$;
- Sample $X(V_R) \sim \mu^X_{V_R}$ and let $\overline{X} \leftarrow X \cup X(V_R)$.

We first compare this chain to the $k \leftrightarrow (k-\ell)$ down-up walk and then compare the Glauber dynamics for μ to this Markov chain. This gives the relaxation time of Glauber dynamics. For the mixing time, we can first obtain a near-linear time sampling algorithm for μ_L, since hardcore model in bipartite graph is a monotone system, we then compare the algorithm to the Glauber dynamics for μ via the censoring inequality.

References

1. Vedat Levi Alev and Lap Chi Lau. Improved analysis of higher order random walks and applications. In *STOC*, pages 1198–1211, 2020. doi:10.1145/3357713.3384317.
2. Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. Entropic independence: optimal mixing of down-up random walks. In *STOC*, pages 1418–1430, 2022. doi:10.1145/3519935.3520048.
3. Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. In *FOCS*, pages 1319–1330, 2020. doi:10.1109/FOCS46700.2020.00125.
4. 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. doi:10.1145/3313276.3316385.
5. Antonio Blanca, Pietro Caputo, Zongchen Chen, Daniel Parisi, Daniel Štefankovič, and Eric Vigoda. On mixing of markov chains: Coupling, spectral independence, and entropy factorization. *arXiv preprint arXiv:2103.07459*, 2021. arXiv:2103.07459.
6. Russ Bubley and Martin Dyer. Path coupling: A technique for proving rapid mixing in markov chains. In *FOCS*, pages 223–231, 1997. doi:10.1109/SFCS.1997.646111.
7. Jin-Yi Cai, Andreas Galanis, Leslie Ann Goldberg, Heng Guo, Mark Jerrum, Daniel Štefankovič, and Eric Vigoda. #BIS-hardness for 2-spin systems on bipartite bounded degree graphs in the tree non-uniqueness region. *J. Comput. System Sci.*, 82(5):690–711, 2016. doi:10.1016/j.jcss.2015.11.009.
8. Sarah Cannon and Will Perkins. Counting independent sets in unbalanced bipartite graphs. In *SODA*, pages 1456–1466, 2020. doi:10.1137/1.9781611975994.88.
9. 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. doi:10.1137/1.9781611975482.134.
10. Xiaoyu Chen and Weiming Feng. Rapid mixing via coupling independence for spin systems with unbounded degree. *arXiv:2407.04672*, 2024. doi:10.48550/arXiv.2407.04672.
11. Xiaoyu Chen, Weiming Feng, Yitong Yin, and Xinyuan Zhang. Rapid mixing of glauber dynamics via spectral independence for all degrees. In *FOCS*, pages 137–148, 2021. doi:10.1109/FOCS52979.2021.00022.
12. Xiaoyu Chen, Weiming Feng, Yitong Yin, and Xinyuan Zhang. Optimal mixing for two-state anti-ferromagnetic spin systems. In *FOCS*, pages 588–599, 2022. doi:10.1109/FOCS54457.2022.00062.
13. Xiaoyu Chen, Jingcheng Liu, and Yitong Yin. Uniqueness and rapid mixing in the bipartite hardcore model. In *FOCS*, pages 1991–2005. IEEE, 2023. doi:10.1109/FOCS57990.2023.00121.

14 Xiaoyu Chen and Xinyuan Zhang. A near-linear time sampler for the Ising model with external field. In *SODA*, pages 4478–4503, 2023. doi:10.1137/1.9781611977554.CH170.

15 Yuansi Chen and Ronen Eldan. Localization schemes: A framework for proving mixing bounds for markov chains. In *FOCS*, pages 110–122, 2022.

16 Zongchen Chen, Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Rapid mixing for colorings via spectral independence. In *SODA*, pages 1548–1557, 2021.

17 Zongchen Chen, Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Sampling colorings and independent sets of random regular bipartite graphs in the non-uniqueness region. In *SODA*, pages 2198–2207, 2022.

18 Zongchen Chen and Yuzhou Gu. Fast sampling of b-matchings and b-edge covers. In *SODA*, pages 4972–4987, 2024. doi:10.1137/1.9781611977912.178.

19 Zongchen Chen, Kuikui Liu, Nitya Mani, and Ankur Moitra. Strong spatial mixing for colorings on trees and its algorithmic applications. In *FOCS*, pages 810–845, 2023. doi:10.1109/FOCS57990.2023.00053.

20 Zongchen Chen, Kuikui Liu, and Eric Vigoda. Rapid mixing of Glauber dynamics up to uniqueness via contraction. In *FOCS*, pages 1307–1318, 2020. doi:10.1109/FOCS46700.2020.00124.

21 Zongchen Chen, Kuikui Liu, and Eric Vigoda. Optimal mixing of Glauber dynamics: entropy factorization via high-dimensional expansion. In *STOC*, pages 1537–1550, 2021. doi:10.1145/3406325.3451035.

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

23 Charilaos Efthymiou, Thomas P. Hayes, Daniel Štefankovič, Eric Vigoda, and Yitong Yin. Convergence of MCMC and loopy BP in the tree uniqueness region for the hard-core model. *SIAM J. Comput.*, 48(2):581–643, 2019. doi:10.1137/17M1127144.

24 Weiming Feng, Heng Guo, Yitong Yin, and Chihao Zhang. Rapid mixing from spectral independence beyond the boolean domain. In *SODA*, pages 1558–1577, 2021. doi:10.1137/1.9781611976465.95.

25 Andreas Galanis, Qi Ge, Daniel Štefankovič, Eric Vigoda, and Linji Yang. Improved inapproximability results for counting independent sets in the hard-core model. *Random Structures Algorithms*, 45(1):78–110, 2014. (conference version in *RANDOM'11)*. doi:10.1002/RSA.20479.

26 David Gamarnik, Dmitriy Katz, and Sidhant Misra. Strong spatial mixing of list coloring of graphs. *Random Structures Algorithms*, 2013.

27 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. doi:10.1137/S0097539704445470.

28 Thomas P. Hayes and Eric Vigoda. Coupling with the stationary distribution and improved sampling for colorings and independent sets. *Ann. Appl. Probab.*, 16(3):1297–1318, 2006.

29 Vishesh Jain, Huy Tuan Pham, and Thuy Duong Vuong. Spectral independence, coupling with the stationary distribution, and the spectral gap of the Glauber dynamics. *Inf. Process. Lett.*, 177:106268, 2022.

30 Vishesh Jain, Ashwin Sah, and Mehtaab Sawhney. Perfectly sampling $k \geq (8/3 + o(1))\Delta$-colorings in graphs. In *STOC*, pages 1589–1600, 2021.

31 Matthew Jenssen, Peter Keevash, and Will Perkins. Algorithms for #BIS-hard problems on expander graphs. *SIAM J. Comput.*, 49(4):681–710, 2020. doi:10.1137/19M1286669.

32 Matthew Jenssen, Will Perkins, and Aditya Potukuchi. Approximately counting independent sets in bipartite graphs via graph containers. *Random Structures Algorithms*, 63(1):215–241, 2023. doi:10.1002/RSA.21145.

33 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. doi:10.1002/RSA.3240070205.

34 Mark Jerrum. Glauber dynamics for the hard-core model on bounded-degree H-free graphs. *CoRR*, abs/2404.07615, 2024. doi:10.48550/arXiv.2404.07615.

35 Tali Kaufman and Izhar Oppenheim. High order random walks: beyond spectral gap. *Combinatorica*, 40(2):245–281, 2020. (conference version in *RANDOM*'18). doi:10.1007/S00493-019-3847-0.

36 Frank P Kelly. Stochastic models of computer communication systems. *Journal of the Royal Statistical Society: Series B (Methodological)*, 47(3):379–395, 1985.

37 David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2017.

38 Chao Liao, Jiabao Lin, Pinyan Lu, and Zhenyu Mao. An FPTAS for the hardcore model on random regular bipartite graphs. *Theoret. Comput. Sci.*, 929:174–190, 2022. doi:10.1016/J.TCS.2022.07.001.

39 Jingcheng Liu and Pinyan Lu. FPTAS for #BIS with degree bounds on one side. In *STOC*, pages 549–556, 2015. doi:10.1145/2746539.2746598.

40 Kuikui Liu. From coupling to spectral independence and blackbox comparison with the down-up walk. *arXiv preprint arXiv:2103.11609*, 2021. arXiv:2103.11609.

41 Michael Luby and Eric Vigoda. Fast convergence of the glauber dynamics for sampling independent sets. *Random Structures Algorithms*, 15(3-4):229–241, 1999. doi:10.1002/(SICI)1098-2418(199910/12)15:3/4\%3C229::AID-RSA3\%3E3.0.CO;2-X.

42 Marc Mezard and Andrea Montanari. *Information, physics, and computation*. Oxford University Press, 2009.

43 Yuval Peres and Peter Winkler. Can extra updates delay mixing? *Comm. Math. Phys.*, 323(3):1007–1016, 2013.

44 Allan Sly. Computational transition at the uniqueness threshold. In *FOCS*, pages 287–296, 2010. doi:10.1109/FOCS.2010.34.

45 Eric Vigoda. Improved bounds for sampling colorings. *J. Math. Phys.*, 41(3):1555–1569, 2000.

46 Dror Weitz. Counting independent sets up to the tree threshold. In *STOC*, pages 140–149, 2006. doi:10.1145/1132516.1132538.