Optimal Mixing of Glauber Dynamics: Entropy Factorization via High-Dimensional Expansion

Zongchen Chen
Georgia Institute of Technology
Atlanta, GA, USA
chenzongchen@gatech.edu

Kuikui Liu
University of Washington
Seattle, WA, USA
liukui17@cs.washington.edu

Eric Vigoda
Georgia Institute of Technology
Atlanta, GA, USA
vigoda@gatech.edu

ABSTRACT

We prove an optimal mixing time bound for the single-site update Markov chain known as the Glauber dynamics or Gibbs sampling in a variety of settings. Our work presents an improved version of the spectral independence approach of Anari et al. (2020) and shows $O(n \log n)$ mixing time on any n-vertex graph of bounded degree when the maximum eigenvalue of an associated influence matrix is bounded. As an application of our results, for the hard-core model on independent sets weighted by a fugacity λ, we establish $O(n \log n)$ mixing time for the Glauber dynamics on any n-vertex graph of constant maximum degree Δ when $\lambda < \lambda_c(\Delta)$ where $\lambda_c(\Delta)$ is the critical point for the uniqueness/non-uniqueness phase transition on the Δ-regular tree. More generally, for any antiferromagnetic 2-spin system we prove $O(n \log n)$ mixing time of the Glauber dynamics on any bounded degree graph in the corresponding tree uniqueness region. Our results apply more broadly; for example, we also obtain $O(n \log n)$ mixing for q-colorings of triangle-free graphs of maximum degree Δ when the number of colors satisfies $q > \alpha\Delta$ where $\alpha \approx 1.763$, and $O(m \log n)$ mixing for generating random matchings of any graph with bounded degree and m edges.

Our approach is based on two steps. First, we show that the approximate tensorization of entropy (i.e., factorizing entropy into single vertices), which is a key step for establishing the modified log-Sobolev inequality in many previous works, can be deduced from entropy factorization into blocks of fixed linear size. Second, we adapt the local-to-global scheme of Alev and Lau (2020) to establish such block factorization of entropy in a more general setting of pure weighted simplicial complexes satisfying local spectral expansion; this also substantially generalizes the result of Cryan et al. (2019).

CCS CONCEPTS

• **Theory of computation → Random walks and Markov chains**; **Generating random combinatorial structures**; *Expander graphs and randomness extractors*; • **Mathematics of computing → Markov processes**.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
STOC '21, June 21–25, 2021, Virtual, Italy
© 2021 Association for Computing Machinery.
ACM ISBN 978-1-4503-8053-9/21/06…$15.00
https://doi.org/10.1145/3406325.3451035

KEYWORDS

Glauber dynamics, spectral independence, high-dimensional expander, modified log-Sobolev inequality, spin system

ACM Reference Format:
Zongchen Chen, Kuikui Liu, and Eric Vigoda. 2021. Optimal Mixing of Glauber Dynamics: Entropy Factorization via High-Dimensional Expansion. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing (STOC '21), June 21–25, 2021, Virtual, Italy*. ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3406325.3451035

1 INTRODUCTION

This paper establishes a well-known conjecture that the Glauber dynamics converges very quickly to its stationary distribution in the tree uniqueness region, i.e., decay of correlations region. The Glauber dynamics is the quintessential example of a local Markov chain, and its convergence rate is of great interest due to its simplicity and wide applicability.

Our setting is the general framework of spin systems. Spin systems capture many combinatorial models of interest, including the hard-core model on weighted independent sets, the Ising model, and colorings, and are equivalent to undirected graphical models. For integer $q \geq 2$, a q-state spin system is defined by a $q \times q$ interaction matrix A. For a given graph $G = (V, E)$ with $n = |V|$ vertices, the configurations of the model are the collection Ω of assignments $\sigma : V \to [q]$ of spins to the vertices of the graph. Each configuration $\sigma \in \Omega$ has an associated weight $w(\sigma)$ defined by the pairwise interactions weighted by the interaction matrix A, see Section 1.1 for a detailed definition.

The Gibbs distribution μ is the probability distribution over the collection Ω of configurations and is defined as $\mu(\sigma) = w(\sigma)/Z$ where $Z = \sum_\sigma w(\sigma)$ is the normalizing factor known as the partition function. Approximately sampling from the Gibbs distribution is polynomial-time equivalent to approximating the partition function [34, 50]. Given an $\varepsilon > 0$ and $\delta > 0$, an FPRAS for the partition function outputs a $(1 \pm \varepsilon)$-relative approximation of the partition function with probability $\geq 1 - \delta$, whereas an FPTAS is the deterministic analog (i.e., it achieves $\delta = 0$).

The canonical example of a spin system in statistical physics is the Ising model. The Ising model is a 2-spin system (i.e., $q = 2$); the spin space is denoted as $\{+, -\}$ and the configurations of the model are the 2^n assignments of spins $\{+, -\}$ to the vertices of the underlying graph. In the simpler case without an external field the Ising model has a single parameter $\beta > 0$ corresponding to the inverse temperature. A configuration $\sigma \in \Omega$ has weight $w(\sigma) = \beta^{m(\sigma)}$ where $m(\sigma) = |\{(u, v) \in E : \sigma(u) = \sigma(v)\}|$ is the number of monochromatic edges in σ. When $\beta > 1$ then the model is *ferromagnetic* as the two fully monochromatic configurations

have maximum weight, whereas when $\beta < 1$ then the model is *antiferromagnetic*.

The hard-core model is a natural combinatorial example of an antiferromagnetic 2-spin system. The model is parameterized by a fugacity $\lambda > 0$. For a graph $G = (V, E)$, configurations of the model are the collection Ω of independent sets of G, and the weight of an independent set σ is $w(\sigma) = \lambda^{|\sigma|}$.

In general, a 2-spin system is defined by three parameters $\beta, \gamma \geq 0$ and $\lambda > 0$. A spin configuration $\sigma \in \{0,1\}^V$ is assigned weight: $w(\sigma) = \beta^{m_1(\sigma)}\gamma^{m_0(\sigma)}\lambda^{n_1(\sigma)}$, where, for $s \in \{0,1\}$, $m_s(\sigma)$ is the number of edges where both endpoints receive spin s and $n_s(\sigma)$ is the number of vertices assigned spin s. Note the Ising model corresponds to the case $\beta = \gamma$ where λ is the external field, and the hard-core model corresponds to $\beta = 0, \gamma = 1$. The model is ferromagnetic when $\beta\gamma > 1$ and antiferromagnetic when $\beta\gamma < 1$ (the model is trivial when $\beta\gamma = 1$).

The Glauber dynamics is a simple Markov chain (X_t) designed for sampling from the Gibbs distribution μ. The transitions $X_t \to X_{t+1}$ update a randomly chosen vertex as follows: (i) select a vertex v uniformly at random; (ii) for all $u \neq v$, set $X_{t+1}(u) = X_t(u)$; and (iii) choose $X_{t+1}(v)$ from the marginal distribution for the spin at v conditional on the configuration $X_{t+1}(N(v))$ on the neighbors $N(v)$ of v. It is straightforward to verify that the chain is ergodic (in the cases considered here, see the definition of totally-connected in Section 1.1) and the unique stationary distribution is the Gibbs distribution.

The *mixing time* is the number of transitions, for the worst initial state X_0, to guarantee that X_t is within total variation distance $\leq 1/4$ of the Gibbs distribution; for a formal statement, see Eq. (1). We say the chain is *rapidly mixing* when the mixing time is polynomial in $n = |V|$. Hayes and Sinclair [29] established that the mixing time of the Glauber dynamics is $\Omega(n \log n)$ for a family of bounded-degree graphs, and hence we say that the Glauber dynamics has optimal mixing time when the mixing time is $O(n \log n)$.

The computational complexity of approximating the partition function is closely connected to statistical physics phase transitions. For $\Delta \geq 3$, consider the tree T_ℓ of height ℓ where all of the internal vertices have degree Δ, and let r denote its root. The uniqueness/non-uniqueness phase transition captures whether the leaves influence the root, in the limit as the height grows.

The uniqueness/non-uniqueness phase transition is nicely illustrated for the Ising model which has two extremal boundaries: the all + boundary and all − boundary. For $s \in \{+,-\}$, let p_ℓ^s denote the marginal probability that the root has spin + in the Gibbs distribution on T_ℓ conditional on all leaves having spin s. The model is in the uniqueness phase iff $\lim_{\ell\to\infty} p_\ell^+ = \lim_{\ell\to\infty} p_\ell^-$. For the Ising model (without an external field) the uniqueness/non-uniqueness phase transition occurs at $\beta_c(\Delta) = (\Delta-2)/\Delta$ for the antiferromagnetic case and $\overline{\beta}_c(\Delta) = \Delta/(\Delta-2)$ for the ferromagnetic case. For the hard-core model, the critical fugacity is $\lambda_c(\Delta) := (\Delta-1)^{\Delta-1}/(\Delta-2)^\Delta$. This phase transition on the Δ-regular tree is connected to the complexity of approximating the partition function on graphs of maximum degree Δ.

For the hard-core model, for constant Δ, for any $\delta > 0$, Weitz [53] presented an FPTAS for the partition function on graphs of maximum degree Δ when $\lambda < (1-\delta)\lambda_c(\Delta)$. In contrast, when $\lambda > \lambda_c$, Sly [48] (see also [25, 49], unless NP = RP, there is no FPRAS for approximating the partition function on graphs of maximum degree Δ. Li, Lu, and Yin [39] generalized Weitz's correlation decay algorithmic approach to all antiferromagnetic 2-spin systems when the system is up-to-Δ unique. One important caveat to these correlation decay approaches is that the running time depends exponentially on $\log\Delta$ and $1/\delta$.

Despite the algorithmic successes of the correlation decay approach, establishing rapid mixing of the Glauber dynamics in the same tree uniqueness region was a vexing open problem. Anari, Liu, and Oveis Gharan [2] introduced the spectral independence approach based on the theory of high-dimensional expanders [1, 19, 35, 37, 45], and established rapid mixing of the Glauber dynamics for the hard-core model on any graph of maximum degree Δ when $\lambda < (1-\delta)\lambda_c(\Delta)$ for $\delta > 0$. However, while the mixing time had polynomial dependence on Δ, it also had doubly exponential dependence on $1/\delta$. In [16] the authors established rapid mixing for all antiferromagnetic 2-spin systems when the system is up-to-Δ-unique with gap δ and improved the mixing time to an exponential dependence on $1/\delta$. Here, roughly speaking, up-to-Δ uniqueness with gap δ means (multiplicative) gap δ from the uniqueness threshold on the Δ-regular tree for all $d \leq \Delta$; see Definition 5.1 for a precise statement, and [39] for more discussion.

In this work, we not only establish a fixed polynomial upper bound on the mixing time, but we also prove optimal mixing of the Glauber dynamics. Our approach holds for general spin systems. The spectral independence approach, first introduced for 2-spins in [2] and subsequently extended to q-spins in [14, 23], considers the $qn \times qn$ influence matrix. For spins $i, j \in [q]$ and vertices $u, v \in V$, the entry $((u,i),(v,j))$ of the influence matrix measures the effect of vertex u having spin i on the marginal probability that vertex v has spin j, see Definition 1.7 for a precise statement. Here we prove that if the maximum eigenvalue of the influence matrix is upper bounded and the marginal probabilities are lower bounded then the mixing time is $O(n \log n)$ where the only dependence on $1/\delta$ and Δ is in the constant factor captured by the big-O notation. Our main result is stated in Theorem 1.9 in Section 1.1 after presenting the necessary definitions.

We establish optimal mixing time of $O(n \log n)$ by proving that the Glauber dynamics contracts relative entropy (with respect to the Gibbs distribution) at a constant rate. This is analogous to establishing a modified log-Sobolev constant for the Glauber dynamics; there are several recent results in other contexts also proving entropy decay for various Markov chains [6, 11, 17]. In contrast, previous works utilizing the spectral independence approach [2, 16] and related works on high-dimensional expanders [1, 19, 35, 37, 45] consider the spectral gap (or analogously, decay of variance); such an approach is unable to establish optimal mixing time. Our proof approach is outlined in Section 2.2.

The application of our results is nicely illustrated for the particular case of antiferromagnetic 2-spin systems. We prove $O(n \log n)$ mixing time of the Glauber dynamics when the system is up-to-Δ-unique. This is the same region where the correlation decay results of [39] and the rapid mixing results of [16] hold, which matches the hardness results of [49]. Note, a mixing time of $O(n \log n)$ implies an $\widetilde{O}(n^2)$ time FPRAS for approximating the partition function [38, 50].

Theorem 1.1. *For all $\Delta \geq 3$, all $\delta \in (0,1)$, and all parameters (β, γ, λ) specifying an antiferromagnetic 2-spin system which is up-to-Δ unique with gap δ, there exists $C = C(\Delta, \delta, \beta, \gamma, \lambda)$ such that for every n-vertex graph $G = (V, E)$ of maximum degree at most Δ, the mixing time of the Glauber dynamics for the 2-spin system on G with parameters (β, γ, λ) is at most $Cn \log n$.*

For the case of the hard-core model our theorem yields the following result.

Theorem 1.2. *For all $\Delta \geq 3$ and all $\delta \in (0,1)$, there exists $C = C(\Delta, \delta)$ such that for every n-vertex graph $G = (V, E)$ of maximum degree at most Δ and every $\lambda \leq (1-\delta)\lambda_c(\Delta)$, the mixing time of the Glauber dynamics for the hard-core model on G with fugacity λ is at most $Cn \log n$.*

For the case of the Ising model in both the antiferromagnetic and ferromagnetic case, our theorem yields optimal mixing whenever β is between $\beta_c(\Delta) = \frac{\Delta-2}{\Delta}$ and $\overline{\beta}_c(\Delta) = \frac{\Delta}{\Delta-2}$.

Theorem 1.3. *For all $\Delta \geq 3$ and all $\delta \in (0,1)$, there exists $C = C(\Delta, \delta)$ such that for every n-vertex graph $G = (V, E)$ of maximum degree at most Δ, every $\beta \in [\frac{\Delta-2+\delta}{\Delta-\delta}, \frac{\Delta-\delta}{\Delta-2+\delta}]$, and every $\lambda > 0$, the mixing time of the Glauber dynamics for the Ising model on G with inverse temperature β and external field λ is at most $Cn \log n$.*

Remark 1. We can actually show that specifically for the Ising model, $C = \Delta^{O(1/\delta)}$ suffices when n is large enough and so we obtain polynomial mixing time even when the graph has unbounded degree.

Recall that the above results are tight as there is no efficient approximation algorithm in the tree non-uniqueness region which corresponds to $\lambda > \lambda_c(\Delta)$ for the hard-core model and $\beta < \beta_c(\Delta)$ for the antiferromagnetic Ising model. The only analog of the above results establishing optimal mixing time in the entire tree uniqueness region was the work of Mossel and Sly [44] for the ferromagnetic Ising model. Their proof utilizes the monotonicity properties of the ferromagnetic Ising model which allows the use of the censoring inequality of Peres and Winkler [46]. The algorithm of Jerrum and Sinclair [33] gives an FPRAS for the ferromagnetic Ising model for any β and any G, but the polynomial exponent is a large constant.

Our results hold for multi-spin systems as well. The most notable example of a multi-spin system is the q-colorings problem, namely, proper vertex q-colorings. Given a graph $G = (V, E)$ of maximum degree Δ, can we approximate the number of q-colorings of G? Jerrum [30] proved $O(n \log n)$ mixing time of the Glauber dynamics whenever $q > 2\Delta$. This was further improved in [13, 51] to $O(n^2)$ mixing time when $q > (11/6 - \varepsilon)\Delta$ for some small $\varepsilon > 0$. There are several further improvements with various assumptions on the girth or maximum degree, c.f. [22]. On the hardness side, Galanis et al. [24] proved that unless NP = RP there is no FPRAS for approximating the number of q-colorings when q is even and $q < \Delta$.

For triangle-free graphs, a recent pair of works [14, 23] extended the spectral independence approach to establish rapid mixing of the Glauber dynamics when $q > (\alpha^* + \delta)\Delta$ for any $\delta > 0$ where $\alpha^* \approx 1.763$; however the polynomial exponent in the mixing time depends on $1/\delta$ in these results. Using our main result we prove $O(n \log n)$ mixing time of the Glauber dynamics under the same conditions.

Theorem 1.4. *Let $\alpha^* \approx 1.763$ denote the unique solution to $x = \exp(1/x)$. For all $\Delta \geq 3$ and all $\delta > 0$, there exists $C = C(\Delta, \delta)$ such that for every n-vertex triangle-free graph $G = (V, E)$ of maximum degree at most Δ and every $q \geq (\alpha^* + \delta)\Delta$, the mixing time of the Glauber dynamics for sampling random q-colorings on G is at most $Cn \log n$.*

We prove spectral independence bounds for the monomer-dimer model on all matchings of a graph; no nontrivial bounds were previously known. Given a graph $G = (V, E)$ and a fugacity $\lambda > 0$, the Gibbs distribution μ for the monomer-dimer model is defined on the collection $\mathcal{M}$ of all matchings of G where $\mu(M) = w(M)/Z$ for $w(M) = \lambda^{|M|}$. The Glauber dynamics for the monomer-dimer model adds or deletes a random edge in each step. In particular, from $X_t \in \mathcal{M}$, choose an edge e uniformly at random from E and let $X' = X_t \oplus e$. If $X' \in \mathcal{M}$ then let $X_{t+1} = X'$ with probability $w(X')/(w(X') + w(X_t))$ and otherwise let $X_{t+1} = X_t$.

We prove $O(m \log n)$ mixing time for the Glauber dynamics for sampling matchings on bounded-degree graphs with n vertices and m edges. A classical result of Jerrum and Sinclair [32] yields rapid mixing of the Glauber dynamics for any graph, but the best mixing time bound was $O(n^2 m \log n)$ [31].

Theorem 1.5. *For all $\Delta \geq 3$ and all $\lambda > 0$, there exists $C = C(\Delta, \lambda)$ such that for every n-vertex graph $G = (V, E)$ of maximum degree at most Δ, the mixing time of the Glauber dynamics for the monomer-dimer model on G with fugacity λ is at most $Cm \log n$.*

For general ferromagnetic 2-spin systems the existing picture is not as clear as for antiferromagnetic systems. Our work extends to ferromagnetic 2-spin systems, proving $O(n \log n)$ mixing time for the same range of parameters as the previously best known bounds [16, 28, 47]. In particular, we recover Theorems 26 and 27 in [16] with $O(n \log n)$ mixing time.

Finally, we mention that our techniques imply asymptotically optimal bounds (up to constant factors) on both the standard and modified log-Sobolev constants of the Glauber dynamics for spin systems on bounded degree graphs in all of the regimes mentioned above. This also applies for certain problems where prior works have obtained rapid mixing via other techniques such as path coupling and canonical paths.

1.1 Result for General Spin Systems

Our main results will follow from a general statement regarding the Glauber dynamics for an arbitrary spin system satisfying marginal bounds and spectral independence. We first proceed with a few definitions.

Let $q \geq 2$ be an integer and $[q] = \{1, \ldots, q\}$. Given a graph $G = (V, E)$, we consider the q-spin system on G parameterized by a symmetric interaction matrix $A \in \mathbb{R}_{\geq 0}^{q \times q}$ representing "interaction strengths" and a field vector $h \in \mathbb{R}_{>0}^{q}$ representing "external fields". A configuration $\sigma \in [q]^V$ is an assignment of spins to vertices. The *Gibbs distribution* $\mu = \mu_{G,A,h}$ over all configurations is given by

$$\mu(\sigma) = \frac{1}{Z_G(A,h)} \prod_{\{u,v\} \in E} A(\sigma_u, \sigma_v) \prod_{v \in V} h(\sigma_v), \qquad \forall \sigma \in [q]^V$$

where

$$Z_G(A,h)=\sum_{\sigma\in[q]^V}\prod_{\{u,v\}\in E}A(\sigma_u,\sigma_v)\prod_{v\in V}h(\sigma_v)$$

is called the *partition function*. The hard-core model, Ising model, random colorings, and monomer-dimer model (equivalent to hard-core model on line graphs) all belong to the family of spin systems.

Let μ be an arbitrary distribution over $[q]^V$. A configuration $\sigma\in[q]^V$ is said to be *feasible* with respect to μ if $\mu(\sigma)>0$. Let $\Omega=\Omega(\mu)$ denote the collection of all feasible configurations (we omit μ when it is clear from the context); namely, Ω is the support of μ. Furthermore, for $\Lambda\subseteq V$ let $\Omega_\Lambda=\{\tau\in[q]^\Lambda:\mu_\Lambda(\tau)>0\}$ denote the collection of all feasible (partial) configurations on Λ, with the convention that $\Omega_v=\Omega_{\{v\}}$ for a single vertex v. Observe that $\Omega_V=\Omega$. For any subset $\Lambda\subseteq V$ and boundary condition $\tau\in\Omega_\Lambda$, we often consider the conditional distribution $\mu^\tau_S(\cdot)=\mu(\cdot|\sigma_\Lambda=\tau)$ over configurations on $S=V\setminus\Lambda$, and we shall write Ω^τ_U for the set of feasible (partial) configurations on $U\subseteq S$ under this conditional measure.

For a subset $S\subseteq V$, the Hamming graph $\mathcal{H}_S$ is defined to be the graph with vertex set $[q]^S$ of all configurations on S such that two configurations are adjacent iff they differ at exactly one vertex. A collection $\Omega_0\subseteq[q]^S$ of configurations on S is said to be connected if the induced subgraph $\mathcal{H}[\Omega_0]$ is connected. A distribution μ over $[q]^V$ is said to be *totally-connected* if for every nonempty subset $S\subseteq V$ and every boundary condition $\tau\in\Omega_{V\setminus S}$, the set Ω^τ_S is connected.

Assumption. Throughout the paper, we always assume that the distribution μ we are interested in is totally-connected.

We remark that all soft-constraint models (i.e., $A(i,j)>0$ for all $i,j\in[q]$) satisfy this assumption and common hard-constraint models, including the hardcore model, q-colorings when $q\geq\Delta+2$, and matchings, all satisfy this assumption as well.

The *Glauber dynamics*, also known as the *Gibbs sampling*, is a simple, natural, and popular Markov chain for sampling from a distribution μ over $[q]^V$. The dynamics starts with some (possibly random) configuration X_0. For every $t\geq1$, a new random configuration X_{t+1} is generated from X_t as follows: pick a coordinate $v\in V$ uniformly at random, set $X_{t+1}(u)=X_t(u)$ for all $u\in V\setminus\{v\}$, and sample $X_{t+1}(v)$ from the conditional distribution $\mu(\sigma_v=\cdot\mid\sigma_{V\setminus\{v\}}=X_t(V\setminus\{v\}))$. Denote the transition matrix of the Glauber dynamics by P_{GL}. If μ is totally-connected, then the Glauber dynamics is ergodic (i.e., irreducible and aperiodic) and has stationary distribution μ.

Let P be the transition matrix of an ergodic Markov chain (X_t) on a finite state space Ω with stationary distribution μ. For $t\geq0$ and $\sigma\in\Omega$, let $P^t(\sigma,\cdot)$ denote the distribution of X_t when starting the chain with $X_0=\sigma$. For $\varepsilon\in(0,1)$, the *mixing time* of P is defined as

$$T_{\mathrm{mix}}(P,\varepsilon)=\max_{\sigma\in\Omega}\min\left\{t\geq0:\left\|P^t(\sigma,\cdot)-\mu\right\|_{\mathrm{TV}}\leq\varepsilon\right\}. \tag{1}$$

We will require two conditions of the distribution μ. The first is that the marginal probability of each vertex is bounded away from 0 under any conditioning.

Definition 1.6 (Bounded Marginals)**.** We say a distribution μ over $[q]^V$ is *b-marginally bounded* if for every $\Lambda\subsetneq V$ and $\tau\in\Omega_\Lambda$, it holds for every $v\in V\setminus\Lambda$ and $i\in\Omega^\tau_v$ that,

$$\mu(\sigma_v=i\mid\sigma_\Lambda=\tau)\geq b.$$

The second condition is the notion of spectral independence, first given by [2] and later generalized to multi-spin systems in [14, 23]. Here we use the definitions from [14].

Definition 1.7 (Influence Matrix)**.** Given $\Lambda\subsetneq V$ and $\tau\in\Omega_\Lambda$, let

$$\widetilde{V}_\tau=\{(u,i):u\in V\setminus\Lambda,i\in\Omega^\tau_u\}.$$

For every $(u,i),(v,j)\in\widetilde{V}_\tau$ with $u\neq v$, we define the *(pairwise) influence* of (u,i) on (v,j) conditioned on τ by

$$\Psi^\tau_\mu\big((u,i),(v,j)\big)=\mu(\sigma_v=j\mid\sigma_u=i,\sigma_\Lambda=\tau)-\mu(\sigma_v=j\mid\sigma_\Lambda=\tau).$$

Furthermore, let $\Psi^\tau_\mu\big((v,i),(v,j)\big)=0$ for all $(v,i),(v,j)\in\widetilde{V}_\tau$. We call Ψ^τ_μ the *(pairwise) influence matrix* conditioned on τ.

Note that all eigenvalues of the influence matrix Ψ^τ_μ are real; see [2, 5, 14].

Definition 1.8 (Spectral Independence)**.** We say a distribution μ over $[q]^V$ is *η-spectrally independent* if for every $\Lambda\subsetneq V$ and $\tau\in\Omega_\Lambda$, the largest eigenvalue $\lambda_1(\Psi^\tau_\mu)$ of the influence matrix Ψ^τ_μ satisfies

$$\lambda_1(\Psi^\tau_\mu)\leq\eta.$$

The work of [23] defined another version of influence matrix by

$$\Psi^\tau_\mu(u,v)=\max_{i,j\in\Omega^\tau_u}\|\mu(\sigma_v=\cdot\mid\sigma_u=i,\sigma_\Lambda=\tau)-\mu(\sigma_v=\cdot\mid\sigma_u=j,\sigma_\Lambda=\tau)\|_{\mathrm{TV}},$$

and the spectral independence correspondingly. We remark that Definition 1.8 is weaker than the notion of spectral independence given in [23], and for all current applications as in [14, 23] or here in this paper, both definitions work.

Our main result is that if the Gibbs distribution on a bounded-degree graph is both marginally bounded and spectrally independent, then the Glauber dynamics satisfies the modified log-Sobolev inequality with constant $\Omega(1/n)$ (see Definition 3.3) and mixes in $O(n\log n)$ steps, where n is the number of vertices of the graph.

Theorem 1.9. *Let $\Delta\geq3$ be an integer and $b,\eta>0$ be reals. Suppose that $G=(V,E)$ is an n-vertex graph of maximum degree at most Δ and μ is a totally-connected Gibbs distribution of some spin system on G. If μ is both b-marginally bounded and η-spectrally independent, then the Glauber dynamics for sampling from μ satisfies the modified log-Sobolev inequality with constant $\frac{1}{C_1 n}$ where*

$$C_1=\left(\frac{\Delta}{b}\right)^{O\left(\frac{\eta}{b^2}+1\right)}.$$

Furthermore, the mixing time of the Glauber dynamics satisfies

$$T_{\mathrm{mix}}(P_{GL},\varepsilon)=\left(\frac{\Delta}{b}\right)^{O\left(\frac{\eta}{b^2}+1\right)}\times O\left(n\log\left(\frac{n}{\varepsilon}\right)\right).$$

Remark 2. More specifically, when $n\geq\frac{24\Delta}{b^2}(\frac{4\eta}{b^2}+1)$ we can choose

$$C_1=\frac{18\log(1/b)}{b^4}\left(\frac{24\Delta}{b^2}\right)^{\frac{4\eta}{b^2}+1},$$

and the mixing time is bounded by

$$T_{\mathrm{mix}}(P_{\mathrm{GL}}, \varepsilon) \leq \left\lceil \frac{18\log(1/b)}{b^4}\left(\frac{24\Delta}{b^2}\right)^{\frac{4\eta}{b^2}+1} n\left(\log n + \log\log\frac{1}{b} + \log\frac{1}{2\varepsilon^2}\right)\right\rceil.$$

Previous results [2, 14, 16, 23] could obtain $\mathrm{poly}(\Delta)\times n^{O(\eta)}$ mixing but without the assumption of marginal boundedness. In the setting of spin systems, we always have b-marginal boundedness with b depending only on the parameters A, h of the spin system and the maximum degree Δ of the graph, and so our results supersede those of [2, 14, 16, 23] in the bounded degree regime.

Remark 3. After the first version of this paper, the work [5] reformulates the proof of Theorem 1.9 without using simplicial complexes; in particular, the constant C_1 is brought down to $C_1 = (\Delta/b)^{O((\eta/b)+1)}$. The proof approach in this paper can also be modified to achieve the same bound, by considering Lemma 5.8 from the full version [15] of this paper specified to simplicial complexes corresponding to spin systems.

1.2 Result for General Simplicial Complexes

The recent work [2] studied spin systems, and more generally any distribution over $[q]^V$, in a novel way by viewing full and partial configurations as a high dimensional simplicial complex and utilizing tools such as high-dimensional expansion. Subsequent works [14, 16, 23] follow the same path as well. In this paper we also study spin systems in the framework of simplicial complexes. Moreover, we obtain new bounds on the mixing time and modified log-Sobolev constant of the global down-up and up-down walks for arbitrary pure weighted simplicial complexes. Before presenting our results, we first review some standard notation.

A *simplicial complex* $\mathfrak{X}$ is a collection of subsets (called faces) of a ground set $\mathcal{U}$ which is downwards closed; that is, if $\sigma \in \mathfrak{X}$ and $\tau \subseteq \sigma$ then $\tau \in \mathfrak{X}$. The dimension of a face is its size, and the dimension of $\mathfrak{X}$ is defined to be the maximum dimension of its faces. We say an n-dimensional simplicial complex $\mathfrak{X}$ is *pure* if every face is contained in a maximal face of size n. We write $\mathfrak{X}(k)$ for the collection of faces of size k. For a k-dimensional face $\tau \in \mathfrak{X}(k)$, we can define a pure $(n-k)$-dimensional simplicial subcomplex $\mathfrak{X}_\tau$ by taking $\mathfrak{X}_\tau = \{\xi \subseteq \mathcal{U} \setminus \tau : \tau \cup \xi \in \mathfrak{X}\}$.

For a pure n-dimensional simplicial complex $\mathfrak{X}$, consider a positive weight function $w : \mathfrak{X}(n) \to \mathbb{R}_{>0}$, which induces a distribution π_n on $\mathfrak{X}(n)$ with $\pi_n(\sigma) \propto w(\sigma)$. Furthermore, we can also define a distribution π_k over $\mathfrak{X}(k)$ for each nonnegative integer $k < n$ via the following process: sample σ from π_n, and select a uniformly random subset of size k. For $\tau \in \mathfrak{X}(k)$, the weight function w induces the weights for the simplicial subcomplex $\mathfrak{X}_\tau$ by $w_\tau(\xi) = w(\tau \cup \xi)$ for each $\xi \in \mathfrak{X}_\tau(n-k)$. The distribution $\pi_{\tau,j}$ is also defined accordingly for each nonnegative integer $j \leq n-k$.

As noticed in [2], there is a natural way to represent every distribution μ over $[q]^V$ with $|V| = n$ as a pure n-dimensional weighted simplicial complex $(\mathfrak{X} = \mathfrak{X}^\Omega, \mu)$, which is defined as follows. The ground set of $\mathfrak{X}$ consists of pairs

$$\widetilde{V} = \{(v, i) : v \in V, i \in \Omega_v\}.$$

The maximal faces of $\mathfrak{X}$ consist of collections of n pairs forming a valid configuration $\sigma \in \Omega$; i.e., every configuration $\sigma \in \Omega$ corresponds to a maximal face $\{(v, \sigma_v) : v \in V\}$. The rest of $\mathfrak{X}$ is generated by taking downwards closure so that $\mathfrak{X}$ is pure by construction. Namely, every $U \subseteq V$ and $\tau \in \Omega_U$ corresponds to a face $\{(v, \tau_v) : v \in U\}$; we shall denote it by (U, τ) for simplicity. Note that the faces of intermediate dimension can be thought of as partial configurations. Now, if there is a weight function $w : \Omega \to \mathbb{R}_{>0}$ associated with μ such that $\mu(\sigma) \propto w(\sigma)$ for each $\sigma \in \Omega$, then it also gives a weight function $w : \mathfrak{X}(n) \to \mathbb{R}_{>0}$ by the one-to-one correspondence between Ω and $\mathfrak{X}(n)$, and thus induces the associated distribution π_n on $\mathfrak{X}(n)$. Observe that π_n is exactly the distribution μ. Moreover, for each $k < n$, the distribution π_k on $\mathfrak{X}(k)$ is given by

$$\pi_k(U, \tau) = \frac{1}{\binom{n}{k}}\mu(\sigma_U = \tau)$$

for every $U \subseteq V$ and $\tau \in \Omega_U$.

For simplicial complexes, the global down-up and up-down walks between faces of distinct dimensions have attracted a lot of attention in recent years [1, 3, 17, 19, 35, 37, 45]. For integers $0 \leq r < s \leq n$, define the *order-(s, r) (global) down-up* walk with transition matrix denoted by $P^{\vee}_{s,r}$ to be the following random walk over $\mathfrak{X}(s)$: in each step we remove $s-r$ elements, chosen uniformly at random, from the current face $\sigma_t \in \mathfrak{X}(s)$ to obtain a face $\tau_t \in \mathfrak{X}(r)$, and then pick $\xi_{t+1} \in \mathfrak{X}_{\tau_t}(s-r)$ from the distribution $\pi_{\tau_t, s-r}$ and set $\sigma_{t+1} = \tau_t \cup \xi_{t+1}$. The stationary distribution of $P^{\vee}_{s,r}$ is π_s. In particular, observe that the Glauber dynamics for a distribution μ over $[q]^V$ is the same as the order-$(n, n-1)$ down-up walk for the weighted simplicial complex $(\mathfrak{X}, \mu)$. Similarly, the *order-(r, s) (global) up-down* walk with transition matrix $P^{\wedge}_{r,s}$ is a random walk over $\mathfrak{X}(r)$ with stationary distribution π_r: given the current face $\tau_t \in \mathfrak{X}(r)$, sample $\xi_{t+1} \in \mathfrak{X}_{\tau_t}(s-r)$ from $\pi_{\tau_t, s-r}$, set $\sigma_{t+1} = \tau_t \cup \xi_{t+1}$, and finally remove $s-r$ elements from σ_{t+1} uniformly at random to obtain $\tau_{t+1} \in \mathfrak{X}(r)$.

We establish the modified log-Sobolev inequality and give meaningful bounds on the mixing time for the down-up and up-down walks for arbitrary weighted simplicial complexes. Our proof utilizes the local-to-global scheme as in [1] and establishes contraction of entropy extending the result of [17]. Before stating our main result, we first give the definitions of marginal boundedness and local spectral expansion for simplicial complexes. As we shall see from Claims 1.11 and 1.13 below, our requirements of marginal boundedness and spectral independence in Theorem 1.9 is translated from the corresponding conditions needed for simplicial complexes.

Definition 1.10 (Bounded Marginal)**.** We say a pure n-dimensional weighted simplicial complex $(\mathfrak{X}, w)$ is $(b_0, \ldots, b_{n-1})$-*marginally bounded* if for all $0 \leq k \leq n-1$, every $\tau \in \mathfrak{X}(k)$, and every $i \in \mathfrak{X}_\tau(1)$, we have

$$\pi_{\tau,1}(i) \geq b_k.$$

Claim 1.11. *If a distribution μ over $[q]^V$ is b-marginally bounded, then the weighted simplicial complex $(\mathfrak{X}, \mu)$ for μ is $(b_0, \ldots, b_{n-1})$-marginally bounded with $b_k = \frac{b}{n-k}$ for each k.*

The proof of Claim 1.11 can be found in the full version [15] of this paper.

The global walks in simplicial complexes can be studied by decomposition into local walks which we define now. For every $0 \le k \le n-2$ and every face $\tau \in \mathfrak{X}(k)$, the *local walk* at τ with transition matrix P_τ is the following random walk over $\mathfrak{X}_\tau(1)$: given the current element $i \in \mathfrak{X}_\tau(1)$, the next element is generated from the distribution $\pi_{\tau\cup\{i\},1}$. One can relate mixing properties of the local walks to the mixing properties of the global walks; see [1, 3, 17, 37]. In nearly all prior works, such a relation was quantified using the spectral gap of the walks. Like in [17], while our ultimate goal is to show the modified log-Sobolev inequality of the global walks, we will still need the notion of local spectral expansion for local walks. Let us now capture this idea using the following definition, taking after [1, 19, 35–37, 45].

Definition 1.12 (Local Spectral Expansion [1])**.** We say a pure n-dimensional weighted simplicial complex $(\mathfrak{X}, w)$ is a $(\zeta_0, \dots, \zeta_{n-2})$-*local spectral expander* if for every $0 \le k \le n-2$ and every $\tau \in \mathfrak{X}(k)$, we have

$$\lambda_2(P_\tau) \le \zeta_k.$$

Claim 1.13. *If a distribution μ over $[q]^V$ is η-spectrally independent, then the weighted simplicial complex $(\mathfrak{X}, \mu)$ is a $(\zeta_0, \dots, \zeta_{n-2})$-local spectral expander with $\zeta_k = \frac{\eta}{n-k-1}$ for each k.*

Proof. This is Theorem 8 from [14]. □

We then show that for any pure weighted simplicial complexes, the modified log-Sobolev inequality (see Definition 3.3) holds for down-up and up-down walks if the marginal probabilities of the simplicial complex are bounded away from zero and all local walks have good expansion properties. This also bounds the mixing times of these random walks.

Theorem 1.14. *Let $(\mathfrak{X}, w)$ be a pure n-dimensional weighted simplicial complex. If $(\mathfrak{X}, w)$ is $(b_0, \dots, b_{n-1})$-marginally bounded and has $(\zeta_0, \dots, \zeta_{n-2})$-local spectral expansion, then for every $0 \le r < s \le n$, both the order-(s, r) down-up walk and the order-(r, s) up-down walk satisfy the modified log-Sobolev inequality with constant $\kappa = \kappa(r, s)$ defined as*

$$\kappa = \frac{\sum_{k=r}^{s-1} \Gamma_k}{\sum_{k=0}^{s-1} \Gamma_k}$$

where: $\Gamma_0 = 1$; for $1 \le k \le s-1$, $\Gamma_k = \prod_{j=0}^{k-1} \alpha_j$; and for $0 \le k \le s-2$,

$$\alpha_k = \max\left\{1 - \frac{4\zeta_k}{b_k^2(s-k)^2}, \frac{1-\zeta_k}{4 + 2\log(\frac{1}{2b_k b_{k+1}})}\right\}.$$

Furthermore, the mixing time of the order-(s, r) down-up walk is bounded by

$$T_{\mathrm{mix}}(P^\vee_{s,r}, \varepsilon) \le \left\lceil \frac{1}{\kappa}\left(\log\log\frac{1}{\pi^*_s} + \log\frac{1}{2\varepsilon^2}\right)\right\rceil \tag{2}$$

where $\pi^*_s = \min_{\sigma\in\mathfrak{X}(s)} \pi_s(\sigma)$. The mixing time of the order-(r, s) up-down walk is also bounded by Eq. (2) with π^*_s replaced by π^*_r.

Theorem 1.14 generalizes both the result of [17] for simplicial complexes with respect to strongly log-concave distributions and the result of [1] for the Poincaré inequality (i.e., bounding the spectral gap). It in some sense answers a question of [17] on local-to-global modified log-Sobolev inequalities in high-dimensional expanders, at least in the bounded marginals setting.

Even though Theorem 1.14 can give a bound on the mixing time of the Glauber dynamics, which is the order-$(n, n-1)$ down-up walk in the corresponding weighted simplicial complex, our main result Theorem 1.9 does *not* follow directly from Theorem 1.14. In fact, we will only consider the order-$(n, n-\ell)$ down-up walk for $\ell = \Theta(n)$, which corresponds to the heat-bath block dynamics that updates a uniformly random subset of ℓ vertices in every step. One of our main technical contributions is to compare this block dynamics with the single-site Glauber dynamics; we shall detail this in Section 2.1 below. Nevertheless, we find Theorem 1.14 interesting of its own and possible for future applications in other problems.

2 PROOF OUTLINE

In this section, we outline our proofs of Theorems 1.9 and 1.14.

2.1 Approximate Tensorization and Uniform Block Factorization

One way of establishing rapid mixing of the Glauber dynamics is to show that the Gibbs distribution satisfies the *approximate tensorization of entropy*. This approach has been (implicitly) used in many literature to establish the log-Sobolev inequalities, from which one can deduce an optimal bound on the mixing time. Before giving the formal definition, we first review some standard definitions.

Consider a distribution μ supported on $\Omega \subseteq [q]^V$. For every $f : \Omega \to \mathbb{R}_{\ge 0}$, we denote the expectation of f under μ by $\mu(f) = \sum_{\sigma\in\Omega} \mu(\sigma) f(\sigma)$ and the entropy of f by $\mathrm{Ent}_\mu(f) = \mu(f \log \frac{f}{\mu(f)})$. We often simply write $\mathrm{Ent}(f)$ for the entropy and drop the subscript μ when it is clear from the context. More generally, given $S \subseteq V$ and $\tau \in \Omega_{V\setminus S}$, for every $f : \Omega^\tau_S \to \mathbb{R}_{\ge 0}$ we use $\mu^\tau_S(f)$ to denote the expectation of f under the conditional distribution μ^τ_S and $\mathrm{Ent}^\tau_S(f)$ for the corresponding entropy. For most of the time we are actually given a function $f : \Omega \to \mathbb{R}_{\ge 0}$, and we will still write $\mu^\tau_S(f)$ and $\mathrm{Ent}^\tau_S(f)$ where we think of f as restricted to the space Ω^τ_S and implicitly assume that the configuration outside S is given by τ; i.e., for an argument $\sigma \in \Omega^\tau_S$ the value of f is $f(\sigma \cup \tau)$. It is helpful to think of $\mu^\tau_S(f)$ and $\mathrm{Ent}^\tau_S(f)$ as a function of the boundary condition τ. In this sense, the notation $\mu[\mathrm{Ent}_S(f)]$, for example, represents the expectation of the function $\mathrm{Ent}^\tau_S(f)$ where $\tau \in \Omega_{V\setminus S}$ is distributed as the marginal of μ on $V \setminus S$.

The notion of approximate tensorization of entropy is formally defined as follows.

Definition 2.1 (Approximate Tensorization)**.** We say that a distribution μ over $[q]^V$ satisfies the *approximate tensorization of entropy* (with constant C_1) if for all $f : \Omega \to \mathbb{R}_{\ge 0}$ we have

$$\mathrm{Ent}(f) \le C_1 \sum_{v\in V} \mu[\mathrm{Ent}_v(f)]. \tag{3}$$

Approximate tensorization can be understood as closeness of μ to a product distribution, or weak dependency of variables. In fact, if μ is exactly a product distribution (e.g., the Gibbs distribution on an empty graph), then approximate tensorization holds with constant $C_1 = 1$; e.g., see [10, 12]. If μ satisfies approximate tensorization with a constant C_1 independent of n, then the Glauber dynamics for sampling from μ mixes in $O(n \log n)$ steps. In fact, given approximate tensorization, one can deduce tight bounds on all of the following quantities: the spectral gap, both standard and modified

log-Sobolev constants, relative entropy decay rate, mixing time, and concentration bounds. See Fact 3.5 for a detailed summary.

In many cases, especially on the integer lattice $\mathbb{Z}^d$, log-Sobolev inequalities for the Glauber dynamics are established through the approximate tensorization of entropy, which is more intuitive and easier to handle; e.g., see [11, 12, 27, 40]. Despite the success on $\mathbb{Z}^d$, there is not much study for spin systems on bounded-degree graphs. The works of [10, 42] considered approximate tensorization for general discrete product spaces, and gave sufficient conditions to derive it; however, for spin systems these results do not cover the whole uniqueness region.

One can regard approximate tensorization of entropy as factorizing entropy into all single vertices. Motivated by tools from high dimensional simplicial complexes [1, 2] and study on general block factorization of entropy [11], we consider in this paper a more general notion of entropy factorization, where the entropy is factorized into subsets of vertices of a fixed size. The formal definition is given as follows.

Definition 2.2 (Uniform Block Factorization)**.** We say that a distribution μ over $[q]^V$ satisfies the *ℓ-uniform block factorization of entropy* (with constant C) if for all $f : \Omega \to \mathbb{R}_{\geq 0}$ we have

$$\frac{\ell}{n}\operatorname{Ent}(f) \leq C \cdot \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \mu[\operatorname{Ent}_S(f)]. \tag{4}$$

We remark that uniform block factorization of entropy is a special case of block factorization given by equation (1.3) in [11]; there, the entropy factorizes into arbitrary blocks with arbitrary weights. Also observe that 1-uniform block factorization is the same as approximate tensorization of entropy. Just as the approximate tensorization corresponds to the single-site Glauber dynamics, the ℓ-uniform block factorization corresponds to the heat-bath block dynamics where in each step a subset of vertices of size ℓ is chosen uniformly at random and gets updated. Moreover, similar results as in Fact 3.5 can be deduced for this block dynamics.

Our first key result is a reduction from approximate tensorization to uniform block factorization. For b-marginally bounded Gibbs distributions on graphs with maximum degree $\leq \Delta$, we show that approximate tensorization is implied by ℓ-uniform block factorization for $\ell = \lceil \theta n \rceil$ and an appropriate constant θ depending on b and Δ. This is given by the following lemma.

Lemma 2.3. *Let $\Delta \geq 3$ be an integer and $b > 0$ be a real. Consider the Gibbs distribution μ on an n-vertex graph G of maximum degree at most Δ and assume that μ is b-marginally bounded. Suppose there exist positive reals $\theta \leq \frac{b^2}{12\Delta}$ and C such that μ satisfies the $\lceil \theta n \rceil$-uniform block factorization of entropy with constant C. Then μ satisfies the approximate tensorization of entropy with constant*

$$C_1 = \frac{18\log(1/b)}{b^4}\, C.$$

Remark 4. The notion of approximate tensorization and uniform block factorization with respect to variance is also meaningful. In fact, for variance these definitions are equivalent to bounding the spectral gap of the corresponding chains. Moreover, Lemma 2.3 holds for variance as well, which can already provide a tight bound on the spectral gap of the Glauber dynamics combining results from [2, 14, 16, 23]. See the full version [15] of this paper for more details.

2.2 Simplicial Complexes and Entropy Contraction

Our next goal is to establish ℓ-uniform block factorization of entropy for $\ell = \Theta(n)$, which relies on the spectral independence property. The following lemma holds for all distributions over $[q]^V$, not only Gibbs distributions.

Lemma 2.4. *Let $b, \eta > 0$ be reals. Then for every real $\theta \in (0,1)$ and every integer $n \geq \frac{2}{\theta}(\frac{4\eta}{b^2} + 1)$ the following holds.*

Let V be a set of size n and μ be a distribution over $[q]^V$. If μ is both b-marginally bounded and η-spectrally independent, then μ satisfies $\lceil \theta n \rceil$-uniform block factorization of entropy with constant

$$C = \left(\frac{2}{\theta}\right)^{\frac{4\eta}{b^2}+1}.$$

Recall that there is a natural correspondence between a distribution μ over $[q]^V$ and the weighted simplicial complex $(\mathfrak{X}, \mu)$. For general weighted simplicial complexes, one property studied in [17] is how the entropy of a function defined on faces contracts when it projects down from higher dimensions to lower. This can be captured by the definition below. For a pure n-dimensional weighted simplicial complex $(\mathfrak{X}, w)$ and a nonnegative integer $k < n$, let $P_k^{\uparrow}$ denote the $|\mathfrak{X}(k)| \times |\mathfrak{X}(k+1)|$ dimensional transition matrix corresponding to adding a random element $i \notin \tau$ to some $\tau \in \mathfrak{X}(k)$ where i is distributed as $\pi_{\tau,1}$. Also for any $0 \leq r < s \leq n$ and any function $f^{(s)} : \mathfrak{X}(s) \to \mathbb{R}_{\geq 0}$, define $f^{(r)} : \mathfrak{X}(r) \to \mathbb{R}_{\geq 0}$ by $f^{(r)} = P_r^{\uparrow} \cdots P_{s-1}^{\uparrow} f^{(s)}$.

Definition 2.5 (Global Entropy Contraction)**.** We say a pure n-dimensional weighted simplicial complex $(\mathfrak{X}, w)$ satisfies the *order-(r,s) global entropy contraction* with rate $\kappa = \kappa(r,s)$ if for all $f^{(s)} : \mathfrak{X}(s) \to \mathbb{R}_{\geq 0}$ we have

$$\operatorname{Ent}_{\pi_r}(f^{(r)}) \leq (1-\kappa)\operatorname{Ent}_{\pi_s}(f^{(s)}).$$

It turns out, as a remarkable fact, that uniform block factorization of entropy for a distribution μ over $[q]^V$ is *equivalent* to global entropy contraction for the weighted simplicial complex $(\mathfrak{X}, \mu)$.

Lemma 2.6. *A distribution μ over $[q]^V$ satisfies the ℓ-uniform block factorization of entropy with some constant C if and only if the corresponding weighted simplicial complex $(\mathfrak{X}, \mu)$ satisfies order-$(n-\ell, n)$ global entropy contraction with rate κ, where $C\kappa = \ell/n$.*

The proof of Lemma 2.6 can be found in the full version [15]. As a consequence, to prove Lemma 2.4, it suffices to establish global entropy contraction for the weighted simplicial complex $(\mathfrak{X}, \mu)$.

Just like approximate tensorization and uniform block factorization having many implications for the corresponding single-site and block dynamics (e.g., see Fact 3.5), the notion of global entropy contraction can provide for weighted simplicial complexes meaningful bounds on the spectral gap, modified log-Sobolev constant, relative entropy decay rate, mixing time, and concentration bounds; see the full version [15] for details. In Lemma 11 of [17], the authors established order-(r,s) global entropy contraction with rate $\kappa = \frac{s-r}{s}$ for simplicial complexes with respect to homogeneous strongly

log-concave distributions. From this, they deduced the modified log-Sobolev inequality for the down-up and up-down walks and showed rapid mixing of it.

We then show that for an arbitrary weighted simplicial complex $(\mathfrak{X}, w)$, one can deduce global entropy contraction from local spectral expansion whenever the marginals of the induced distributions are nicely bounded. For this, we prove a local-to-global result for entropy contraction in the spirit of [1]. If we additionally know that the marginals are nicely bounded, we can further reduce the local entropy contraction to local spectral expansion.

Lemma 2.7. *Let $(\mathfrak{X}, w)$ be a pure n-dimensional weighted simplicial complex. Suppose that $(\mathfrak{X}, w)$ is $(b_0, \dots, b_{n-1})$-marginally bounded and has $(\zeta_0, \dots, \zeta_{n-2})$-local spectral expansion. Then for all $0 \le r < s \le n$, $(\mathfrak{X}, w)$ satisfies order-(r, s) global entropy contraction with rate $\kappa = \kappa(r, s)$ given as in Theorem 1.14.*

Theorem 1.14 follows immediately from Lemma 2.7 and [15, Fact 5.2]. We remark that Lemma 2.7 recovers Lemma 11 of [17] for simplicial complexes corresponding to discrete log-concave distributions, since there one has $\zeta_k = 0$ for all k as shown in [3].

We present next the proof of Lemma 2.4, which follows directly from Lemmas 2.6 and 2.7.

Proof of Lemma 2.4. From Claims 1.11 and 1.13 we know that the weighted simplicial complex $(\mathfrak{X}, \mu)$ regarding μ is $(b_0, \dots, b_{n-1})$-marginally bounded with $b_k = \frac{b}{n-k}$ and has $(\zeta_0, \dots, \zeta_{n-2})$-local spectral expansion with $\zeta_k = \frac{\eta}{n-k-1}$. Then, Lemma 2.7 implies that $(\mathfrak{X}, \mu)$ satisfies order-$(n-\ell, n)$ global entropy contraction for $\ell = \lceil \theta n \rceil$ with rate

$$\kappa = \frac{\sum_{k=n-\ell}^{n-1} \Gamma_k}{\sum_{k=0}^{n-1} \Gamma_k}$$

where $\Gamma_0 = 1$, $\Gamma_k = \prod_{j=0}^{k-1} \alpha_j$, and

$$\alpha_k = \max\left\{1 - \frac{4\eta}{b^2(n-k-1)}, \frac{1-\eta/(n-k-1)}{4 + 2\log((n-k)(n-k-1)/(2b^2))}\right\}.$$

Define an integer $R = \left\lceil \frac{4\eta}{b^2} \right\rceil$ and observe that $n \ge \ell \ge \theta n \ge 2R$ by our assumption. Thus, we have

$$\alpha_k \ge \hat{\alpha}_k := \max\left\{1 - \frac{R}{n-k-1}, 0\right\}.$$

Notice that κ, when viewed as a function of α_k's, is monotone increasing with each α_k. Thus, to lower bound κ, we can plug in the lower bounds $\hat{\alpha}_k$'s and get

$$\kappa \ge \frac{\sum_{k=n-\ell}^{n-1} \hat{\Gamma}_k}{\sum_{k=0}^{n-1} \hat{\Gamma}_k}$$

where $\hat{\Gamma}_0 = 1$ and $\hat{\Gamma}_k = \prod_{j=0}^{k-1} \hat{\alpha}_j$ for each $k \ge 1$. We will show that for every $0 \le k \le n-1$ one actually has

$$\hat{\Gamma}_k = \frac{(n-k-1)(n-k-2)\cdots(n-k-R)}{(n-1)(n-2)\cdots(n-R)}. \tag{5}$$

For $k = 0$ we have $\hat{\Gamma}_0 = 1$ and Eq. (5) holds. For $1 \le j \le n-R-2$ we have

$$\hat{\alpha}_j = \max\left\{\frac{n-j-1-R}{n-j-1}, 0\right\} = \frac{n-j-1-R}{n-j-1}$$

and thus for $1 \le k \le n-R-1$

$$\hat{\Gamma}_k = \prod_{j=0}^{k-1} \frac{n-j-1-R}{n-j-1} = \frac{(n-k-1)(n-k-2)\cdots(n-k-R)}{(n-1)(n-2)\cdots(n-R)}.$$

Finally, since $\hat{\alpha}_j = 0$ when $n-R-1 \le j \le n-2$, we have $\hat{\Gamma}_k = 0$ for $n-R \le k \le n-1$. Therefore, Eq. (5) is true for all k. It then follows that

$$\kappa \ge \frac{\sum_{k=n-\ell}^{n-1}(n-k-1)(n-k-2)\cdots(n-k-R)}{\sum_{k=0}^{n-1}(n-k-1)(n-k-2)\cdots(n-k-R)} = \frac{\sum_{j=0}^{\ell-1} j(j-1)\cdots(j-R+1)}{\sum_{j=0}^{n-1} j(j-1)\cdots(j-R+1)}.$$

The following is a standard equality which can be proved by induction:

$$\sum_{j=0}^{N-1} j(j-1)\cdots(j-R+1) = \frac{1}{R+1} N(N-1)\cdots(N-R).$$

Hence, we obtain

$$\kappa \ge \frac{\ell(\ell-1)\cdots(\ell-R)}{n(n-1)\cdots(n-R)}.$$

Finally, we deduce from Lemma 2.6 that

$$C \le \frac{\ell}{n} \cdot \frac{1}{\kappa} \le \frac{(n-1)\cdots(n-R)}{(\ell-1)\cdots(\ell-R)} \le \left(\frac{n-R}{\ell-R}\right)^R \le \left(\frac{2n}{\ell}\right)^R \le \left(\frac{2}{\theta}\right)^{\frac{4\eta}{b^2}+1}$$

where we use our assumption $\ell \ge \theta n \ge 2R$. □

2.3 Wrapping up

Combining Lemmas 2.3 and 2.4, we establish approximate tensorization of entropy with a constant independent of n, when the Gibbs distribution is marginally bounded and spectrally independent. This is stated in the following theorem.

Theorem 2.8. *Let $\Delta \ge 3$ be an integer and $b, \eta > 0$ be reals. Suppose that $G = (V, E)$ is an n-vertex graph of maximum degree at most Δ and μ is a totally-connected Gibbs distribution of some spin system on G. If μ is both b-marginally bounded and η-spectrally independent and $n \ge \frac{24\Delta}{b^2}(\frac{4\eta}{b^2} + 1)$, then μ satisfies the approximate tensorization of entropy with constant*

$$C_1 = \frac{18\log(1/b)}{b^4}\left(\frac{24\Delta}{b^2}\right)^{\frac{4\eta}{b^2}+1}.$$

Theorem 1.9 then follows immediately from Theorem 2.8 and Fact 3.5.

Our main results Theorems 1.1 to 1.5 will follow from Theorem 1.9 by establishing marginal boundedness and spectral independence for each model. The detailed proofs are contained in Section 5, we include here a brief sketch. The marginal boundedness is a trivial bound. The spectral independence was previously established for antiferromagnetic 2-spin systems including the hard-core model and the Ising model in the whole uniqueness region [2, 16], and for random q-colorings when q is sufficiently large [14, 23]. For the monomer-dimer model, spectral independence is not known previously. Following the proof strategy of [16] and utilizing the two-step recursion from [4], we show the following.

Theorem 2.9. *Let $\Delta \geq 3$ be an integer and $\lambda > 0$ be a real. Then for every graph $G = (V, E)$ of maximum degree at most Δ with $m = |E|$, the Gibbs distribution μ of the monomer-dimer model on G with fugacity λ is η-spectrally independent for $\eta = \min\left\{2\lambda\Delta, 2\sqrt{1+\lambda\Delta}\right\}$.*

The rest of the paper is organized as follows. In Section 3, we collect relevant preliminaries. In Section 4, we show how to reduce approximate tensorization to uniform block factorization with linear-sized blocks; specifically, we prove Lemma 2.3. In the full version [15] of this paper, we reduce uniform block factorization and, more generally, global entropy contraction in the setting of weighted simplicial complexes to local entropy contraction; we then further reduce local entropy contraction to local spectral expansion when the simplicial complexes have bounded marginals and thus prove Lemma 2.7. We also bound the spectral independence of the monomer-dimer model on bounded degree graphs and prove Theorem 2.9 in the full version [15]. Finally, we finish off the proofs of our main mixing time results in Section 5 and conclude with some open problems in Section 6. We also discuss analogous results for variance in the full version [15].

3 PRELIMINARIES

In this section we review some standard definitions.

In the following definition, we assume the underlying distribution μ is fixed and omit it from the subscript.

Definition 3.1. Let Ω be a finite set and μ be a distribution over Ω. For all functions $f, g : \Omega \to \mathbb{R}$:

(a) The *expectation* of f is defined as
$$\mu(f) = \sum_{x\in\Omega} \mu(x) f(x);$$

(b) The *variance* of f is defined as
$$\mathrm{Var}(f) = \mu[(f - \mu(f))^2] = \mu(f^2) - \mu(f)^2;$$

(c) The *covariance* of f and g is defined as
$$\mathrm{Cov}(f, g) = \mu[(f - \mu(f))(g - \mu(g))] = \mu(fg) - \mu(f)\mu(g);$$

(d) If $f \geq 0$, the *entropy* of f is defined as
$$\mathrm{Ent}(f) = \mu\left[f \log\left(\frac{f}{\mu(f)}\right)\right] = \mu(f \log f) - \mu(f)\log\mu(f)$$
with the convention that $0 \log 0 = 0$.

For two distributions μ, ν over a finite set Ω, the Kullback–Leibler divergence (KL divergence), also called relative entropy, is defined as
$$D_{\mathrm{KL}}(\nu \,\|\, \mu) = \sum_{x\in\Omega} \nu(x) \log\left(\frac{\nu(x)}{\mu(x)}\right).$$
Let $f = \nu/\mu$ be the relative density of ν with respect to μ; i.e., $f(x) = \nu(x)/\mu(x)$ for all $x \in \Omega$. Then $\mathrm{Ent}(f) = D_{\mathrm{KL}}(\nu \,\|\, \mu)$. The following is a well-known fact; see, e.g., [21].

Fact 3.2 (Donsker-Varadhan's Variational Representation)**.** *For two distributions μ, ν over a finite set Ω, the KL divergence admits the following variational formula:*
$$D_{\mathrm{KL}}(\nu \,\|\, \mu) = \sup_{f:\Omega\to\mathbb{R}} \left\{\nu(f) - \log\mu(e^f)\right\}.$$

We then review some standard functional inequalities, and refer to [7, 43] for more backgrounds.

Definition 3.3. Let Ω be a finite set and μ be a distribution over Ω. Let P denote the transition matrix of an ergodic, reversible Markov chain on Ω with stationary distribution μ.

(a) The Dirichlet form of P is defined as for every $f, g : \Omega \to \mathbb{R}$,
$$\mathcal{E}_P(f, g) = \frac{1}{2}\sum_{x,y\in\Omega} \mu(x)P(x, y)(f(x) - f(y))(g(x) - g(y)).$$
In particular, if $\Omega \subseteq [q]^V$ and $P = P_{\mathrm{GL}}$ is the Glauber dynamics for μ, then we can write
$$\mathcal{E}_{P_{\mathrm{GL}}}(f, g) = \frac{1}{n}\sum_{v\in V} \mu[\mathrm{Cov}_v(f, g)].$$

(b) We say the *Poincaré inequality* holds with constant λ if for every $f : \Omega \to \mathbb{R}$,
$$\lambda\,\mathrm{Var}(f) \leq \mathcal{E}_P(f, f).$$
The spectral gap of P is
$$\lambda(P) = \inf\left\{\frac{\mathcal{E}_P(f, f)}{\mathrm{Var}(f)} \;\middle|\; f : \Omega \to \mathbb{R}, \mathrm{Var}(f) \neq 0\right\}.$$

(c) We say the *standard log-Sobolev inequality* holds with constant ρ if for every $f : \Omega \to \mathbb{R}_{\geq 0}$,
$$\rho\,\mathrm{Ent}(f) \leq \mathcal{E}_P(\sqrt{f}, \sqrt{f}).$$
The standard log-Sobolev constant of P is
$$\rho(P) = \inf\left\{\frac{\mathcal{E}_P(\sqrt{f}, \sqrt{f})}{\mathrm{Ent}(f)} \middle| f : \Omega \to \mathbb{R}_{\geq 0}, \mathrm{Ent}(f) \neq 0\right\}.$$

(d) We say the *modified log-Sobolev inequality* holds with constant ρ_0 if for every $f : \Omega \to \mathbb{R}_{\geq 0}$,
$$\rho_0\,\mathrm{Ent}(f) \leq \mathcal{E}_P(f, \log f).$$
The modified log-Sobolev constant of P is
$$\rho_0(P) = \inf\left\{\frac{\mathcal{E}_P(f, \log f)}{\mathrm{Ent}(f)} \;\middle|\; f : \Omega \to \mathbb{R}_{\geq 0}, \mathrm{Ent}(f) \neq 0\right\}.$$

(e) We say *the relative entropy decays* with rate α if for every distribution ν over Ω,
$$D_{\mathrm{KL}}(\nu P \,\|\, \mu) \leq (1 - \alpha)\, D_{\mathrm{KL}}(\nu \,\|\, \mu).$$

Next, we consider the case that $\Omega \subseteq [q]^V$ for a finite set V. Let $S \subseteq V$ and $\tau \in \Omega_{V\setminus S}$. Recall that for every function $f : \Omega \to \mathbb{R}_{\geq 0}$, we write $\mu_S^\tau(f)$ and $\mathrm{Ent}_S^\tau(f) = \mathrm{Ent}_{\mu_S^\tau}(f)$ for the expectation and entropy of f under the conditional distribution $\mu_S^\tau(\cdot) = \mu(\sigma_S = \cdot \mid \sigma_{V\setminus S} = \tau)$, where $f = f_\tau$ is understood as a function of the configuration on S with τ fixed outside S. We think of $\mu_S^\tau(f)$ and $\mathrm{Ent}_S^\tau(f)$ as functions of τ, and we will use, for example, $\mathrm{Ent}[\mu_S(f)]$ to represent the entropy of $\mu_S^\tau(f)$ under μ, and $\mu[\mathrm{Ent}_S(f)]$ for the expectation of $\mathrm{Ent}_S^\tau(f)$. We give below a useful property of the expectation and entropy; see, e.g., [41] for proofs.

Fact 3.4. *Let $S \subseteq V$ and $\tau \in \Omega_{V\setminus S}$. For every function $f : \Omega \to \mathbb{R}_{\geq 0}$, we have*
$$\mu(f) = \mu[\mu_S(f)] \qquad \text{and} \qquad \mathrm{Ent}(f) = \mu[\mathrm{Ent}_S(f)] + \mathrm{Ent}[\mu_S(f)].$$

Implications of Approximate Tensorization. We summarize here a few corollaries of approximate tensorization of entropy for arbitrary distributions over discrete product spaces.

Fact 3.5. *Let V be a set of size n and μ be a distribution over $[q]^V$. If μ satisfies the approximate tensorization of entropy with constant C_1, then the Glauber dynamics for μ satisfies all of the following:*

(1) The Poincaré inequality holds with constant $\lambda = \frac{1}{C_1 n}$;
(2) The modified log-Sobolev inequality holds with constant $\rho_0 = \frac{1}{C_1 n}$;
(3) The relative entropy decays with rate $\alpha = \frac{1}{C_1 n}$;
(4) The mixing time of the Glauber dynamics satisfies

$$T_{\mathrm{mix}}(P_{GL}, \varepsilon) \le \left\lceil C_1 n \left(\log\log \frac{1}{\mu_{\min}} + \log \frac{1}{2\varepsilon^2} \right) \right\rceil$$

where $\mu_{\min} = \min_{\sigma \in \Omega} \mu(\sigma)$; If furthermore μ is b-marginally bounded, then we have $\mu_{\min} \ge b^n$ and thus

$$T_{\mathrm{mix}}(P_{GL}, \varepsilon) \le \left\lceil C_1 n \left(\log n + \log\log \frac{1}{b} + \log \frac{1}{2\varepsilon^2} \right) \right\rceil;$$

(5) For every $f : \Omega \to \mathbb{R}$ which is c-Lipschitz with respect to the Hamming distance on $[q]^V$ and every $a \ge 0$, we have the concentration inequality

$$\Pr_{\sigma \sim \mu} [|f(\sigma) - \mu(f)| \ge a] \le 2 \exp\left(-\frac{a^2}{2c^2 C_1 n} \right);$$

(6) If furthermore μ is b-marginally bounded, then the standard log-Sobolev inequality holds with constant $\rho = \frac{1-2b}{\log(1/b-1)} \cdot \frac{1}{C_1 n}$ when $b < \frac{1}{2}$, or $\rho = \frac{1}{2C_1 n}$ when $b = \frac{1}{2}$.

The ℓ-uniform block factorization of entropy implies similar results for the heat-bath block dynamics that updates a random subset of vertices of size ℓ in each step.

The implications in Fact 3.5 are all known and have been widely used, often implicitly. In the proof below, we give references where explicit statements or direct proofs are available.

Proof of Fact 3.5. (1) and (2) are proved in [10, Proposition 1.1]. To show (3), let P_v be the transition matrix corresponding to updating the spin at v conditioned on all other vertices. Thus, we have the decomposition

$$P_{\mathrm{GL}} = \frac{1}{n} \sum_{v \in V} P_v.$$

Let $f = \nu/\mu$ be the relative density of ν with respect to μ. Then we get

$$\begin{aligned} D_{\mathrm{KL}}(\nu P_{\mathrm{GL}} \,\|\, \mu) &= D_{\mathrm{KL}}\left(\frac{1}{n} \sum_{v \in V} \nu P_v \,\middle\|\, \mu \right) \le \frac{1}{n} \sum_{v \in V} D_{\mathrm{KL}}(\nu P_v \,\|\, \mu) \\ &= \frac{1}{n} \sum_{v \in V} \mathrm{Ent}(P_v f) = \frac{1}{n} \sum_{v \in V} \mathrm{Ent}[\mu_v(f)] \\ &= \frac{1}{n} \sum_{v \in V} \mathrm{Ent}(f) - \mu[\mathrm{Ent}_v(f)] \\ &= \mathrm{Ent}(f) - \frac{1}{n} \sum_{v \in V} \mu[\mathrm{Ent}_v(f)] \\ &\le \left(1 - \frac{1}{C_1 n}\right) \mathrm{Ent}(f) = \left(1 - \frac{1}{C_1 n}\right) D_{\mathrm{KL}}(\nu \,\|\, \mu). \end{aligned}$$

(4) can be deduced from (3) as shown by [6, Lemma 2.4]; see also [7, Corollary 2.8] for the continuous time setting. (5) follows from (2) and [17, Lemma 15]. Finally, (6) follows by an application of [18, Theorem A.1]. □

4 APPROXIMATE TENSORIZATION VIA UNIFORM BLOCK FACTORIZATION

Fix a graph G on n vertices of maximum degree at most Δ, and assume that μ is a b-marginally bounded Gibbs distribution defined on G satisfying the $\lceil \theta n \rceil$-uniform block factorization of entropy with constant C where $\theta \le b^2/(4e\Delta)$; i.e., for $\ell = \lceil \theta n \rceil$ and all $f : \Omega \to \mathbb{R}_{\ge 0}$ it holds that

$$\frac{\ell}{n} \mathrm{Ent}(f) \le C \cdot \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \mu[\mathrm{Ent}_S(f)].$$

We will show that μ also satisfies the approximate tensorization of entropy with constant $\Theta(C)$, which establishes Lemma 2.3.

The intuition behind our approach is that for ℓ as large as θn, if one picks a uniformly random subset $S \subseteq V$ satisfying $|S| = \ell$, then the induced subgraph $G[S]$ of G on vertex set S is disconnected into many small connected components, each of which has constant size in expectation and at most $O(\log n)$ with high probability. Since the conditional Gibbs distribution μ_S^τ is a product distribution of each connected component, we can use entropy factorization for product distributions to reduce approximate tensorization on G to that on small connected subgraphs of G. This allows us to upper bound the optimal approximate tensorization constant with a converging series.

Towards fulfilling this intuition, for any $S \subseteq V$, let $\mathcal{C}(S)$ denote the set of connected components of $G[S]$, with each connected component being viewed as a subset of vertices of S. Note that $\mathcal{C}(S)$ is a partition of S. For any $v \in S$, let S_v denote the (unique) connected component in $\mathcal{C}(S)$ containing v; for $v \notin S$, take $S_v = \emptyset$. The following is a well-known fact regarding the factorization of entropy for product measures; see, e.g., [10, 12].

Lemma 4.1. *For every subset $S \subseteq V$, every boundary condition $\tau \in \Omega_{V \setminus S}$, and every function $f : \Omega_S^\tau \to \mathbb{R}_{\ge 0}$, we have*

$$\mathrm{Ent}_S^\tau(f) \le \sum_{U \in \mathcal{C}(S)} \mu_S^\tau[\mathrm{Ent}_U(f)].$$

Recall that $\mathrm{Ent}_U(f) = \mathrm{Ent}_U^\phi(f)$ is regarded as a function of the boundary condition $\phi \in \Omega_{S \setminus U}^\tau$ on $S \setminus U$, and $\mu_S^\tau[\mathrm{Ent}_U(f)]$ is the expectation of it under the conditional Gibbs measure μ_S^τ.

We also need the following crude exponential upper bound on the approximate tensorization constant for a Gibbs distribution with bounded marginals.

Lemma 4.2. *If μ is b-marginally bounded, then for every subset $U \subseteq V$, every boundary condition $\xi \in \Omega_{V \setminus U}$, and every function $f : \Omega_U^\xi \to \mathbb{R}_{\ge 0}$, we have*

$$\mathrm{Ent}_U^\xi(f) \le \frac{3|U|^2 \log(1/b)}{2b^{2|U|+2}} \sum_{v \in U} \mu_U^\xi[\mathrm{Ent}_v(f)].$$

Finally, the lemma below shows that when a uniformly random and sufficiently small subset of vertices is selected, the size of the

connected component containing a given vertex is small with high probability.

Lemma 4.3. *Let $G = (V, E)$ be an n-vertex graph of maximum degree at most Δ. Then for every $k \in \mathbb{N}^+$ we have*

$$\mathbb{P}_S(|S_v| = k) \le \frac{\ell}{n} \cdot (2e\Delta\theta)^{k-1},$$

where the probability $\mathbb{P}$ is taken over a uniformly random subset $S \subseteq V$ of size $\ell = \lceil \theta n \rceil$.

We postpone the proofs of Lemmas 4.2 and 4.3. We are now ready to prove Lemma 2.3.

Proof of Lemma 2.3. Combining everything in this section, we deduce that

$$\begin{aligned}
\mathrm{Ent}(f) &\le C \cdot \frac{n}{\ell} \cdot \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \mu[\mathrm{Ent}_S(f)] \\
&\qquad (\ell\text{-uniform block factorization}) \\
&\le C \cdot \frac{n}{\ell} \cdot \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \sum_{U \in C(S)} \mu[\mathrm{Ent}_U(f)] \qquad \text{(Lemma 4.1)} \\
&\le C \cdot \frac{n}{\ell} \cdot \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \sum_{U \in C(S)} \frac{3|U|^2 \log(1/b)}{2b^{2|U|+2}} \sum_{v \in U} \mu[\mathrm{Ent}_v(f)] \\
&\qquad \text{(Lemma 4.2)} \\
&= \frac{3C\log(1/b)}{2b^4} \cdot \frac{n}{\ell} \sum_{v \in V} \mu[\mathrm{Ent}_v(f)] \sum_{k=1}^{\ell} \mathbb{P}_S(|S_v| = k) \cdot \frac{k^2}{b^{2(k-1)}} \\
&\qquad \text{(rearranging)} \\
&\le \frac{3C\log(1/b)}{2b^4} \sum_{v \in V} \mu[\mathrm{Ent}_v(f)] \sum_{k=1}^{\ell} k^2 \left(\frac{2e\Delta\theta}{b^2}\right)^{k-1} \qquad \text{(Lemma 4.3)} \\
&\le \frac{3C\log(1/b)}{2b^4} \sum_{k=1}^{\ell} \frac{k^2}{2^{k-1}} \sum_{v \in V} \mu[\mathrm{Ent}_v(f)] \qquad (\theta \le \tfrac{b^2}{12\Delta}) \\
&\le \frac{18C\log(1/b)}{b^4} \sum_{v \in V} \mu[\mathrm{Ent}_v(f)]. \qquad (\textstyle\sum_{k=1}^{\infty} \frac{k^2}{2^{k-1}} = 12)
\end{aligned}$$

This establishes the lemma. □

Proof of Technical Lemmas. We first prove Lemma 4.2 which gives a crude bound on the approximate tensorization constant for any subset and boundary condition.

Proof of Lemma 4.2. Fix a subset $U \subseteq V$ of size $k \ge 1$ and some boundary condition $\xi \in \Omega_{V \setminus U}$. Let $C_1 = C_1(U, \xi)$ be the optimal constant of approximate tensorization for μ_U^ξ; hence, for every function $f : \Omega_U^\xi \to \mathbb{R}_{\ge 0}$ one has

$$\mathrm{Ent}_U^\xi(f) \le C_1 \sum_{v \in U} \mu_U^\xi[\mathrm{Ent}_v(f)].$$

Let $\lambda = \lambda(U, \xi)$ be the spectral gap of the Glauber dynamics for μ_U^ξ, and let $\rho = \rho(U, \xi)$ be the standard log-Sobolev constant. Thus, for every function $f : \Omega_U^\xi \to \mathbb{R}_{\ge 0}$ it holds that

$$\lambda \,\mathrm{Var}_U^\xi(f) \le \frac{1}{k} \sum_{v \in U} \mu_U^\xi[\mathrm{Var}_v(f)];$$

$$\rho \,\mathrm{Ent}_U^\xi(f) \le \frac{1}{k} \sum_{v \in U} \mu_U^\xi[\mathrm{Var}_v(\sqrt{f})].$$

Since $\mathrm{Var}_v(\sqrt{f}) \le \mathrm{Ent}_v(f)$, we have

$$C_1 \le \frac{1}{\rho k}; \tag{6}$$

see also [10, Proposition 1.1]. Next, [18, Corollary A.4] gives a comparison between the standard log-Sobolev constant and the spectral gap:

$$\rho \ge \frac{(1 - 2\mu^*)}{\log(1/\mu^* - 1)} \lambda$$

where $\mu^* = \min_{\sigma \in \Omega_U^\xi} \mu_U^\xi(\sigma)$. Since μ is b-marginally bounded, we have $\mu^* \ge b^k$. Also, notice that $|\Omega_U^\xi| = 1$ and $|\Omega_U^\xi| = 2$ corresponds to trivial cases where we have $C_1 \le 1$, so we may assume that $|\Omega_U^\xi| \ge 3$ which makes $\mu^* \le 1/3$. It follows that

$$\rho \ge \frac{\lambda}{3k\log(1/b)}. \tag{7}$$

Finally, Cheeger's inequality yields

$$\lambda \ge \frac{\Phi^2}{2} \tag{8}$$

where Φ is the conductance of the Glauber dynamics defined by

$$\begin{aligned}
\Phi &= \min_{\substack{\Omega_0 \subseteq \Omega_U^\xi \\ \mu_U^\xi(\Omega_0) \le \frac{1}{2}}} \Phi_{\Omega_0}, \\
\Phi_{\Omega_0} &= \frac{P_{\mathrm{GL}}(\Omega_0, \Omega_U^\xi \setminus \Omega_0)}{\mu_U^\xi(\Omega_0)} \\
&= \frac{1}{\mu_U^\xi(\Omega_0)} \sum_{\sigma \in \Omega_0} \sum_{\tau \in \Omega_U^\xi \setminus \Omega_0} \mu_U^\xi(\sigma) P_{\mathrm{GL}}(\sigma, \tau).
\end{aligned}$$

Our assumption that μ is totally-connected guarantees $\Phi_{\Omega_0} > 0$ for every $\Omega_0 \subseteq \Omega_U^\xi$ with $\mu_U^\xi(\Omega_0) \le \frac{1}{2}$. Furthermore, since μ is b-marginally bounded, for every $\sigma \in \Omega_0$ and $\tau \in \Omega_U^\xi \setminus \Omega_0$ such that $P_{\mathrm{GL}}(\sigma, \tau) > 0$ we have

$$\mu_U^\xi(\sigma) P_{\mathrm{GL}}(\sigma, \tau) \ge b^k \cdot \frac{b}{k} = \frac{b^{k+1}}{k}.$$

This gives

$$\Phi \ge \frac{2b^{k+1}}{k}. \tag{9}$$

Combining Eqs. (6) to (9), we finally conclude that

$$C_1 \le \frac{3k^2 \log(1/b)}{2b^{2k+2}},$$

as claimed. □

Next we establish Lemma 4.3. We use the following lemma concerning the number of connected induced subgraphs in a bounded degree graph.

Lemma 4.4 ([8, Lemma 2.1]). *Let $G = (V, E)$ be a graph with maximum degree at most Δ, and $v \in V$. Then for every $k \in \mathbb{N}^+$, the number of connected induced subgraphs of G containing v with k vertices is at most $(e\Delta)^{k-1}$.*

We then prove Lemma 4.3.

Proof of Lemma 4.3. If $\mathcal{A}_v(k)$ denotes the collection of subsets of vertices $U \subseteq V$ such that $|U| = k$, $v \in U$, and $G[U]$ is connected, then by the union bound, we have

$$\begin{aligned}\mathbb{P}_S(|S_v| = k) &\le \mathbb{P}_S(\exists U \in \mathcal{A}_v(k) : U \subseteq S)\\ &\le \sum_{U \in \mathcal{A}_v(k)} \mathbb{P}_S(U \subseteq S)\\ &= |\mathcal{A}_v(k)| \cdot \frac{\ell}{n} \cdot \frac{\ell-1}{n-1} \cdots \frac{\ell-k+1}{n-k+1}\\ &\le |\mathcal{A}_v(k)| \cdot \frac{\ell}{n} \cdot \left(\frac{\ell-1}{n-1}\right)^{k-1}.\end{aligned}$$

We may assume that $n \ge 2$ (when $n = 1$ the lemma holds trivially), and thus

$$\frac{\ell-1}{n-1} \le \frac{\theta n}{n-1} \le 2\theta.$$

The lemma then follows immediately from $|\mathcal{A}_v(k)| \le (e\Delta)^{k-1}$ by Lemma 4.4. □

5 PROOFS OF MAIN RESULTS

In this section we prove our main results Theorems 1.1 to 1.5.

By Theorem 1.9, to establish optimal mixing time bound it suffices to show marginal boundedness and spectral independence for the corresponding Gibbs distribution.

We first consider antiferromagnetic 2-spin systems. Let β, γ, λ be reals such that $0 \le \beta \le \gamma$, $\gamma > 0$, $\beta\gamma < 1$ and $\lambda > 0$ so the triple (β, γ, λ) specifies parameters of an antiferromagnetic 2-spin system. We state here the formal definition of up-to-Δ uniqueness with gap δ given in [39].

Definition 5.1 (Up-to-Δ uniqueness with gap δ, [39])**.** For each $1 \le d < \Delta$ define

$$f_d(R) = \lambda\left(\frac{\beta R + 1}{R + \gamma}\right)^d$$

and denote the unique fixed point of f_d by R_d^*. We say the parameters (β, γ, λ) are *up-to-Δ unique with gap δ* if $|f_d'(R_d^*)| < 1 - \delta$ for all $1 \le d < \Delta$.

Proof of Theorem 1.1. The proof of [16, Theorem 3] showed that for antiferromagnetic 2-spin systems that are up-to-Δ unique with gap δ, the Gibbs distribution μ is $O(1/\delta)$-spectrally independent. Also, by considering the worst configuration of the neighborhood for soft-constraint models (i.e., $0 < \beta \le \gamma$) or 2-hop neighborhood for hard-constraint models (i.e., $0 = \beta < \gamma$), one can check that μ is b-marginally bounded for some constant $b = b(\Delta, \beta, \gamma, \lambda)$. The theorem then follows from Theorem 1.9. □

Though in general the constant $C = C(\Delta, \delta, \beta, \gamma, \lambda)$ for bounding the mixing time depends on the parameters (β, γ, λ) of the model, in most applications such as the hard-core model (Theorem 1.2) and the Ising model (Theorem 1.3) we can make the constant C independent of all parameters. This is achieved by considering separately when the parameters are pretty far away from the uniqueness threshold, in which case we can deduce rapid mixing under the *Dobrushin uniqueness condition* [20], see also [9, 52].

Lemma 5.2. *Consider an arbitrary distribution μ over $[q]^V$. For two distinct vertices $u, v \in V$, define*

$$R(u,v) = \max_{\substack{\tau,\xi \in \Omega_{V\setminus\{v\}}\\ \mathrm{Dif}(\tau,\xi)=\{u\}}} \left\|\mu(\sigma_v = \cdot \mid \sigma_{V\setminus\{v\}} = \tau) - \mu(\sigma_v = \cdot \mid \sigma_{V\setminus\{v\}} = \xi)\right\|_{\mathrm{TV}}$$

where $\mathrm{Dif}(\tau, \xi) = \{w \in V : \tau_w \ne \xi_w\}$. If there exists $c \in (0, 1)$ such that for every vertex $v \in V$ we have

$$\sum_{u \in V\setminus\{v\}} R(u, v) \le 1 - c$$

(in which case we say the Dobrushin uniqueness condition *holds with constant c), then the mixing time of the Glauber dynamics for sampling from μ satisfies*

$$T_{\mathrm{mix}}(P_{GL}, \varepsilon) \le \frac{n}{c} \log\left(\frac{n}{\varepsilon}\right).$$

We present next the proofs of Theorems 1.2 and 1.3.

Proof of Theorem 1.2. By Theorem 1.1, for every $\lambda \le (1 - \delta)\lambda_c(\Delta)$ there exists $C = C(\Delta, \delta, \lambda)$ such that the Glauber dynamics mixes in $Cn\log(n/\varepsilon)$ steps. Meanwhile, it is easy to check that, when $\lambda \le \frac{1}{2\Delta}$ the Dobrushin uniqueness condition holds with $c = 1/2$, and thus the mixing time is upper bounded by $2n\log(n/\varepsilon)$. If we take

$$C' = C'(\Delta, \delta) := \max\left\{2, \sup_{\frac{1}{2\Delta} < \lambda \le (1-\delta)\lambda_c(\Delta)} C(\Delta, \delta, \lambda)\right\},$$

then the mixing time of the Glauber dynamics is upper bounded by $C'n\log(n/\varepsilon)$, as claimed. □

Proof of Theorem 1.3. Consider the antiferromagnetic Ising model ($\beta = \gamma < 1$) and by symmetry we may assume $\lambda \le 1$. It is shown in [16] that the Gibbs distribution μ is $O(1/\delta)$-spectrally independent in this case, and by considering the worst neighborhood configuration one can check that μ is b-marginally bounded for

$$\begin{aligned}b &= \min\left\{\frac{\lambda\beta^\Delta}{\lambda\beta^\Delta + 1}, \frac{\lambda^{-1}\gamma^\Delta}{\lambda^{-1}\gamma^\Delta + 1}\right\}\\ &= \frac{\lambda\beta^\Delta}{\lambda\beta^\Delta + 1} > \frac{\lambda\left(\frac{\Delta-2}{\Delta}\right)^\Delta}{\lambda\left(\frac{\Delta-2}{\Delta}\right)^\Delta + 1} \ge \frac{\lambda}{28}.\end{aligned}$$

Thus, when $\lambda \ge 1/500$, Theorem 1.9 implies that the mixing time of the Glauber dynamics is at most $\Delta^{O(1/\delta)}n\log(n/\varepsilon)$ for large enough n. Meanwhile, if $\lambda < 1/500$ then one can check that the Dobrushin uniqueness condition holds with $c = 1/2$, and thus the mixing time is upper bounded by $2n\log(n/\varepsilon)$. This proves the theorem for the antiferromagnetic case.

Next, consider the ferromagnetic Ising model ($\beta = \gamma > 1$). Assume $\lambda \ge 1$ for convenience. The Gibbs distribution μ is $O(1/\delta)$-spectrally independent by Theorem 26 of [16] and b-marginally bounded for

$$\begin{aligned}b &= \min\left\{\frac{1}{\lambda\beta^\Delta + 1}, \frac{1}{\lambda^{-1}\gamma^\Delta + 1}\right\}\\ &= \frac{1}{\lambda\beta^\Delta + 1} > \frac{1}{\lambda(\frac{\Delta}{\Delta-2})^\Delta + 1} \ge \frac{1}{28\lambda}.\end{aligned}$$

If $\lambda \le 500$ the mixing time is $\le \Delta^{O(1/\delta)} n \log(n/\varepsilon)$ by Theorem 1.9, and if $\lambda > 500$ the mixing time is $\le 2n \log(n/\varepsilon)$ by the Dobrushin uniqueness condition. This shows the ferromagnetic case, and completes the proof of the theorem. □

For random colorings, we can use the same argument.

Proof of Theorem 1.4. [23] showed that the uniform distribution μ of colorings is $O(1/\delta)$-spectrally independent under our assumption. (Note that the notion of spectral independence in [23] implies the one in [14] which is Definition 1.8; see Lemma 3.6 of [23] and Theorem 8 of [14]; also, [23] gave a better bound on the spectral independence constant and applicable to a slightly larger parameter region). Also, the proof of Lemma 3 from [26] can be adapted to show that μ is $\Omega(1/q)$-marginally bounded. Hence, Theorem 1.9 implies that the mixing time of the Glauber dynamics is at most $Cn \log(n/\varepsilon)$ for some $C = C(\Delta, \delta, q)$. Notice that when $q \ge 3\Delta$, the Dobrushin uniqueness condition holds with $c = 1/2$ and thus the mixing time is at most $2n \log(n/\varepsilon)$. By taking

$$C' = C'(\Delta, \delta) := \max\left\{2, \max_{(\alpha^*+\delta)\Delta \le q < 3\Delta} C(\Delta, \delta, q)\right\},$$

we get an upper bound $C'n \log(n/\varepsilon)$ for the mixing time. □

Finally, we give the proof for the monomer-dimer model.

Proof of Theorem 1.5. Notice that the monomer-dimer model on G is equivalent to the hard-core model on the line graph of G; so Theorem 1.9 is still applicable. Theorem 2.9 shows that the Gibbs distribution μ of the monomer-dimer model is η-spectrally independent for

$$\eta = \min\left\{2\lambda\Delta, 2\sqrt{1+\lambda\Delta}\right\}.$$

Meanwhile, by considering the worst configuration on the 2-hop neighborhood one can show that μ is b-marginally bounded for some $b = b(\Delta, \lambda)$. Thus, the theorem follows from Theorem 1.9. □

6 OPEN PROBLEMS

- Can we improve the approximate tensorization constant C_1 in Theorem 2.8 and the mixing time bound in Theorem 1.9 with a better dependence on the maximum degree Δ and on the spectral independence η? For example, for the hard-core model when $\lambda \le (1-\delta)\lambda_c(\Delta)$, currently our mixing time bound scales as $\Delta^{O(\Delta^2/\delta)} \times O(n \log n)$. Can we improve it and get $\mathrm{poly}(\Delta, 1/\delta)\, n \log n$?
- One can show the spectral independence of the monomer-dimer model on the infinite Δ-regular tree $\mathbb{T}_\Delta$ is exactly $\frac{2x}{1-x}$ where
$$x = \frac{1}{\Delta - 1}\left(1 - \frac{2}{\sqrt{1+4\lambda(\Delta-1)}+1}\right)$$
is the (unsigned) pairwise influence between edges of $\mathbb{T}_\Delta$ sharing a vertex. Note that for $\Delta = 2$, this is $\sqrt{1+4\lambda} - 1 = \Theta(\sqrt{\lambda})$, while for $\Delta \ge 3$, the spectral independence $O(1/\Delta)$ is independent of λ.
This suggests that while the bound on the total influence of an edge is tight, the bound in Theorem 2.9 on the maximum eigenvalue obtained by controlling the ∞-norm of the influence matrix is not tight, in contrast to the upper bound in [16] for vertex two-spin systems (which has a matching lower bound [2]). It would be interesting to obtain improved bounds on the spectral independence for the monomer-dimer model.
- Finally, we ask if spectral independence is equivalent to the notion of strong spatial mixing which has been studied extensively.

ACKNOWLEDGMENTS

Zongchen Chen and Eric Vigoda are supported in part by NSF grant CCF-2007022. Kuikui Liu is supported in part by NSF grant CCF-1907845 and ONR-YIP grant N00014-17-1-2429.

REFERENCES

[1] Vedat Levi Alev and Lap Chi Lau. 2020. Improved analysis of higher order random walks and applications. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC)*. 1198–1211.
[2] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. 2020. Spectral Independence in High-Dimensional Expanders and Applications to the Hardcore Model. In *Proceedings of the 61st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 1319–1330.
[3] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. 2019. Log-Concave Polynomials II: High-Dimensional Walks and an FPRAS for Counting Bases of a Matroid. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing (STOC)*. 1–12.
[4] Mohsen Bayati, David Gamarnik, Dimitriy Katz, Chandra Nair, and Prasad Tetali. 2007. Simple Deterministic Approximation Algorithms for Counting Matchings. In *Proceedings of the 39th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*. 122–127.
[5] Antonio Blanca, Pietro Caputo, Zongchen Chen, Daniel Parisi, Daniel Štefankovič, and Eric Vigoda. 2021. On Mixing of Markov Chains: Coupling, Spectral Independence, and Entropy Factorization. *ArXiv preprint, arXiv:2103.07459* (2021).
[6] Antonio Blanca, Pietro Caputo, Daniel Parisi, Alistair Sinclair, and Eric Vigoda. 2021. Entropy decay in the Swendsen-Wang dynamics. In *Proceedings of the 53rd Annual ACM Symposium on Theory of Computing (STOC)*.
[7] Sergey G Bobkov and Prasad Tetali. 2006. Modified logarithmic Sobolev inequalities in discrete settings. *Journal of Theoretical Probability* 19, 2 (2006), 289–336.
[8] Christian Borgs, Jennifer Chayes, Jeff Kahn, and László Lovász. 2013. Left and right convergence of graphs with bounded degree. *Random Structures & Algorithms* 42, 1 (2013), 1–28.
[9] Russ Bubley and Martin E. Dyer. 1997. Path Coupling: A Technique for Proving Rapid Mixing in Markov Chains. In *Proceedings of the 38th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 223–231.
[10] Pietro Caputo, Georg Menz, and Prasad Tetali. 2015. Approximate tensorization of entropy at high temperature. In *Annales de la Faculté des sciences de Toulouse: Mathématiques*, Vol. 24. 691–716.
[11] Pietro Caputo and Daniel Parisi. 2020. Block factorization of the relative entropy via spatial mixing. *ArXiv preprint, arXiv:2004.10574* (2020).
[12] Filippo Cesi. 2001. Quasi-factorization of the entropy and logarithmic Sobolev inequalities for Gibbs random fields. *Probability Theory and Related Fields* 120, 4 (2001), 569–584.
[13] Sitan Chen, Michelle Delcourt, Ankur Moitra, Guillem Perarnau, and Luke Postle. 2019. Improved Bounds for Randomly Sampling Colorings via Linear Programming. In *Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2216–2234.
[14] Zongchen Chen, Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. 2021. Rapid Mixing for Colorings via Spectral Independence. In *Proceedings of the 32nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 1548–1557.
[15] Zongchen Chen, Kuikui Liu, and Eric Vigoda. 2020. Optimal Mixing of Glauber Dynamics: Entropy Factorization via High-Dimensional Expansion. *Full version at arXiv:2011.02075* (2020).
[16] Zongchen Chen, Kuikui Liu, and Eric Vigoda. 2020. Rapid Mixing of Glauber Dynamics up to Uniqueness via Contraction. In *Proceedings of the 61st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 1307–1318.
[17] Mary Cryan, Heng Guo, and Giorgos Mousa. 2019. Modified log-Sobolev inequalities for strongly log-concave distributions. In *Proceedings of the 60th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 1358–1370.
[18] Persi Diaconis and Laurent Saloff-Coste. 1996. Logarithmic Sobolev inequalities for finite Markov chains. *The Annals of Applied Probability* 6, 3 (1996), 695–750.

[19] Irit Dinur and Tali Kaufman. 2017. High Dimensional Expanders Imply Agreement Expanders. In *Proceedings of the 58th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 974–985.
[20] R. L. Dobrushin. 1970. Prescribing a System of Random Variables by Conditional Distributions. *Theory of Probability and its Applications* 15, 3 (1970), 458–486.
[21] John Duchi. 2019. Information Theory and Statistics. *Lecture Notes for Statistics 311 and Electrical Engineering 377* (2019), 304.
[22] Martin E. Dyer, Alan M. Frieze, Thomas P. Hayes, and Eric Vigoda. 2013. Randomly coloring constant degree graphs. *Random Struct. Algorithms* 43, 2 (2013), 181–200.
[23] Weiming Feng, Heng Guo, Yitong Yin, and Chihao Zhang. 2021. Rapid mixing from spectral independence beyond the Boolean domain. In *Proceedings of the 32nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 1558–1577.
[24] Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. 2015. Inapproximability for Antiferromagnetic Spin Systems in the Tree Nonuniqueness Region. *J. ACM* 62, 6 (2015), 1–60.
[25] Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. 2016. Inapproximability of the Partition Function for the Antiferromagnetic Ising and Hard-Core Models. *Combinatorics, Probability and Computing* 25, 4 (2016), 500–559.
[26] David Gamarnik, Dmitriy Katz, and Sidhant Misra. 2015. Strong spatial mixing of list coloring of graphs. *Random Structures & Algorithms* 46, 4 (2015), 599–613.
[27] A. Guionnet and B. Zegarlinksi. 2003. *Lectures on Logarithmic Sobolev Inequalities*. Springer, Berlin, 1–134.
[28] Heng Guo and Pinyan Lu. 2018. Uniqueness, Spatial Mixing, and Approximation for Ferromagnetic 2-Spin Systems. *ACM Transactions on Computation Theory* 10, 4, Article 17 (2018).
[29] Thomas P. Hayes and Alistair Sinclair. 2007. A general lower bound for mixing of single-site dynamics on graphs. *The Annals of Applied Probability* 17, 3 (2007), 931–952.
[30] Mark Jerrum. 1995. A very simple algorithm for estimating the number of k-colorings of a low-degree graph. *Random Structures & Algorithms* 7, 2 (1995), 157–165.
[31] Mark Jerrum. 2003. *Counting, Sampling and Integrating: Algorithms and Complexity*. Birkhäuser Basel.
[32] Mark Jerrum and Alistair Sinclair. 1989. Approximating the Permanent. *SIAM J. Comput.* 18, 6 (1989), 1149–1178.
[33] Mark Jerrum and Alistair Sinclair. 1993. Polynomial-time approximation algorithms for the Ising model. *SIAM J. Comput.* 22, 5 (1993), 1087–1116.
[34] Mark Jerrum, Leslie G. Valiant, and Vijay V. Vazirani. 1986. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science* 43 (1986), 169–188.
[35] Tali Kaufman and David Mass. 2017. High Dimensional Random Walks and Colorful Expansion. In *Proceedings of the 8th Innovations in Theoretical Computer Science Conference (ITCS)*. 4:1–4:27.
[36] Tali Kaufman and David Mass. 2020. Local-To-Global Agreement Expansion via the Variance Method. In *Proceedings of the 11th Innovations in Theoretical Computer Science Conference (ITCS)*. 74:1–74:14.
[37] Tali Kaufman and Izhar Oppenheim. 2020. High Order Random Walks: Beyond Spectral Gap. *Combinatorica* 40 (2020), 245–281.
[38] Vladimir Kolmogorov. 2018. A Faster Approximation Algorithm for the Gibbs Partition Function. In *Proceedings of the 31st Conference On Learning Theory (COLT)*, Vol. 75. 228–249.
[39] Liang Li, Pinyan Lu, and Yitong Yin. 2013. Correlation Decay Up to Uniqueness in Spin Systems. In *Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 67–84.
[40] Fabio Martinelli. 1999. *Lectures on Glauber Dynamics for Discrete Spin Models*. Springer, Berlin, 93–191.
[41] Fabio Martinelli, Alistair Sinclair, and Dror Weitz. 2003. The Ising model on trees: boundary conditions and mixing time. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 628–639.
[42] Katalin Marton. 2019. Logarithmic Sobolev inequalities in discrete product spaces. *Combinatorics, Probability & Computing* 28, 6 (2019), 919–935.
[43] Ravi Montenegro and Prasad Tetali. 2006. Mathematical aspects of mixing times in Markov chains. *Foundations and Trends in Theoretical Computer Science* 1, 3 (2006), 236–236.
[44] Elchanan Mossel and Allan Sly. 2013. Exact thresholds for Ising-Gibbs samplers on general graphs. *The Annals of Probability* 41, 1 (2013), 294–328.
[45] Izhar Oppenheim. 2018. Local spectral expansion approach to high dimensional expanders part I: Descent of spectral gaps. *Discrete and Computational Geometry* 59, 2 (2018), 293–330.
[46] Yuval Peres and Peter Winkler. 2013. Can Extra Updates Delay Mixing? *Communications in Mathematical Physics* 323 (2013), 1007–1016.
[47] Shuai Shao and Yuxin Sun. 2020. Contraction: A Unified Perspective of Correlation Decay and Zero-Freeness of 2-Spin Systems. In *Proceedings of the 47th International Colloquium on Automata, Languages, and Programming (ICALP)*. 96:1–15.
[48] Allan Sly. 2010. Computational Transition at the Uniqueness Threshold. In *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 287–296.
[49] Allan Sly and Nike Sun. 2014. The Computational Hardness of Counting in Two-Spin Models on d-Regular Graphs. *The Annals of Probability* 42, 6 (2014), 2383–2416.
[50] Daniel Štefankovič, Santosh Vempala, and Eric Vigoda. 2009. Adaptive simulated annealing: A near-optimal connection between sampling and counting. *J. ACM* 56, 3 (2009), 1–36.
[51] Eric Vigoda. 1999. Improved Bounds for Sampling Colorings. In *Proceedings of the 40th Annual Symposium on Foundations of Computer Science (FOCS)*. 51–59.
[52] Dror Weitz. 2005. Combinatorial criteria for uniqueness of Gibbs measures. *Random Structures & Algorithms* 27, 4 (2005), 445–475.
[53] Dror Weitz. 2006. Counting Independent Sets Up to the Tree Threshold. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC)*. 140–149.