Rapid Mixing of Global Markov Chains via Spectral Independence: The Unbounded Degree Case

Antonio Blanca
Pennsylvania State University, University Park, PA, USA

Xusheng Zhang
Pennsylvania State University, University Park, PA, USA

Abstract

We consider spin systems on general n-vertex graphs of unbounded degree and explore the effects of spectral independence on the rate of convergence to equilibrium of global Markov chains. Spectral independence is a novel way of quantifying the decay of correlations in spin system models, which has significantly advanced the study of Markov chains for spin systems. We prove that whenever spectral independence holds, the popular Swendsen–Wang dynamics for the q-state ferromagnetic Potts model on graphs of maximum degree Δ, where Δ is allowed to grow with n, converges in $O((\Delta \log n)^c)$ steps where $c > 0$ is a constant independent of Δ and n. We also show a similar mixing time bound for the block dynamics of general spin systems, again assuming that spectral independence holds. Finally, for *monotone* spin systems such as the Ising model and the hardcore model on bipartite graphs, we show that spectral independence implies that the mixing time of the systematic scan dynamics is $O(\Delta^c \log n)$ for a constant $c > 0$ independent of Δ and n. Systematic scan dynamics are widely popular but are notoriously difficult to analyze. This result implies optimal $O(\log n)$ mixing time bounds for any systematic scan dynamics of the ferromagnetic Ising model on general graphs up to the tree uniqueness threshold. Our main technical contribution is an improved factorization of the entropy functional: this is the common starting point for all our proofs. Specifically, we establish the so-called k-partite factorization of entropy with a constant that depends polynomially on the maximum degree of the graph.

2012 ACM Subject Classification Theory of computation → Random walks and Markov chains; Mathematics of computing → Gibbs sampling

Keywords and phrases Markov chains, Spectral Independence, Potts model, Mixing Time

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2023.53

Category RANDOM

Related Version *Full Version*: https://arxiv.org/abs/2307.00683

Funding The research is supported in part by NSF grant CCF-2143762.

Acknowledgements The authors thank the anonymous referees for their constructive comments during the manuscript review process.

1 Introduction

Spectral independence is a powerful new approach for quantifying the decay of correlations in spin system models. Initially introduced in [4], this condition has revolutionized the study of Markov chains for spin systems. In a series of important and recent contributions, spectral independence has been shown to be instrumental in determining the convergence rate of the Glauber dynamics, the simple single-site update Markov chain that updates the spin at a randomly chosen vertex in each step.

The first efforts in this series (see [4, 24, 25]) showed that spectral independence implies optimal $O(n \log n)$ mixing of the Glauber dynamics on n-vertex graphs of bounded degree for general spin systems. The unbounded degree case was studied in [20, 19, 3, 44], while [6] explored the effects of this condition on the speed of convergence of global Markov chains (i.e., Markov chains that update the spins of a large number of vertices in each step) in the bounded degree setting. Research exploring the applications of spectral independence is ongoing. We contribute to this line of work by investigating how spectral independence affects the speed of convergence of *global Markov chains* for general spin systems on graphs of unbounded degree.

A *spin system* is defined on a graph $G = (V, E)$. There is a set $\mathcal{S} = \{1, \dots, q\}$ of spins or colors, and configurations are assignments of spin values from $\mathcal{S}$ to each vertex of G. The probability of a configuration $\sigma \in \mathcal{S}^V$ is given by the Gibbs distribution:

$$\mu(\sigma) = \frac{e^{-H(\sigma)}}{Z}, \tag{1}$$

where the normalizing factor Z is known as the partition function, and the Hamiltonian $H : \mathcal{S}^V \to \mathbb{R}$ contains terms that depend on the spin values at each vertex (a "vertex potential" or "external field") and at each pair of adjacent vertices (an "edge potential"); see Definition 24. A widely studied spin system, and one that we will pay close attention to in this paper, is the ferromagnetic Potts model, where for a real parameter $\beta > 0$, associated with inverse temperature in physical applications, the Hamiltonian is given by:

$$H(\sigma) = -\beta \sum_{\{u,v\} \in E} \mathbb{1}(\sigma_u = \sigma_v).$$

The classical ferromagnetic Ising model corresponds to the $q = 2$ case. (In this variant of the Potts model, the Hamiltonian only includes edge potentials, and there is no external field.) We shall use μ_{Ising} and μ_{Potts} for the Gibbs distributions corresponding to the Ising and Potts models. Other well-known, well-studied spin systems include uniform proper colorings and the hardcore model.

Spin systems provide a robust framework for studying interacting systems of simple elements and have a wide range of applications in computer science, statistical physics, and other fields. In such applications, generating samples from the Gibbs distribution (1) is a fundamental computational task and one in which Markov chain-based algorithms have been quite successful. A long line of work dating back to the 1980s relates the speed of convergence of Markov chains to various forms of decay of correlations in the model. Spectral independence, defined next, captures the decay of correlations in a novel way.

Roughly speaking, spectral independence holds when the spectral norm of a "pairwise" influence matrix is bounded. To formally define it, let us begin by introducing some notations. Let $\Omega \subseteq \mathcal{S}^V$ be the support of μ: the set of configurations σ such that $\mu(\sigma) > 0$. A *pinning* τ on a subset of vertices $\Lambda \subseteq V$ is a fixed partial configuration on Λ; i.e., a spin assignment from $\mathcal{S}^\Lambda$ to the vertices of Λ. For a pinning τ on $\Lambda \subseteq V$ and $U \subseteq V \setminus \Lambda$, we let $\Omega_U^\tau = \{\sigma_U \in \mathcal{S}^U : \mu(\sigma_U \mid \sigma_\Lambda = \tau) > 0\}$ be the set of partial configurations on U that are consistent with the pinning τ. We write $\Omega_u^\tau = \Omega_{\{u\}}^\tau$ if u is a single vertex. Let

$$\mathcal{P}^\tau := \{(u, s) : u \notin \Lambda, s \in \Omega_u^\tau\}$$

denote the set of consistent vertex-spin pairs in $\Omega_{V \setminus \Lambda}^\tau$ under μ. For each $\Lambda \subseteq V$ and pinning τ on Λ, we define the *signed pairwise influence matrix* $\Psi_\mu^\tau \in \mathbb{R}^{\mathcal{P}^\tau \times \mathcal{P}^\tau}$ to be the matrix with entries:

$$\Psi_\mu^\tau((u, a), (v, b)) = \mu(\sigma_v = b \mid \sigma_u = a, \sigma_\Lambda = \tau) - \mu(\sigma_v = b \mid \sigma_\Lambda = \tau)$$

for $u \neq v$, and $\Psi_\mu^\tau((u, a), (u, b)) = 0$ otherwise.

▶ **Definition 1** (Spectral Independence). *A distribution μ satisfies η-*spectral independence *if for every subset of vertices $\Lambda \subseteq V$ and every pinning $\tau \in \Omega_\Lambda$, the largest eigenvalue of the signed pairwise influence matrix Ψ_μ^τ, denoted $\lambda_1(\Psi_\mu^\tau)$, satisfies $\lambda_1(\Psi_\mu^\tau) \le \eta$.*

There are several definitions of spectral independence in the literature; we use here the one from [22].

We show that spectral independence implies new upper bounds on the mixing time of several well-studied global Markov chains in the case where the maximum degree Δ of the underlying graph $G = (V, E)$ is unbounded; i.e., $\Delta \to \infty$ with n. The mixing time is defined as the number of steps required for a Markov chain to reach a distribution close in total variation distance to its stationary distribution, assuming a worst possible starting state; a formal definition is given in Section 2. The global Markov chains we consider include the Swendsen–Wang dynamics for the ferromagnetic q-state Potts, the systematic scan dynamics for monotone spin systems, and the block dynamics for general spin systems. These three dynamics are among the most popular and well-studied global Markov chains and present certain advantages (e.g., faster convergence and amenability to parallelization) to the Glauber dynamics.

1.1 The Swendsen–Wang dynamics

A canonical example of a global Markov chain is the Swendsen–Wang (SW) dynamics for the ferromagnetic q-state Potts model. The SW dynamics transitions from a configuration σ_t to σ_{t+1} by:

1. For each edge $e = \{u, v\} \in E$, if $\sigma_t(u) = \sigma_t(v)$, independently include e in the set A_t with probability $p = 1 - e^{-\beta}$;
2. Then, independently for each connected component $\mathcal{C}$ in (V, A_t), draw a spin $s \in \{1, \dots, q\}$ uniformly at random and set $\sigma_{t+1}(v) = s$ for all $v \in \mathcal{C}$.

The SW dynamics is ergodic and reversible with respect to μ_{Potts} and thus converges to it. This Markov chain originated in the late 1980s [53] as an alternative to the Glauber dynamics, which mixes exponentially slowly at low temperatures (large β). The SW dynamics bypasses key barriers that cause the slowdown of the Glauber dynamics at low temperatures. For the Ising model ($q = 2$), for instance, it was recently shown to converge in poly(n) steps on any n-vertex graph for any value of $\beta > 0$ [39]. (The conjectured mixing time is $\Theta(n^{1/4})$, but we seem to be far from proving such a conjecture.) For $q \ge 3$, on the other hand, the SW dynamics can converge exponentially slowly at certain "intermediate" temperatures regimes corresponding to first-order phase transitions; see [38, 15, 36, 37, 26].

Recently, η-spectral independence (with $\eta = O(1)$) was shown to imply that the mixing time of the SW dynamics is $O(\log n)$ on graphs of maximum degree $\Delta = O(1)$, i.e., bounded degree graphs [6]. This mixing time bound is optimal since the SW dynamics requires $\Omega(\log n)$ steps to mix in some cases where η and Δ are both $O(1)$ [7, 9]. However, it does not extend to the unbounded degree setting since the constant factor hidden by the big-O notation depends exponentially on the maximum degree Δ; this is the case even when $\eta = O(1)$ and $\beta\Delta = O(1)$. Our first result provides a mixing time bound that depends only polynomially on Δ.

▶ **Theorem 2.** *Let $q \ge 2$, $\beta > 0$, $\eta > 0$ and $\Delta \ge 3$. Suppose $G = (V, E)$ is an n-vertex graph of maximum degree Δ. Let μ_{Potts} be the Gibbs distribution of the q-state ferromagnetic Potts model on G with parameter β. If μ_{Potts} is η-spectrally independent with $\eta = O(1)$ and $\beta\Delta = O(1)$, then there exists a constant $c > 0$ such that the mixing time of the SW dynamics satisfies $T_{mix}(P_{SW}) = O\big((\Delta \log n)^c\big)$.*

The constant c has a near linear dependency on η and $\beta\Delta$; a more precise statement of Theorem 2 with a precise expression for c is given in Theorem 11.

Despite the expectation that the SW dynamics mixes in $O(\log n)$ steps in weakly correlated systems (i.e., when $\beta\Delta$ is small), proving sub-linear upper bounds on its mixing time has been difficult. Recently, various forms of decay of correlation (e.g., strong spatial mixing, entropy mixing, and spectral independence) have been used to obtain $O(\log n)$ bounds for the mixing time of the SW dynamics on cubes of the integer lattice graph $\mathbb{Z}^d$, regular trees, and general graphs of bounded degree (see [7, 9, 6]). However, for graphs of large degree, i.e., with $\Delta \to \infty$ with n, the only sub-linear mixing time bounds known either hold for the very distinctive mean-field model, where G is the complete graph [35, 11], or hold for very small values of β; i.e., $\beta \lesssim 1/(3\Delta)$ [43]. Our results provide new sub-linear mixing time bounds for graph families of sub-linear maximum degree, provided $\eta = O(1)$ and $\beta\Delta = O(1)$. These last two conditions go hand-in-hand: in all known cases where $\eta = O(1)$, we also have $\beta\Delta = O(1)$.

On graphs of degree at most Δ, η-spectral independence is supposed to hold with $\eta = O(1)$ whenever $\beta < \beta_u(q,\Delta)$, where $\beta_u(q,\Delta)$ is the threshold for the uniqueness/non-uniqueness phase transition on Δ-regular trees. This has been confirmed for the Ising model ($q = 2$) but not for the Potts model. Specifically, for the ferromagnetic Ising model, we have $\beta_u(2,\Delta) = \ln \frac{\Delta}{\Delta-2}$, and when $\beta \leq (1-\delta)\beta_u(2,\Delta)$ for some $\delta \in (0,1)$, μ_{Ising} is η-spectrally independent with $\eta = O(1/\delta)$; see [24, 25]. In contrast, for the ferromagnetic Potts model with $q \geq 3$, there is no closed-form expression for $\beta_u(q,\Delta)$ (it is defined as the threshold value where an equation starts to have a double root), and for graphs of unbounded degree η-spectral independence is only known to hold when $\beta \leq \frac{2(1-\delta)}{\Delta}$. As a result, we obtain the following corollary of Theorem 2.

► **Corollary 3.** *Let $\delta \in (0,1)$, $\Delta \geq 3$. Suppose that either $q = 2$ and $\beta < (1-\delta)\beta_u(2,\Delta)$, or $q \geq 3$ and $\beta \leq \frac{2(1-\delta)}{\Delta}$. Then, there exists a constant $c = c(\delta) > 0$ such that the mixing time of the SW dynamics for the q-state ferromagnetic Potts model on any n-vertex graph of maximum degree Δ satisfies $T_{mix}(P_{SW}) = O\big((\Delta \log n)^c\big)$.*

We mention that other conditions known to imply spectral independence (e.g., those in [14]) are not well-suited for the unbounded degree setting since under those conditions, the best known bound for η depends polynomially on Δ. For another application of Theorem 2, see Section 3.3.1 where we provide a bound on the mixing of the SW dynamics on random graphs.

We comment briefly on our proof approach for Theorem 2. A mixing time bound for the SW dynamics can be deduced from the so-called *edge-spin* factorization of the entropy functional introduced in [7]. It was noted there that this factorization, in turn, follows from a different factorization of entropy known as *k-partite factorization*, or KPF. Spectral independence is known to imply KPF but with a loss of a multiplicative constant that depends exponentially on the maximum degree of the graph. Our proof of Theorem 2 follows this existing framework, but pays closer attention to establishing KPF with an optimized constant with a better dependence on the model parameters. This is done through a multi-scale analysis of the entropy functional; in each scale, we apply spectral independence to achieve a tighter KPF condition. Our new results for KPF not only hold for the Potts model, but also for a general class of spin systems, and we use it to establish new mixing time bounds for the systematic scan dynamics and block dynamics.

1.2 The systematic scan dynamics

Our next contribution pertains the *systematic scan dynamics*, which is a family of Markov chains closely related to the Glauber dynamics in the sense that updates occur at single vertices sequentially. The key difference is that the vertex updates happen according to a predetermined ordering ϕ of the vertices instead of at random vertices. These dynamics offer practical advantages since there is no need to randomly select vertices at each step, thereby reducing computation time.

There is a folklore belief that the mixing time of the systematic scan dynamics (properly scaled) is closely related to that of the Glauber dynamics. However, analyzing this type of dynamics has proven very challenging (see, e.g., [28, 41, 30, 29, 49, 40, 8]), and the best general condition under which the systematic scan dynamics is known to be optimally mixing is a Dobrushin-type condition due to Dyer, Goldberg, and Jerrum [30]. The new developments on Markov chain mixing stemming from spectral independence have not yet provided new results for this dynamics, even for the bounded degree case where much progress has already been made. We show that spectral independence implies optimal mixing of the systematic scan dynamics for *monotone* spin systems with *bounded marginals*; we define both of these notions next.

▶ **Definition 4** (Monotone spin system). *In a monotone system, there is a linear ordering of the spins at each vertex which induces a partial order $\preceq_q$ over the state space. A spin system is* monotone *with respect to the partial order $\preceq_q$ if for every $\Lambda \subseteq V$ and every pair of pinnings $\tau_1 \succeq_q \tau_2$ on $V \setminus \Lambda$, the conditional distribution $\mu(\cdot \mid \sigma_\Lambda = \tau_1)$ stochastically dominates $\mu(\cdot \mid \sigma_\Lambda = \tau_2)$.*

Canonical examples of monotone spin systems include the ferromagnetic Ising model and the hardcore model on bipartite graphs. As in earlier work (see [24, 25, 6]), our bounds on the mixing time will depend on a lower bound on the marginal probability of any vertex-spin pair. This is formalized as follows.

▶ **Definition 5** (Bounded marginals). *The distribution μ is said to be b-*marginally bounded *if for every $\Lambda \subseteq V$ and pinning $\tau \in \Omega_\Lambda$, and each $(v, s) \in \mathcal{P}^\tau$, we have $\mu(\sigma_v = s \mid \sigma_\Lambda = \tau) \geq b$.*

Before stating our result for the systematic scan dynamics of b-marginally bounded monotone spin systems, we note that this Markov chain updates in a single step each vertex once in the order prescribed by ϕ. Under a minimal assumption on the spin system (the same one required to ensure the ergodicity of the Glauber dynamics), the systematic scan dynamics is ergodic. Specifically, when the spin system is totally-connected (see Definition 25), the systematic scan dynamics is ergodic. Moreover, the systematic scan dynamics is not necessarily reversible with respect to μ, so, as in earlier works, we work with the symmetrized version of the dynamics in which, in each step, the vertices are updated according to ϕ first, and subsequently in the reverse order of ϕ. The resulting dynamics, which we denote by P_ϕ, is reversible with respect to μ. Our main result for the systematic scan dynamics is the following.

▶ **Theorem 6.** *Let $b > 0$, $\eta > 0$, and $\Delta \geq 3$. Suppose $G = (V, E)$ is an n-vertex graph of maximum degree Δ. Let μ be the distribution of a totally-connected monotone spin system on G. If μ is η-spectrally independent and b-marginally bounded, then for any ordering ϕ,*

$$T_{mix}(P_\phi) = \left(\frac{e^2\Delta}{b}\right)^{9+4\lceil\frac{2\eta}{b}\rceil} \cdot O(\log n).$$

The bound in this theorem is tight: for a particular ordering ϕ, we prove an $\Omega(\log n)$ mixing time lower bound that applies to settings where Δ, b and η are all $\Theta(1)$; see Lemma 26.

We present next several interesting consequences of Theorem 6. First, we obtain the following corollary using the known results about spectral independence for the ferromagnetic Ising model.

▶ **Corollary 7.** *Let $\delta \in (0,1), \Delta \geq 3$ and $0 < \beta < (1-\delta)\beta_u(2,\Delta)$. Suppose $G = (V,E)$ is an n-vertex graph of maximum degree Δ. For any ordering ϕ of the vertices of G, the mixing time of P_ϕ for the Ising model on G with parameter β satisfies $T_{mix}(P_\phi) = O(\log n)$.*

The constant hidden by the big-O notation is an absolute constant that depends only on the constant δ, even when Δ depends on n. This result, compared to the earlier conditions in [28, 41, 30], extends the parameter regime where the $O(\log n)$ mixing time bound applies; in fact, the parameter regime in Corollary 7 is tight, as the systematic scan dynamics undergoes an exponential slowdown when $\beta > \beta_u(2,\Delta)$ [49]. We also derive analogous results for the hardcore model on bipartite graphs; see Section 4.1.

Our next application concerns the specific but relevant case where the underlying graph is an n-vertex cube of the integer lattice graph $\mathbb{Z}^d$. In this context, it was proved in [8] that all systematic scan dynamics converge in $O(\log n(\log\log n)^2)$ steps whenever a well-known condition known as *strong spatial mixing (SSM)* holds. A pertinent open question is whether SSM implies spectral independence. In fact, spectral independence is often proved by adapting earlier arguments for establishing SSM (see, e.g., [4, 24]). Recently, it was proved in [23] that SSM on trees implies spectral independence on large-girth graphs. We show that for *general* spin systems on $\mathbb{Z}^d$, SSM implies η-spectral independence with $\eta = O(1)$.

▶ **Lemma 8.** *For a spin system on a d-dimensional cube $V \subseteq \mathbb{Z}^d$, SSM implies η-spectral independence, where $\eta = O(1)$.*

The formal definition of SSM is given later in Section 4. Lemma 8 does not assume monotonicity for the spin system and could be of independent interest. An interesting consequence of this lemma, when combined with Theorem 6 is the following.

▶ **Corollary 9.** *Let $d \geq 2$. For a b-marginally bounded monotone spin system on a d-dimensional cube $V \subseteq \mathbb{Z}^d$, SSM implies that the mixing time of any systematic scan P_ϕ is $O(\log n)$.*

For the ferromagnetic Ising model on $\mathbb{Z}^2$, SSM is known to hold for all $\beta < \beta_c(2) = \ln(1+\sqrt{2})$ (see [17, 45, 2, 5]), so by Corollary 9 we deduce that when $\beta < \beta_c(2)$, the mixing time of any systematic scan P_ϕ on an n-vertex square box of $\mathbb{Z}^2$ is $O(\log n)$; note that $\beta_c(2) > \beta_u(2,2d)$, the corresponding tree uniqueness threshold.

We comment briefly on the techniques used to establish our results for the systematic scan dynamics. Our starting point is again the k-partite factorization of entropy (KPF). Our improved bounds for KPF imply that a global Markov chain that updates a random independent set of vertices in each step is rapidly mixing. We then use the censoring technique from [34, 10] to relate the mixing time of this Markov chain to that of the systematic scan dynamics. To establish Lemma 8, we use SSM to construct a contractive coupling for a particular Markov chain. Our Markov chain is similar to the one from [31], but modified to update rectangles instead of balls, and thus match the variant of SSM that holds up to the critical threshold for the Ising model on $\mathbb{Z}^2$. This contractive coupling is then used to establish spectral independence using the machinery from [6].

1.3 The block dynamics

Our final result concerns a family of Markov chains known as the *block dynamics*. They are a natural generalization of the Glauber dynamics where a random subset of vertices (instead of a random vertex) is updated in each step. More precisely, let $\mathcal{B} := \{B_1, \dots, B_K\}$ be a collection of subsets of vertices (called blocks) such that $V = \cup_{i=1}^K B_i$. Let α be a distribution over $\mathcal{B}$. The *(heat-bath) block dynamics* with respect to $(\mathcal{B}, \alpha)$ is the Markov chain that, in each step, given a spin configuration σ_t, selects $B_i \in \mathcal{B}$ according to the distribution α and updates the configuration on B_i with a sample from the $\mu(\cdot \mid \sigma_t(V \setminus B_i))$; that is, from the conditional distribution on B_i given the spins of σ_t in $V \setminus B_i$. We denote this Markov chain (and its transition matrix) by $P_{\mathcal{B},\alpha}$. When the B_i's are each single vertices, and α is a uniform distribution over the blocks in $\mathcal{B}$, we obtain the Glauber dynamics. Our result for the mixing time of the block dynamics is the following.

▶ **Theorem 10.** *Let $b > 0$, $\eta > 0$ and $\Delta \geq 3$. Suppose $G = (V, E)$ is an n-vertex graph of maximum degree Δ. Let μ be a Gibbs distribution of a totally-connected spin system on G. Let $\mathcal{B} := \{B_1, \dots, B_K\}$ be any collection of blocks such that $V = \cup_{i=1}^K B_i$, and let α be a distribution over $\mathcal{B}$. If μ is η-spectrally independent and b-marginally bounded, then there exists a constant $C > 0$ such that the mixing time of block dynamics $P_{\mathcal{B},\alpha}$ satisfies:*

$$T_{mix}(P_{\mathcal{B},\alpha}) = O\Big(\alpha_{min}^{-1} \cdot \Big(\frac{C\Delta \log n \log\log n}{b^7}\Big)^{2+\lceil \frac{2\eta}{b} \rceil}\Big),$$

where $\alpha_{min} = \min_{v \in V} \sum_{B \in \mathcal{B}} \alpha_B$.

Previous results for the block dynamics only apply to the bounded degree case [9, 17, 6], so Theorem 10 provides the first bounds for its mixing time in the unbounded degree setting.

Organization. The rest of the paper is organized as follows. In Section 2, we provide a number of definitions and background results. In Sections 3 and 4, we provide proof sketches for our results for the SW dynamics and the systematic scan dynamics; that is, Theorems 2 and 6, respectively. Some of our proofs are deferred to the full version of the paper [12].

2 Mixing times and modified log-Sobolev inequalities

Let P be an irreducible and aperiodic (i.e., ergodic) Markov chain with state space Ω and stationary distribution μ. Let us assume that P is reversible with respect to μ, and let

$$d(t) := \max_{x\in\Omega} \|P^t(x,\cdot) - \mu\|_{TV} := \max_{x\in\Omega} \max_{A\subseteq\Omega} |P^t(x, A) - \mu(A)|,$$

where $P^t(x,\cdot)$ denotes the distribution of the chain at time t assuming $x \in \Omega$ as the starting state; $\|\cdot\|_{TV}$ denotes the total variation distance. Note that with a slight abuse of notation we use P for both the Markov chain and its transition matrix. For $\varepsilon > 0$, let

$$T_{mix}(P, \varepsilon) := \min\{t > 0 : d(t) \leq \varepsilon\},$$

and the *mixing time of* P is defined as $T_{mix}(P) = T_{mix}(P, 1/4)$.

For functions $f, g : \Omega \to \mathbb{R}$, the *Dirichlet form* of a reversible Markov chain P with stationary distribution μ is defined as

$$\mathcal{E}_P(f, g) = \langle f, (I - P)g\rangle_\mu = \frac{1}{2} \sum_{x,y\in\Omega} \mu(x)P(x, y)(f(x) - f(y))(g(x) - g(y)),$$

where $\langle f, g\rangle_\mu := \sum_{x\in\Omega} f(x)g(x)\mu(x)$.

The spectrum of the ergodic and reversible Markov chain P is real, and we let $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_{|\Omega|} \geq -1$ denote its eigenvalues. The (absolute) spectral gap of P is defined by $\mathrm{GAP}(P) = 1 - \max\{|\lambda_2|, |\lambda_{|\Omega|}|\}$. When P is positive semidefinite, we have

$$\mathrm{GAP}(P) = 1 - \lambda_2 = \inf\left\{\frac{\mathcal{E}_P(f,f)}{\langle f,f\rangle_\mu} \mid f:\Omega \to \mathbb{R}, \langle f,f\rangle_\mu \neq 0\right\}.$$

For P reversible and ergodic, we have the following standard comparison between the spectral gap and the mixing time

$$T_{mix}(P,\varepsilon) = \frac{1}{\mathrm{GAP}(P)} \cdot \log\left(\frac{1}{\varepsilon\mu_{min}}\right), \tag{2}$$

where $\mu_{min} := \min_{x\in\Omega}\mu(x)$.

The expected value of a function $f:\Omega \to \mathbb{R}_{\geq 0}$ with respect to μ is defined as $\mathrm{E}_\mu[f] = \sum_{x\in\Omega} f(x)\mu(x)$. Similarly, the entropy of the function with respect to μ is given by

$$\mathrm{Ent}_\mu(f) := \mathrm{E}_\mu\left[f \log\frac{f}{\mathrm{E}_\mu[f]}\right] = \mathrm{E}_\mu[f\log f] - \mathrm{E}_\mu[f \log(\mathrm{E}_\mu[f])].$$

We say that the Markov chain P satisfies a *modified log-Sobolev inequality* (MLSI) with constant ρ if for every function $f:\Omega\to\mathbb{R}_{\geq 0}$,

$$\rho \cdot \mathrm{Ent}_\mu(f) \leq \mathcal{E}_P(f, \log f).$$

The smallest ρ satisfying the inequality above is called the *modified log-Sobolev constant* of P and is denoted by $\rho(P)$. A well-known general relationship (see [27, 13]) shows that

$$\frac{1-2\mu_{min}}{\log(1/\mu_{min}-1)}\mathrm{GAP}(P) \leq \rho(P) \leq 2\mathrm{GAP}(P). \tag{3}$$

For distributions μ and ν over Ω, the relative entropy of ν with respect to μ, denoted as $\mathcal{H}(\nu \mid \mu)$, is defined as $\mathcal{H}(\nu \mid \mu) := \sum_{x\in\Omega}\nu(x)\log\frac{\nu(x)}{\mu(x)}$. A Markov chain P with stationary distribution μ is said to satisfy discrete *relative entropy decay* with rate $r > 0$ if for all distributions ν:

$$\mathcal{H}(\nu P \mid \mu) \leq (1-r)\mathcal{H}(\nu \mid \mu). \tag{4}$$

It is a standard fact (see, e.g., Lemma 2.4 in [7]) that when (4) holds, then $\rho(P) \geq r$, and

$$T_{mix}(P,\varepsilon) \leq \frac{1}{r}\cdot\left(\log\log\left(\frac{1}{\mu_{min}}\right) + \log\left(\frac{1}{2\varepsilon}\right)\right). \tag{5}$$

3 Swendsen-Wang dynamics on general graphs

In this section, we consider the SW dynamics for the q-state ferromagnetic Potts models on general graphs. In particular, we establish Theorem 2 from the introduction, which is a direct corollary of the following more general result.

▶ **Theorem 11.** *Let $q \geq 2$, $\beta > 0$, $\eta > 0$, $b > 0$, $\Delta \geq 3$, and $\chi \geq 2$. Suppose $G = (V,E)$ is an n-vertex graph of maximum degree Δ and chromatic number χ. Let μ_{Potts} be the Gibbs distribution of the q-state ferromagnetic Potts model on G with parameter β. If μ_{Potts} is η-spectrally independent and b-marginally bounded, then there exists a universal constant $C > 1$ such that the modified log-Sobolev constant of the SW dynamics satisfies:*

$$\rho(P_{SW}) = \Omega\left(\frac{b^{7+6\kappa}}{\chi \cdot (C\Delta \log n)^{\kappa} \cdot (\log\log n)^{\kappa+1}}\right),$$

where $\kappa = 1 + \lceil \frac{2\eta}{b} \rceil$, *and*

$$T_{mix}(P_{SW}) = O\big(b^{-(7+6\kappa)} \cdot \chi \cdot (C\Delta \log n)^{\kappa}(\log\log n)^{\kappa+1} \cdot \log n\big).$$

Theorem 2 follows from this theorem by noting that $\chi \leq \Delta$ and that under the assumptions $\eta = O(1)$ and $\beta\Delta = O(1)$, we have $b = O(1)$ and $\kappa = O(1)$.

▶ Remark 12. When Δ is small, i.e., $\Delta = o(\log n)$, we can obtain slightly better bounds on $\rho(P_{SW})$ and $T_{mix}(P_{SW})$ and replace the $(C\Delta \log n \cdot \log\log n)^{\kappa}$ factor by a factor of $(C\Delta)^{6+4\lceil \frac{2\eta}{b} \rceil}$. This result is included in the full version of this paper [12].

Before proving Theorem 11, we provide a number of definitions and required background results in Section 3.1. We then sketch the proof of Theorem 11 in Sections 3.2 and include some applications of this result in Section 3.3.

3.1 Factorization of entropy

We present next several factorizations of the entropy functional $\mathrm{Ent}_\mu(f)$, which are instrumental in establishing the decay of the relative entropy for the SW dynamics. We introduce some useful notations first. For a pinning τ in $V \setminus \Lambda$ (i.e., $\tau \in \Omega_{V\setminus\Lambda}$), we let $\mu^\tau_\Lambda(\cdot) := \mu(\cdot \mid \sigma_{V\setminus\Lambda} = \tau)$. Given a function $f : \Omega \to \mathbb{R}_{\geq 0}$, subsets of vertices $B \subseteq \Lambda \subset V$, and $\tau \in \Omega_{V\setminus\Lambda}$, the function $f^\tau_B : \Omega^\tau_B \to \mathbb{R}_{\geq 0}$ is defined by:

$$f^\tau_B(\sigma) = \mathrm{E}_{\xi\sim\mu^\tau_{\Lambda\setminus B}}[f(\tau \cup \xi \cup \sigma)].$$

If $B = \Lambda$, we often write f^τ for f^τ_B, and if $\tau = \emptyset$, then we use f_B for f^τ_B. We use $\mathrm{Ent}^\tau_B(f^\tau)$ to denote $\mathrm{Ent}_{\mu^\tau_B}(f^\tau)$, and if the pinning τ on $V \setminus B$ is from a distribution π over $\Omega_{V\setminus B}$, we use $\mathrm{E}_{\tau\sim\pi}[\mathrm{Ent}^\tau_B(f^\tau)]$ to denote the expected value of the function f on S over the random pinning τ.

Various forms of entropy factorization arise from bounding $\mathrm{Ent}_\mu(f)$ by different (weighted) sums of restricted entropies of the function f. The first one we introduced, is the so-called *ℓ-uniform block factorization of entropy* of *ℓ-UBF*. For an integer $\ell \leq n$, ℓ-UBF holds for μ with constant C_{UBF} if for all functions $f : \Omega \to \mathbb{R}_{\geq 0}$,

$$\frac{\ell}{n} \cdot \mathrm{Ent}_\mu(f) \leq C_{\mathrm{UBF}} \cdot \frac{1}{\binom{n}{\ell}} \sum_{S\in\binom{V}{\ell}} \mathrm{E}_{\tau\sim\mu_{V\setminus S}}\left[\mathrm{Ent}^\tau_S(f^\tau)\right], \tag{6}$$

where $\binom{V}{\ell}$ denotes the collection of all subsets of V of size ℓ. An important special case is when $\ell = 1$, in which case (6) is called *approximate tensorization of entropy (AT)*; this special case has been quite useful for establishing optimal mixing time bounds for the Glauber dynamics in various settings (see, e.g., [47, 16, 18, 46]). The following result will be useful for us.

▶ **Theorem 13** ([25, 6]). *Let b and η be fixed. For $\theta \in (0,1)$ and $n \geq \frac{2}{\theta}(\frac{4\eta}{b^2}+1)$, the following holds. If the Gibbs distribution μ of a spin system on an n-vertex graph is η-spectrally independent and b-marginally bounded, then $\lceil\theta n\rceil$-UBF holds with $C_{\mathrm{UBF}} = (e/\theta)^{\lceil \frac{2\eta}{b} \rceil}$. In addition, if $\theta < b^2/(12\Delta)$, then:*

$$\mathrm{Ent}_\mu(f) \leq C_{\mathrm{UBF}} \cdot \frac{18}{b^5\theta} \sum_{i=1}^{n} \mathrm{E}_{\tau\sim\mu_{V\setminus\{i\}}}\left[\mathrm{Ent}^\tau_i(f^\tau)\right].$$

Note that the inequality in the theorem corresponds to AT with constant $C_{\mathrm{AT}} = C_{\mathrm{UBF}} \cdot \frac{18}{b^5\theta}$.

Another useful notion is *k-partite factorization of entropy* or KPF. Let $U_1, \ldots, U_k$ be k disjoint independent sets of V such that $\bigcup_{i=1}^k U_i = V$. We say μ satisfies KPF with constant C_{KPF} if for all functions $f : \Omega \to \mathbb{R}_{\geq 0}$,

$$\mathrm{Ent}_\mu(f) \leq C_{\mathrm{KPF}} \sum_{i=1}^{k} \mathrm{E}_{\tau \sim \mu_{V \setminus U_i}} \left[\mathrm{Ent}_{U_i}^{\tau}(f^\tau)\right].$$

KPF was introduced in [6], where it was used to analyze global Markov chains. The interplay between KPF and UBF is intriguing and is further explored in this paper.

3.2 Proof of main result for the SW dynamics: Theorem 11

The main technical contribution in the proof of Theorem 11 is establishing KPF with a better (i.e., smaller) constant C_{KPF}. As in [6], KPF is then used to derive an improved "edge-spin" factorization of entropy which is known to imply the desired bounds on the modified log-Sobolev constant and on the mixing time of the SW dynamics.

▶ **Theorem 14.** *For a b-marginally bounded Gibbs distribution μ that satisfies η-spectral independence on an n-vertex graph $G = (V, E)$ of maximum degree Δ, if b and η are constants independent of Δ and n, and $\Delta \in [3, \frac{b^4 n}{10e(4\eta+b^2)}]$, then there exists an absolute constant $c > 0$ such that k-partite factorization of entropy holds for μ with constant $C_{\mathrm{KPF}} = (\Delta \log n)^c$. Specifically, for a set of k disjoint independent sets $V_1, \ldots, V_k$ such that $\bigcup_{j=1}^k V_j = V$,*

$$\mathrm{Ent}_\mu(f) \leq 54 \cdot \frac{e^{13\kappa}}{b^{5+6\kappa}} \cdot (\Delta \log n)^\kappa \cdot (\log\log n)^{1+\kappa} \sum_{j=1}^{k} \mathrm{E}_{\tau \sim \mu_{V \setminus V_j}} [\mathrm{Ent}_{V_j}^{\tau}(f^\tau)], \tag{7}$$

where $\kappa = 1 + \lceil \frac{2\eta}{b} \rceil$. Moreover, if $\Delta^2 \leq \frac{b^4 n}{10e(4\eta+b^2)}$, then the following also holds

$$\mathrm{Ent}_\mu(f) \leq 72 \cdot \frac{e^{8\kappa}}{b^{5+4\kappa}} \cdot \Delta^{2+4\kappa} \sum_{j=1}^{k} \mathrm{E}_{\tau \sim \mu_{V \setminus V_j}} [\mathrm{Ent}_{V_j}^{\tau}(f^\tau)]. \tag{8}$$

This Theorem is proved in the full version [12].

▶ Remark 15. Let $\mathcal{B} = \{B_1, \ldots, B_k\}$ be a collection of disjoint independent sets such that $V = \bigcup_{i=1}^k B_i$. The independent set dynamics $P_{\mathcal{B}}$ is a heat-bath block dynamics w.r.t. $\mathcal{B}$ and a uniform distribution over $\mathcal{B}$. If μ satisfies k-partite factorization of entropy with C_{KPF}, then $P_{\mathcal{B}}$ satisfies a relative entropy decay with rate $r \geq 1/(k \cdot C_{\mathrm{KPF}})$.

As mentioned, KPF was first studied in [6]; the constant proved there was

$$C_{\mathrm{KPF}} = b^{O(\Delta)} \cdot (b\Delta)^{O(\eta/b)},$$

so our new bound improves the dependence on Δ from exponential to polynomial.

With KPF on hand, the next step in the proof of Theorem 11 relies on the so-called edge-spin factorization of entropy. Let $\Omega_J := \Omega \times \{0,1\}^E$ be the set of joint configurations (σ, A) corresponding to pairs of a spin configuration $\sigma \in \Omega$ and an *edge configuration* (a subset of edges in a graph) $A \subseteq E$. For a q-state Potts model μ_{Potts} with parameter $p = 1 - e^{-\beta}$, we use ν to denote the *Edwards-Sokal* measure on Ω_J given by

$$\nu(\sigma, A) := \frac{1}{Z_J}(1-p)^{|E|-|A|} p^{|A|} \mathbf{1}(\sigma \sim A),$$

where $\sigma \sim A$ is the event that every edge in A has its two endpoints with the same spin in σ, and $Z_J := \sum_{(A,\sigma)\in\Omega_J}(1-p)^{|E|-|A|}p^{|A|}\mathbf{1}(\sigma \sim A)$ is a normalizing constant. Let $\nu(\cdot \mid \sigma)$ and $\nu(\cdot \mid A)$ denote the conditional measures obtained from ν by fixing the spin configuration to be σ or fixing the edge configuration to be A respectively. For a function $f : \Omega_J \to \mathbb{R}_{\geq 0}$, let $f^\sigma : \{0,1\}^{|E|} \to \mathbb{R}_{\geq 0}$ be the function given by $f^\sigma(A) = f(\sigma \cup A)$, and let $f^A : \Omega \to \mathbb{R}_{\geq 0}$ be the function given by $f^A(\sigma) = f(\sigma \cup A)$. We say that *edge-spin factorization of entropy* holds with constant C_{ES} if for all functions $f : \Omega_J \to \mathbb{R}_{\geq 0}$,

$$\mathrm{Ent}_\nu(f) \leq C_{\mathrm{ES}}\left(\mathrm{E}_{(\sigma,A)\sim\nu}\left[\mathrm{Ent}_{A\sim\nu(\cdot|\sigma)}(f^\sigma)\right] + \mathrm{E}_{(\sigma,A)\sim\nu}\left[\mathrm{Ent}_{\sigma\sim\nu(\cdot|A)}(f^A)\right]\right). \tag{9}$$

The following result from [6] will be useful for us.

▶ **Lemma 16** (Theorem 6.1 [6]). *Suppose the q-state ferromagnetic Potts model with parameter β on a graph G of maximum degree is $\Delta \geq 3$ satisfies KPF with constant C_{KPF}. Then, the edge-spin factorization of entropy holds with constant $C_{\mathrm{ES}} = O(\beta\Delta k e^{\beta\Delta}) \cdot C_{\mathrm{KPF}}$.*

The final ingredient in the proof of Theorem 11 is the following.

▶ **Lemma 17** (Lemma 1.8 [7]). *Suppose edge-spin factorization of entropy holds with constant C_{ES}. Then, the SW dynamics P_{SW} satisfies the relative entropy decay with rate $\Omega\left(\frac{1}{C_{\mathrm{ES}}}\right)$.*

We are now ready to prove Theorem 11. Since Theorem 14 requires an upper bound on the maximum degree Δ, when $\Delta = \Omega(n)$ we use a crude comparison argument to obtain a polynomial bound for the modified log-Sobolev constant and mixing time of the SW dynamics.

Proof of Theorem 11. First, we assume $\Delta \in [3, \frac{b^4 n}{10e(4\eta+b^2)}]$. By Theorem 14, μ_{Potts} satisfies χ-partite factorization of entropy with constant

$$C_{\mathrm{KPF}} = (\Delta \log n)^\kappa (\log\log n)^{1+\kappa} \cdot O\left(\frac{e^{13\kappa}}{b^{5+6\kappa}}\right).$$

It follows from Lemma 16 and Lemma 17 that the SW dynamics satisfies (4) with

$$r = \Omega\left(\frac{b^{5+6\kappa}}{\chi\beta\Delta e^{\beta\Delta} \cdot (\Delta\log n)^\kappa (\log\log n)^{1+\kappa} \cdot e^{13\kappa}}\right).$$

Note that $b \leq q^{-1}e^{-\beta\Delta}$, and so $\beta\Delta e^{\beta\Delta} \leq e^{2\beta\Delta} \leq b^{-2}$. Therefore, the mixing time bound follows from (5).

Next, let us consider the case when $\Delta = \Omega(n)$. In this case, it suffices to provide a $1/\mathrm{poly}(n)$ lower bound on the modified log-Sobolev constant of the SW dynamics, which can be obtained in a straightforward manner using the known bounds for the Potts Glauber dynamics and the comparison technology from [8].

Recall that $P_{\mathcal{B}}$ is the independent set dynamics; that is, the block dynamics with respect to a collection of disjoint independent sets $\{B_1, \dots, B_k\}$; see Remark 15. From Theorem 3.2 in [32], we know that $\mathrm{GAP}(P_{GD}) \geq n^{-(2\eta+1)}$, where P_{GD} denotes the Potts Glauber dynamics. Since $\mathcal{E}_{P_{GD}}(f,f) \leq \mathcal{E}_{P_{\mathcal{B}}}(f,f)$ for any function f, it follows that $\mathrm{GAP}(P_{\mathcal{B}}) \geq n^{-(2\eta+1)}$. In addition, the comparison inequalities from [8] imply that

$$\mathrm{GAP}(P_{SW}) \geq \mathrm{GAP}(P_{\mathcal{B}}) \cdot \min_{i=1,\dots,k}\ \min_{\tau\in\Omega_{V\setminus B_i}}\ \min_{v\in B_i} \mathrm{GAP}(P_v^\tau),$$

where P_v^τ is the transition matrix for the update at vertex v, with τ as the fixed boundary condition, that adds each monochromatic edge between v and its neighbors independently with probability $p := 1 - e^{-\beta}$, and assigns a new random spin to v only if no edge is added. From a simple coupling argument it follows that for any $v \in B_i$, $\mathrm{GAP}(P_v^\tau) \geq (1-p)^\Delta = e^{-\beta\Delta} \geq qb$. Thus, $\mathrm{GAP}(P_{SW}) \geq n^{-(2\eta+1)}qb$, and $\rho(P_{SW}) = \Omega(n^{-(2\eta+2)}b)$ by (3). The mixing time bound follows from (2). ◀

3.3 Applications of Theorem 11

In this section, we prove Corollary 3 from the introduction and present another application of Theorem 11 concerning the SW dynamics on a random graph generated from the classical Erdős-Rényi $G(n,p)$ model. For this, we first define Dobrushin's influence matrix.

▶ **Definition 18.** *The* Dobrushin influence matrix $A \in \mathbb{R}^{n\times n}$ *is defined by* $A(u,u) = 0$ *and for* $u \neq v$,

$$A(u,v) = \max_{(\sigma,\tau)\in S_{u,v}} d_{TV}(\mu_v(\cdot \mid \sigma), \mu_v(\cdot \mid \tau)),$$

where $S_{u,v}$ *contains the set of all pairs of partial configurations* (σ,τ) *in* $\Omega_{V\setminus\{v\}}$ *that can only disagree at* u*, namely,* $\sigma_w = \tau_w$ *if* $w \neq u$*.*

It is known that an upper bound on the spectral norm of A implies spectral independence. In particular, we have the following result from [6].

▶ **Proposition 19** (Theorem 1.13, [6])**.** *If the Dobrushin influence matrix* A *of a distribution* μ *satisfies* $\|A\| \leq 1-\varepsilon$ *for some* $\varepsilon > 0$*, then* μ *is spectral independent with constant* $\eta = 2/\varepsilon$*.*

For the ferromagnetic Ising model, $\beta_u(\Delta) := \ln \frac{\Delta}{\Delta-2}$ corresponds to the threshold value of the parameter β for the uniqueness/non-uniqueness phase transition on the Δ-regular tree. For the anti-ferromagnetic Ising model, the phase transition occurs at $\bar{\beta}_u(\Delta) := -\ln \frac{\Delta}{\Delta-2}$. If $\bar{\beta}_u(\Delta)(1-\delta) < \beta < \beta_u(\Delta)(1-\delta)$, we say the Ising model satisfies the δ*-uniqueness condition*. On a bounded degree graph, $\|A\| \leq 1-\delta$ for the Ising model is a strictly stronger condition than δ-uniqueness condition. However, due to the observation made in [3], if $\Delta \to \infty$, the two conditions are roughly equivalent.

▶ **Proposition 20.** *The Ising model with parameter* $\bar{\beta}_u(\Delta)(1-\delta) < \beta < \beta_u(\Delta)(1-\delta)$ *and* $\Delta \to \infty$ *satisfies* $\|A\| \leq 1-\delta/2$*.*

This proposition is proved in the full version [12]; we show next that Corollary 3 follows from Theorem 11. For this, we first restate the corollary in a more precise manner.

▶ **Corollary 21.** *Let* $\delta \in (0,1)$ *and* $\Delta \geq 3$*. For the ferromagnetic Ising model with* $\beta \leq (1-\delta)\beta_u(\Delta)$ *on any graph* G *of maximum degree* Δ *and chromatic number* χ*, or for the ferromagnetic* q*-state Potts model with* $q \geq 3$ *and* $\beta \leq \frac{2(1-\delta)}{\Delta}$ *on the same graph, the mixing time of the SW dynamics satisfies*

$$T_{mix}(P_{SW}) = O\big(\chi \cdot \Delta^{\kappa} \cdot (\log n \log\log n)^{1+\kappa}\big),$$

where $\kappa = 1 + \lceil \frac{4qe^2}{\delta} \rceil$*.*

Proof. If $\Delta = O(1)$, then the corollary was proved in a stronger form in [6]. Thus, we assume $\Delta \to \infty$.

We first show spectral independence. Let $q = 2$. Under the δ-uniqueness condition $0 < \beta < (1-\delta)\beta_u(\Delta)$, by Proposition 20 and Proposition 19, the Ising model μ_{Ising} satisfies $(4/\delta)$-spectral independence. For the q-state Potts model with $q \geq 3$, the Dobrushin influence matrix corresponding to μ_{Potts} satisfies $\|A\| \leq \frac{1}{2}\beta\Delta$; see proof of Theorem 2.13 in [54]. Thus, if $\beta \leq \frac{2(1-\delta)}{\Delta}$, then $\|A\| \leq 1-\delta$, and by Proposition 19, μ_{Potts} satisfies $(2/\delta)$-spectral independence.

Letting $N(v)$ denote the neighborhood of v, and noting that for any configuration η on $N(v)$ we have $\mu(\sigma_v = c \mid \sigma_{N(v)} = \eta) \geq 1/(qe^2)$, we deduce that μ_{Potts} and μ_{Ising} are both $(1/(qe^2))$-marginally bounded. Therefore, by noting that $\kappa = 1 + \lceil \frac{4qe^2}{\delta} \rceil$ is a constant that only depends on δ, the mixing time bound follows from Theorem 11

$$T_{mix}(P_{SW}) = O\big(b^{-(7+6\kappa)} \cdot \chi \cdot (C\Delta \log n)^{\kappa} (\log\log n)^{\kappa+1} \cdot \log n\big) = O\big(\chi \cdot \Delta^{\kappa} \cdot (\log n \log\log n)^{1+\kappa}\big),$$

as desired. ◀

3.3.1 The SW dynamics on random graphs

As another application of Theorem 11, we consider the SW dynamics on a random graph generated from the classical $G(n, \frac{d}{n})$ model in which each edge is included independently with probability $p = d/n$; we consider the case where d is a constant independent of n. In this setting, while a typical graph has $\tilde{O}(n)$ edges, its maximum degree is of order $\Theta(\frac{\log n}{\log\log n})$ with high probability. Our results imply that the SW dynamics has polylogarithmic mixing on this type of graph provided β is small enough.

▶ **Corollary 22.** *Let $\delta \in (0,1)$ and $d \in \mathbb{R}_{\geq 0}$ be constants independent of n. Suppose that $G \sim G(n, d/n)$ and G has maximum degree Δ. For the ferromagnetic Ising model with parameter $\beta < (1-\delta)\beta_u(\Delta)$ on G or the ferromagnetic q-Potts model with $q \geq 2$ and $\beta \leq \frac{2(1-\delta)}{\Delta}$ on the same graph, the SW dynamics has $(\log n)^{3+2\lceil \frac{4qe^2}{\delta} \rceil} \cdot O(\log\log n)$ mixing time, with high probability over the choice of the random graph G.*

Corollary 22 is established using Corollary 21 and the following fact about random graphs. The full proof is provided in the full version [12].

▶ **Proposition 23** ([1]). *Let $G \sim G(n, \frac{d}{n})$ for a fixed $d \in \mathbb{R}_{\geq 0}$, and let χ be the chromatic number of G. With high probability over the choice of G, $\chi = k_d$ or $\chi = k_d + 1$, where k_d is the smallest integer k such that $d < 2k \log k$.*

4 Systematic scan dynamics

In this section, we study the systematic scan dynamics for general spin systems, which we define next.

▶ **Definition 24** (Spin system). *Let $G = (V, E)$ be a graph and $\mathcal{S} = \{1, \dots, q\}$ a set of spins. Let $\Omega \subseteq \mathcal{S}^V$ be the set of possible spin configurations on G. We write σ_v for the spin assigned to v by σ. Given a configuration $\sigma \in \Omega$ and a subset Λ of V, we write $\sigma_\Lambda \in \mathcal{S}^\Lambda$ for the configuration of σ restricted to Λ. For a subset of vertices $\Lambda \subseteq V$, a* boundary condition *τ is an assignment of spins to (some) vertices in outer vertex boundary $\partial\Lambda \subseteq V \setminus \Lambda$ of Λ; namely, $\tau : (\partial\Lambda)_\tau \to \mathcal{S}$, with $(\partial\Lambda)_\tau \subseteq \partial\Lambda$. Note that a boundary condition is simply a pinning of a subset of vertices identified as being in the boundary of G. Given a boundary condition $\tau : (\partial V)_\tau \to \mathcal{S}$, the* Hamiltonian *$H : \Omega \to \mathbb{R}$ of a spin system is defined as*

$$H(\sigma) = -\sum_{\{v,u\}\in E} K(\sigma_v, \sigma_u) - \sum_{\{v,u\}\in E: u\in V, v\in(\partial V)_\tau} K(\sigma_v, \tau_v) - \sum_{v\in V} U(\sigma_v), \tag{10}$$

where $K : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ and $U : \mathcal{S} \to \mathbb{R}$ are respectively the symmetric edge interaction potential function *and the* spin potential function *of the system. The* Gibbs distribution *of a spin system with Hamiltonian H is defined as*

$$\mu(\sigma) = \frac{1}{Z_H} e^{-H(\sigma)},$$

where $Z_H := \sum_{\sigma\in\Omega} e^{-H(\sigma)}$. We use Ω for the set of configurations σ satisfying $\mu(\sigma) > 0$.

The Potts model, as defined in the introduction, corresponds to the spin system with $q \geq 2$, $K(x,y) = \beta \cdot \mathbb{1}(x = y)$, and $U(\sigma_v) = 0$ for all $v \in V$. In this section, we focus on the ferromagnetic Ising model where $\beta > 0$ and $\mathcal{S} = \{-1, +1\}$. Another important spin system is the hardcore model that can be defined by setting $\mathcal{S} = \{1, 0\}$, $K(x,y) = \infty$ if $x = y = 1$ and $K(x,y) = 0$ otherwise, and $U(x) = \mathbb{1}(x = 1) \cdot \ln \lambda$, where $\lambda > 0$ is referred to as the *fugacity* parameter of the model.

We restrict attention to *totally-connected* spin systems, as this ensures that the Glauber dynamics, the systematic scan dynamics, and the block dynamics are all irreducible Markov chains (and thus ergodic).

▶ **Definition 25.** *For a subset $\mathcal{C}_U$ of partial configurations on $U \subseteq V$, let $H[\mathcal{C}_U] = (\mathcal{C}_U, E[\mathcal{C}_U])$ be the induced subgraph where $E[\mathcal{C}_U]$ consists of all pairs of configurations on $\mathcal{C}_U$ that differ at exactly one vertex. We say that $\mathcal{C}_U$ is connected when $H[\mathcal{C}_U]$ is connected. For a pinning τ on $\Lambda \subseteq V$, we say $\Omega^\tau_{V \setminus \Lambda}$ is connected if $H[\Omega^\tau_{V \setminus \Lambda}]$ is connected. A distribution μ over $\mathcal{S}^V$ is* totally-connected *if for every $\Lambda \subseteq V$ and every pinning τ on Λ, $\Omega^\tau_{V \setminus \Lambda}$ is connected.*

Given an ordering $\phi = [v_1, \dots, v_n]$ of the vertices, a systematic scan dynamics performs heat-bath updates on $v_1, \dots, v_n$ sequentially in this order. Recall that a heat-bath update on v_i simply means the replacement of the spin on v_i by a new spin assignment generated according to the conditional distribution in v_i given the configuration in $V \setminus \{v_i\}$. Let $P_i \in \mathbb{R}^{|\Omega| \times |\Omega|}$ be the transition matrix corresponding to a heat-bath update on the vertex v_i. The transition matrix of the systematic scan dynamics for the ordering ϕ can be written as $\mathcal{S}_\phi := P_n \dots P_1$. In general, $\mathcal{S}_\phi$ is not reversible, so as in earlier works we work with the symmetrized version of the scan dynamics that updates the spins in the order ϕ and in addition updates the spins in the reverse order of ϕ [33, 48]. The transition matrix of the symmetrized systematic scan dynamics can then be written as

$$P_\phi := \prod_{i=1}^{n} P_i \prod_{i=0}^{n-1} P_{n-i}.$$

Henceforth, we only consider the symmetrized version of the dynamics. Since P_ϕ is a symmetrized product of reversible transition matrices, one can straightforwardly verify its reversibility with respect to μ; its ergodicity follows from the assumption that the spin system is totally-connected (see Definition 25).

We show tight mixing time bounds for P_ϕ for monotone spin systems (see Definition 4). Our main result for the systematic scan dynamics is Theorem 6 from the introduction; its proof is included in the full version of this paper [12]. We complement Theorem 6 with a lower bound for the mixing time of systematic scan dynamics for a particular ordering ϕ. Specifically, on a bipartite graph $G = (V_E \cup V_O, E)$, an *even-odd scan dynamics P_{EOE}* is a systematic scan dynamics with respect to an ordering ϕ such that v_e appears before v_o in ϕ for all $v_e \in V_E$ and $v_o \in V_O$. In other words,

$$P_\phi = \prod_{i: v_i \in V_E} P_i \prod_{i: v_i \in V_O} P_i \prod_{i: v_i \in V_O} P_i \prod_{i: v_i \in V_E} P_i.$$

The above expression is well-defined without specifying the ordering in which the vertices in V_E and V_O are updated since the updates commute.

▶ **Lemma 26.** *Let Δ be a constant and let G be an n-vertex connected bipartite graph with maximum degree Δ. The even-odd scan dynamics P_{EOE} for the ferromagnetic Ising model on G has mixing time $T_{mix}(P_{EOE}) = \Omega(\log n)$.*

The lower bound in Lemma 26 is proved in the full version of this paper [12] using the machinery from [42] and the fact that even-odd scan dynamics does not propagate disagreements quickly (under a standard coupling). Our proof can thus be extended to other scan orderings that propagate disagreements slowly; however, there are orderings that do propagate disagreements quickly (think of a box in $\mathbb{Z}^2$ with the vertices sorted in a "spiral" from the boundary of the box to its center). For this type of ordering, the technique does not provide the $\Omega(\log n)$ lower bound. In addition, while we focus on the ferromagnetic Ising model to ensure clarity in the proof, the established lower bound is expected to apply to a broader class of spin systems.

4.1 Applications of Theorem 6

We discuss next some applications of Theorem 6. As a first application, we can establish *optimal* mixing for the systematic scan dynamics on the ferromagnetic Ising model under the δ-uniqueness condition, improving the best known results that hold under the Dobrushin-type conditions [51, 28, 41]. This result was stated in Corollary 7 in the introduction and is proved next. For this, we recall that under δ-uniqueness condition, the Ising distribution μ_{Ising} satisfies spectral independence and the bounded marginals condition.

▶ **Proposition 27** ([24, 25])**.** *The ferromagnetic Ising model with parameter β such that $\bar{\beta}_u(\Delta)(1-\delta) < \beta < \beta_u(\Delta)(1-\delta)$ is $O(1/\delta)$-spectrally independent and b-marginally bounded with $b = O(1)$.*

Proof of Corollary 7. We fix $\delta \in (0,1)$ and first assume that Δ is a constant. By Proposition 27, the ferromagnetic Ising model with parameter $\beta < (1-\delta)\beta_u(\Delta)$ satisfies η-spectral independence and b-bounded marginals, where $\eta = O(1/\delta)$ and b is a constant. Since the ferromagnetic Ising model is a monotone system, it follows from Theorem 6 that $T_{mix} = O(\log n)$ for any ordering ϕ.

Now, when $\Delta \to \infty$ as $n \to \infty$, by Proposition 20, the Dobrushin's influence matrix A of ferromagnetic Ising model satisfies that $\|A\| \leq 1 - \delta/2$. Under this assumption, it is known that $T_{mix} = O(\log n)$ for any ordering ϕ; see [41]. ◀

We can similarly show mixing time bound for the systematic scan dynamics of the hardcore model on bipartite graphs under δ-uniqueness condition.

▶ **Corollary 28.** *Let $\delta \in (0,1)$ be a constant. Suppose G is an n-vertex bipartite graph of maximum degree $\Delta \geq 3$. For the hardcore model on G with fugacity λ such that $0 < \lambda < (1-\delta)\lambda_u(\Delta)$, where $\lambda_u(\Delta) = \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta}}$ is the tree uniqueness threshold on the Δ-regular tree, the systematic scan with respect to any ordering ϕ satisfies*

$$T_{mix}(P_\phi) = \Delta^{O(1/\delta)} \cdot O(\log n).$$

Proof of Corollary 28. The hardcore model on a bipartite graph $(V_1 \cup V_2, E)$ with fugacity $0 < \lambda < (1-\delta)\lambda_u(\Delta)$ is monotone, and [25, 3, 21] show that it satisfies $O(1/\delta)$-spectral independence and the $O(\lambda)$-bounded marginals condition. Theorem 6 then implies $\Delta^{O(1/\delta)} \cdot O(\log n)$ mixing of systematic scan for any ordering. ◀

We consider next the application of Theorem 6 to the special case where the underlying graph is a cube of the d-dimensional lattice graph $\mathbb{Z}^d$. We show that strong spatial mixing implies optimal $O(\log n)$ mixing of any systematic scan dynamics. Previously, under the same type of condition, [8] gave an $O(\log n (\log\log n)^2)$ mixing time bound for arbitrary orderings, and an $O(\log n)$ mixing time bound for a special class of scans that (deterministically) propagate disagreements slowly under the standard identity coupling. We first provide the definition of our SSM condition.

▶ **Definition 29.** *We say a spin system μ on $\mathbb{Z}^d$ satisfies the* strong spatial mixing (SSM) *condition if there exist constants $\alpha, \gamma, L > 0$ such that for every d-dimensional rectangle $\Lambda \subset \mathbb{Z}^d$ of side length between L and $2L$ and every subset $B \subset \Lambda$, with any pair (τ, τ') of boundary configurations on $\partial\Lambda$ that only differ at a vertex u, we have*

$$\|\mu_B^\tau(\cdot) - \mu_B^{\tau'}(\cdot)\|_{TV} \leq \gamma \cdot \exp(-\alpha \cdot dist(u, B)),$$

where $dist(\cdot, \cdot)$ denotes graph distance.

The definition above differs from other variants of SSM in the literature (e.g., [31, 8, 45]) in that Λ has been restricted to "regular enough" rectangles. In particular, our variant of SSM is easier to satisfy than those in [31, 45] but more restricting than the one in [8] (that only considers squares). Nevertheless, it follows from [17, 45, 2, 5] that for the ferromagnetic Ising model, this form of SSM holds up to a critical threshold temperature $\beta < \beta_c(2) = \ln(1 + \sqrt{2})$ on $\mathbb{Z}^2$.

Corollary 9 from the introduction states that for b-marginally bounded monotone spin system on d-dimensional cubes $V \subseteq \mathbb{Z}^d$, SSM implies that the mixing time of any systematic scan P_ϕ is $O(\log n)$. As mentioned there, this result in turn implies that any systematic scan dynamics for the ferromagnetic Ising model is mixing in $O(\log n)$ steps on boxes of $\mathbb{Z}^2$ when $\beta < \beta_c(2)$. Another interesting consequence of Corollary 9 is that we obtain $O(\log n)$ mixing time for any systematic scan dynamics P_ϕ for the hardcore model on $\mathbb{Z}^2$ when $\lambda < 2.538$, which is the best known condition for ensuring SSM [52, 50].

Our proof of Corollary 9 relies on Lemma 8 that is restated below. The proof of Lemma 8 is provided in the full version of this paper [12]. Remarkably, Lemma 8 generalizes beyond monotone systems and may be of independent interests.

▶ **Lemma 8.** *For a spin system on a d-dimensional cube $V \subseteq \mathbb{Z}^d$, SSM implies η-spectral independence, where $\eta = O(1)$.*

Proof of Corollary 9. Assume a monotone spin system satisfies SSM condition. Then the spin system satisfies η-spectral independence, where $\eta = O(1)$ by Lemma 8. By noting that $\Delta = 2^d$ the corollary follows from Theorem 6. ◀

References

1 Dimitris Achlioptas and Assaf Naor. The two possible values of the chromatic number of a random graph. *Annals of Mathematics*, 162(3):1335–1351, 2005.

2 Kenneth S. Alexander. On weak mixing in lattice models. *Probab. Theory Relat. Fields*, 110(441-471), 1998.

3 Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. Entropic independence: Optimal mixing of down-up random walks. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2022, pages 1418–1430, New York, NY, USA, 2022. Association for Computing Machinery. doi:10.1145/3519935.3520048.

4 Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1319–1330, 2020. doi:10.1109/FOCS46700.2020.00125.

5 Vincent Beffara and Hugo Duminil-Copin. The self-dual point of the two-dimensional random-cluster model is critical for $q \geq 1$. *Probab. Theory Relat. Fields*, 153(511-542), 2012.

6 Antonio Blanca, Pietro Caputo, Zongchen Chen, Daniel Parisi, Daniel Stefankovic, and Eric Vigoda. On mixing of Markov chains: coupling, spectral independence, and entropy factorization. *Electronic Journal of Probability*, 27:1–42, 2022. `doi:10.1214/22-EJP867`.

7 Antonio Blanca, Pietro Caputo, Daniel Parisi, Alistair Sinclair, and Eric Vigoda. Entropy decay in the Swendsen–Wang dynamics on $\mathbb{Z}^d$. *The Annals of Applied Probability*, 32(2):1018–1057, 2022. `doi:10.1214/21-AAP1702`.

8 Antonio Blanca, Pietro Caputo, Alistair Sinclair, and Eric Vigoda. Spatial mixing and nonlocal markov chains. *Random Structures & Algorithms*, 55(3):584–614, 2019. `doi:10.1002/rsa.20844`.

9 Antonio Blanca, Zongchen Chen, Daniel Stefankovic, and Eric Vigoda. The Swendsen–Wang dynamics on trees. *Random Structures & Algorithms*, 2023.

10 Antonio Blanca, Zongchen Chen, and Eric Vigoda. Swendsen-wang dynamics for general graphs in the tree uniqueness region. *Random Structures & Algorithms*, 56(2):373–400, 2020.

11 Antonio Blanca and Alistair Sinclair. Dynamics for the mean-field random-cluster model. In *Proceedings of APPROX/RANDOM*, 2015.

12 Antonio Blanca and Xusheng Zhang. Rapid mixing of global markov chains via spectral independence: the unbounded degree case. *arXiv preprint arXiv:2307.00683*, 2023.

13 Sergey Bobkov and Prasad Tetali. Modified log-sobolev inequalities, mixing and hypercontractivity. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, STOC '03, pages 287–296, New York, NY, USA, 2003. Association for Computing Machinery. `doi:10.1145/780542.780586`.

14 Magnus Bordewich, Catherine Greenhill, and Viresh Patel. Mixing of the glauber dynamics for the ferromagnetic potts model. *Random Structures & Algorithms.*, 48(1):21–52, January 2016. URL: `http://dro.dur.ac.uk/15252/`.

15 Christian Borgs, Jennifer T. Chayes, and Prasad Tetali. Tight bounds for mixing of the Swendsen-Wang algorithm at the Potts transition point. *Probab. Theory Relat. Fields*, 152(3-4):509–557, 2012. `doi:10.1007/s00440-010-0329-0`.

16 Pietro Caputo, Georg Menz, and Prasad Tetali. Approximate tensorization of entropy at high temperature, 2014. `doi:10.48550/arXiv.1405.0608`.

17 Pietro Caputo and Daniel Parisi. Block factorization of the relative entropy via spatial mixing. *Communications in Mathematical Physics*, 388(2):793–818, October 2021. `doi:10.1007/s00220-021-04237-1`.

18 Filippo Cesi. Quasi-factorization of the entropy and logarithmic Sobolev inequalities for Gibbs random fields. *Probability Theory and Related Fields*, 120:569–584, 2001.

19 Xiaoyu Chen, Weiming Feng, Yitong Yin, and Xinyuan Zhang. Optimal mixing for two-state anti-ferromagnetic spin systems. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 588–599. IEEE, 2022.

20 Xiaoyu Chen, Weiming Feng, Yitong Yin, and Xinyuan Zhang. Rapid mixing of glauber dynamics via spectral independence for all degrees. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 137–148, 2022. `doi:10.1109/FOCS52979.2021.00022`.

21 Xiaoyu Chen, Jingcheng Liu, and Yitong Yin. Uniqueness and rapid mixing in the bipartite hardcore model, 2023. `arXiv:2305.00186`.

22 Zongchen Chen, Andreas Galanis, Daniel Stefankovic, and Eric Vigoda. Rapid mixing for colorings via spectral independence. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1548–1557, 2021.

23 Zongchen Chen, Kuikui Liu, Nitya Mani, and Ankur Moitra. Strong spatial mixing for colorings on trees and its algorithmic applications, 2023. `arXiv:2304.01954`.

24 Zongchen Chen, Kuikui Liu, and Eric Vigoda. Rapid mixing of Glauber dynamics up to uniqueness via contraction. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1307–1318. IEEE Computer Society, 2020.

25 Zongchen Chen, Kuikui Liu, and Eric Vigoda. 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 2021, pages 1537–1550, New York, NY, USA, 2021. Association for Computing Machinery. doi:10.1145/3406325.3451035.

26 Amin Coja-Oghlan, Andreas Galanis, Leslie Ann Goldberg, Jean Bernoulli Ravelomanana, Daniel Štefankovič, and Eric Vigoda. Metastability of the Potts ferromagnet on random regular graphs. *Communications in Mathematical Physics*, 2023. doi:10.1007/s00220-023-04644-6.

27 P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *The Annals of Applied Probability*, 6(3):695–750, 1996. doi:10.1214/aoap/1034968224.

28 Martin Dyer, Leslie Ann Goldberg, and Mark Jerrum. Dobrushin conditions and systematic scan. In Josep Díaz, Klaus Jansen, José D. P. Rolim, and Uri Zwick, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 327–338, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

29 Martin Dyer, Leslie Ann Goldberg, and Mark Jerrum. Systematic scan for sampling colorings. *The Annals of Applied Probability*, 16(1):185–230, 2006.

30 Martin Dyer, Leslie Ann Goldberg, and Mark Jerrum. Matrix norms and rapid mixing for spin systems. *The Annals of Applied Probability*, 19(1):71–107, 2009.

31 Martin Dyer, Alistair Sinclair, Eric Vigoda, and Dror Weitz. Mixing in time and space for lattice spin systems: A combinatorial view. *Random Structures & Algorithms*, 24(4):461–479, 2004. doi:10.1002/rsa.20004.

32 Weiming Feng, Heng Guo, Yitong Yin, and Chihao Zhang. Rapid mixing from spectral independence beyond the boolean domain. *ACM Trans. Algorithms*, 18(3), October 2022. doi:10.1145/3531008.

33 James Allen Fill. Eigenvalue Bounds on Convergence to Stationarity for Nonreversible Markov Chains, with an Application to the Exclusion Process. *The Annals of Applied Probability*, 1(1):62–87, 1991. doi:10.1214/aoap/1177005981.

34 James Allen Fill and Jonas Kahn. Comparison inequalities and fastest-mixing Markov chains. *The Annals of Applied Probability*, 23(5):1778–1816, 2013. doi:10.1214/12-AAP886.

35 Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Swendsen-Wang algorithm on the mean-field Potts model. In *Proceedings of APPROX/RANDOM*, 2015.

36 Reza Gheissari and Eyal Lubetzky. Mixing times of critical two-dimensional Potts models. *Comm. Pure Appl. Math*, 71(5):994–1046, 2018.

37 Reza Gheissari, Eyal Lubetzky, and Yuval Peres. Exponentially slow mixing in the mean-field Swendsen–Wang dynamics. *Annales de l'Institut Henri Poincare (B)*, 2019. Extended abstract appeared in Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2018), pp. 1981–1988.

38 Vivek K. Gore and Mark R. Jerrum. The Swendsen-Wang process does not always mix rapidly. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '97, pages 674–681, New York, NY, USA, 1997. Association for Computing Machinery. doi:10.1145/258533.258662.

39 Heng Guo and Mark Jerrum. Random cluster dynamics for the Ising model is rapidly mixing. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017*, pages 1818–1827, 2017.

40 Heng Guo, Kaan Kara, and Ce Zhang. Layerwise systematic scan: Deep Boltzmann machines and beyond. In *International Conference on Artificial Intelligence and Statistics*, pages 178–187. PMLR, 2018.

41 Thomas P. Hayes. A simple condition implying rapid mixing of single-site dynamics on spin systems. In *The 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, pages 39–46, 2006. doi:10.1109/FOCS.2006.6.

42 Thomas P. Hayes and Alistair Sinclair. A general lower bound for mixing of single-site dynamics on graphs. *The Annals of Applied Probability*, 17(3), June 2007. doi:10.1214/105051607000000104.

43 Mark Huber. A bounding chain for Swendsen-Wang. *Random Structures & Algorithms*, 22(1):43–59, 2003.

44 Vishesh Jain, Huy Tuan Pham, and Thuy-Duong Vuong. Spectral independence, coupling, and the spectral gap of the glauber dynamics. *Information Processing Letters*, 177:106268, 2022.

45 F. Martinelli, E. Olivieri, and R. H. Schonmann. For 2-D lattice spin systems weak mixing implies strong mixing. *Communications in Mathematical Physics*, 165(1):33–47, 1994. `doi:cmp/1104271032`.

46 Fabio Martinelli. Lectures on Glauber dynamics for discrete spin models. *Lectures on probability theory and statistics (Saint-Flour, 1997)*, 1717:93–191, 1999.

47 Katalin Marton. Logarithmic sobolev inequalities in discrete product spaces. *Combinatorics, Probability and Computing*, 28(6):919–935, 2019. `doi:10.1017/S0963548319000099`.

48 Ravi Montenegro and Prasad Tetali. Mathematical aspects of mixing times in markov chains. *Foundations and Trends® in Theoretical Computer Science*, 1(3):237–354, 2006. `doi:10.1561/0400000003`.

49 Yuval Peres and Peter Winkler. Can extra updates delay mixing? *Communications in Mathematical Physics*, 323(3):1007–1016, 2013.

50 Ricardo Restrepo, Jinwoo Shin, Prasad Tetali, Eric Vigoda, and Linji Yang. Improved mixing condition on the grid for counting and sampling independent sets. *Probab. Theory Relat. Fields*, 2013. `doi:10.1007/s00440-012-0421-8`.

51 BARRY SIMON. *The Statistical Mechanics of Lattice Gases, Volume I.* Princeton University Press, 1993. URL: `http://www.jstor.org/stable/j.ctt7ztwsn`.

52 Alistair Sinclair, Piyush Srivastava, Daniel Stefankovic, and Yitong Yin. Spatial mixing and the connective constant: Optimal bounds. *Probab. Theory Relat. Fields*, 168:153–197, 2017. `doi:10.1007/s00440-016-0708-2`.

53 Robert H. Swendsen and Jian-Sheng Wang. Nonuniversal critical dynamics in Monte Carlo simulations. *Phys. Rev. Lett.*, 58:86–88, 1987. `doi:10.1103/PhysRevLett.58.86`.

54 Mario Ullrich. Rapid mixing of swendsen–wang dynamics in two dimensions. *Dissertationes Mathematicae*, 502:1–64, 2014. `doi:10.4064/dm502-0-1`.