SIAM J. COMPUT. © 2023 Society for Industrial and Applied Mathematics
Vol. 52, No. 1, pp. 196-237

RAPID MIXING OF GLAUBER DYNAMICS UP TO UNIQUENESS VIA CONTRACTION*

ZONGCHEN CHEN†, KUIKUI LIU†, AND ERIC VIGODA‡

Abstract. For general antiferromagnetic 2-spin systems, including the hardcore model on weighted independent sets and the antiferromagnetic Ising model, there is an FPTAS for the partition function on graphs of maximum degree Δ when the infinite regular tree lies in the uniqueness region by Li, Lu, and Yin [*Correlation Decay up to Uniqueness in Spin Systems*, preprint, https://arxiv.org/abs/1111.7064, 2021]. Moreover, in the tree nonuniqueness region, Sly in [*Computational transition at the uniqueness threshold*, in Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science, 2010, pp. 287–296] showed that there is no FPRAS to estimate the partition function unless $\mathsf{NP} = \mathsf{RP}$. The algorithmic results follow from the correlation decay approach due to Weitz [*Counting independent sets up to the tree threshold*, in Proceedings of the 38th Annual ACM Symposium on Theory of Computing, 2006, pp. 140–149] or the polynomial interpolation approach developed by Barvinok [*Combinatorics and Complexity of Partition Functions*, Springer, 2016]. However, the running time is only polynomial for constant Δ. For the hardcore model, recent work of Anari, Liu, and Oveis Gharan [*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, 2020, pp. 1319–1330] establishes rapid mixing of the simple single-site Markov chain, known as the Glauber dynamics, in the tree uniqueness region. Our work simplifies their analysis of the Glauber dynamics by considering the total pairwise influence of a fixed vertex v on other vertices, as opposed to the total influence of other vertices on v, thereby extending their work to all 2-spin models and improving the mixing time. More important, our proof ties together the three disparate algorithmic approaches: we show that contraction of the so-called tree recursions with a suitable potential function, which is the primary technique for establishing efficiency of Weitz's correlation decay approach and Barvinok's polynomial interpolation approach, also establishes rapid mixing of the Glauber dynamics. We emphasize that this connection holds for all 2-spin models (both antiferromagnetic and ferromagnetic), and existing proofs for the correlation decay and polynomial interpolation approaches immediately imply rapid mixing of the Glauber dynamics. Our proof utilizes the fact that the graph partition function is a divisor of the partition function for Weitz's self-avoiding walk tree. This fact leads to new tools for the analysis of the influence of vertices and may be of independent interest for the study of complex zeros.

Key words. approximate counting, Glauber dynamics, spectral independence, phase transitions, correlation decay

MSC codes. 60J10, 68Q87, 68W20

DOI. 10.1137/20M136685X

1. Introduction.

A remarkable connection has been established between the computational complexity of approximate counting problems in general graphs of maximum degree Δ and the statistical physics phase transition on infinite, regular trees of degree Δ (or up to Δ in the more general case). This connection holds for

*Received by the editors September 14, 2020; accepted for publication (in revised form) September 23, 2022; published electronically February 28, 2023. A preliminary short version of these results appeared in *Proceedings of the 2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)* [11].
https://doi.org/10.1137/20M136685X
Funding: The first and third authors are supported by NSF grant CCF-2007022. The second author is supported by NSF grant CCF-1907845 and ONR-YIP grant N00014-17-1-2429.
†Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 01239 USA (zongchen@mit.edu,liukui@csail.mit.edu).
‡University of California Santa Barbara, Santa Barbara, CA 93106 USA (vigoda@ucsb.edu).

2-state antiferromagnetic spin systems: the hardcore model on independent sets and the Ising model are the most interesting examples of such systems.

Given an n-vertex graph $G = (V, E)$, configurations of the 2-spin model are the 2^n assignments of spins $0, 1$ to the vertices. A 2-spin system is defined by three parameters: two edge weights $\beta \geq 0, \gamma > 0$ and one vertex weight $\lambda > 0$. Edge parameter β controls the (relative) strength of interaction between neighboring 1-spins, γ corresponds to neighboring 0-spins, and λ is the external field applied to vertices with 1-spins.

Every spin configuration $\sigma \in \{0,1\}^V$ is assigned a weight

$$w_G(\sigma) = \beta^{m_1(\sigma)}\gamma^{m_0(\sigma)}\lambda^{n_1(\sigma)},$$

where, for spin $s \in \{0,1\}$, $m_s(\sigma) = \#\{uv \in E : \sigma_u = \sigma_v = s\}$ is the number of monochromatic edges with spin s, and $n_1(\sigma) = \#\{v \in V : \sigma_v = 1\}$ is the number of vertices with spin 1 (as is standard, the parameters are normalized, so we can avoid two additional parameters). The Gibbs distribution over spin configurations is given by $\mu_G(\sigma) = \frac{w_G(\sigma)}{Z_G(\beta,\gamma,\lambda)}$, where $Z_G(\beta, \gamma, \lambda) = \sum_{\sigma \in \{0,1\}^V} \beta^{m_1(\sigma)}\gamma^{m_0(\sigma)}\lambda^{n_1(\sigma)}$ is the partition function.

There are two examples of particular interest: the hardcore model and the Ising model. When $\beta = 0$ and $\gamma = 1$, the only configurations with nonzero weight are independent sets of G, and the weight of an independent set σ is $w(\sigma) = \lambda^{|\sigma|}$; this example is known as the *hardcore model,* where the parameter λ corresponds to the fugacity.

In the case when $\beta = \gamma$, the important quantity is the total number of monochromatic edges $m(\sigma) = m_0(\sigma) + m_1(\sigma)$, and the weight of a configuration σ is $w(\sigma) = \beta^{m(\sigma)}\lambda^{n_1(\sigma)}$; this is the classical *Ising model,* where the parameter β corresponds to the inverse temperature, and λ is the external field ($\lambda = 1$ means no external field). Note that when $\beta > 1$, the model is *ferromagnetic* as neighboring vertices prefer to have the same spin, and $\beta < 1$ is the *antiferromagnetic* Ising model. In the general 2-spin system, the model is ferromagnetic when $\beta\gamma > 1$ and antiferromagnetic when $\beta\gamma < 1$. (When $\beta\gamma = 1$ we get a trivial product distribution.)

The fundamental algorithmic tasks are to sample from the Gibbs distribution and to estimate the partition function. For the approximate sampling problem, we are given a graph G and an $\epsilon > 0$, and our goal is to generate a sample from a distribution π which is within total variation distance $\leq \epsilon$ of the Gibbs distribution μ_G in time $\mathrm{poly}(n, \log(1/\epsilon))$. An efficient approximate sampling algorithm implies an FPRAS (fully polynomial randomized approximation scheme) for the approximate counting problem [17, 33]. Recall that, given an n-vertex graph G and $\epsilon, \delta > 0$, an FPRAS outputs a $(1 \pm \epsilon)$-approximation of Z_G with probability $\geq 1 - \delta$ in time $\mathrm{poly}(n, 1/\epsilon, \log(1/\delta))$, whereas an fully polynomial-time approximation scheme (FPTAS) is the deterministic analogue (i.e., $\delta = 0$).

A standard approach to the approximate sampling problem is the Markov Chain Monte Carlo (MCMC) method; in fact, there is a simple Markov chain known as the *Glauber dynamics*. The Glauber dynamics works as follows: from a configuration X_t at time t, choose a random vertex v; we then set $X_{t+1}(w) = X_t(w)$ for all $w \neq v$, and finally we choose $X_{t+1}(v)$ from the conditional distribution of $\mu(\sigma_v | \sigma_w = X_{t+1}(w)$ for all $w \neq v)$. For the case of the hardcore model, $X_{t+1}(v)$ is set to "occupied" (i.e., spin 1) with probability $\lambda/(1 + \lambda)$ if no neighbors are currently occupied; otherwise, it is set to "unoccupied."

It is straightforward to verify that the Glauber dynamics is ergodic, with the Gibbs distribution as the unique stationary distribution. The *mixing time* is the minimum number of steps needed to guarantee, from the worst initial state X_0, that the distribution of X_t is within total variation distance $\leq 1/4$ of the Gibbs distribution. The goal is to prove that the mixing time is polynomial in n, where $n = |V|$ is the number of vertices, in which case the chain is said to be *rapidly mixing*.

For the case of the ferromagnetic Ising model (with or without an external field), a classical result of Jerrum and Sinclair [16] gives an FPRAS for all graphs via the MCMC method. This is the only case with an efficient algorithm for general graphs. For antiferromagnetic 2-spin models, the picture is closely tied to statistical physics phase transitions on the regular tree.

The uniqueness/nonuniqueness phase transition is nicely illustrated for the case of the hardcore model. Consider the infinite Δ-regular tree T rooted at r, and let T_h denote the tree truncated at the first h levels. This phase transition captures whether the configuration at the leaves of T_h "influences" the root in the limit $h \to \infty$. For the hardcore model we can consider even height trees (corresponding to the all even boundary condition) versus odd height trees. Let p_h denote the marginal probability that the root is occupied in the Gibbs distribution μ_{T_h}. Let $p_{\mathsf{even}} = \lim_{h\to\infty} p_{2h}$ and $p_{\mathsf{odd}} = \lim_{h\to\infty} p_{2h+1}$. We say that *tree uniqueness* holds if $p_{\mathsf{even}} = p_{\mathsf{odd}}$ and that *tree nonuniqueness* holds if they are not equal. For all $\Delta \geq 3$ there exists a critical fugacity $\lambda_c(\Delta) = (\Delta-1)^{\Delta-1}/(\Delta-2)^{\Delta}$ [18], where tree uniqueness holds if and only if $\lambda \leq \lambda_c(\Delta)$.

The remarkable connection is that an algorithmic phase transition for general graphs of maximum degree Δ occurs at this same tree critical point. For all constant Δ, all $\delta > 0$, all $\lambda < (1-\delta)\lambda_c(\Delta)$, and all graphs of maximum degree Δ, [34] presented an FPTAS for approximating the partition function. On the other hand, for all $\delta > 0$ and all $\lambda > (1+\delta)\lambda_c(\Delta)$, [31, 32, 12] proved that, unless $\mathsf{NP} = \mathsf{RP}$, there is no FPRAS for estimating the partition function.

One important caveat is that the running time of Weitz's algorithm is $(n/\epsilon)^{C\log\Delta}$, where the approximation factor is $(1 \pm \epsilon)$, and the constant C depends polynomially on the gap δ (recall that $\lambda < (1-\delta)\lambda_c$). Weitz's correlation decay algorithm was extended to the antiferromagnetic Ising model in the tree uniqueness region by Sinclair, Srivastava, and Thurley [30], and to all antiferromagnetic 2-spin systems in the corresponding tree uniqueness region (as we detail below) by Li, Lu, and Yin [21].

An intriguing new algorithmic approach was presented by Barvinok [4], and refined by Patel and Regts [26], utilizing the absence of zeros of the partition function in the complex plane to efficiently approximate a suitable transformation of the logarithm of the partition function using Taylor approximation. This polynomial interpolation approach was shown to be efficient in the same tree uniqueness region as that for Weitz's result by Peters and Regts [27], although the exponent in the running time depends exponentially on Δ.

It was long conjectured that the simple Glauber dynamics is rapidly mixing in the tree uniqueness region. This was recently proved by Anari, Liu, and Oveis Gharan [3]; they proved that, for all $\delta > 0$, the mixing time is $n^{O(\exp(1/\delta))}$ whenever $\lambda < (1-\delta)\lambda_c(\Delta)$. We improve this result. First, we improve the mixing time from $n^{O(\exp(1/\delta))}$ to $n^{O(1/\delta)}$ as detailed in the following theorem.

THEOREM 1.1 (hardcore model). *Let $\Delta \geq 3$ be an integer, and let $\delta \in (0,1)$. For every n-vertex graph G of maximum degree Δ and every $0 < \lambda \leq (1-\delta)\lambda_c(\Delta)$, the mixing time of the Glauber dynamics for the hardcore model on G with fugacity λ is $O(n^{2+32/\delta})$.*

This bound is optimal barring further improvements in the local-to-global arguments from [1]. Our improved result follows from a simpler, cleaner proof approach, which enables us to extend our result to a wide variety of 2-spin models, matching the key results for the correlation decay algorithm with vastly improved running times.

Our proof approach unifies the three major algorithmic tools for approximate counting: correlation decay, polynomial interpolation, and MCMC. Most known results for both the correlation decay and polynomial interpolation approaches are proved by showing contraction of a suitably defined potential function on the so-called tree recursions; the tree recursions arise as a result of Weitz's self-avoiding walk tree that we will describe in more detail later in this paper. Recent works of Liu [23] and Shao and Sun [29] unify these two approaches by showing that the contraction which is normally used to prove efficiency of the correlation decay algorithm also implies (under some additional analytic conditions) that the polynomial interpolation approach is efficient.

Here we prove that this same contraction of a potential function also implies rapid mixing of the Glauber dynamics, with our improved running time that is independent of Δ; see Definition 1.4 and Theorem 1.5 for detailed statements. Our proof utilizes several new tools concerning Weitz's self-avoiding walk tree, which are detailed in section 3. In particular, we show that the partition function of a graph G divides the partition function of Weitz's self-avoiding walk tree; see Lemma 3.3. This result is potentially of independent interest for establishing absence of zeros for the partition function with complex parameters, as it enables one to consider the self-avoiding walk tree. This result also yields a new, useful equivalence for bounding the influence in a graph in terms of the self-avoiding tree, which strengthens the previously known connection by Weitz [34]; see Lemma 3.3 for details.

As an easy consequence we obtain rapid mixing for the Glauber dynamics for the antiferromagnetic Ising model in the tree uniqueness region. In terms of the edge activity, the two critical points for the Ising model on the Δ-regular tree are at $\beta_c(\Delta) = \frac{\Delta-2}{\Delta}$ and $\overline{\beta}_c(\Delta) = \frac{1}{\beta_c(\Delta)} = \frac{\Delta}{\Delta-2}$; the first lies in the antiferromagnetic regime, while the second lies in the ferromagnetic regime. If $\beta_c(\Delta) < \beta < \overline{\beta}_c(\Delta)$, then uniqueness holds for all external field λ on the Δ-regular tree.

As mentioned earlier, for the ferromagnetic Ising model, an FPRAS was known for general graphs [16]. Furthermore, Mossel and Sly [25] proved $O(n \log n)$ mixing time of the Glauber dynamics for the ferromagnetic Ising model when $1 \le \beta < \overline{\beta}_c(\Delta)$. However, rapid mixing for the antiferromagnetic Ising model in the tree uniqueness region was not known.

We provide the following mixing result for the case $\beta > \beta_c(\Delta)$. Note that when $\beta \le \beta_c$ there is an additional uniqueness region for certain values of the external field λ; this region is covered by Theorem 1.3.

THEOREM 1.2 (antiferromagnetic Ising model). *Let $\Delta \ge 3$ be an integer, and let $\delta \in (0,1)$. Assume that $1 > \beta \ge \beta_c(\Delta) + \delta(1 - \beta_c(\Delta))$ and $\lambda > 0$. Then for every n-vertex graph G of maximum degree Δ, the mixing time of the Glauber dynamics for the Ising model on G with edge weight β and external field λ is $O(n^{2+1.5/\delta})$.*

Our results for the hardcore and Ising models fit within a larger framework of general antiferromagnetic 2-spin systems. Recall that we have the antiferromagnetic case when $\beta\gamma < 1$.

For general 2-spin systems the appropriate tree phase transition is more complicated, as there are models where the tree uniqueness threshold is not monotone in Δ. Hence the appropriate notion is "up-to-Δ uniqueness" as considered by [21]. Roughly speaking, we say uniqueness with gap $\delta \in (0,1)$ holds on the d-regular tree if for every integer $\ell \ge 1$, all vertices at distance ℓ from the root have total "influence" $\lesssim (1-\delta)^\ell$ on the marginal of the root. We say up-to-Δ uniqueness with gap δ holds if uniqueness with gap δ holds on the d-regular tree for all $1 \le d \le \Delta$; see section 2 for the precise definition.

Both Theorems 1.1 and 1.2 are corollaries of the following general rapid mixing result which holds for general antiferromagnetic 2-spin systems in the entire tree uniqueness region.

THEOREM 1.3 (general antiferromagnetic 2-spin system). *Let $\Delta \geq 3$ be an integer, and let $\delta \in (0,1)$. Let β,γ,λ be reals such that $0 \leq \beta \leq \gamma$, $\gamma > 0$, $\beta\gamma < 1$, and $\lambda > 0$. Assume that the parameters (β,γ,λ) are up-to-Δ unique with gap δ. Then for every n-vertex graph G of maximum degree Δ, the mixing time of the Glauber dynamics for the antiferromagnetic 2-spin system on G with parameters (β,γ,λ) is $O(n^{2+72/\delta})$.*

We also match existing correlation decay results [15, 29] for ferromagnetic 2-spin models; see section 8 for results and Appendix F for proofs.

1.1. Mixing by the potential method. The tree recursion is very useful in the study of approximating counting. Consider a tree rooted at r. Suppose that r has d children, denoted by $v_1,\dots,v_d$. For $1 \leq i \leq \Delta_i$ we define T_{v_i} to be the subtree of T rooted at v_i that contains all descendants of v_i. Let $R_r = \mu_T(\sigma_r = 1)/\mu_T(\sigma_r = 0)$ denote the marginal ratio of the root, and let $R_{v_i} = \mu_{T_{v_i}}(\sigma_{v_i} = 1)/\mu_{T_{v_i}}(\sigma_{v_i} = 0)$ for each subtree. The *tree recursion* is a formula that computes R_r given $R_{v_1},\dots,R_{v_d}$, due to the independence of T_{v_i}'s. More specifically, we can write $R_r = F_d(R_{v_1},\dots,R_{v_d})$, where $F_d : [0,+\infty]^d \to [0,+\infty]$ is a multivariate function such that for $(x_1,\dots,x_d) \in [0,+\infty]^d$,

$$F_d(x_1,\dots,x_d) = \lambda \prod_{i=1}^{d} \frac{\beta x_i + 1}{x_i + \gamma}.$$

In this paper, however, we pay particular attention to the log of marginal ratios because we will carefully study the *pairwise influence matrix* $\mathcal{I}_G$ of the Gibbs distribution μ_G, introduced in [3] and defined, for every $r, v \in V$, as

$$\mathcal{I}_G(r \to v) = \mu_G(\sigma_v = 1 \mid \sigma_r = 1) - \mu_G(\sigma_v = 1 \mid \sigma_r = 0).$$

In [3], the authors show that if the maximum eigenvalue of $\mathcal{I}_G$ is bounded appropriately, then the Glauber dynamics is rapid mixing. One crucial observation we make in this paper is that the influence $\mathcal{I}_G(r \to v)$ of r on v can be viewed as the derivative of $\log R_r$ with respect to the log external field at v (see Lemma 4.3). Thus, it is more convenient for us to work with the log ratios. To this end, we rewrite the tree recursion as $\log R_v = H_d(\log R_{v_1},\dots,\log R_{v_d})$, where $H_d : [-\infty,+\infty]^d \to [-\infty,+\infty]$ is a function such that for $(y_1,\dots,y_d) \in [-\infty,+\infty]^d$,

$$H_d(y_1,\dots,y_d) = \log\lambda + \sum_{i=1}^{d} \log\left(\frac{\beta e^{y_i} + 1}{e^{y_i} + \gamma}\right).$$

Observe that $H = \log \circ F \circ \exp$. Moreover, we define

$$h(y) = -\frac{(1-\beta\gamma)e^y}{(\beta e^y + 1)(e^y + \gamma)}$$

for $y \in [-\infty,+\infty]$, so that $\frac{\partial}{\partial y_i} H_d(y_1,\dots,y_d) = h(y_i)$ for each i.

To prove our main results, we use the potential method, which has been widely used to establish the decay of correlation. By choosing a suitable potential function for the log ratios, we show that the total influence from a given vertex decays exponentially with the distance, and thus we establish rapid mixing of the Glauber dynamics. Let us first specify our requirements on the potential. For every integer $d \geq 0$, we define a bounded interval J_d which contains all log ratios at a vertex of degree d. More specifically, we let $J_d = [\log(\lambda\beta^d), \log(\lambda/\gamma^d)]$ when $\beta\gamma < 1$, and $J_d = [\log(\lambda/\gamma^d), \log(\lambda\beta^d)]$ when $\beta\gamma > 1$. Furthermore, define $J = \bigcup_{d=0}^{\Delta-1} J_d$ to be the interval containing all log ratios with degree less than Δ.

DEFINITION 1.4 ((α, c)-potential function). *Let $\Delta \geq 3$ be an integer. Let β, γ, λ be reals such that $0 \leq \beta \leq \gamma$, $\gamma > 0$, and $\lambda > 0$. Let $\Psi : [-\infty, +\infty] \to (-\infty, +\infty)$ be a strictly increasing function with image $S = \Psi[-\infty, +\infty]$, which is differentiable on $(-\infty, +\infty)$ with derivative $\psi = \Psi'$. For any $\alpha \in (0, 1)$ and $c > 0$, we say Ψ is an* (α, c)-potential function *with respect to Δ and (β, γ, λ) if it satisfies the following conditions:*

1. *(Contraction) For every integer d such that $1 \leq d < \Delta$, and every tuple $(\tilde{y}_1, \ldots, \tilde{y}_d) \in S^d$, we have*
$$\left\| \nabla H_d^\Psi(\tilde{y}_1, \ldots, \tilde{y}_d) \right\|_1 = \sum_{i=1}^{d} \frac{\psi(y)}{\psi(y_i)} \cdot |h(y_i)| \leq 1 - \alpha,$$
where $H_d^\Psi = \Psi \circ H_d \circ \Psi^{-1}$, $y_i = \Psi^{-1}(\tilde{y}_i)$ for $1 \leq i \leq d$, and $y = H_d(y_1, \ldots, y_d)$.
2. *(Boundedness) For every $y_1, y_2 \in J$, we have*
$$\frac{\psi(y_2)}{\psi(y_1)} \cdot |h(y_1)| \leq \frac{c}{\Delta}.$$

In the definition of (α, c)-potential, one should think of y as the log marginal ratio at a vertex, and the potential function is of $\log R$. The following theorem establishes rapid mixing of the Glauber dynamics given an (α, c)-potential function.

THEOREM 1.5. *Let $\Delta \geq 3$ be an integer. Let β, γ, λ be reals such that $0 \leq \beta \leq \gamma$, $\gamma > 0$, and $\lambda > 0$. Suppose that there is an (α, c)-potential with respect to Δ and (β, γ, λ) for some $\alpha \in (0, 1)$ and $c > 0$. Then for every n-vertex graph G of maximum degree Δ, the mixing time of the Glauber dynamics for the 2-spin system on G with parameters (β, γ, λ) is $O(n^{2+c/\alpha})$.*

We outline our proofs in section 3. Note that in both Definition 1.4 and Theorem 1.5, the constant c is allowed to depend on the maximum degree Δ and parameters (β, γ, λ) in general. For example, a straightforward black-box application of the potential in [21] would give $c = \Theta(\Delta)$ for the boundedness condition, resulting in $n^{\Theta(\Delta)}$ mixing. However, this is undesirable for graphs with potentially unbounded degrees. One of our contributions is showing that the boundedness condition holds for a universal constant c *independent* of Δ and (β, γ, λ). Thus, our mixing time is $O(n^{2+c/\delta})$ with no parameters in the exponent except for $1/\delta$.

In section 7, we give a slightly more general definition of (α, c)-potentials, which relaxes the boundedness condition, and is necessary for our analysis of antiferromagnetic 2-spin systems with $0 \leq \beta < 1 < \gamma$. Theorem 1.5 still holds for this larger class of potentials.

We remark that in all previous works on the potential method, results and proofs are always presented in terms of F_d, the tree recursion of R, and Φ, a potential function of R. In fact, our results can also be translated into the language of (F_d, Φ). To see this, note that since $H_d = \log \circ F_d \circ \exp$, it is straightforward to check that $H_d^\Psi = \Psi \circ H_d \circ \Psi^{-1} = \Phi \circ F_d \circ \Phi^{-1} = F_d^\Phi$ if we pick $\Phi = \Psi \circ \log$, and thereby $\nabla H_d^\Psi = \nabla F_d^\Phi$. This implies that the contraction condition in Definition 1.4 holds for

(H_d, Ψ) if and only if the corresponding contraction condition holds for (F_d, Φ). The boundedness condition can also be stated equivalently for (F_d, Φ). Nevertheless, in this paper we choose to work with (H_d, Ψ) for the following two reasons. First, as mentioned earlier, the fact that $\mathcal{I}_G(r \to v)$ is a derivative of $\log R_r$ makes it natural to consider the tree recursion for the log ratios. Indeed, it is easier and cleaner to present our results and proofs using (H_d, Ψ) directly rather than by switching to (F_d, Φ). Second, the potential function Ψ we will use is obtained from the exact potential Φ in [21] by the transformation $\Psi = \Phi \circ \exp$.[1] It is intriguing to notice that the derivative of this potential is simply $\psi = \sqrt{|h|}$. Then the contraction condition has the nice form $\sum_{i=1}^{d} \sqrt{h(y)h(y_i)} \leq 1 - \alpha$, and the boundedness condition only involves an upper bound on $h(y)$. This seems to shed some light on the mysterious potential function Φ from [21], and it also indicates that H_d is a meaningful variant of the tree recursion to consider. To add one more piece of evidence, we note that for a lot of cases (e.g., $\frac{\Delta-2}{\Delta} < \sqrt{\beta\gamma} < \frac{\Delta}{\Delta-2}$) where the potential $\Phi = \log$ is picked, we can take Ψ to be the identity function, in which case H_d itself is contracting without any nontrivial potential (see, e.g.. [35, 29]).

Revision in July 2021. After the publication of this paper in FOCS 2020 [10], a small error was found in [21] regarding descriptions of the uniqueness region for antiferromagnetic 2-spin systems. The error was fixed in the latest, preprint version [22]. In the current revision, we update corresponding results and proofs in section 7 and Appendix E that are affected by the changes in [22]; in particular, Lemma E.2 is adjusted in accordance with the current description of uniqueness regions. We remark that these changes are purely technical and do not affect the validity of our main results, such as Theorem 1.5.

2. Preliminaries.

Mixing time and spectral gap. Let P be the transition matrix of an ergodic (i.e., irreducible and aperiodic) Markov chain on a finite state space Ω with stationary distribution μ. Let $P^t(x_0, \cdot)$ denote the distribution of the chain after t steps starting from $x_0 \in \Omega$. The *mixing time* of P is defined as

$$T_{\mathrm{mix}}(P) = \max_{x_0 \in \Omega} \min \left\{ t \geq 0 : \left\| P^t(x_0, \cdot) - \mu(\cdot) \right\|_{\mathrm{TV}} \leq \frac{1}{4} \right\},$$

where $\|\pi(\cdot) - \mu(\cdot)\|_{\mathrm{TV}} = \frac{1}{2}\sum_{x\in\Omega} |\pi(x) - \mu(x)|$ is the total variation distance between two probability distributions π, μ on a common state space Ω.

We say P is *reversible* if $\mu(x)P(x,y) = \mu(y)P(y,x)$ for all $x, y \in \Omega$. If P is reversible, then P has only real eigenvalues which can be denoted by $1 = \lambda_1 \geq \cdots \geq \lambda_{|\Omega|} \geq -1$. The *spectral gap* of P is defined to be $1 - \lambda_2$, and the *absolute spectral gap* of P is defined as $\lambda^*(P) = 1 - \max\{|\lambda_2|, |\lambda_{|\Omega|}|\}$. If P is also positive semidefinite with respect to the inner product $\langle \cdot, \cdot \rangle_\mu$, then all eigenvalues of P are nonnegative, and thus $\lambda^*(P) = 1 - \lambda_2$. Finally, the mixing time and the absolute spectral gap are related by

$$T_{\mathrm{mix}}(P) \leq \frac{1}{\lambda^*(P)} \log \left(\frac{4}{\min_{x\in\Omega} \mu(x)} \right). \tag{2.1}$$

See [19] for more background on Markov chains and mixing times.

[1] To be more precise, we also multiply a constant factor, which only simplifies our calculation and does not matter much; also notice that [21] denotes the potential function by φ and its derivative by $\Phi = \varphi'$.

Uniqueness. Let $\Delta \geq 3$ be an integer or $\Delta = \infty$. Let β, γ, λ be reals such that $0 \leq \beta \leq \gamma$, $\gamma > 0$, $\beta\gamma < 1$, and $\lambda > 0$. For $1 \leq d < \Delta$, 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^*. For $\delta \in (0,1)$, we say the parameters (β, γ, λ) are *up-to-*Δ *unique with gap* δ if $|f'_d(R_d^*)| < 1 - \delta$ for all $1 \leq d < \Delta$.

Ratio and influence. Consider the 2-spin system on a graph $G = (V, E)$. Let $\Lambda \subseteq V$ and $\sigma_\Lambda \in \{0,1\}^\Lambda$. For all $v \in V \setminus \Lambda$, we define the *marginal ratio* at v to be

$$R_G^{\sigma_\Lambda}(v) = \frac{\mu_G(\sigma_v = 1 \mid \sigma_\Lambda)}{\mu_G(\sigma_v = 0 \mid \sigma_\Lambda)}.$$

For all $u, v \in V \setminus \Lambda$, we define the *(pairwise) influence* of u on v by

$$\mathcal{I}_G^{\sigma_\Lambda}(u \to v) = \mu_G(\sigma_v = 1 \mid \sigma_u = 1, \sigma_\Lambda) - \mu_G(\sigma_v = 1 \mid \sigma_u = 0, \sigma_\Lambda).$$

Write $\mathcal{I}_G^{\sigma_\Lambda}$ for the *(pairwise) influence matrix* whose entries are given by $\mathcal{I}_G^{\sigma_\Lambda}(u \to v)$. Note that unlike in [3], our influence matrix has 1 on the diagonal as opposed to 0.

Weitz's self-avoiding walk tree. Let $G = (V, E)$ be a connected graph, and let $r \in V$ be a vertex of G. The *self-avoiding walk (SAW) tree* is defined as follows. Suppose that there is a total ordering of the vertex set V. A self-avoiding walk from r is a path $r = v_0 - v_1 - \cdots - v_\ell$ such that $v_i \neq v_j$ for all $0 \leq i < j \leq \ell$. The SAW tree $T_{\mathrm{SAW}}(G, r)$ is a tree rooted at r, consisting of all self-avoiding walks $r = v_0 - v_1 - \cdots - v_\ell$ with $\deg(v_\ell) = 1$, and those appended with one more vertex that closes the cycle (i.e., $r = v_0 - v_1 - \cdots - v_\ell - v_i$ for some $0 \leq i \leq \ell - 2$ such that $\{v_\ell, v_i\} \in E$). Note that a vertex of G might have many copies in the SAW tree, and the degrees of vertices are preserved except for leaves. See Figure 1 for an example.

We can define a 2-spin system on $T_{\mathrm{SAW}}(G, r)$ with the same parameters (β, γ, λ), in which some of the leaves are fixed to a particular spin. More specifically, for a self-avoiding walk $r = v_0 - v_1 - \cdots - v_\ell$ appended with v_i, we fix v_i to be spin 1 if $v_{i+1} < v_\ell$,

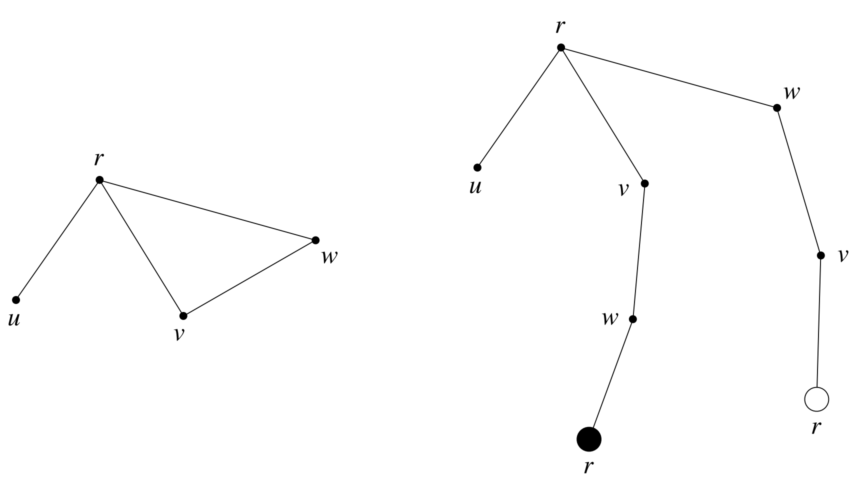

FIG. 1. *A graph G and the self-avoiding walk tree $T_{SAW}(G, r)$ rooted at r. Vertices with the same label in $T_{SAW}(G, r)$ are copies of the same vertex from G. (•/○: fixed to spin 1/0.)*

with respect to the total ordering on V, and spin 0 if $v_{i+1} > v_\ell$. For each $v \in V$, we denote the set of all free (unfixed) copies of v in $T_{\mathrm{SAW}}(G,r)$ by $\mathcal{C}_v$. For $\Lambda \subseteq V$ and a partial configuration $\sigma_\Lambda \in \{0,1\}^\Lambda$, we define the SAW tree with conditioning σ_Λ by assigning the spin σ_v to every copy $\hat{v}$ of v from $\mathcal{C}_v$ and removing all descendants of $\hat{v}$, for each $v \in \Lambda$. Finally, in the case when every vertex v has a distinct field λ_v, all copies of v from $\mathcal{C}_v$ will have the same field λ_v in the SAW tree.

3. Proof outline for main results.

Step 1 [3]: Spectral Independence implies rapid mixing. Our proof builds on the work of Anari, Liu, and Oveis Gharan [3] who showed that the Glauber dynamics for sampling from the hardcore distribution on graphs of maximum degree at most Δ mixes in $O(n^{\exp(O(1/\delta))})$ steps whenever $\lambda \le (1-\delta)\lambda_c(\Delta)$. One of the key ingredients of their proof is a notion they call spectral independence. The authors of [3] show that the spectral independence property implies rapid mixing. Note that the -1 in the definition of spectral independence below is due to $\mathcal{I}_G^{\sigma_\Lambda}$ having diagonal entries equal to 1, as opposed to 0 in [3].

Definition 3.1 (spectral Independence [3]). *We say the Gibbs distribution μ_G on an n-vertex graph G is $(\eta_0,\dots,\eta_{n-2})$-spectrally independent if for every $0 \le k \le n-2$, $\Lambda \subseteq V$ of size k, and $\sigma_\Lambda \in \{0,1\}^\Lambda$, one has $\lambda_{\max}(\mathcal{I}_G^{\sigma_\Lambda}) - 1 \le \eta_k$.*

Theorem 3.2 ([3]). *If μ is an $(\eta_0,\dots,\eta_{n-2})$-spectrally independent distribution, then the Glauber dynamics for sampling from μ has spectral gap at least*

$$\frac{1}{n}\prod_{i=0}^{n-2}\left(1 - \frac{\eta_i}{n-i-1}\right).$$

Our primary goal now is to bound the maximum eigenvalue of $\mathcal{I}_G^{\sigma_\Lambda}$.

Step 2: Self-avoiding walk trees preserve influences. From standard linear algebra, we know that the maximum eigenvalue of $\mathcal{I}_G^{\sigma_\Lambda}$ is upper bounded by both the 1-norm $\|\mathcal{I}_G^{\sigma_\Lambda}\|_1 = \max_{r\in V\setminus\Lambda}\sum_{v\in V\setminus\Lambda}|\mathcal{I}_G^{\sigma_\Lambda}(v\to r)|$, which corresponds to total influences on a vertex r, and the infinity-norm $\|\mathcal{I}_G^{\sigma_\Lambda}\|_\infty = \max_{r\in V\setminus\Lambda}\sum_{v\in V\setminus\Lambda}|\mathcal{I}_G^{\sigma_\Lambda}(r\to v)|$, corresponding to total influences of r. In [3] the authors use $\|\mathcal{I}_G^{\sigma_\Lambda}\|_1$ as an upper bound on $\lambda_{\max}(\mathcal{I}_G^{\sigma_\Lambda})$. Roughly speaking, they show that the sum of absolute influences on a fixed vertex r is upper bounded by the maximum absolute influences on r in the self-avoiding walk tree rooted at r, over all boundary conditions. In this paper, we will use $\|\mathcal{I}_G^{\sigma_\Lambda}\|_\infty$ to upper bound $\lambda_{\max}(\mathcal{I}_G^{\sigma_\Lambda})$ instead. In fact, we see that much more is true when we look at the influences from r in the self-avoiding tree. We show that for every vertex $v \in V\setminus\Lambda$, the influence $\mathcal{I}_G^{\sigma_\Lambda}(r\to v)$ in G is preserved in the self-avoiding walk tree $T = T_{\mathrm{SAW}}(G,r)$ rooted at r in the form of sum of influences $\mathcal{I}_T^{\sigma_\Lambda}(r\to\hat{v})$ over all copies $\hat{v}$ of v.

The way we establish this fact is by viewing the partition function as a polynomial in λ. In fact, it will be useful to consider the more general case with an arbitrary external field λ_v for every $v \in V$. Let $\boldsymbol{\lambda} = \{\lambda_v : v \in V\}$ denote the fields. For $\Lambda \subseteq V$ and $\sigma_\Lambda \in \{0,1\}^\Lambda$, the weight of $\sigma \in \{0,1\}^{V\setminus\Lambda}$ conditional on σ_Λ is defined to be $w_G(\sigma \mid \sigma_\Lambda) = \beta^{m_1(\sigma|\sigma_\Lambda)}\gamma^{m_0(\sigma|\sigma_\Lambda)}\prod_{v\in V\setminus\Lambda}\lambda_v^{\sigma_v}$, where $m_i(\cdot \mid \sigma_\Lambda)$ is the number of i-i edges with at least one endpoint in $V\setminus\Lambda$ for $i = 0,1$. Furthermore, $Z_G^{\sigma_\Lambda} = \sum_{\sigma\in\{0,1\}^{V\setminus\Lambda}} w_G(\sigma \mid \sigma_\Lambda)$ is the partition function conditioned on σ_Λ. We may assume $\gamma, \boldsymbol{\lambda} > 0$ (since otherwise, the model is vacuous) and hence $Z_G^{\sigma_\Lambda} > 0$.

We shall view β and γ as some fixed constants and think of $\boldsymbol{\lambda}$ as $n = |V|$ variables. In this sense, we regard the weights $w_G(\sigma \mid \sigma_\Lambda)$ as monomials in $\boldsymbol{\lambda}$ and the partition function $Z_G^{\sigma_\Lambda}$ as a polynomial in $\boldsymbol{\lambda}$. Moreover, the marginal ratios $R_G^{\sigma_\Lambda}(v)$ and the influences $\mathcal{I}_G^{\sigma_\Lambda}(r \to v)$ for $r, v \in V$ are all functions in $\boldsymbol{\lambda}$. Our main result is that the partition function of G divides that of $T_{\mathrm{SAW}}(G, r)$ for each $r \in V$. From that, we show that the SAW tree preserves influences of the root, as well as re-establishing Weitz's celebrated result [34]; see Lemma 4.4.

LEMMA 3.3. *Let $G = (V, E)$ be a connected graph, $r \in V$ be a vertex, and $\Lambda \subseteq V \setminus \{r\}$ be such that $G \setminus \Lambda$ is connected. Let $T = T_{\mathrm{SAW}}(G, r)$ be the self-avoiding walk tree of G rooted at r. Then for every $\sigma_\Lambda \in \{0,1\}^\Lambda$, $Z_G^{\sigma_\Lambda}$ divides $Z_T^{\sigma_\Lambda}$. More precisely, there exists a polynomial $P_{G,r}^{\sigma_\Lambda} = P_{G,r}^{\sigma_\Lambda}(\boldsymbol{\lambda})$ independent of λ_r such that*

$$Z_T^{\sigma_\Lambda} = Z_G^{\sigma_\Lambda} \cdot P_{G,r}^{\sigma_\Lambda}. \tag{3.1}$$

As a corollary, for each vertex $v \in V \setminus \Lambda$, we have

$$\mathcal{I}_G^{\sigma_\Lambda}(r \to v) = \sum_{\hat{v} \in \mathcal{C}_v} \mathcal{I}_T^{\sigma_\Lambda}(r \to \hat{v}), \tag{3.2}$$

where $\mathcal{C}_v$ is the set of all free (unfixed) copies of v in T.

Remark 3.4. We emphasize that for the purposes of bounding the total influence of a vertex in G, only (3.2) of Lemma 3.3 is needed, which can be proved in a purely combinatorial fashion. However, we believe the divisibility property (3.1) of the multivariate partition function of G and its self-avoiding walk tree may be of independent interest.

We note that a univariate version of the divisibility statement (3.1) has already appeared in [13] for the monomer-dimer model (matchings), in [5] for the hard-core model (independent sets), and in [24] for the zero-field Ising model in the study of complex roots of the partition function. From Lemma 3.3, we can get $\sum_{v \in V \setminus \Lambda} |\mathcal{I}_G^{\sigma_\Lambda}(r \to v)| \le \sum_{v \in V_T} |\mathcal{I}_T^{\sigma_\Lambda}(r \to v)|$ for any fixed r. That means we only need to upper bound the sum of all influences for trees in order to get an upper bound on $\lambda_{\max}(\mathcal{I}_G^{\sigma_\Lambda})$.

Step 3: Decay of influences given a good potential. The tree recursion provides a great tool for computing the (log) ratios of vertices recursively for trees. As we show in Lemma 4.3, the influence $\mathcal{I}_G^{\sigma_\Lambda}(r \to v)$ is, in fact, a version of the derivative of the log marginal ratio at r. Thus, the tree recursion can be used naturally to relate these influences. We then apply the potential method, which has been widely used in the literature to establish the decay of correlations (strong spatial mixing). The following lemma shows that the sum of absolute influences to distance k has exponential decay with k, which can be thought of as the decay of pairwise influences.

LEMMA 3.5. *If there exists an (α, c)-potential function Ψ with respect to Δ and (β, γ, λ) where $\alpha \in (0,1)$ and $c > 0$, then for every $\Lambda \subseteq V_T \setminus \{r\}$, $\sigma_\Lambda \in \{0,1\}^\Lambda$, and all integers $k \ge 1$,*

$$\sum_{v \in L_r(k)} |\mathcal{I}_T^{\sigma_\Lambda}(r \to v)| \le c \cdot (1-\alpha)^{k-1},$$

where $L_r(k)$ denote the set of all free vertices at distance k away from r.

Theorem 1.5 is then proved by combining Theorem 3.2 and Lemmas 3.3 and 3.5. We defer its proof to Appendix A.

Step 4: Find a good potential. As our final step, we need to find an (α, c)-potential function as defined in Definition 1.4. The potential Ψ we choose is exactly the one from [21], adapted to the log marginal ratios and the tree recursion H (see section 6 for more details). We show that if the parameters (β, γ, λ) are up-to-Δ unique with gap $\delta \in (0,1)$ and either $\sqrt{\beta\gamma} > \frac{\Delta-2}{\Delta}$ or $\gamma \le 1$, then Ψ is an (α, c)-potential.

LEMMA 3.6. *Let $\Delta \ge 3$ be an integer. Let β, γ, λ be reals such that $0 \le \beta \le \gamma$, $\gamma > 0$, $\beta\gamma < 1$, and $\lambda > 0$. Assume that (β, γ, λ) is up-to-Δ unique with gap $\delta \in (0,1)$. Define the function Ψ implicitly by*

$$\Psi'(y) = \psi(y) = \sqrt{\frac{(1-\beta\gamma)e^y}{(\beta e^y + 1)(e^y + \gamma)}} = \sqrt{|h(y)|}, \qquad \Psi(0) = 0. \tag{3.3}$$

If $\sqrt{\beta\gamma} > \frac{\Delta-2}{\Delta}$, then Ψ is an (α, c)-potential function with $\alpha \ge \delta/2$ and $c \le 1.5$. If $\sqrt{\beta\gamma} \le \frac{\Delta-2}{\Delta}$ and $\gamma \le 1$, then Ψ is an (α, c)-potential with $\alpha \ge \delta/2$ and $c \le 18$; we can further take $c \le 4$ if $\beta = 0$.

We deduce Theorem 1.3 from the case when $\sqrt{\beta\gamma} > \frac{\Delta-2}{\Delta}$ or the case when $\gamma \le 1$ from Theorem 1.5 and Lemma 3.6. The proof of Theorem 1.3 can be found in Appendix A. The case when $\sqrt{\beta\gamma} \le \frac{\Delta-2}{\Delta}$ and $\gamma > 1$ is trickier. As discussed in section 5 of [21], when $\sqrt{\beta\gamma} \le \frac{\Delta-2}{\Delta}$ and $\gamma > 1$, for some $\lambda > 0$ the spin system lies in the uniqueness region for arbitrary graphs, even with unbounded degrees (i.e., up-to-∞ unique). Thus, in this case the total influences of a vertex can be as large as $\Theta(\Delta/\delta)$, resulting in $n^{\Theta(\Delta/\delta)}$ mixing time. To deal with this, we consider a suitably weighted sum of absolute influences of a fixed vertex, which also upper bounds the maximum eigenvalue of the influence matrix. Definition 1.4 and Theorem 1.5 are then modified to a slightly stronger version. The statements and proofs for this case are presented in section 7 and Appendix D.

The rest of the paper is organized as follows. In section 4 we prove Lemma 3.3 about properties of the SAW tree. In section 5 we establish Lemma 3.5 regarding the decay of influences by the potential method. We verify the contraction condition in section 6 for our choice of potential. In section 7, we focus on the case $\sqrt{\beta\gamma} \le \frac{\Delta-2}{\Delta}$ and $\gamma > 1$, where more general versions of Definition 1.4 and Theorem 1.5 are required; missing proofs can be found in Appendix D. In Appendix E we verify the boundedness condition and its generalization for our potential in all cases. We consider ferromagnetic spin systems in section 8, and the proofs are deferred to Appendix F. We prove all of our main results in Appendix A.

4. Preservation of influences for self-avoiding walk trees. In this section we show that the self-avoiding walk (SAW) tree, introduced in [34] (see also [28]), maintains all the influence of the root and thus establishes Lemma 3.3. To do this, we show that the partition function of G, viewed as a polynomial of the external fields $\boldsymbol{\lambda}$, divides that of the SAW tree. From there we prove that the influence of the root vertex r on another vertex v in G is exactly equal to that on all copies of v in the SAW tree. Using our proof approach, we show that the marginal of the root is maintained in the SAW tree, re-establishing Weitz's celebrated result [34]; also, all pairwise covariances concerned with v are preserved.

THEOREM 4.1. *Let $G = (V, E)$ be a connected graph, $r \in V$ be a vertex, and $\Lambda \subseteq V \setminus \{r\}$ be such that $G \setminus \Lambda$ is connected. Let $T = T_{\mathrm{SAW}}(G, r)$ be the self-avoiding* *walk tree of G rooted at r. Then for every $\sigma_\Lambda \in \{0,1\}^\Lambda$, $Z_G^{\sigma_\Lambda}$ divides $Z_T^{\sigma_\Lambda}$. More precisely, there exists a polynomial $P_{G,r}^{\sigma_\Lambda} = P_{G,r}^{\sigma_\Lambda}(\boldsymbol{\lambda})$ such that*

$$Z_T^{\sigma_\Lambda} = Z_G^{\sigma_\Lambda} \cdot P_{G,r}^{\sigma_\Lambda}.$$

Moreover, the polynomial $P_{G,r}^{\sigma_\Lambda}$ is independent of λ_r.

Remark 4.2. The proof of Theorem 4.1 can be adapted to give a purely combinatorial proof of (3.2) in Lemma 3.3. Like in the proof of [34, Theorem 3.1], one can proceed via vertex splitting and telescoping, where instead of telescoping a product of marginal ratios, one telescopes a sum of single-vertex influences.

We remark that [5] proved a univariate version of Theorem 4.1 for the hardcore model, and [24] showed a similar result for the zero-field Ising model with a uniform edge weight. Our result holds for all 2-spin systems and arbitrary fields for each vertex. We can also generalize it to arbitrary edge weights for each edge in a straightforward fashion. It is crucial that the quotient polynomial $P_{G,r}^{\sigma_\Lambda}$ is independent of the field λ_r at the root, from which we can immediately deduce the preservation of the marginal and the influences of the root.

Before proving Theorem 4.1, we first give a few consequences of it. For all $u, v \in V \setminus \Lambda$, we define the *marginal* at v as $M_G^{\sigma_\Lambda}(v) = \mu_G(v = 1 \mid \sigma_\Lambda)$ (henceforth we write $v = i$ for the event $\sigma_v = i$ for convenience) and define the *covariance* of u and v as

$$K_G^{\sigma_\Lambda}(u,v) = \mu_G(u = v = 1 \mid \sigma_\Lambda) - \mu_G(u = 1 \mid \sigma_\Lambda)\mu_G(v = 1 \mid \sigma_\Lambda).$$

The following lemma relates the quantities we are interested in with appropriate derivatives of the (log) partition function. Parts 1 and 2 of the lemma are folklore.

LEMMA 4.3. *For every graph $G = (V, E)$, $\Lambda \subseteq V$, and $\sigma_\Lambda \in \{0,1\}^\Lambda$, the following hold:*

1. *For all $v \in V$,*

$$\left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) \log Z_G^{\sigma_\Lambda} = M_G^{\sigma_\Lambda}(v).$$

2. *For all $u, v \in V$,*

$$\left(\lambda_v \frac{\partial}{\partial \lambda_v}\right)\left(\lambda_u \frac{\partial}{\partial \lambda_u}\right) \log Z_G^{\sigma_\Lambda} = \left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) M_G^{\sigma_\Lambda}(u) = K_G^{\sigma_\Lambda}(u,v).$$

3. *For all $u, v \in V$,*

$$\left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) \log R_G^{\sigma_\Lambda}(u) = \mathcal{I}_G^{\sigma_\Lambda}(u \to v).$$

Proof. Parts 1 and 2 are standard. We include the proofs of these two facts in Appendix B for completeness. For part 3, we deduce from part 1 that

$$\begin{aligned}\left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) \log R_G^{\sigma_\Lambda}(u) &= \left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) \log\left(\frac{Z_G^{\sigma'_\Lambda}}{Z_G^{\sigma''_\Lambda}}\right) \\ &= M_G^{\sigma'_\Lambda}(v) - M_G^{\sigma''_\Lambda}(v) \\ &= \mathcal{I}_G^{\sigma_\Lambda}(u \to v),\end{aligned}$$

where ${\sigma'}_\Lambda$ (resp., ${\sigma''}_\Lambda$) denotes the pinning which is the union of σ_Λ and $u=+$ (resp., $u=-$). ∎

We deduce Lemma 3.3 from Theorem 4.1 and the second item of the following lemma. The proof of Theorem 4.1 is presented in subsection 4.1.

LEMMA 4.4. *Let $G=(V,E)$ be a connected graph, $r \in V$ be a vertex, and $\Lambda \subseteq V\setminus\{r\}$ be such that $G\setminus\Lambda$ is connected. Let $T=T_{\mathrm{SAW}}(G,r)$ be the self-avoiding walk tree of G rooted at r. Then for every $\sigma_\Lambda \in \{0,1\}^\Lambda$ we have the following:*

1. *(See [34, Theorem 3.1].) Preservation of marginal of the root r:*

$$M_G^{\sigma_\Lambda}(r) = M_T^{\sigma_\Lambda}(r) \qquad \text{and} \qquad R_G^{\sigma_\Lambda}(r) = R_T^{\sigma_\Lambda}(r).$$

2. *Preservation of covariances and influences of r: for every $v \in V$,*

$$K_G^{\sigma_\Lambda}(r,v) = \sum_{\hat{v}\in\mathcal{C}_v} K_T^{\sigma_\Lambda}(r,\hat{v}) \qquad \text{and} \qquad \mathcal{I}_G^{\sigma_\Lambda}(r\to v) = \sum_{\hat{v}\in\mathcal{C}_v} \mathcal{I}_T^{\sigma_\Lambda}(r\to\hat{v}),$$

where $\mathcal{C}_v$ is the set of all free (unfixed) copies of v in T.

Proof. By Theorem 4.1, there exists a polynomial $P_{G,r}^{\sigma_\Lambda} = P_{G,r}^{\sigma_\Lambda}(\boldsymbol{\lambda})$ such that $Z_T^{\sigma_\Lambda} = Z_G^{\sigma_\Lambda}\cdot P_{G,r}^{\sigma_\Lambda}$, and $P_{G,r}^{\sigma_\Lambda}$ is independent of λ_r. Then it follows from Lemma 4.3 that

$$\begin{aligned} M_T^{\sigma_\Lambda}(r) &= \left(\lambda_r \frac{\partial}{\partial \lambda_r}\right)\log Z_T^{\sigma_\Lambda} = \left(\lambda_r \frac{\partial}{\partial \lambda_r}\right)\left(\log Z_G^{\sigma_\Lambda} + \log P_{G,r}^{\sigma_\Lambda}\right) \\ &= \left(\lambda_r \frac{\partial}{\partial \lambda_r}\right)\log Z_G^{\sigma_\Lambda} = M_G^{\sigma_\Lambda}(r), \end{aligned}$$

and therefore $R_T^{\sigma_\Lambda}(r) = R_G^{\sigma_\Lambda}(r)$. For the second item, again from Lemma 4.3 we get

$$K_G^{\sigma_\Lambda}(r,v) = \left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) M_G^{\sigma_\Lambda}(r) = \left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) M_T^{\sigma_\Lambda}(r).$$

Recall that for the spin system on the SAW tree T, every free copy $\hat{v}$ of v from $\mathcal{C}_v$ has the same external field $\lambda_{\hat{v}} = \lambda_v$. Then, by the chain rule of derivatives and Lemma 4.3, we deduce that

$$K_G^{\sigma_\Lambda}(r,v) = \sum_{\hat{v}\in\mathcal{C}_v}\left(\lambda_{\hat{v}} \frac{\partial}{\partial \lambda_{\hat{v}}}\right) M_T^{\sigma_\Lambda}(r)\cdot\frac{\partial \lambda_{\hat{v}}}{\partial \lambda_v}\cdot\frac{\lambda_v}{\lambda_{\hat{v}}} = \sum_{\hat{v}\in\mathcal{C}_v} K_T^{\sigma_\Lambda}(r,\hat{v}).$$

Finally, we have

$$\mathcal{I}_G^{\sigma_\Lambda}(r\to v) = \left(\lambda_v \frac{\partial}{\partial \lambda_v}\right)\log R_G^{\sigma_\Lambda}(r) = \left(\lambda_v \frac{\partial}{\partial \lambda_v}\right)\log R_T^{\sigma_\Lambda}(r) = \sum_{\hat{v}\in\mathcal{C}_v}\mathcal{I}_T^{\sigma_\Lambda}(r\to\hat{v}),$$

where the last equality follows as above. ∎

4.1. Proof of Theorem 4.1. Before presenting our proof, let us first review the notation and definitions introduced earlier. Denote the set of fields at all vertices by $\boldsymbol{\lambda} = \{\lambda_v : v\in V\}$. For $\Lambda\subseteq V$ and $\sigma_\Lambda\in\{0,1\}^\Lambda$, the weight of $\sigma\in\{0,1\}^{V\setminus\Lambda}$ conditional on σ_Λ is given by

$$w_G(\sigma\mid\sigma_\Lambda) = \beta^{m_1(\sigma|\sigma_\Lambda)}\gamma^{m_0(\sigma|\sigma_\Lambda)}\prod_{v\in V\setminus\Lambda}\lambda_v^{\sigma_v},$$

where for $i=0,1$, $m_i(\cdot \mid \sigma_\Lambda)$ denotes the number of edges such that both endpoints receive the spin i and at least one of them is in $V\setminus\Lambda$. The partition function conditional on σ_Λ is defined as $Z_G^{\sigma_\Lambda} = \sum_{\sigma\in\{0,1\}^{V\setminus\Lambda}} w_G(\sigma \mid \sigma_\Lambda)$. For the SAW tree, we define the conditional weights and partition function in the same way. In particular, recall that when we fix a conditioning σ_Λ on the SAW tree, we also remove all descendants of $\hat{v}\in\mathcal{C}_v$ for each $v\in\Lambda$.

For every $v\in V\setminus\Lambda$ and $i\in\{0,1\}$, we shall write $v=i$ to represent the set of configurations such that $\sigma_v=i$ (i.e., $\{\sigma\in\{0,1\}^{V\setminus\Lambda} : \sigma_v=i\}$) and let $Z_G^{\sigma_\Lambda}(v=i)$ be the sum of weights of all configurations with $v=i$. We further extend this notation and write $Z_G^{\sigma_\Lambda}(U=\sigma_U)$ for every $U\subseteq V\setminus\Lambda$ and $\sigma_U\in\{0,1\}^U$. For the SAW tree we adopt the same notation as well.

Proof of Theorem 4.1. We show that there exists a polynomial $P_{G,r}^{\sigma_\Lambda} = P_{G,r}^{\sigma_\Lambda}(\boldsymbol{\lambda})$, independent of λ_r, such that

$$Z_T^{\sigma_\Lambda}(r=1) = Z_G^{\sigma_\Lambda}(r=1)\cdot P_{G,r}^{\sigma_\Lambda} \quad \text{and} \quad Z_T^{\sigma_\Lambda}(r=0) = Z_G^{\sigma_\Lambda}(r=0)\cdot P_{G,r}^{\sigma_\Lambda}. \tag{4.1}$$

The high-level proof idea of (4.1) is similar to the corresponding result in [34, Theorem 3.1]. Let m be the number of edges with at least one endpoint in $V\setminus\Lambda$. We use induction on m. When $m=0$ the statement is trivial since $T=G$. Assume that (4.1) holds for all graphs and all conditioning with less than m edges. Suppose that the root r has d neighbors $v_1,\dots,v_d$. Define G' to be the graph obtained by replacing the vertex r with d vertices $r_1,\dots,r_d$ and then connecting $\{r_i,v_i\}$ for $1\le i\le d$.

Consider first the case where $(G\setminus\{r\})\setminus\Lambda$ is still connected. For each i, let $G_i = G' - r_i$. Define the 2-spin system on G_i with the same parameters $(\beta,\gamma,\boldsymbol{\lambda})$ plus the additional conditioning that the vertices $r_1,\dots,r_{i-1}$ are fixed to spin 0 while $r_{i+1},\dots,r_d$ are fixed to spin 1; we denote this conditioning by σ_{U_i} with $U_i=\{r_1,\dots,r_d\}\setminus\{r_i\}$. Then, $T=T_{\mathrm{SAW}}(G,r)$ can be generated by the following recursive procedure. Also see Figure 2 for an illustration.

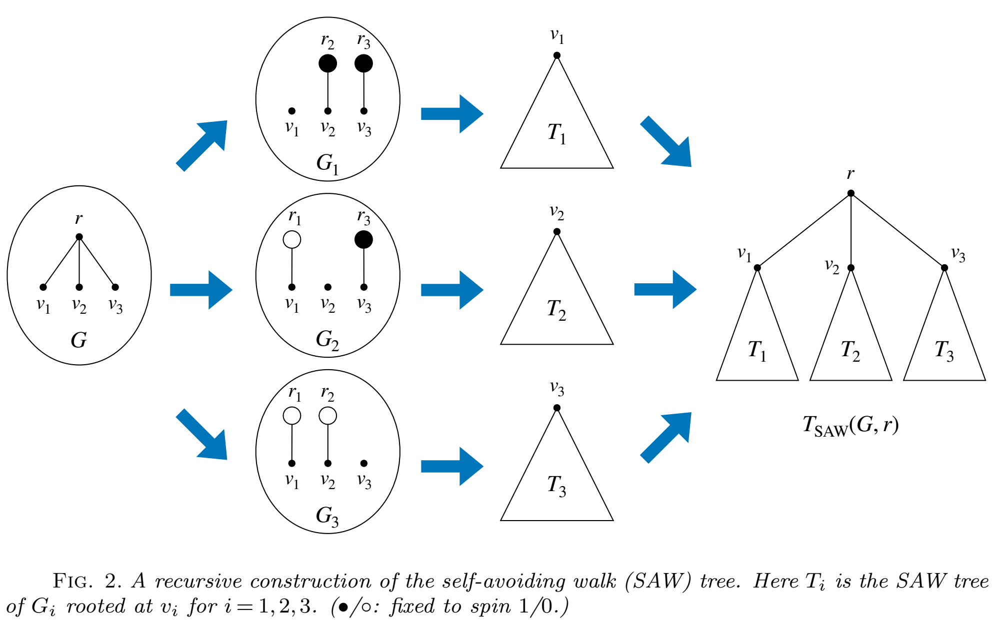

FIG. 2. *A recursive construction of the self-avoiding walk (SAW) tree. Here T_i is the SAW tree of G_i rooted at v_i for $i=1,2,3$. (●/○: fixed to spin 1/0.)*

Algorithm. $T_{\text{SAW}}(G,r)$.

1. For each i, let $T_i = T_{\text{SAW}}(G_i, v_i)$ with the conditioning σ_{U_i};
2. Let $T = T_{\text{SAW}}(G,r)$ be the union of r and $T_1,\dots,T_d$ by connecting $\{r, v_i\}$ for $1 \le i \le d$; output T.

For the purposes of the proof, we also consider the 2-spin system on G' with the same parameters $(\beta,\gamma,\boldsymbol{\lambda})$, with the exception that we let the vertices $r_1,\dots,r_d$ have no fields (i.e., setting $\lambda_{r_i}=1$ for $1\le i\le d$ instead of λ_r). We then observe that

$$Z_G^{\sigma_\Lambda}(r=1) = \lambda_r \cdot Z_{G'}^{\sigma_\Lambda}(r_1=1,\dots,r_d=1),$$

and the same holds with spin 1 replaced by 0. For $1\le i\le d$, let σ_{Λ_i} denote the union of the conditioning σ_Λ and σ_{U_i}, where $\Lambda_i = \Lambda\cup U_i$. Then for every $1\le i\le d$ we have

$$Z_{G'}^{\sigma_\Lambda}(r_1=0,\dots,r_{i-1}=0, r_i=1,\dots,r_d=1) = \beta\cdot Z_{G_i}^{\sigma_{\Lambda_i}}(v_i=1) + Z_{G_i}^{\sigma_{\Lambda_i}}(v_i=0).$$

Notice that both sides are independent of the field λ_r: for the left side, all r_i's do not have a field for the spin system on G'; for the right side, recall that we do not count the weight of fixed vertices for the conditional partition function for each G_i. Now define $Q_{G,r}^{\sigma_\Lambda} = Q_{G,r}^{\sigma_\Lambda}(\boldsymbol{\lambda})$ by

$$Q_{G,r}^{\sigma_\Lambda} = \prod_{i=2}^{d} Z_{G'}^{\sigma_\Lambda}(r_1=0,\dots,r_{i-1}=0, r_i=1,\dots,r_d=1),$$

which is independent of λ_r. Then we get

$$\begin{aligned} Z_G^{\sigma_\Lambda}(r=1)\cdot Q_{G,r}^{\sigma_\Lambda} &= \lambda_r\cdot\prod_{i=1}^{d} Z_{G'}^{\sigma_\Lambda}(r_1=0,\dots,r_{i-1}=0, r_i=1,\dots,r_d=1)\\ &= \lambda_r\cdot\prod_{i=1}^{d}\left(\beta\cdot Z_{G_i}^{\sigma_{\Lambda_i}}(v_i=1) + Z_{G_i}^{\sigma_{\Lambda_i}}(v_i=0)\right). \end{aligned}$$

Using a similar argument, we also have

$$\begin{aligned} Z_G^{\sigma_\Lambda}(r=0)\cdot Q_{G,r}^{\sigma_\Lambda} &= \prod_{i=1}^{d} Z_{G'}^{\sigma_\Lambda}(r_1=0,\dots,r_i=0, r_{i+1}=1,\dots,r_d=1)\\ &= \prod_{i=1}^{d}\left(Z_{G_i}^{\sigma_{\Lambda_i}}(v_i=1) + \gamma\cdot Z_{G_i}^{\sigma_{\Lambda_i}}(v_i=0)\right). \end{aligned}$$

Since we assume that $(G\setminus\{r\})\setminus\Lambda$ is connected, the graph $G_i\setminus\Lambda$ is also connected for each i. Then, by the induction hypothesis, for each i there exists a polynomial $P_{G_i,v_i}^{\sigma_{\Lambda_i}} = P_{G_i,v_i}^{\sigma_{\Lambda_i}}(\boldsymbol{\lambda})$ such that

$$Z_{T_i}^{\sigma_{\Lambda_i}}(r=1) = Z_{G_i}^{\sigma_{\Lambda_i}}(r=1)\cdot P_{G_i,v_i}^{\sigma_{\Lambda_i}} \quad\text{and}\quad Z_{T_i}^{\sigma_{\Lambda_i}}(r=0) = Z_{G_i}^{\sigma_{\Lambda_i}}(r=0)\cdot P_{G_i,v_i}^{\sigma_{\Lambda_i}};$$

these polynomials are independent of λ_r since the conditional partition functions for the G_i's do not involve λ_r. Now if we let

$$P_{G,r}^{\sigma_\Lambda} = Q_{G,r}^{\sigma_\Lambda}\cdot\prod_{i=1}^{d} P_{G_i,v_i}^{\sigma_{\Lambda_i}},$$

then it follows from the tree recursion that

$$\begin{aligned}
Z_T^{\sigma_\Lambda}(r=1) &= \lambda_r \cdot \prod_{i=1}^{d} \left(\beta \cdot Z_{T_i}^{\sigma_{\Lambda_i}}(v_i=1) + Z_{T_i}^{\sigma_{\Lambda_i}}(v_i=0)\right) \\
&= \lambda_r \cdot \prod_{i=1}^{d} \left(\beta \cdot Z_{G_i}^{\sigma_{\Lambda_i}}(v_i=1) \cdot P_{G_i,v_i}^{\sigma_{\Lambda_i}} + Z_{G_i}^{\sigma_{\Lambda_i}}(v_i=0) \cdot P_{G_i,v_i}^{\sigma_{\Lambda_i}}\right) \\
&= Z_G^{\sigma_\Lambda}(r=1) \cdot Q_{G,r}^{\sigma_\Lambda} \cdot \prod_{i=1}^{d} P_{G_i,v_i}^{\sigma_{\Lambda_i}} \\
&= Z_G^{\sigma_\Lambda}(r=1) \cdot P_{G,r}^{\sigma_\Lambda}.
\end{aligned}$$

The other equality $Z_T^{\sigma_\Lambda}(r=0) = Z_G^{\sigma_\Lambda}(r=0) \cdot P_{G,r}^{\sigma_\Lambda}$ is established in the same way. This completes the proof for the case when $(G \setminus \{r\}) \setminus \Lambda$ is connected.

If $(G \setminus \{r\}) \setminus \Lambda$ has two or more connected components, then we can construct $T_{\text{SAW}}(G,r)$ by the SAW tree of each component. Recall that G' is defined by splitting the vertex r into d copies in the graph G. Suppose that $G' \setminus \Lambda$ has k connected components for an integer $k \geq 2$. Let ${G'}_{(1)}, \dots, {G'}_{(k)}$ be the subgraphs induced by each component, together with vertices from Λ that are adjacent to it. For each j, let $G_{(j)}$ be the graph obtained from ${G'}_{(j)}$ by contracting all copies of r into one vertex $r_{(j)}$, and let $T_{(j)} = T_{\text{SAW}}({G'}_{(j)}, r_{(j)})$. Observe that once we contract the roots $r_{(1)}, \dots, r_{(k)}$ of $T_{(1)}, \dots, T_{(k)}$, the resulting tree is $T_{\text{SAW}}(G,r)$.

We define the 2-spin system on each $G_{(j)}$ with the same parameters $(\beta, \gamma, \boldsymbol{\lambda})$, except that the vertex $r_{(j)}$ does not have a field (i.e., $\lambda_{r_{(j)}} = 1$ instead of λ_r). For $1 \leq j \leq k$, let $\Lambda_{(j)} = \Lambda \cap V(G_{(j)})$, and let $\sigma_{\Lambda_{(j)}}$ be the configuration σ_Λ restricted on $\Lambda_{(j)}$. Then $G_{(j)} \setminus \Lambda_{(j)}$ is connected for every j, and since $k \geq 2$, each $G_{(j)}$ with conditioning $\sigma_{\Lambda_{(j)}}$ has fewer than m edges. Thus, we can apply the induction hypothesis; namely, for $1 \leq j \leq k$ there exists a polynomial $P_{G_{(i)},r_{(i)}}^{\sigma_{\Lambda_{(j)}}} = P_{G_{(i)},r_{(i)}}^{\sigma_{\Lambda_{(j)}}}(\boldsymbol{\lambda})$, which is independent of λ_r, such that

$$\begin{aligned}
& Z_{T_{(j)}}^{\sigma_{\Lambda_{(j)}}}(r_{(j)}=1) = Z_{G_{(j)}}^{\sigma_{\Lambda_{(j)}}}(r_{(j)}=1) \cdot P_{G_{(j)},r_{(j)}}^{\sigma_{\Lambda_{(j)}}} \\
\text{and} \quad & Z_{T_{(j)}}^{\sigma_{\Lambda_{(j)}}}(r_{(j)}=0) = Z_{G_{(j)}}^{\sigma_{\Lambda_{(j)}}}(r_{(j)}=0) \cdot P_{G_{(j)},r_{(j)}}^{\sigma_{\Lambda_{(j)}}}.
\end{aligned}$$

We define the polynomial $P_{G,r}^{\sigma_\Lambda} = P_{G,r}^{\sigma_\Lambda}(\boldsymbol{\lambda})$ to be

$$P_{G,r}^{\sigma_\Lambda} = \prod_{j=1}^{k} P_{G_{(j)},r_{(j)}}^{\sigma_{\Lambda_{(j)}}}.$$

It is then easy to check that

$$\begin{aligned}
Z_T^{\sigma_\Lambda}(r=1) &= \lambda_r \cdot \prod_{j=1}^{k} Z_{T_{(j)}}^{\sigma_{\Lambda_{(j)}}}(r_{(j)}=1) = \lambda_r \cdot \prod_{j=1}^{k} \left(Z_{G_{(j)}}^{\sigma_{\Lambda_{(j)}}}(r_{(j)}=1) \cdot P_{G_{(j)},r_{(j)}}^{\sigma_{\Lambda_{(j)}}} \right) \\
&= Z_G^{\sigma_\Lambda}(r=1) \cdot \prod_{j=1}^{k} P_{G_{(j)},r_{(j)}}^{\sigma_{\Lambda_{(j)}}} = Z_G^{\sigma_\Lambda}(r=1) \cdot P_{G,r}^{\sigma_\Lambda},
\end{aligned}$$

and similarly, $Z_T^{\sigma_\Lambda}(r=0) = Z_G^{\sigma_\Lambda}(r=0) \cdot P_{G,r}^{\sigma_\Lambda}$. The theorem then follows. ∎

5. Influence bound for trees. In this section, we study the influences of the root on other vertices in a tree. We give an upper bound on the total influences of the root on all vertices at a fixed distance away. To do this, we apply the potential method, which has been used to establish the correlation decay property (see, e.g., [20, 21, 15]). Given an arbitrary potential function Ψ, our upper bound is in terms of properties of Ψ, involving bounds on $\|\nabla H_d^\Psi\|_1$ and $|\psi|$ where $\psi = \Psi'$. We then deduce Lemma 3.5 in the case where Ψ is an (α, c)-potential.

Assume that $T = (V_T, E_T)$ is a tree rooted at r of maximum degree at most Δ. Let $\Lambda \subseteq V_T \setminus \{r\}$ and $\sigma_\Lambda \in \{0,1\}^\Lambda$ be arbitrary and fixed. Consider the 2-spin system on T, with parameters (β, γ, λ), conditioned on σ_Λ. We need to bound the influence $\mathcal{I}_T^{\sigma_\Lambda}(r \to v)$ from the root r to another vertex $v \in V_T$. Notice that if v is disconnected from r when Λ is removed, then $\mathcal{I}_T^{\sigma_\Lambda}(r \to v) = 0$ by the Markov property of spin systems. Therefore, we may assume that, by removing all such vertices, Λ contains only leaves of T.

For a vertex $v \in V_T$, let $T_v = (V_{T_v}, E_{T_v})$ be the subtree of T rooted at v that contains all descendants of v; note that $T_r = T$. We will write $L_v(k) \subseteq V_T \setminus \Lambda$ for the set of all free vertices at distance k away from v in T_v. We pay particular attention to the marginal ratio at v in the subtree T_v and write $R_v = R_{T_v}^{\sigma_\Lambda}(v)$ for simplicity. The $\log R_v$'s are related by the tree recursion H. If a vertex v has d children, denoted by $u_1, \dots, u_d$, then the tree recursion is given by

$$\log R_v = H_d(\log R_{u_1}, \dots, \log R_{u_d}),$$

where for $1 \le d \le \Delta$ and $(y_1, \dots, y_d) \in [-\infty, +\infty]^d$,

$$H_d(y_1, \dots, y_d) = \log \lambda + \sum_{i=1}^{d} \log\left(\frac{\beta e^{y_i} + 1}{e^{y_i} + \gamma}\right).$$

Also recall that for $y \in [-\infty, +\infty]$, we define

$$h(y) = -\frac{(1-\beta\gamma)e^y}{(\beta e^y + 1)(e^y + \gamma)}$$

and $\frac{\partial}{\partial y_i} H_d(y_1, \dots, y_d) = h(y_i)$ for all $1 \le i \le d \le \Delta$.

The following lemma allows us to bound the sum of all influences from the root to distance k using an arbitrary potential function.

Lemma 5.1. *Let $\Psi : [-\infty, +\infty] \to (-\infty, +\infty)$ be a differentiable and increasing (potential) function with image $S = \Psi[-\infty, +\infty]$ and derivative $\psi = \Psi'$. Denote the degree of the root r by Δ_r. Then for every integer $k \ge 1$,*

$$\sum_{v \in L_r(k)} |\mathcal{I}_T^{\sigma_\Lambda}(r \to v)| \le \Delta_r A_\Psi B_\Psi \left(\max_{1 \le d < \Delta} \sup_{\tilde{\boldsymbol{y}} \in S^d} \left\|\nabla H_d^\Psi(\tilde{\boldsymbol{y}})\right\|_1\right)^{k-1},$$

where

$$A_\Psi = \max_{u \in L_r(1)} \left\{\frac{|h(\log R_u)|}{\psi(\log R_u)}\right\} \quad \text{and} \quad B_\Psi = \max_{v \in L_r(k)} \{\psi(\log R_v)\}.$$

Before proving Lemma 5.1, we first present two useful properties of the influences on trees. First, it was shown in [3] that the influences satisfy the following form of chain rule on trees.

LEMMA 5.2 ([3, Lemma B.2]). *Suppose that $u, v, w \in V_T$ are three distinct vertices such that u is on the unique path from v to w. Then*

$$\mathcal{I}_T^{\sigma_\Lambda}(v \to w) = \mathcal{I}_T^{\sigma_\Lambda}(v \to u) \cdot \mathcal{I}_T^{\sigma_\Lambda}(u \to w).$$

Second, for two adjacent vertices on a tree, the influence from one to the other is given by the function h.

LEMMA 5.3. *Let $v \in V_T$, and let u be a child of v in the subtree T_v. Then*

$$\mathcal{I}_T^{\sigma_\Lambda}(v \to u) = h(\log R_u).$$

Proof. The lemma can be proved through an explicit computation of the influence. Here we present a more delicate proof utilizing Lemma 4.3, which gives some insight into the relation between the influence and the function h. We assume that v has d children in the subtree T_v, denoted by $u_1 = u$ and $u_2, \dots, u_d$, respectively. We also assume, as a more general setting than uniform fields, that each vertex w is attached to a field λ_w of its own. Then Lemma 4.3 and the tree recursion imply that

$$\begin{aligned}
\mathcal{I}_T^{\sigma_\Lambda}(v \to u) = \mathcal{I}_{T_v}^{\sigma_\Lambda}(v \to u) &= \left(\lambda_u \frac{\partial}{\partial \lambda_u}\right) \log R_v \\
&= \left(\lambda_u \frac{\partial}{\partial \lambda_u}\right) H_d(\log R_{u_1}, \dots, \log R_{u_d}) \\
&= \sum_{i=1}^d \frac{\partial}{\partial \log R_{u_i}} H_d(\log R_{u_1}, \dots, \log R_{u_d}) \cdot \left(\lambda_u \frac{\partial}{\partial \lambda_u}\right) \log R_{u_i} \\
&= \sum_{i=1}^d h(\log R_{u_i}) \cdot \mathcal{I}_{T_{u_i}}^{\sigma_\Lambda}(u_i \to u) = h(\log R_u),
\end{aligned}$$

where the last equality is due to $\mathcal{I}_{T_{u_i}}^{\sigma_\Lambda}(u_i \to u) = 0$ for $u_i \neq u$ since $u \notin T_{u_i}$, and $\mathcal{I}_{T_u}^{\sigma_\Lambda}(u \to u) = 1$. Note that the argument still holds even if some children u_i's are fixed to certain spins. ∎

We are now ready to prove Lemma 5.1.

Proof of Lemma 5.1. For a vertex $v \in V_T$, denote the number of its children by d_v; note that $d_r = \Delta_r$. Let $u_1, \dots, u_{\Delta_r}$ be the children of the root r. We may assume that all these children of r are free, since if u_i is fixed, then $\mathcal{I}_T^{\sigma_\Lambda}(r \to u_i) = 0$ by definition. Then by Lemmas 5.2 and 5.3, we get

$$\begin{aligned}
\sum_{v \in L_r(k)} |\mathcal{I}_T^{\sigma_\Lambda}(r \to v)| &= \sum_{i=1}^{\Delta_r} |\mathcal{I}_T^{\sigma_\Lambda}(r \to u_i)| \cdot \left(\sum_{v \in L_{u_i}(k-1)} |\mathcal{I}_T^{\sigma_\Lambda}(u_i \to v)|\right) \\
&= \sum_{i=1}^{\Delta_r} |h(\log R_{u_i})| \cdot \left(\sum_{v \in L_{u_i}(k-1)} |\mathcal{I}_T^{\sigma_\Lambda}(u_i \to v)|\right) \\
&= \sum_{i=1}^{\Delta_r} \frac{|h(\log R_{u_i})|}{\psi(\log R_{u_i})} \cdot \left(\sum_{v \in L_{u_i}(k-1)} \psi(\log R_{u_i})\, |\mathcal{I}_T^{\sigma_\Lambda}(u_i \to v)|\right).
\end{aligned}$$

Hence, we obtain that

$$\sum_{v\in L_r(k)} |\mathcal{I}_T^{\sigma_\Lambda}(r\to v)| \le \Delta_r \cdot \max_{1\le i\le \Delta_r}\left\{\frac{|h(\log R_{u_i})|}{\psi(\log R_{u_i})}\right\} \cdot \max_{1\le i\le \Delta_r}\left\{\sum_{v\in L_{u_i}(k-1)} \psi(\log R_{u_i})\,|\mathcal{I}_T^{\sigma_\Lambda}(u_i\to v)|\right\}. \tag{5.1}$$

Next, we show by induction that for every vertex $u\in V_T\setminus\{r\}$ and every integer $k\ge 0$ we have

$$\sum_{v\in L_u(k)} \psi(\log R_u)\,|\mathcal{I}_T^{\sigma_\Lambda}(u\to v)| \le \max_{v\in L_u(k)}\{\psi(\log R_v)\}\cdot\left(\max_{w\in V_{T_u}}\sup_{\tilde{\boldsymbol{y}}\in S^{d_w}}\left\|\nabla H_{d_w}^{\Psi}(\tilde{\boldsymbol{y}})\right\|_1\right)^k. \tag{5.2}$$

Observe that once we establish (5.2), the lemma follows immediately by plugging (5.2) into (5.1). We will use induction on k to prove (5.2). When $k=0$, if $u\in\Lambda$ is fixed, then $L_u(0)=\emptyset$ and there is nothing to show; otherwise, (5.2) becomes

$$\psi(\log R_u)\,|\mathcal{I}_T^{\sigma_\Lambda}(u\to u)| \le \psi(\log R_u),$$

which holds with equality since $\mathcal{I}_T^{\sigma_\Lambda}(u\to u)=1$. Now suppose that (5.2) holds for some integer $k-1\ge 0$ (and for every vertex $u\in V_T\setminus\{r\}$). Let $u\in V_T\setminus\{r\}$ be arbitrary, and denote the children of u by $w_1,\dots,w_d$, where $1\le d<\Delta$ (if $d=0$, then $L_u(k)=\emptyset$ and (5.2) holds trivially). Again by Lemmas 5.2 and 5.3 we have

$$\begin{aligned}
&\sum_{v\in L_u(k)} \psi(\log R_u)\,|\mathcal{I}_T^{\sigma_\Lambda}(u\to v)| \\
&\quad= \sum_{i=1}^d \psi(\log R_u)\,|\mathcal{I}_T^{\sigma_\Lambda}(u\to w_i)| \sum_{v\in L_{w_i}(k-1)} |\mathcal{I}_T^{\sigma_\Lambda}(w_i\to v)| \\
&\quad= \sum_{i=1}^d \frac{\psi(\log R_u)}{\psi(\log R_{w_i})}\,|h(\log R_{w_i})| \sum_{v\in L_{w_i}(k-1)} \psi(\log R_{w_i})\,|\mathcal{I}_T^{\sigma_\Lambda}(w_i\to v)|.
\end{aligned}$$

Using the induction hypothesis, we get

$$\begin{aligned}
&\sum_{v\in L_u(k)} \psi(\log R_u)\,|\mathcal{I}_T^{\sigma_\Lambda}(u\to v)| \\
&\quad\le \sum_{i=1}^d \frac{\psi(\log R_u)}{\psi(\log R_{w_i})}\,|h(\log R_{w_i})| \\
&\qquad\cdot \max_{v\in L_{w_i}(k-1)}\{\psi(\log R_v)\}\cdot\left(\max_{w\in V_{T_{w_i}}}\sup_{\tilde{\boldsymbol{y}}\in S^{d_w}}\left\|\nabla H_{d_w}^{\Psi}(\tilde{\boldsymbol{y}})\right\|_1\right)^{k-1} \\
&\quad\le \max_{v\in L_u(k)}\{\psi(\log R_v)\} \\
&\qquad\cdot\left(\max_{w\in V_{T_u}\setminus\{u\}}\sup_{\tilde{\boldsymbol{y}}\in S^{d_w}}\left\|\nabla H_{d_w}^{\Psi}(\tilde{\boldsymbol{y}})\right\|_1\right)^{k-1}\cdot\sum_{i=1}^d \frac{\psi(\log R_u)}{\psi(\log R_{w_i})}\,|h(\log R_{w_i})| \\
&\quad\le \max_{v\in L_u(k)}\{\psi(\log R_v)\}\cdot\left(\max_{w\in V_{T_u}}\sup_{\tilde{\boldsymbol{y}}\in S^{d_w}}\left\|\nabla H_{d_w}^{\Psi}(\tilde{\boldsymbol{y}})\right\|_1\right)^{k},
\end{aligned}$$

where the last inequality follows from the following calculation:

$$\begin{aligned}\sum_{i=1}^{d} \frac{\psi(\log R_u)}{\psi(\log R_{w_i})} |h(\log R_{w_i})| &= \sum_{i=1}^{d} \left| \frac{\partial}{\partial \Psi(\log R_{w_i})} H_d^{\Psi}\left(\Psi(\log R_{w_1}), \dots, \Psi(\log R_{w_d})\right) \right| \\ &= \left\| \nabla H_d^{\Psi}\left(\Psi(\log R_{w_1}), \dots, \Psi(\log R_{w_d})\right) \right\|_1 .\end{aligned}$$

This establishes (5.2) and thus completes the proof of the lemma. ◻

We then derive Lemma 3.5 as a corollary.

Proof of Lemma 3.5. Since Ψ is an (α, c)-potential, the contraction condition implies that

$$\max_{1 \le d < \Delta} \sup_{\tilde{\boldsymbol{y}} \in S^d} \left\| \nabla H_d^{\Psi}(\tilde{\boldsymbol{y}}) \right\|_1 \le 1 - \alpha.$$

Meanwhile, since the degree of a vertex $v \in V_T \setminus \{r\}$ in the subtree T_v is less than Δ, we have $\log R_v \in J$. Then the boundedness condition implies that for all $u \in L_r(1)$ and $v \in L_r(k)$,

$$\frac{\psi(\log R_v)}{\psi(\log R_u)} \cdot |h(\log R_u)| \le \frac{c}{\Delta}.$$

Therefore, we get

$$\Delta_r A_\Psi B_\Psi = \Delta_r \cdot \max_{u \in L_r(1)} \left\{ \frac{|h(\log R_u)|}{\psi(\log R_u)} \right\} \cdot \max_{v \in L_r(k)} \{\psi(\log R_v)\} \le c.$$

The lemma then follows immediately from Lemma 5.1. ◻

6. Verifying a good potential: Contraction. In this section, we make a first step for proving Lemma 3.6. Let $\Delta \ge 3$ be an integer. Let β, γ, λ be reals such that $0 \le \beta \le \gamma$, $\gamma > 0$, $\beta\gamma < 1$, and $\lambda > 0$. Recall that we define our potential function $\Psi : [-\infty, +\infty] \to (-\infty, +\infty)$ through its derivative by

$$\Psi'(y) = \psi(y) = \sqrt{\frac{(1 - \beta\gamma)e^y}{(\beta e^y + 1)(e^y + \gamma)}}, \qquad \Psi(0) = 0.$$

We include a short proof in Appendix C to show that Ψ is well-defined. If (β, γ, λ) is up-to-Δ unique with gap $\delta \in (0, 1)$, then we show that Ψ satisfies the contraction condition for $\alpha = \delta/2$. This holds for all parameters (β, γ, λ) in the uniqueness region, *without* requiring that $\gamma \le 1$. Later, in Appendix E we establish the boundedness condition for Ψ when $\gamma \le 1$, completing the proof of Lemma 3.6. The case of $\gamma > 1$ is more complicated and is deferred to section 7. We refer the reader to [21] for a heuristic justification of this potential function.

Before giving our proof, we first point out that the potential function Ψ is essentially the same potential function Φ used in [21] (notice that [21] uses φ as the notation of the potential function and $\Phi = \varphi'$ for its derivative). Recall that the tree recursion for the marginal ratios is given by the function $F_d : [0, +\infty]^d \to [0, +\infty]$ where $1 \le d \le \Delta$ such that for all $(x_1, \dots, x_d) \in [0, +\infty]^d$,

$$F_d(x_1, \dots, x_d) = \lambda \prod_{i=1}^{d} \frac{\beta x_i + 1}{x_i + \gamma}.$$

The potential function $\Phi : [0,+\infty] \to (-\infty,+\infty)$ from [21] is defined implicitly via its derivative as

$$\Phi'(x) = \varphi(x) = \frac{1}{\sqrt{x(\beta x+1)(x+\gamma)}}, \qquad \Phi(1) = 0.$$

The following lemma explains how we obtain our potential Ψ from Φ.

Lemma 6.1. *We have* $\Psi = \sqrt{1-\beta\gamma} \cdot (\Phi \circ \exp)$*; namely,* $\Psi(y) = \sqrt{1-\beta\gamma} \cdot \Phi(e^y)$ *for all* $y \in [-\infty,+\infty]$.

Proof. It is straightforward to check that

$$\begin{aligned}\psi(y) &= \sqrt{\frac{(1-\beta\gamma)e^y}{(\beta e^y+1)(e^y+\gamma)}}\\ &= \sqrt{1-\beta\gamma} \cdot e^y \cdot \sqrt{\frac{1}{e^y(\beta e^y+1)(e^y+\gamma)}} = \sqrt{1-\beta\gamma} \cdot e^y \varphi(e^y).\end{aligned}$$

Therefore,

$$\begin{aligned}\Psi(y) &= \int_0^y \psi(t)\,\mathrm{d}t = \sqrt{1-\beta\gamma} \cdot \int_0^y e^t \varphi(e^t)\,\mathrm{d}t\\ &= \sqrt{1-\beta\gamma} \cdot \int_1^{e^y} \varphi(s)\,\mathrm{d}s = \sqrt{1-\beta\gamma} \cdot \Phi(e^y).\end{aligned}$$ ∎

Combining the results of Lemmas 12, 13, and 14 from [21], we get that the potential function Φ satisfies the following gradient bound when the parameters (β,γ,λ) are in the uniqueness region. Note that this can be regarded as the contraction condition but rather for Φ and F_d.

Theorem 6.2 ([21]). *Let* $S_\Phi = \Phi[0,+\infty]$ *be the image of* Φ*. If the parameters* (β,γ,λ) *are up-to-*Δ *unique with gap* $\delta \in (0,1)$*, then for every integer* d *such that* $1 \le d < \Delta$ *and every* $(\tilde{x}_1,\dots,\tilde{x}_d) \in S_\Phi^d$*,*

$$\left\|\nabla F_d^\Phi(\tilde{x}_1,\dots,\tilde{x}_d)\right\|_1 \le \sqrt{1-\delta},$$

where $F_d^\Phi = \Phi \circ F_d \circ \Phi^{-1}$*.*

Recall our definition from subsection 1.1. The tree recursion, in terms of the log marginal ratios, is described by the function $H_d : [-\infty,+\infty]^d \to [-\infty,+\infty]$ where $1 \le d \le \Delta$ such that for every $(y_1,\dots,y_d) \in [-\infty,+\infty]^d$,

$$H_d(y_1,\dots,y_d) = \log\lambda + \sum_{i=1}^d \log\left(\frac{\beta e^{y_i}+1}{e^{y_i}+\gamma}\right).$$

Observe that $H_d = \log \circ F_d \circ \exp$, since we move from ratios to log ratios. We are now ready to establish the contraction condition for Ψ.

Lemma 6.3. *Let* $S_\Psi = \Psi[-\infty,+\infty]$ *be the image of* Ψ*. If the parameters* (β,γ,λ) *are up-to-*Δ *unique with gap* $\delta \in (0,1)$*, then for every integer* d *such that* $1 \le d < \Delta$ *and every* $(\tilde{y}_1,\dots,\tilde{y}_d) \in S_\Psi^d$*,*

$$\left\|\nabla H_d^\Psi(\tilde{y}_1,\dots,\tilde{y}_d)\right\|_1 \le \sqrt{1-\delta},$$

where $H_d^\Psi = \Psi \circ H_d \circ \Psi^{-1}$*.*

Proof. Define the linear function $a:\mathbb{R}\to\mathbb{R}$ to be $a(x)=\sqrt{1-\beta\gamma}\cdot x$ for $x\in\mathbb{R}$. Then Lemma 6.1 gives $\Psi=a\circ\Phi\circ\exp$, and thereby $\Psi\circ\log=a\circ\Phi$. It follows that for every $1\le d<\Delta$,

$$H_d^{\Psi}=\Psi\circ H_d\circ\Psi^{-1}=\Psi\circ\log\circ F_d\circ\exp\circ\Psi^{-1}=a\circ\Phi\circ F_d\circ\Phi^{-1}\circ a^{-1}=a\circ F_d^{\Phi}\circ a^{-1}.$$

That means for every $(\tilde{y}_1,\dots,\tilde{y}_d)\in S_{\Psi}^d$ we have

$$H_d^{\Psi}(\tilde{y}_1,\dots,\tilde{y}_d)=\sqrt{1-\beta\gamma}\cdot F_d^{\Phi}(\tilde{x}_1,\dots,\tilde{x}_d),$$

where $\tilde{x}_i=\tilde{y}_i/\sqrt{1-\beta\gamma}$ for $1\le i\le d$. Then, for each i,

$$\frac{\partial}{\partial\tilde{y}_i}H_d^{\Psi}(\tilde{y}_1,\dots,\tilde{y}_d)=\sqrt{1-\beta\gamma}\cdot\frac{\partial}{\partial\tilde{x}_i}F_d^{\Phi}(\tilde{x}_1,\dots,\tilde{x}_d)\cdot\frac{\mathrm{d}\tilde{x}_i}{\mathrm{d}\tilde{y}_i}=\frac{\partial}{\partial\tilde{x}_i}F_d^{\Phi}(\tilde{x}_1,\dots,\tilde{x}_d).$$

This implies that $\nabla H_d^{\Psi}(\tilde{y}_1,\dots,\tilde{y}_d)=\nabla F_d^{\Phi}(\tilde{x}_1,\dots,\tilde{x}_d)$ for all $(\tilde{y}_1,\dots,\tilde{y}_d)\in S_{\Psi}^d$, and the lemma then follows from Theorem 6.2. ∎

7. Remaining antiferromagnetic cases: $\sqrt{\beta\gamma}\le\frac{\Delta-2}{\Delta}$ and $\gamma>1$. In this section, we discuss the case where $\sqrt{\beta\gamma}\le\frac{\Delta-2}{\Delta}$ and $\gamma>1$. As studied in [21], in this case the uniqueness region is more complicated. For example, there exists a critical $\lambda_c^*>0$ such that the 2-spin system with $\lambda<\lambda_c^*$ is in the uniqueness region for arbitrary graphs; namely, (β,γ,λ) is up-to-∞ unique. To deal with large degrees, we need to relax the boundedness condition in Definition 1.4 and define a more general version of (α,c)-potentials. We shall see that Theorem 1.5 still holds for this general (α,c)-potential. The reason behind this is that in order to bound the maximum eigenvalue of the influence matrix, it suffices to consider a vertex-weighted sum of absolute influences of a vertex with large degree.

Remark 7.1. We give more background on the uniqueness region in Appendix E.1. Note that in a recent revision of [21], the authors updated the descriptions of the uniqueness region for the case where $\sqrt{\beta\gamma}\le\frac{\Delta-2}{\Delta}$ and $\gamma>1$, fixing a small error in the previous version. Statements and proofs in this section and Appendix E of this paper are also adjusted accordingly based on the updated preprint [22].

Recall that our goal is to bound the maximum eigenvalue of the matrix $\mathcal{I}_G^{\sigma_\Lambda}$. We can do this by upper bounding the absolute row sum $\sum_{v\in V\setminus\Lambda}|\mathcal{I}_G^{\sigma_\Lambda}(r\to v)|$ for fixed r, thereby giving a valid upper bound on $\lambda_{\max}(\mathcal{I}_G^{\sigma_\Lambda})$. However, this approach does not work when $\sqrt{\beta\gamma}\le\frac{\Delta-2}{\Delta}$ and $\gamma>1$. In this case, the potential Ψ fails to be an (α,c)-potential for a universal constant c independent of Δ. In fact, no such (α,c)-potentials exist, as the absolute row sum $\sum_{v\in V\setminus\Lambda}|\mathcal{I}_G^{\sigma_\Lambda}(r\to v)|$ can be as large as $\Theta(\Delta)$. Especially, if the parameters (β,γ,λ) are up-to-∞ unique, which means the spin system has uniqueness for arbitrary graphs, then the absolute row sum $\sum_{v\in V\setminus\Lambda}|\mathcal{I}_G^{\sigma_\Lambda}(r\to v)|$ can be $\Theta(n)$ where $n=|V|$. We give a specific example where this is the case.

Example 7.2. Consider the antiferromagnetic 2-spin system specified by parameters $\beta=0$, $\gamma>1$, and $\lambda>0$ on the star graph centered at r with $\Delta<\infty$ leaves. A simple calculation reveals that $|\mathcal{I}_G(r\to v)|=\frac{\lambda}{\lambda+\gamma}$ for any leaf vertex $v\ne r$. Hence, $\sum_{v\ne r}|\mathcal{I}_G(r\to v)|=\Delta\cdot\frac{\lambda}{\lambda+\gamma}$. Now, since $\gamma>1$, we have

$$\lambda_c=\lambda_c(\gamma,\Delta)=\min_{1<d<\Delta}\frac{\gamma^{d+1}d^d}{(d-1)^{d+1}}=\Theta_\gamma(1),$$

forcing $\sum_{v\neq r}|\mathcal{I}_G(r\to v)| = \Theta_\gamma(\Delta)$ even when $\lambda < \lambda_c$ lies in the uniqueness region. However, we still have $\lambda_{\max}(\mathcal{I}_G) = O(1)$ since $\sum_{v\neq r}|\mathcal{I}_G(v\to r)| = O(1)$.

To solve this issue, one might want to consider the absolute column sum, involving the sum of absolute influences on a fixed vertex. However, this will not allow us to use the beautiful connection between graphs and SAW trees as shown in Lemma 3.3. Instead, we consider here a vertex-weighted version of the absolute row sum of $\mathcal{I}_G^{\sigma_\Lambda}$, which also upper bounds the maximum eigenvalue.

LEMMA 7.3. *Let $\rho : V \to \mathbb{R}^+$ be a positive weight function of vertices. If there is a constant $\xi > 0$ such that for every $r \in V$ we have*

$$\sum_{v\in V\setminus\Lambda} \rho_v \cdot |\mathcal{I}_G^{\sigma_\Lambda}(r\to v)| \le \xi\cdot\rho_r, \tag{7.1}$$

then $\lambda_{\max}(\mathcal{I}_G^{\sigma_\Lambda}) \le \xi$.

Proof. Let $\mathcal{P} = \mathrm{diag}\{\rho_v : v \in V\setminus\Lambda\}$. Then the assumption is equivalent to $\|\mathcal{P}^{-1}\mathcal{I}_G^{\sigma_\Lambda}\mathcal{P}\|_\infty \le \xi$. It follows that $\lambda_{\max}(\mathcal{I}_G^{\sigma_\Lambda}) = \lambda_{\max}(\mathcal{P}^{-1}\mathcal{I}_G^{\sigma_\Lambda}\mathcal{P}) \le \xi$. ∎

We then modify our definition of (α, c)-potentials from Definition 1.4 which allows a weaker boundedness condition. We remark that the only two differences between Definitions 1.4 and 7.4 is that we allow $\Delta=\infty$ and that the boundedness condition is relaxed to what we call General Boundedness. Recall that for every $0\le d<\Delta$, we let $J_d = [\log(\lambda\beta^d), \log(\lambda/\gamma^d)]$ when $\beta\gamma<1$, and $J_d = [\log(\lambda/\gamma^d), \log(\lambda\beta^d)]$ when $\beta\gamma>1$.

DEFINITION 7.4 (general (α, c)-potential function). *Let $\Delta \ge 3$ be an integer, or let $\Delta=\infty$. Let β,γ,λ be reals such that $0\le\beta\le\gamma$, $\gamma>0$, and $\lambda>0$. Let $\Psi : [-\infty,+\infty] \to (-\infty,+\infty)$ be a differentiable and increasing function with image $S = \Psi[-\infty,+\infty]$ and derivative $\psi = \Psi'$. For any $\alpha\in(0,1)$ and $c>0$, we say Ψ is a* general (α, c)-potential function *with respect to Δ and (β,γ,λ) if it satisfies the following conditions:*

1. *(Contraction) For every integer d such that $1\le d<\Delta$, and every tuple $(\tilde{y}_1,\dots,\tilde{y}_d)\in S^d$, we have*
$$\left\|\nabla H_d^\Psi(\tilde{y}_1,\dots,\tilde{y}_d)\right\|_1 = \sum_{i=1}^d \frac{\psi(y)}{\psi(y_i)}\cdot|h(y_i)| \le 1-\alpha,$$
where $H_d^\Psi = \Psi\circ H_d\circ\Psi^{-1}$, $y_i = \Psi^{-1}(\tilde{y}_i)$ for $1\le i\le d$, and $y = H_d(y_1,\dots,y_d)$.
2. *(General Boundedness) For all integers d_1, d_2 such that $0\le d_1,d_2<\Delta$, and all reals $y_1\in J_{d_1}, y_2\in J_{d_2}$, we have*
$$\frac{\psi(y_2)}{\psi(y_1)}\cdot|h(y_1)| \le \frac{2c}{d_1+d_2+2}.$$

Notice that General Boundedness is a weaker condition than Boundedness. To see this, if a potential function Ψ satisfies Boundedness with parameter c, then for every $0\le d_i<\Delta$ and every $y_i\in J_{d_i}$ where $i=1,2$ we have

$$\frac{\psi(y_2)}{\psi(y_1)}\cdot|h(y_1)| \le \frac{c}{\Delta} \le \frac{2c}{d_1+d_2+2}.$$

The following theorem generalizes Theorem 1.5 and shows that a general (α, c)-potential function is sufficient to establish rapid mixing of the Glauber dynamics.

THEOREM 7.5. *Let $\Delta \geq 3$ be an integer, or let $\Delta = +\infty$. Let β, γ, λ be reals such that $0 \leq \beta \leq \gamma$, $\gamma > 0$, and $\lambda > 0$. Suppose that there is a general (α, c)-potential with respect to Δ and (β, γ, λ) for some $\alpha \in (0,1)$ and $c > 0$. Then for every n-vertex graph G of maximum degree Δ, the mixing time of the Glauber dynamics for the 2-spin system on G with parameters (β, γ, λ) is $O(n^{2+2c/\alpha})$.*

We then give a counterpart of Lemma 3.6, showing that Ψ is a general (α, c)-potential when $\sqrt{\beta\gamma} \leq \frac{\Delta-2}{\Delta}$ and $\gamma > 1$. Theorem 1.3 for this case is then obtained from Theorem 7.5 and Lemma 7.6.

LEMMA 7.6. *Let $\Delta \geq 3$ be an integer. Let β, γ, λ be reals such that $0 \leq \beta < 1 < \gamma$ and $\sqrt{\beta\gamma} \leq \frac{\Delta-2}{\Delta}$. Assume that (β, γ, λ) is up-to-Δ unique with gap $\delta \in (0,1)$. Then the function Ψ defined implicitly by (3.3) is a general (α, c)-potential function with $\alpha \geq \delta/2$ and $c \leq 18$; we can further take $c \leq 4$ if $\beta = 0$.*

The proof of Theorem 7.5 can be found in Appendix D. For Lemma 7.6, the contraction condition of Ψ follows from Lemma 6.3, and General Boundedness is proved in Appendix E together with all other cases.

8. Ferromagnetic cases.

In the ferromagnetic case, the best known correlation decay results are given in [15, 29]. Using the potential functions in [15] and [16, 30], we show the following two results, which match the known correlation decay results. See [15, 14] for further discussion on the tightness of these results.

To establish our next result, we use the potential function from [29], which turns out to be an (α, c)-potential function for constants $\alpha = \Theta(\delta)$ and $c = O(1)$.

THEOREM 8.1. *Fix an integer $\Delta \geq 3$, real numbers $\beta, \gamma, \lambda > 0$, and $0 < \delta < 1$, and assume (β, γ, λ) satisfies one of the following three conditions:*

1. $\frac{\Delta-2+\delta}{\Delta-\delta} \leq \sqrt{\beta\gamma} \leq \frac{\Delta-\delta}{\Delta-2+\delta}$, *and $\lambda > 0$ is arbitrary;*
2. $\sqrt{\beta\gamma} \geq \frac{\Delta}{\Delta-2}$ *and* $0 < \lambda \leq (1-\delta)\frac{\gamma}{\max\{1, \beta^{\Delta-1}\} \cdot ((\Delta-2)\beta\gamma - \Delta)}$;
3. $\sqrt{\beta\gamma} \geq \frac{\Delta}{\Delta-2}$ *and* $\lambda \geq \frac{1}{1-\delta} \cdot \frac{(\Delta-2)\beta\gamma - \Delta}{\beta \cdot \min\{1, 1/\gamma^{\Delta-1}\}}$.

Then the identity function $\Psi(y) = y$ (based on the potential given in [29]) is an (α, c)-potential function for $\alpha = \Theta(\delta)$ and $c \leq O(1)$. Furthermore, for every n-vertex graph G 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 $O(n^{2+c/\delta})$ for a universal constant $c > 0$.

Remark 8.2. Condition 1 includes both the ferromagnetic case $1 < \sqrt{\beta\gamma} \leq \frac{\Delta-\delta}{\Delta-2+\delta}$ and the antiferromagnetic case $\frac{\Delta-2+\delta}{\Delta-\delta} \leq \sqrt{\beta\gamma} < 1$. Note that in both cases (β, γ, λ) is up-to-Δ unique with gap δ. For the antiferromagnetic case, the identity function Ψ is an (α, c)-potential with $c \leq 1.5$ and a better contraction rate $\alpha \geq \delta$, compared with the bound $\alpha \geq \delta/2$ of the potential Ψ given by (3.3) in Lemma 3.6. For the ferromagnetic case with $\beta = \gamma > 1$ (Ising model), a stronger result by [25] was known, which gives $O(n \log n)$ mixing.

The potential function from [15] is indeed an (α, c)-potential, but c must, unfortunately, depend on Δ. We have the following result, which is weaker than the correlation decay algorithm in [15] for unbounded degree graphs.

THEOREM 8.3. *Fix an integer $\Delta \geq 3$ and nonnegative real numbers β, γ, λ satisfying $\beta \leq 1 \leq \gamma$, $\sqrt{\beta\gamma} \geq \frac{\Delta}{\Delta-2}$, and $\lambda < \left(\frac{\gamma}{\beta}\right)^{\frac{\sqrt{\beta\gamma}}{\sqrt{\beta\gamma}-1}}$. Then for every n-vertex graph G with maximum degree at most Δ, the mixing time of the Glauber dynamics for the ferromagnetic 2-spin system on G with parameters (β, γ, λ) is $O(n^C)$ for a constant C depending only on $\beta, \gamma, \lambda, \Delta$ but not n.*

Proofs of these theorems are provided in Appendix F.

9. Further remarks. In this work, we essentially showed that if the tree recursion for a 2-spin system contracts under a suitable potential function with rate $1-\delta$, then $n^{O(1/\delta)}$ mixing holds. However, we believe the following folklore conjecture holds.

CONJECTURE 9.1. *Fix $\Delta \geq 3$ and nonnegative real numbers β, γ, λ satisfying $\beta\gamma < 1$ and $0 < \delta < 1$, and assume (β, γ, λ) is up-to-Δ unique with gap δ. Then there exists a $C = C(\beta, \gamma, \lambda, \delta)$ such that the Glauber dynamics mixes in at most $Cn \log n$ steps for all n-vertex graphs of maximum degree at most Δ.*

Since the preliminary proceedings version of this paper [11], great progress has been made towards the conjecture and it has been proved for most cases including both the Ising and the hardcore model; see [11, 8, 7, 3, 10, 9].

Appendix A. Proof of main results. In this section we give the proofs of Theorems 1.1–1.3 and 1.5.

Proof of Theorem 1.5. Note that since the transition matrix P for the Glauber dynamics has all nonnegative eigenvalues, we have that $\lambda^*(P) = 1 - \lambda_2(P)$, and so in order to deduce the mixing time, it suffices to lower bound $1 - \lambda_2(P)$. We do this by employing Theorem 3.2. It suffices to show $(\eta_0, \dots, \eta_{n-2})$-spectral independence for sufficiently small η_i.

To bound η_i, it suffices to bound $\sum_{v \in V \setminus \{r\}} |\mathcal{I}_G^{\sigma_\Lambda}(r \to v)|$ for all graphs $G = (V, E)$ with $n = |V|$ vertices and all boundary conditions σ_Λ on a subset Λ of i vertices. We claim the following:

$$\sum_{v \in V \setminus \{r\}} |\mathcal{I}_G^{\sigma_\Lambda}(r \to v)| \leq \min\left\{\frac{c}{\alpha}, C(n-i-1)\right\}, \tag{A.1}$$

where $C \in (0,1)$ is a constant depending only on $\beta, \gamma, \lambda, \Delta$. The first upper bound $\frac{c}{\delta}$ is deduced by

$$\begin{aligned}
&(\text{Lemma 3.3; } T = T_{\text{SAW}}(G,r)) && \textstyle\sum_{v \in V \setminus \{r\}} |\mathcal{I}_G^{\sigma_\Lambda}(r \to v)| \leq \sum_{v \in V_T \setminus \{r\}} |\mathcal{I}_T^{\sigma_\Lambda}(r \to v)| \\
&(\text{split the sum by levels}) && = \sum_{k=1}^{\infty} \sum_{v \in L_r(k)} \left|\mathcal{I}_T^{\sigma_{\Lambda_\Lambda}}(r \to v)\right| \\
&(\text{Lemma 3.5}) && \leq c \sum_{k=1}^{\infty} (1-\alpha)^{k-1} \\
& && = \frac{c}{\alpha}.
\end{aligned}$$

The second upper bound $C(n-i-1)$ is more trivial. Intuitively, it means each absolute pairwise influence $|\mathcal{I}_G^{\sigma_\Lambda}(r \to v)|$ is at most some constant C, and hence the sum of absolute influences is upper bounded by $C(n-i-1)$. The following two claims, whose proofs are provided in Appendix A.2, give a more precise statement.

CLAIM A.1 (antiferromagnetic case). *Fix an integer $\Delta \geq 3$ and real numbers β, γ, λ, and assume $0 \leq \beta \leq \gamma$, $\gamma > 0$, $\beta\gamma < 1$, and $\lambda > 0$. Then for every n-vertex graph G of maximum degree at most Δ, the antiferromagnetic 2-spin system on G with parameters (β, γ, λ) is Cn-spectrally independent for a constant $0 < C < 1$ depending only on $\beta, \gamma, \lambda, \Delta$. Furthermore, if (β, γ, Δ) is up-to-Δ unique, then we can drop the dependence on Δ.*

CLAIM A.2 (ferromagnetic case). *Fix an integer $\Delta \geq 3$ and positive real numbers β, γ, λ, and assume $\beta \leq \gamma$ and $\beta\gamma > 1$. Then for every n-vertex graph G of maximum* *degree at most* Δ*, the ferromagnetic* 2*-spin system on* G *with parameters* (β, γ, λ) *is* Cn*-spectrally independent for a constant* $0 < C < 1$ *depending only on* $\beta, \gamma, \lambda, \Delta$*.*

With (A.1) in hand, we immediately see that by Theorem 3.2,

$$1 - \lambda_2(P) \geq \frac{1}{n} \prod_{i=0}^{n-2} \left(1 - \frac{\eta_i}{n-i-1}\right)$$
$$\geq \frac{1}{n} \cdot (1-C)^{2\lceil c/\alpha \rceil - 1} \cdot \prod_{i=0}^{n-2\lceil c/\alpha \rceil - 1} \left(1 - \frac{c}{\alpha} \cdot \frac{1}{n-i-1}\right).$$

Using the fact that $1 - x \geq \exp(-x - x^2)$ for all $0 \leq x \leq \frac{1}{2}$ (which can be proved straightforwardly by calculus), we get

$$\prod_{i=0}^{n-2\lceil c/\alpha \rceil - 1} \left(1 - \frac{c}{\alpha} \cdot \frac{1}{n-i-1}\right) = \prod_{j=2\lceil c/\alpha \rceil}^{n-1} \left(1 - \frac{c}{\alpha} \cdot \frac{1}{j}\right)$$
$$\geq \exp\left(-\frac{c}{\alpha} \sum_{j=2\lceil c/\alpha \rceil}^{n-1} \frac{1}{j} - \frac{c^2}{\alpha^2} \sum_{j=2\lceil c/\alpha \rceil}^{n-1} \frac{1}{j^2}\right).$$

Now since

$$\sum_{j=2\lceil c/\alpha \rceil}^{n-1} \frac{1}{j} \leq \sum_{j=2}^{n} \frac{1}{j} \leq \int_1^n \frac{dx}{x} = \log n$$

and

$$\sum_{j=2\lceil c/\alpha \rceil}^{n-1} \frac{1}{j^2} \leq \sum_{j=2}^{\infty} \frac{1}{j(j-1)} = 1,$$

we deduce that

$$1 - \lambda_2(P) \geq (1-C)^{2\lceil c/\alpha \rceil - 1} \cdot e^{-(c/\alpha)^2} \cdot n^{-(1+c/\alpha)}.$$

The theorem then follows from (2.1).

Proof of Theorem 1.3. We leverage Theorems 1.5 and 7.5, which show $O(n^{2+\frac{c}{\alpha}})$ mixing as long as there is an (α, c)-potential, or $O(n^{2+\frac{2c}{\alpha}})$ mixing if there is a general (α, c)-potential. We use the potential given by (3.3), which is an adaptation of the potential function in [21] to the log marginal ratios. When (β, γ, λ) is up-to-Δ unique with gap $\delta \in (0, 1)$, it is an (α, c)-potential or a general (α, c)-potential by Lemmas 3.6 and 7.6, with $\alpha \geq \delta/2$ and c a universal constant specified by the range of parameters. The theorem then follows. ∎

Proof of Theorem 1.1. By Claim A.3 later in Appendix A.1, $\lambda \leq (1-\delta)\lambda_c(\Delta)$ implies up-to-Δ uniqueness with gap $\geq \delta/4$. Since $\gamma \leq 1$, we can again appeal to Lemma 3.6 to obtain an (α, c)-potential with $\alpha \geq \delta/8$ and $c \leq 4$. Theorem 1.1 then follows by Theorem 1.5 with $O(n^{2+32/\delta})$ mixing. ∎

Proof of Theorem 1.2. By Claim A.4 later in Appendix A.1, $\beta \geq \beta_c(\Delta) + \delta(1 - \beta_c(\Delta))$ implies up-to-Δ uniqueness with gap δ. Again, appealing to Lemma 3.6, we obtain an (α, c)-potential with $\alpha \geq \delta/2$ and $c \leq 1.5$. Theorem 1.2 then follows by Theorem 1.5 with $O(n^{2+3/\delta})$ mixing.

Though we technically get $O(n^{2+3/\delta})$ by using the potential function in [21], we can improve it to $O(n^{2+1.5/\delta})$ mixing by using the trivial identity function as the potential. See the first case of Theorem 8.1 (proved in Appendix F.1) and Remark 8.2. ∎

A.1. Uniqueness gaps in terms of parameter gaps. In this subsection we state and prove Claims A.3 and A.4, which relate the parameter gaps with the uniqueness gaps.

CLAIM A.3 (hardcore model; Lemma C.1 from [3]). *Fix an integer $\Delta \geq 3$, $0 < \delta < 1$, and $\beta = 0, \gamma > 0$. If $\lambda \leq (1-\delta)\lambda_c(\gamma, \Delta)$, then (β, γ, λ) is up-to-Δ unique with gap $\delta/4$.*

CLAIM A.4 (large $\sqrt{\beta\gamma}$). *Fix an integer $\Delta \geq 3$ and $0 < \delta < 1$. If $\sqrt{\beta\gamma} \geq \frac{\Delta-2}{\Delta} + \delta\left(1 - \frac{\Delta-2}{\Delta}\right) = \frac{\Delta-2(1-\delta)}{\Delta}$, then (β, γ, λ) is up-to-Δ unique with gap $0 < \delta < 1$ for all λ. Note that if $\beta = \gamma$, this is precisely the condition $\beta \geq \beta_c(\Delta) + \delta(1 - \beta_c(\Delta))$.*

Proof. Consider the univariate recursion for the marginal ratios with $d < \Delta$ children $f_d(R) = \lambda\left(\frac{\beta R+1}{R+\gamma}\right)^d$. Differentiating, we have

$$\begin{aligned} {f_d}'(R) &= d\lambda\left(\frac{\beta R+1}{R+\gamma}\right)^{d-1} \cdot \left(\frac{\beta}{R+\gamma} - \frac{\beta R+1}{(R+\gamma)^2}\right) \\ &= -d(1-\beta\gamma)\lambda\left(\frac{\beta R+1}{R+\gamma}\right)^{d} \cdot \frac{1}{(\beta R+1)(R+\gamma)} \\ &= -d(1-\beta\gamma) \cdot \frac{f_d(R)}{(\beta R+1)(R+\gamma)}. \end{aligned}$$

At the unique fixed point R_d^*, we have $f_d(R_d^*) = R_d^*$, so

$$|{f_d}'(R_d^*)| = d(1-\beta\gamma)\frac{R_d^*}{(\beta R_d^*+1)(R_d^*+\gamma)}.$$

By Lemma E.3, we have the upper bound

$$|{f_d}'(R_d^*)| \leq d \cdot \frac{1-\beta\gamma}{(1+\sqrt{\beta\gamma})^2} = d \cdot \frac{1-\sqrt{\beta\gamma}}{1+\sqrt{\beta\gamma}}.$$

Since we assumed $\sqrt{\beta\gamma} \geq \frac{\Delta-2(1-\delta)}{\Delta}$, we obtain

$$d \cdot \frac{1-\sqrt{\beta\gamma}}{1+\sqrt{\beta\gamma}} \leq d \cdot \frac{\Delta-(\Delta-2(1-\delta))}{\Delta+(\Delta-2(1-\delta))} = d \cdot \frac{1-\delta}{\Delta-1+\delta} \leq (1-\delta)\frac{d}{\Delta-1}.$$

As this is at most $1-\delta$ for all $d < \Delta$, we have up-to-Δ uniqueness with gap δ. ∎

A.2. Spectral independence bounds for constant-size graphs In this subsection, we prove spectral independence bounds for graphs with fewer than $O(c/\alpha)$-many vertices, since for graphs with such few vertices, our bounds based on contraction of the tree recursions become trivial.

Proof of Claim A.1. If R_v denotes the marginal ratio of a vertex $v \in G$, then $R_v \geq \lambda\beta^{\Delta}$. In the case $\gamma \leq 1$, we have $R_v \leq \lambda/\gamma^{\Delta}$; however, if $\gamma > 1$, we have

$R_v \le \lambda$ where the equality holds for v isolated. It follows that we immediately have the bounds

$$|\mathcal{I}_G(u \to v)| \le \begin{cases} \left|\frac{\lambda}{\lambda+\gamma^\Delta} - \frac{\lambda\beta^\Delta}{1+\lambda\beta^\Delta}\right| = \frac{\lambda(1-\beta^\Delta\gamma^\Delta)}{(\lambda+\gamma^\Delta)(1+\lambda\beta^\Delta)} & \text{if } \gamma \le 1 \\ \left|\frac{\lambda}{1+\lambda} - \frac{\lambda\beta^\Delta}{1+\lambda\beta^\Delta}\right| = \frac{\lambda(1-\beta^\Delta)}{(\lambda+1)(1+\lambda\beta^\Delta)} & \text{o.w.} \end{cases}$$

for all $u, v \in G$. Note that these constants are less than 1, and only depend on $\beta, \gamma, \lambda, \Delta$, yielding the first claim.

Now, we proceed to remove the dependence on Δ when up-to-Δ uniqueness holds. We have the following cases:

1. If $\gamma > 1$, we immediately obtain a bound of $\frac{\lambda}{1+\lambda}$ which is independent of Δ.
2. If $\beta = 0$ and $\gamma \le 1$, then $\frac{\lambda(1-\beta^\Delta\gamma^\Delta)}{(\lambda+\gamma^\Delta)(1+\lambda\beta^\Delta)} = \frac{\lambda}{\lambda+\gamma^\Delta} \le \frac{\lambda}{\gamma^\Delta}$. Since (β, γ, λ) is up-to-Δ unique, we must have $\lambda \le \lambda_c(\gamma, \Delta) = \min_{1<d<\Delta} \frac{\gamma^{d+1} d^d}{(d-1)^{d+1}} \le \frac{\gamma^\Delta(\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta} \le \gamma^\Delta \cdot O(1/\Delta)$. It follows that $\frac{\lambda}{\gamma^\Delta} \le O(1/\Delta)$.
3. If $\sqrt{\beta\gamma} > \frac{\Delta-2}{\Delta}$ and $\gamma \le 1$, then
$$\frac{\lambda(1-\beta^\Delta\gamma^\Delta)}{(\lambda+\gamma^\Delta)(1+\lambda\beta^\Delta)} \le 1 - \beta^\Delta\gamma^\Delta \approx 1 - e^{-2}.$$
4. If $\sqrt{\beta\gamma} \le \frac{\Delta-2}{\Delta}$, then let Δ_0 be the maximal $1 < d < \Delta$ such that $\sqrt{\beta\gamma} > \frac{d-2}{d}$. If $\lambda \le \lambda_c(\beta, \gamma, \Delta)$, then by Lemma E.1, we have
$$\frac{\lambda(1-\beta^\Delta\gamma^\Delta)}{(\lambda+\gamma^\Delta)(1+\lambda\beta^\Delta)} \le \frac{\lambda}{\gamma^\Delta} \le O(\Delta_0/\Delta).$$
If $\lambda \ge \overline{\lambda}_c(\beta, \gamma, \Delta)$, then again by Lemma E.1, we have
$$\frac{\lambda(1-\beta^\Delta\gamma^\Delta)}{(\lambda+\gamma^\Delta)(1+\lambda\beta^\Delta)} \le \frac{1}{\lambda\beta^\Delta} \le O(\Delta_0/\Delta).$$ □

Proof of Claim A.2. The proof is identical to the antiferromagnetic case and is omitted here. □

Appendix B. Proof of Lemma 4.3 (parts 1 and 2).

Proof of Lemma 4.3 *(parts* 1 *and* 2*).* To see the first equality, we compute directly and get

$$\begin{aligned}\left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) \log Z_G^{\sigma_\Lambda} &= \frac{1}{Z_G^{\sigma_\Lambda}} \cdot \left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) Z_G^{\sigma_\Lambda} \\ &= \frac{1}{Z_G^{\sigma_\Lambda}} \sum_{\sigma \in \{0,1\}^{V \setminus \Lambda}} \left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) \left(\beta^{m_1(\sigma)} \gamma^{m_0(\sigma)} \prod_{w \in V} \lambda_w^{\sigma_w}\right) \\ &= \frac{1}{Z_G^{\sigma_\Lambda}} \sum_{\sigma \in \{0,1\}^{V \setminus \Lambda}} \sigma_v \left(\beta^{m_1(\sigma)} \gamma^{m_0(\sigma)} \prod_{w \in V} \lambda_w^{\sigma_w}\right) \\ &= \sum_{\sigma \in \{0,1\}^{V \setminus \Lambda}} \sigma_v \cdot \mu_G(\sigma \mid \sigma_\Lambda) = M_G^{\sigma_\Lambda}(v).\end{aligned}$$

For part 2, using the result above, we can also get

$$
\begin{aligned}
&\left(\lambda_v \frac{\partial}{\partial \lambda_v}\right)\left(\lambda_u \frac{\partial}{\partial \lambda_u}\right) \log Z_G^{\sigma_\Lambda} \\
&= \left(\lambda_v \frac{\partial}{\partial \lambda_v}\right)\left(\frac{1}{Z_G^{\sigma_\Lambda}} \cdot \left(\lambda_u \frac{\partial}{\partial \lambda_u}\right) Z_G^{\sigma_\Lambda}\right) \\
&= \frac{1}{Z_G^{\sigma_\Lambda}} \cdot \left(\lambda_v \frac{\partial}{\partial \lambda_v}\right)\left(\lambda_u \frac{\partial}{\partial \lambda_u}\right) Z_G^{\sigma_\Lambda} - \frac{1}{(Z_G^{\sigma_\Lambda})^2} \cdot \left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) Z_G^{\sigma_\Lambda} \cdot \left(\lambda_u \frac{\partial}{\partial \lambda_u}\right) Z_G^{\sigma_\Lambda} \\
&= \frac{1}{Z_G^{\sigma_\Lambda}} \cdot \left(\lambda_v \frac{\partial}{\partial \lambda_v}\right)\left(\sum_{\sigma \in \{0,1\}^{V \setminus \Lambda}} \sigma_u \left(\beta^{m_1(\sigma)} \gamma^{m_0(\sigma)} \prod_{w \in V} \lambda_w^{\sigma_w}\right)\right) - M_G^{\sigma_\Lambda}(u) \cdot M_G^{\sigma_\Lambda}(v) \\
&= \frac{1}{Z_G^{\sigma_\Lambda}} \sum_{\sigma \in \{0,1\}^{V \setminus \Lambda}} \sigma_u \cdot \left(\lambda_v \frac{\partial}{\partial \lambda_v}\right)\left(\beta^{m_1(\sigma)} \gamma^{m_0(\sigma)} \prod_{w \in V} \lambda_w^{\sigma_w}\right) - M_G^{\sigma_\Lambda}(u) \cdot M_G^{\sigma_\Lambda}(v) \\
&= \frac{1}{Z_G^{\sigma_\Lambda}} \sum_{\sigma \in \{0,1\}^{V \setminus \Lambda}} \sigma_u \cdot \sigma_v \left(\beta^{m_1(\sigma)} \gamma^{m_0(\sigma)} \prod_{w \in V} \lambda_w^{\sigma_w}\right) - M_G^{\sigma_\Lambda}(u) \cdot M_G^{\sigma_\Lambda}(v) \\
&= \sum_{\sigma \in \{0,1\}^{V \setminus \Lambda}} \sigma_u \cdot \sigma_v \cdot \mu_G(\sigma \mid \sigma_\Lambda) - M_G^{\sigma_\Lambda}(u) \cdot M_G^{\sigma_\Lambda}(v) \\
&= K_G^{\sigma_\Lambda}(u, v). \qquad \square
\end{aligned}
$$

Appendix C. A technical lemma for Ψ. The following lemma implies that the potential Ψ given by (3.3) is well-defined.

LEMMA C.1. *For all $\beta, \gamma > 0$ such that $\beta\gamma < 1$, we have*

$$\int_{-\infty}^{+\infty} \sqrt{\frac{(1-\beta\gamma)e^y}{(\beta e^y + 1)(e^y + \gamma)}}\, \mathrm{d}y < +\infty.$$

Proof. For the $+\infty$ side we have

$$
\begin{aligned}
\int_0^{+\infty} \sqrt{\frac{(1-\beta\gamma)e^y}{(\beta e^y + 1)(e^y + \gamma)}}\, \mathrm{d}y &= \int_0^{+\infty} \sqrt{\frac{1-\beta\gamma}{\beta e^y + \gamma e^{-y} + \beta\gamma + 1}}\, \mathrm{d}y \\
&< \int_0^{+\infty} \frac{1}{\sqrt{\beta e^y}}\, \mathrm{d}y < +\infty.
\end{aligned}
$$

Similarly, for the $-\infty$ side we have

$$\int_{-\infty}^{0} \sqrt{\frac{(1-\beta\gamma)e^y}{(\beta e^y + 1)(e^y + \gamma)}}\, \mathrm{d}y < \int_{-\infty}^{0} \frac{1}{\sqrt{\gamma e^{-y}}}\, \mathrm{d}y < +\infty. \qquad \square$$

Appendix D. Mixing by the potential method: Proof of Theorem 7.5. In this section, we prove Theorem 7.5 in the same way as for Theorem 1.5, as outlined in section 3. The major difference here is that we consider a weighted sum of absolute influences $\sum_{v \in V \setminus \Lambda} \rho_v \cdot |\mathcal{I}_G^{\sigma_\Lambda}(r \to v)|$ where $\rho : V \to \mathbb{R}^+$ is a weight function. This is sufficient for us to bound the eigenvalue of the influence matrix, as indicated by Lemma 7.3. We will choose the weight of a vertex v to be $\rho_v = \Delta_v$, the degree of v. The following lemma provides us an upper bound on the weighted sum of absolute influences to distance k, given a general (α, c)-potential. In particular, it generalizes Lemma 3.5.

LEMMA D.1. *If there exists a general (α, c)-potential function Ψ with respect to Δ and (β, γ, λ) where $\alpha \in (0,1)$ and $c > 0$, then for every $\Lambda \subseteq V_T \setminus \{r\}$, $\sigma_\Lambda \in \{0,1\}^\Lambda$ and all integers $k \geq 1$,*

$$\sum_{v \in L_r(k)} \Delta_v \cdot |\mathcal{I}_T^{\sigma_\Lambda}(r \to v)| \leq 2c \cdot (1-\alpha)^{k-1} \cdot \Delta_r,$$

where $L_r(k)$ denote the set of all free vertices at distance k away from r.

To prove Lemma D.1, we first state the following generalization of Lemma 5.1 for any weight function ρ. The proof of Lemma D.2 is identical to Lemma 5.1 and is omitted here.

LEMMA D.2. *Let $\Psi : [-\infty, +\infty] \to (-\infty, +\infty)$ be a differentiable and increasing (potential) function with image $S = \Psi[-\infty, +\infty]$ and derivative $\psi = \Psi'$. Denote the degree of the root r by Δ_r. Then for every integer $k \geq 1$,*

$$\sum_{v \in L_r(k)} \rho_v \cdot |\mathcal{I}_T^{\sigma_\Lambda}(r \to v)| \leq \Delta_r A_\Psi B_\Psi^\rho \left(\max_{1 \leq d < \Delta} \sup_{\tilde{\boldsymbol{y}} \in S^d} \left\| \nabla H_d^\Psi(\tilde{\boldsymbol{y}}) \right\|_1 \right)^{k-1},$$

where

$$A_\Psi = \max_{u \in L_r(1)} \left\{ \frac{|h(\log R_u)|}{\psi(\log R_u)} \right\} \quad \text{and} \quad B_\Psi^\rho = \max_{v \in L_r(k)} \{\rho_v \cdot \psi(\log R_v)\}.$$

We then prove Lemma D.1 and Theorem 7.5.

Proof of Lemma D.1. Denote the degree of a vertex $v \in V_T \setminus \{r\}$ by Δ_v. Pick the weights of vertices to be $\rho_v = \Delta_v$ for all $v \in V_T$. Since Ψ is a general (α, c)-potential, the contraction condition implies that

$$\max_{1 \leq d < \Delta} \sup_{\tilde{\boldsymbol{y}} \in S^d} \left\| \nabla H_d^\Psi(\tilde{\boldsymbol{y}}) \right\|_1 \leq 1 - \alpha.$$

Since $\log R_v \in J_{\Delta_v - 1}$ by the definition of J_d (notice the degree of v in the subtree T_v is $\Delta_v - 1$), the General Boundedness condition implies that for all $u \in L_r(1)$ and $v \in L_r(k)$,

$$\frac{\psi(\log R_v)}{\psi(\log R_u)} \cdot |h(\log R_u)| \leq \frac{2c}{\Delta_u + \Delta_v}.$$

Therefore, we get

$$\Delta_r A_\Psi B_\Psi^\rho = \Delta_r \cdot \max_{u \in L_r(1)} \left\{ \frac{|h(\log R_u)|}{\psi(\log R_u)} \right\} \cdot \max_{v \in L_r(k)} \{\Delta_v \cdot \psi(\log R_v)\} \leq 2c \cdot \Delta_r.$$

The lemma then follows immediately from Lemma D.2. ∎

Proof of Theorem 7.5. The proof of Theorem 7.5 is almost identical to that of Theorem 1.5. We point out that the only difference here is that we consider the weighted sum of absolute influences of a given vertex. Since the SAW tree preserves degrees of vertices, we can still apply Lemma 3.3 for the sum of absolute influences weighted by the degrees. Then, combining Theorem 3.2 and Lemmas 3.3, 7.3, and D.1, we complete the proof of the theorem. ∎

Appendix E. Verifying a good potential: Boundedness. In this section, we show the Boundedness or General Boundedness condition for our potential function Ψ defined by (3.3) in different ranges of parameters. Combining these and Lemma 6.3, we complete the proofs of Lemmas 3.6 and 7.6.

In Appendix E.1 we give background on the uniqueness region of the parameters (β,γ,λ), based on the work of [21]. We then show Boundedness and General Boundedness in Appendix E.2. Proofs of technical lemmas are deferred to Appendix E.3.

E.1. Preliminaries of the uniqueness region. In this subsection we give a brief description of the uniqueness region of parameters (β,γ,λ). All the results here, and also their proofs, can be found in Lemma 21 of [21].

Let $\Delta \geq 3$ be an integer and β,γ,λ be reals. We assume that $0 \leq \beta \leq \gamma$, $\gamma > 0$, $\beta\gamma < 1$, and $\lambda > 0$. For $1 \leq d \leq \Delta$ 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^*. Recall that the parameters (β,γ,λ) are up-to-Δ unique with gap $\delta \in (0,1)$ if $|f'_d(R_d^*)| < 1-\delta$ for all $1 \leq d < \Delta$.

When $\beta = 0$ the spin system is called a *hard-constraint model*. In this case, there exists a critical threshold for the external field, defined as

$$\lambda_c = \lambda_c(\gamma,\Delta) = \min_{1<d<\Delta} \frac{\gamma^{d+1} d^d}{(d-1)^{d+1}},$$

such that the parameters $(0,\gamma,\lambda)$ are up-to-Δ unique if and only if $\lambda < \lambda_c$. In particular, when $\gamma \leq 1$ the critical field is given by

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

When $\beta > 0$ the spin system is called a *soft-constraint model*. If $\sqrt{\beta\gamma} > \frac{\Delta-2}{\Delta}$, then (β,γ,λ) is up-to-Δ unique for all $\lambda > 0$; in this case we define the uniqueness region for λ to be the interval $\mathcal{A} = (0,\infty)$. If $\sqrt{\beta\gamma} \leq \frac{\Delta-2}{\Delta}$, the uniqueness region is more complicated, which we now describe. Let

$$\overline{\Delta} = \frac{1+\sqrt{\beta\gamma}}{1-\sqrt{\beta\gamma}},$$

so that for every $1 \leq d < \overline{\Delta}$ we have $d \cdot \frac{1-\sqrt{\beta\gamma}}{1+\sqrt{\beta\gamma}} < 1$, and for every $d \geq \overline{\Delta}$ we have $d \cdot \frac{1-\sqrt{\beta\gamma}}{1+\sqrt{\beta\gamma}} \geq 1$. For every $\overline{\Delta} \leq d < \Delta$, we define $x_1(d) \leq x_2(d)$ to be the two positive roots of the quadratic equation

$$\frac{d(1-\beta\gamma)x}{(\beta x+1)(x+\gamma)} = 1.$$

More specifically, $x_1(d)$ and $x_2(d)$ are given by

$$x_1(d) = \frac{\theta(d) - \sqrt{\theta(d)^2 - 4\beta\gamma}}{2\beta} \qquad \text{and} \qquad x_2(d) = \frac{\theta(d) + \sqrt{\theta(d)^2 - 4\beta\gamma}}{2\beta},$$

where

$$\theta(d) = d(1-\beta\gamma) - (1+\beta\gamma).$$

Notice that $\theta(d) \geq 2\sqrt{\beta\gamma}$ for all $d \geq \overline{\Delta}$. For $i = 1, 2$ we let

$$\lambda_i(d) = x_i(d)\left(\frac{x_i(d)+\gamma}{\beta x_i(d)+1}\right)^d.$$

Then, the parameters (β, γ, λ) are up-to-Δ unique if and only if λ belongs to the following region:

$$\mathcal{A} = \bigcap_{\overline{\Delta} \leq d < \Delta} \Big[(0, \lambda_1(d)) \cup (\lambda_2(d), \infty)\Big]. \tag{E.1}$$

In particular, when $\gamma \leq 1$ there are two critical thresholds $0 < \lambda_c < \overline{\lambda}_c$ such that the parameters (β, γ, λ) are up-to-Δ unique if and only if $\lambda < \lambda_c$ or $\lambda > \overline{\lambda}_c$ (i.e., $\mathcal{A} = (0, \lambda_c) \cup (\overline{\lambda}_c, \infty)$), where

$$\lambda_c = \lambda_c(\beta, \gamma, \Delta) = \min_{\overline{\Delta} \leq d < \Delta} \lambda_1(d)$$
$$\text{and} \qquad \overline{\lambda}_c = \overline{\lambda}_c(\beta, \gamma, \Delta) = \max_{\overline{\Delta} \leq d < \Delta} \lambda_2(d) = \lambda_2(\Delta - 1).$$

The following bounds on the critical fields are helpful for our proofs later.

LEMMA E.1.

1. *If $\beta = 0$, then for every integer d such that $1 < d < \Delta$ we have*

$$\lambda_c \leq \frac{4\gamma^{d+1}}{d-1}.$$

2. *If $\beta > 0$ and $\sqrt{\beta\gamma} \leq \frac{\Delta-2}{\Delta}$, then for every integer d such that $\overline{\Delta} \leq d < \Delta$ we have*

$$\lambda_1(d) \leq \frac{18\gamma^{d+1}}{\theta(d)} \qquad \text{and} \qquad \lambda_2(d) \geq \frac{\theta(d)}{18\beta^{d+1}},$$

where $\theta(d) = d(1-\beta\gamma) - (1+\beta\gamma)$.

The proof of Lemma E.1 is postponed to Appendix E.3.

E.2. Proofs of boundedness. In this subsection we complete the proofs of Lemmas 3.6 and 7.6 by establishing Boundedness and General Boundedness in the corresponding range of parameters.

Let $\Delta \geq 3$ be an integer. Let β, γ, λ be reals such that $0 \leq \beta \leq \gamma$, $\gamma > 0$, $\beta\gamma < 1$, and $\lambda > 0$. Recall that the potential function Ψ is defined by

$$\Psi'(y) = \psi(y) = \sqrt{\frac{(1-\beta\gamma)e^y}{(\beta e^y + 1)(e^y + \gamma)}} = \sqrt{|h(y)|}, \qquad \Psi(0) = 0.$$

It is surprising to find that $\psi = \sqrt{|h|}$, as the potential Ψ is exactly the one from [21] as indicated by Lemma 6.1. This seems not to be a coincidence, and it provides some intuition into why the potential from [21] works. More importantly, the fact that $\psi = \sqrt{|h|}$ is helpful in our proof of Boundedness and General Boundedness. Recall that for $0 \leq d < \Delta$ and $\beta\gamma < 1$ we let $J_d = [\log(\lambda\beta^d), \log(\lambda/\gamma^d)]$ be the range of log marginal ratios of a vertex with d children. Then for every $0 \leq d_i < \Delta$ and $y_i \in J_{d_i}$ where $i = 1, 2$, we have

$$\frac{\psi(y_2)}{\psi(y_1)} \cdot |h(y_1)| = \sqrt{|h(y_1)| \cdot |h(y_2)|}. \tag{E.2}$$

The following lemma gives upper bounds on $\sqrt{|h(y_1)| \cdot |h(y_2)|}$, from this and (E.2) we deduce Boundedness and General Boundedness immediately. The brackets in the lemma indicate which lemma the bound is applied to.

LEMMA E.2. *Let $\Delta \geq 3$ be an integer. Let β, γ, λ be reals such that $0 \leq \beta \leq \gamma$, $\gamma > 0$, $\beta\gamma < 1$, and $\lambda > 0$. Assume that the parameters (β, γ, λ) are up-to-Δ unique with gap $\delta \in (0,1)$. Then for all integers d_1, d_2 such that $0 \leq d_1, d_2 < \Delta$, and all reals $y_i \in J_{d_i}$ where $i = 1, 2$, the following hold:*

H. Hard-constraint models: $\beta = 0$ and $\lambda < \lambda_c$.

H.1. (Lemma 3.6) If $\gamma \leq 1$, then

$$|h(y_1)| \leq \frac{4}{\Delta}.$$

H.2. (Lemma 7.6) If $\gamma > 1$, then

$$\sqrt{|h(y_1)| \cdot |h(y_2)|} \leq \frac{8}{d_1 + d_2 + 2}.$$

S. Soft-constraint models: $\beta > 0$ and $\lambda \in \mathcal{A}$.

S.1. (Lemma 3.6) If $\sqrt{\beta\gamma} > \frac{\Delta-2}{\Delta}$, then

$$|h(y_1)| \leq \frac{1.5}{\Delta}.$$

S.2. (Lemma 3.6) If $\sqrt{\beta\gamma} \leq \frac{\Delta-2}{\Delta}$ and $\gamma \leq 1$, then

$$|h(y_1)| \leq \frac{18}{\Delta}.$$

S.3. (Lemma 7.6) If $\sqrt{\beta\gamma} \leq \frac{\Delta-2}{\Delta}$ and $\gamma > 1$, then

$$\sqrt{|h(y_1)| \cdot |h(y_2)|} \leq \frac{36}{d_1 + d_2 + 2}.$$

The following lemma, proved in Appendix E.3, is helpful.

LEMMA E.3. *The function*

$$|h(y)| = \frac{|1 - \beta\gamma| e^y}{(\beta e^y + 1)(e^y + \gamma)}$$

is increasing on $[-\infty, \log\sqrt{\gamma/\beta}]$ and decreasing on $[\log\sqrt{\gamma/\beta}, +\infty]$. In particular, $|h(y)|$ is maximized at $y^* = \log\sqrt{\gamma/\beta}$, and we have the following inequality for all $y \in [-\infty, +\infty]$:

$$|h(y)| \leq |h(y^*)| = \frac{|1 - \sqrt{\beta\gamma}|}{1 + \sqrt{\beta\gamma}}.$$

We present here the proof of Lemma E.2.

Proof of Lemma E.2. We use notation and results from Appendix E.1.

H. Hard-constraint models: $\beta = 0$ and $\lambda < \lambda_c$.

H.1. $\gamma \leq 1$.

For every $y_1 \in J_{d_1}$ we deduce from Lemma E.1 that

$$e^{y_1} \leq \frac{\lambda}{\gamma^{d_1}} \leq \frac{\lambda_c}{\gamma^{\Delta-1}} \leq \frac{4\gamma}{\Delta - 2}.$$

Hence,

$$|h(y_1)| = \frac{e^{y_1}}{e^{y_1}+\gamma} \le \frac{\frac{4\gamma}{\Delta-2}}{\frac{4\gamma}{\Delta-2}+\gamma} = \frac{4}{\Delta+2} \le \frac{4}{\Delta}.$$

H.2. $\gamma > 1$.

Let $\bar{y} = \frac{y_1+y_2}{2}$ and $\bar{d} = \frac{d_1+d_2}{2}$. Then we get

$$\sqrt{|h(y_1)|\cdot|h(y_2)|} = \sqrt{\frac{e^{y_1}}{e^{y_1}+\gamma}}\cdot\sqrt{\frac{e^{y_2}}{e^{y_2}+\gamma}} = \frac{1}{\sqrt{(1+\gamma e^{-y_1})(1+\gamma e^{-y_2})}} \le \frac{1}{1+\gamma e^{-\bar{y}}},$$

where the last inequality follows from the AM-GM inequality by

$$(1+\gamma e^{-y_1})(1+\gamma e^{-y_2}) = 1+\gamma(e^{-y_1}+e^{-y_2})+\gamma^2 e^{-2\bar{y}} \ge 1+2\gamma e^{-\bar{y}}+\gamma^2 e^{-2\bar{y}} = (1+\gamma e^{-\bar{y}})^2.$$

Since $y_i \in J_{d_i}$ for $i=1,2$, we have

$$e^{\bar{y}} = \sqrt{e^{y_1}\cdot e^{y_2}} \le \sqrt{\frac{\lambda}{\gamma^{d_1}}\cdot\frac{\lambda}{\gamma^{d_2}}} = \frac{\lambda}{\gamma^{\bar{d}}}.$$

If $\bar{d} \ge 2$, then we deduce from Lemma E.1 and $\gamma > 1$ that

$$e^{\bar{y}} \le \frac{\lambda_c}{\gamma^{\lfloor \bar{d}\rfloor}} \le \frac{4\gamma}{\lfloor \bar{d}\rfloor - 1}.$$

It follows that

$$\sqrt{|h(y_1)|\cdot|h(y_2)|} \le \frac{1}{1+\gamma e^{-\bar{y}}} \le \frac{1}{1+\frac{\lfloor \bar{d}\rfloor-1}{4}} = \frac{4}{\lfloor \bar{d}\rfloor+3} \le \frac{8}{d_1+d_2+2}.$$

If $\bar{d} < 2$, then it is easy to see that

$$\sqrt{|h(y_1)|\cdot|h(y_2)|} \le 1 \le \frac{8}{d_1+d_2+2}.$$

S. Soft-constraint models: $\beta > 0$ and $\lambda \in \mathcal{A}$.

S.1. $\sqrt{\beta\gamma} > \frac{\Delta-2}{\Delta}$.

For every $y_1 \in J$ we deduce from Lemma E.3 that

$$|h(y_1)| \le \frac{1-\sqrt{\beta\gamma}}{1+\sqrt{\beta\gamma}} \le \frac{1}{\Delta-1} \le \frac{1.5}{\Delta}.$$

S.2. $\sqrt{\beta\gamma} \le \frac{\Delta-2}{\Delta}$ *and* $\gamma \le 1$.

In this case, we have either $\lambda < \lambda_c$ or $\lambda > \overline{\lambda}_c$ where $\lambda_c, \overline{\lambda}_c$ are the two critical fields. Consider first $\lambda > \overline{\lambda}_c$. For every $y_1 \in J_{d_1}$ we deduce from Lemma E.1 and $\beta < 1$ that

$$e^{y_1} \ge \lambda\beta^{d_1} \ge \overline{\lambda}_c\beta^{\Delta-1} \ge \frac{\theta(\Delta-1)}{18\beta},$$

where $\theta(d) = d(1-\beta\gamma) - (1+\beta\gamma)$. Hence,

$$|h(y_1)| = \frac{(1-\beta\gamma)e^{y_1}}{(\beta e^{y_1}+1)(e^{y_1}+\gamma)} = \frac{1-\beta\gamma}{\beta e^{y_1}+\gamma e^{-y_1}+(1+\beta\gamma)}$$
$$\leq \frac{1-\beta\gamma}{\frac{\theta(\Delta-1)}{18}+(1+\beta\gamma)} = \frac{18(1-\beta\gamma)}{(\Delta-1)(1-\beta\gamma)+17(1+\beta\gamma)} \leq \frac{18}{\Delta}.$$

Next we consider $\lambda < \lambda_c$. For every $y_1 \in J_{d_1}$ we deduce from Lemma E.1 and $\gamma \leq 1$ that

$$e^{y_1} \leq \frac{\lambda}{\gamma^{d_1}} \leq \frac{\lambda_c}{\gamma^{\Delta-1}} \leq \frac{18\gamma}{\theta(\Delta-1)}.$$

Hence,

$$|h(y_1)| = \frac{1-\beta\gamma}{\beta e^{y_1}+\gamma e^{-y_1}+(1+\beta\gamma)} \leq \frac{1-\beta\gamma}{\frac{\theta(\Delta-1)}{18}+(1+\beta\gamma)} \leq \frac{18}{\Delta}.$$

S.3. $\sqrt{\beta\gamma} \leq \frac{\Delta-2}{\Delta}$ *and* $\gamma > 1$.

Let $\bar{y} = \frac{y_1+y_2}{2}$, $\bar{d} = \frac{d_1+d_2}{2}$, $d_L = \lfloor \bar{d} \rfloor$, and $d_R = \lceil \bar{d} \rceil$. We first consider some trivial cases. If $\bar{d} \leq 2$, then it is easy to see that

$$\sqrt{|h(y_1)| \cdot |h(y_2)|} \leq 1 \leq \frac{6}{d_1+d_2+2}.$$

If $\bar{d} > 2$ and $d_L \leq \overline{\Delta}$, then we deduce from Lemma E.3 that

$$\sqrt{|h(y_1)| \cdot |h(y_2)|} \leq \frac{1-\sqrt{\beta\gamma}}{1+\sqrt{\beta\gamma}} = \frac{1}{\overline{\Delta}} \leq \frac{2}{d_1+d_2-2} \leq \frac{6}{d_1+d_2+2}.$$

Hence, in the following we may assume that $\bar{d} > 2$ and $d_L > \overline{\Delta}$.

Since the parameters (β, γ, λ) are up-to-Δ unique, we have $\lambda \in \mathcal{A}$ where the regime $\mathcal{A}$ is given by (E.1). Observe that

$$\mathcal{A} \subseteq (0, \lambda_1(d_L)) \cup (\lambda_2(d_R), \infty) \cup (\lambda_2(d_L), \lambda_1(d_R)),$$

where the last interval is nonempty only when $\lambda_2(d_L) < \lambda_1(d_R)$. This means that λ is contained in at least one of the three intervals. We establish the bound by considering these three cases separately.

Case 1*:* $\lambda < \lambda_1(d_L)$. By the Cauchy–Schwarz inequality, we have

$$\sqrt{|h(y_1)| \cdot |h(y_2)|} = \sqrt{\frac{1-\beta\gamma}{\beta e^{y_1}+\gamma e^{-y_1}+(1+\beta\gamma)}} \cdot \sqrt{\frac{1-\beta\gamma}{\beta e^{y_2}+\gamma e^{-y_2}+(1+\beta\gamma)}}$$
$$\leq \frac{1-\beta\gamma}{\sqrt{(\beta e^{y_1}+\gamma e^{-y_1})(\beta e^{y_2}+\gamma e^{-y_2})}+(1+\beta\gamma)}. \tag{E.3}$$

Therefore, we get

$$\sqrt{|h(y_1)| \cdot |h(y_2)|} \leq \frac{1-\beta\gamma}{\gamma e^{-\bar{y}}+(1+\beta\gamma)}.$$

Since $y_i \in J_{d_i}$ for $i = 1, 2$ and $\gamma > 1$, we deduce from Lemma E.1 that

$$e^{\bar{y}} \leq \frac{\lambda}{\gamma^{\bar{d}}} \leq \frac{\lambda_1(d_L)}{\gamma^{d_L}} \leq \frac{18\gamma}{\theta(d_L)},$$

where $\theta(d_L) = d_L(1 - \beta\gamma) - (1 + \beta\gamma)$. It follows that

$$\sqrt{|h(y_1)| \cdot |h(y_2)|} \leq \frac{1 - \beta\gamma}{\gamma e^{-\bar{y}} + (1 + \beta\gamma)} \leq \frac{1 - \beta\gamma}{\frac{\theta(d_L)}{18} + (1 + \beta\gamma)} \leq \frac{36}{d_1 + d_2 + 2}.$$

Case 2*:* $\lambda > \lambda_2(d_R)$. Similarly, we obtain from (E.3) that

$$\sqrt{|h(y_1)| \cdot |h(y_2)|} \leq \frac{1 - \beta\gamma}{\beta e^{\bar{y}} + (1 + \beta\gamma)}.$$

Since $y_i \in J_{d_i}$ for $i = 1, 2$ and $\beta < 1$, we deduce from Lemma E.1 that

$$e^{\bar{y}} \geq \lambda \beta^{\bar{d}} \geq \lambda_2(d_R)\beta^{d_R} \geq \frac{\theta(d_R)}{18\beta},$$

where $\theta(d_R) = d_R(1 - \beta\gamma) - (1 + \beta\gamma)$. It follows that

$$\sqrt{|h(y_1)| \cdot |h(y_2)|} \leq \frac{1 - \beta\gamma}{\beta e^{\bar{y}} + (1 + \beta\gamma)} \leq \frac{1 - \beta\gamma}{\frac{\theta(d_R)}{18} + (1 + \beta\gamma)} \leq \frac{36}{d_1 + d_2 + 2}.$$

Case 3*:* $\lambda_2(d_L) < \lambda < \lambda_1(d_R)$. We may assume that $d_1 \geq d_2$. By (E.3), we obtain

$$\sqrt{|h(y_1)| \cdot |h(y_2)|} \leq \frac{1 - \beta\gamma}{\sqrt{\beta\gamma} e^{\frac{y_2 - y_1}{2}} + (1 + \beta\gamma)}.$$

Since $y_i \in J_{d_i}$ for $i = 1, 2$ and $\beta < 1 < \gamma$, we have

$$e^{y_2 - y_1} \geq \beta^{d_2}\gamma^{d_1} \geq \beta^{d_L}\gamma^{d_R}.$$

Meanwhile, we deduce from Lemma E.1 that

$$\frac{\theta(d_L)}{18\beta^{d_L+1}} \leq \lambda_2(d_L) < \lambda < \lambda_1(d_R) \leq \frac{18\gamma^{d_R+1}}{\theta(d_R)},$$

which implies

$$\sqrt{\beta\gamma} e^{\frac{y_2 - y_1}{2}} \geq \sqrt{\beta^{d_L+1}\gamma^{d_R+1}} \geq \frac{\sqrt{\theta(d_L)\theta(d_R)}}{18} \geq \frac{\theta(d_L)}{18}.$$

It follows that

$$\sqrt{|h(y_1)| \cdot |h(y_2)|} \leq \frac{1 - \beta\gamma}{\sqrt{\beta\gamma} e^{\frac{y_2 - y_1}{2}} + (1 + \beta\gamma)} \leq \frac{1 - \beta\gamma}{\frac{\theta(d_L)}{18} + (1 + \beta\gamma)} \leq \frac{36}{d_1 + d_2 + 2}. \qquad \square$$

E.3. Proofs of technical lemmas.

Proof of Lemma E.1. 1. For every $1 < d < \Delta$ we have

$$\lambda_c \leq \frac{\gamma^{d+1} d^d}{(d-1)^{d+1}} = \frac{\gamma^{d+1}}{d-1}\left(\frac{d}{d-1}\right)^d \leq \frac{4\gamma^{d+1}}{d-1},$$

where the last inequality follows from that $(\frac{d}{d-1})^d \leq 4$ for all integers $d > 1$.

2. For every $\overline{\Delta} \leq d < \Delta$ we have

$$x_1(d) = \frac{2\gamma}{\theta(d) + \sqrt{\theta(d)^2 - 4\beta\gamma}} \leq \frac{2\gamma}{\theta(d)}.$$

Observe that the function $\frac{x+\gamma}{\beta x+1}$ is monotone increasing in x when $\beta\gamma < 1$, and thus we deduce that

$$\frac{x_1(d)+\gamma}{\beta x_1(d)+1} \leq \frac{\frac{2\gamma}{\theta(d)}+\gamma}{\frac{2\beta\gamma}{\theta(d)}+1} = \gamma \cdot \frac{2+d(1-\beta\gamma)-(1+\beta\gamma)}{2\beta\gamma+d(1-\beta\gamma)-(1+\beta\gamma)} = \gamma \cdot \frac{d+1}{d-1}.$$

Therefore,

$$\lambda_1(d) = x_1(d)\left(\frac{x_1(d)+\gamma}{\beta x_1(d)+1}\right)^d \leq \frac{2\gamma}{\theta(d)} \cdot \gamma^d \cdot \left(\frac{d+1}{d-1}\right)^d \leq \frac{18\gamma^{d+1}}{\theta(d)},$$

where the last inequality follows from the fact that $(\frac{d+1}{d-1})^d \leq 9$ for all integers $d > 1$.

The second part can be proved similarly. For every $\overline{\Delta} \leq d < \Delta$ we have

$$x_2(d) = \frac{\theta(d)+\sqrt{\theta(d)^2-4\beta\gamma}}{2\beta} \geq \frac{\theta(d)}{2\beta},$$

and hence,

$$\frac{x_2(d)+\gamma}{\beta x_2(d)+1} \geq \frac{\frac{\theta(d)}{2\beta}+\gamma}{\frac{\theta(d)}{2}+1} = \frac{1}{\beta} \cdot \frac{d(1-\beta\gamma)-(1+\beta\gamma)+2\beta\gamma}{d(1-\beta\gamma)-(1+\beta\gamma)+2} = \frac{1}{\beta} \cdot \frac{d-1}{d+1}.$$

We then conclude that

$$\lambda_2(d) = x_2(d)\left(\frac{x_2(d)+\gamma}{\beta x_2(d)+1}\right)^d \geq \frac{\theta(d)}{2\beta} \cdot \frac{1}{\beta^d} \cdot \left(\frac{d-1}{d+1}\right)^d \geq \frac{\theta(d)}{18\beta^{d+1}},$$

where the last inequality again follows from the fact that $(\frac{d+1}{d-1})^d \leq 9$ for all integers $d > 1$. ◻

Proof of Lemma E.3. For convenience, define $f : [0, +\infty] \to [0, +\infty]$ by $f(x) = \frac{|1-\beta\gamma|\cdot x}{(\beta x+1)(x+\gamma)}$; note that $f(e^y) = |h(e^y)|$. Since e^y is monotone increasing, it suffices to show that f is increasing on $[0, \sqrt{\gamma/\beta}]$ and decreasing on $[\sqrt{\gamma/\beta}, +\infty]$. To this end, we compute the derivative of f as

$$\begin{aligned} f'(x) &= |1-\beta\gamma| \cdot \left(\frac{1}{(\beta x+1)(x+\gamma)} - \frac{x(\beta(x+\gamma)+(\beta x+1))}{(\beta x+1)^2(x+\gamma)^2}\right) \\ &= \frac{|1-\beta\gamma|}{(\beta x+1)^2(x+\gamma)^2}\left((\beta x+1)(x+\gamma) - x(\beta(x+\gamma)+(\beta x+1))\right) \\ &= \frac{|1-\beta\gamma|}{(\beta x+1)^2(x+\gamma)^2} \cdot (\gamma - \beta x^2). \end{aligned}$$

Note that this is nonnegative on $[0, \sqrt{\gamma/\beta}]$ and nonpositive on $[\sqrt{\gamma/\beta}, +\infty]$, so we are done. ◻

Appendix F. Proofs for ferromagnetic cases.

F.1. Proof of Theorem 8.1. Throughout the proof, we use the trivial potential function $\Psi(y) = y$. Note that then, $\psi(y) = 1$ is a constant function.

Now, we prove Contraction and Boundedness. We split our proof into three cases.

1. *Case 1: $\frac{\Delta-2+\delta}{\Delta-\delta} \leq \sqrt{\beta\gamma} \leq \frac{\Delta-\delta}{\Delta-2+\delta}$, and $\lambda > 0$ is arbitrary.* We first prove the contraction part. By Lemma E.3, for all $y \in [-\infty, +\infty]$ we have

$$|h(y)| \leq \frac{|1-\sqrt{\beta\gamma}|}{1+\sqrt{\beta\gamma}} \leq \frac{1-\delta}{\Delta-1}.$$

Now let us prove the boundedness condition. From the above inequality we have

$$|h(y)| \le \frac{1}{\Delta-1} \le \frac{1.5}{\Delta}$$

for $\Delta \ge 3$.

2. *Case* 2*:* $\sqrt{\beta\gamma} \ge \frac{\Delta}{\Delta-2}$ *and* $0 < \lambda \le (1-\delta)\frac{\gamma}{\max\{1,\beta^{\Delta-1}\}\cdot((\Delta-2)\beta\gamma-\Delta)}$. For the contraction part, since $\log(\lambda\min\{1,1/\gamma^{\Delta-1}\}) \le y_i \le \log(\lambda\max\{1,\beta^{\Delta-1}\})$, we have

$$\left|\frac{\partial H_d(\boldsymbol{y})}{\partial y_i}\right| = |h(y_i)| = \frac{\beta\gamma-1}{1+\beta\gamma+\gamma e^{-y_i}+\beta e^{y_i}} \le \frac{\beta\gamma-1}{1+\beta\gamma+\gamma e^{-y_i}} \\ \le \frac{\beta\gamma-1}{1+\beta\gamma+\frac{\gamma}{\lambda\max\{1,\beta^{\Delta-1}\}}}.$$

Since we assumed $\lambda \le (1-\delta)\frac{\gamma}{\max\{1,\beta^{\Delta-1}\}\cdot((\Delta-2)\beta\gamma-\Delta)}$, it follows that we have the upper bound

$$\frac{\beta\gamma-1}{1+\beta\gamma+\frac{(\Delta-2)\beta\gamma-\Delta}{1-\delta}} = (1-\delta)\frac{\beta\gamma-1}{(\Delta-1-\delta)\beta\gamma-(\Delta-1+\delta)} \\ = (1-\delta)\frac{\beta\gamma-1}{(\Delta-1-\delta)(\beta\gamma-1)+2\delta} \\ \le \frac{1-\delta}{\Delta-1-\delta} \le (1-\Theta(\delta))\frac{1}{\Delta-1}.$$

Now, we prove the boundedness condition. Note that since

$$\lambda \le \frac{\gamma}{\max\{1,\beta^{\Delta-1}\}\cdot((\Delta-2)\beta\gamma-\Delta)},$$

it follows that $y \le \log(\lambda\max\{1,\beta^{\Delta-1}\}) \le \log\left(\frac{\gamma}{(\Delta-2)\beta\gamma-\Delta}\right)$. A simple calculation reveals that $\frac{\gamma}{(\Delta-2)\beta\gamma-\Delta} \le \sqrt{\frac{\gamma}{\beta}}$, and so by Lemma E.3 we have

$$|h(y)| \le \left|h\left(\log\left(\frac{\gamma}{(\Delta-2)\beta\gamma-\Delta}\right)\right)\right| \le \frac{(\beta\gamma-1)e^{\log\left(\frac{\gamma}{(\Delta-2)\beta\gamma-\Delta}\right)}}{e^{\log\left(\frac{\gamma}{(\Delta-2)\beta\gamma-\Delta}\right)}+\gamma} \\ = (\beta\gamma-1)\frac{1}{1+(\Delta-2)\beta\gamma-\Delta} = \frac{\beta\gamma-1}{(\Delta-2)(\beta\gamma-1)-1} \le O(1/\Delta).$$

3. *Case* 3*:* $\sqrt{\beta\gamma} \ge \frac{\Delta}{\Delta-2}$ *and* $\lambda \ge \frac{1}{1-\delta}\cdot\frac{(\Delta-2)\beta\gamma-\Delta}{\beta\cdot\min\{1,1/\gamma^{\Delta-1}\}}$. For the contraction part, since $\log(\lambda\min\{1,1/\gamma^{\Delta-1}\}) \le y_i \le \log(\lambda\max\{1,\beta^{\Delta-1}\})$, we have

$$\left|\frac{\partial H_d(\boldsymbol{y})}{\partial y_i}\right| = |h(y_i)| = \frac{\beta\gamma-1}{1+\beta\gamma+\gamma e^{-y_i}+\beta e^{y_i}} \le \frac{\beta\gamma-1}{1+\beta\gamma+\beta e^{y_i}} \\ \le \frac{\beta\gamma-1}{1+\beta\gamma+\beta\lambda\min\{1,1/\gamma^{\Delta-1}\}}.$$

Since we assumed $\lambda \ge \frac{1}{1-\delta}\cdot\frac{(\Delta-2)\beta\gamma-\Delta}{\beta\cdot\min\{1,1/\gamma^{\Delta-1}\}}$, it follows that we have the upper bound

$$\frac{\beta\gamma-1}{1+\beta\gamma+\frac{(\Delta-2)\beta\gamma-\Delta}{1-\delta}}$$

which is again upper bounded by $(1-\Theta(\delta))\frac{1}{\Delta-1}$ as we calculated in Case 2 above. Now, we prove the boundedness condition. Note that since

$$\lambda \ge \frac{(\Delta-2)\beta\gamma-\Delta}{\beta\min\{1,1/\gamma^{\Delta-1}\}},$$

it follows that $y \ge \log(\lambda\min\{1,1/\gamma^{\Delta-1}\}) \ge \log\left(\frac{(\Delta-2)\beta\gamma-\Delta}{\beta}\right)$. A simple calculation reveals that $\frac{(\Delta-2)\beta\gamma-\Delta}{\beta} \ge \sqrt{\frac{\gamma}{\beta}}$, and so by Lemma E.3 we have

$$\begin{aligned}|h(y)| &\le \left|h\left(\log\left(\frac{(\Delta-2)\beta\gamma-\Delta}{\beta}\right)\right)\right| \le (\beta\gamma-1)\frac{1}{\beta\cdot\frac{(\Delta-2)\beta\gamma-\Delta}{\beta}+1}\\ &= \frac{\beta\gamma-1}{(\Delta-2)(\beta\gamma-1)-1} \le O(1/\Delta).\end{aligned}$$

∎

F.2. Proof of Theorem 8.3 In this subsection, we use results from [15] to prove Theorem 8.3. Their potential function is implicitly defined by its derivative for the marginal ratios as

$$\Phi'(R) = \phi(R) = \min\left\{\frac{\beta\gamma-1}{\alpha\gamma\log\frac{\lambda+\gamma}{\beta\lambda+1}}, \frac{1}{R\log\frac{\lambda}{R}}\right\}$$

for a constant $0 \le \alpha \le 1$ depending only on β,γ,λ (see [15] for a precise definition). In our context, the corresponding potential for the log ratios is

$$\Psi'(y) = \psi(y) = e^y\phi(e^y) = \min\left\{\frac{\beta\gamma-1}{\alpha\gamma\log\frac{\lambda+\gamma}{\beta\lambda+1}}e^y, \frac{1}{\log\frac{\lambda}{e^y}}\right\}$$

and is bounded by constants depending on $\beta,\gamma,\lambda,\Delta$ for $\log(\lambda/\gamma^{\Delta-1}) \le y \le \log\lambda$.

One of the main technical results in [15] is that the tree recursion is contracting with the potential function Φ, and the derivative ϕ is bounded in the sense that there exist positive constants C_1, C_2 depending only on β,γ,λ such that $C_1 \le \phi(R) \le C_2$ for all $0 \le R \le \lambda$. Reference [15] refers to such a function as a *universal potential function*.

In our context, we get that Ψ is an (α,c)-potential function, which satisfies Definition 1.4, but with a constant c that depends on γ,Δ. Indeed, worst case, we have

$$\max_{y_1,y_2}\frac{\psi(y_2)}{\psi(y_1)} \ge \frac{\psi(\log\lambda)}{\psi(\log(\lambda/\gamma^{\Delta-1}))} = \frac{\lambda\frac{\beta\gamma-1}{\alpha\gamma\log\frac{\lambda+\gamma}{\beta\lambda+1}}}{\frac{\beta\gamma-1}{\alpha\log\frac{\lambda+\gamma}{\beta\lambda+1}}\cdot\frac{\lambda}{\gamma^{\Delta}}} = \gamma^{\Delta-1}.$$

More precisely, we have the following result from [15], stated in terms of the log marginal ratios.

THEOREM F.1. *Assume β,γ,λ are nonnegative real numbers satisfying $\beta \le 1 \le \gamma$, $\sqrt{\beta\gamma} \ge 1$, and $\lambda < \left(\frac{\gamma}{\beta}\right)^{\frac{\sqrt{\beta\gamma}}{\sqrt{\beta\gamma}-1}}$. Then the function Ψ is an (α,c)-potential function for a constant $0<\alpha<1$ depending on β,γ,λ and a constant $c>0$ depending on $\beta,\gamma,\lambda,\Delta$.*

Combined with Theorem 1.5, this gives $O(n^C)$ mixing with a constant C depending only on $\beta,\gamma,\lambda,\Delta$. We note that this is weaker than the correlation decay result in [15], since there, C does not depend on Δ, and hence is efficient for arbitrary graphs.

Appendix G. Slightly faster mixing. In this appendix, we slightly optimize our mixing time results for certain antiferromagnetic 2-spin systems by more carefully taking into account the trade-off between the (nontrivial) spectral independence bound we prove based on contraction, and the (trivial) spectral independence bound we obtained in Appendix A.2 for handling constant-sized graphs.

PROPOSITION G.1. *Suppose a distribution μ on subsets of $[n]$ is $(\eta_0, \dots, \eta_{n-2})$-spectrally independent for $\eta_i \le \min\{a, (n-i-1)b\}$, for some $a \ge 0$ and $0 \le b \le 1$. Then the Glauber dynamics for sampling from μ has spectral gap at least $\frac{1}{n} \cdot \Omega\left(\frac{a}{bn}\right)^a$.*

Proof. The bound $\eta_i \le (n-i-1)b$ is better when $i \ge n-1-\lfloor a/b \rfloor$, while the bound $\eta_i \le a$ is better when $i \le n-1-\lfloor a/b \rfloor$. It follows that the final spectral gap lower bound is

$$\frac{1}{n} \cdot (1-b)^{\lfloor a/b \rfloor} \cdot \prod_{k=0}^{n-1-\lfloor a/b \rfloor} \left(1 - \frac{a}{n-k-1}\right)$$

Note that $(1-b)^{\lfloor a/b \rfloor} \gtrsim e^{-a}$, while

$$\begin{aligned}
\prod_{k=0}^{n-1-\lfloor a/b \rfloor} \left(1 - \frac{a}{n-k-1}\right) &\gtrsim \exp\left(-a \sum_{k=0}^{n-1-\lfloor a/b \rfloor} \frac{1}{n-k-1}\right) \\
&\gtrsim \exp\left(-a \left(\underbrace{\sum_{k=0}^{n-2} \frac{1}{n-k-1}}_{\approx \log n} - \underbrace{\sum_{k=n-\lfloor a/b \rfloor}^{n-2} \frac{1}{n-k-1}}_{\approx \log \lfloor a/b \rfloor}\right)\right) \\
&\gtrsim \exp\left(-a \log \frac{bn}{a}\right) \\
&\gtrsim \left(\frac{a}{bn}\right)^a.
\end{aligned}$$

Putting these together, we obtain the desired lower bound. ∎

With this result, we can apply this result to the antiferromagnetic models with $\sqrt{\beta\gamma} \le \frac{\Delta-2}{\Delta}, \gamma \le 1$, and $\beta = 0, \gamma \le 1$, since by the proof of Claim A.1, we have such that systems are Cn-spectrally independent roughly with $C \le O(1/\Delta)$.

COROLLARY G.2 (soft constraints). *Fix integers $\Delta \ge 3$ and $1 < \overline{\Delta} < \Delta$. Let $\beta, \gamma, \lambda \ge 0$ be nonnegative real numbers satisfying $\frac{\overline{\Delta}-2}{\overline{\Delta}} \le \sqrt{\beta\gamma} \le \frac{\overline{\Delta}-1}{\overline{\Delta}+1}$ and $\gamma \le 1$. Assume further that (β, γ, λ) is up-to-Δ unique with gap $0 < \delta < 1$. Then for every n-vertex graph G with maximum degree at most Δ, the Glauber dynamics for sampling from the antiferromagnetic 2-spin system with parameters (β, γ, λ) mixes in $O\left(\frac{\overline{\Delta} \cdot n}{\Delta}\right)^{O(1/\delta)}$ steps.*

COROLLARY G.3 (hard constraints). *Fix an integer $\Delta \ge 3$, fix $\beta = 0$, and let $0 \le \gamma \le 1, \lambda \ge 0$ be up-to-Δ unique with gap $0 < \delta < 1$. Then for every n-vertex graph G with maximum degree at most Δ, the Glauber dynamics for sampling from the antiferromagnetic 2-spin system with parameters (β, γ, λ)mixes in $O\left(\frac{n}{\Delta}\right)^{O(1/\delta)}$ steps.*

Acknowledgments. We would like to thank Shayan Oveis Gharan and Nima Anari for stimulating discussions. We also thank the anonymous referees for helpful comments and suggestions. We are grateful to Yitong Yin for communicating with us about the updated preprint [22] and for providing helpful instructions on modifying statements and proofs of results in Appendix E, particularly Lemma E.2.

REFERENCES

[1] V. L. Alev and L. C. Lau, *Improved analysis of higher order random walks and applications*, in Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC), 2020, pp. 1198–1211.

[2] N. Anari, V. Jain, F. Koehler, H. T. Pham, and T.-D. 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, pp. 1418–1430.

[3] N. Anari, K. Liu, and S. Oveis Gharan, *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), 2020, pp. 1319–1330.

[4] A. Barvinok, *Combinatorics and Complexity of Partition Functions*, vol. 30, Springer Algorithms and Combinatorics, 2016.

[5] F. Bencs, *On trees with real-rooted independence polynomial*, Discrete Math., 341 (2018), pp. 3321–3330.

[6] X. Chen, W. Feng, Y. Yin, and X. Zhang, *Optimal Mixing Time for the Ising Model in the Uniqueness Regime*, arXiv preprint arXiv:2111.03034, 2021.

[7] X. Chen, W. Feng, Y. Yin, and X. Zhang, *Rapid mixing of Glauber dynamics via spectral independence for all degrees*, in Proceedings of the 62nd Annual IEEE Symposium on Foundations of Computer Science (FOCS), 2021, pp. 137–148.

[8] X. Chen, W. Feng, Y. Yin, and X. Zhang, *Optimal mixing for two-state anti-ferromagnetic spin systems*, in Proceedings of the 63rd Annual IEEE Symposium on Foundations of Computer Science (FOCS), 2022

[9] Y. Chen and R. Eldan, *Localization schemes: A framework for proving mixing bounds for Markov chains*, in Proceedings of the 63rd Annual IEEE Symposium on Foundations of Computer Science (FOCS), 2022

[10] Z. Chen, K. Liu, and E. Vigoda, *Rapid mixing of Glauber dynamics up to uniqueness via contraction*, in Proceedings of the 2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS), 2020, pp. 1307–1318, https://doi.org/10.1109/FOCS46700.2020.00124.

[11] Z. Chen, K. Liu, and E. 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, pp. 1537–1550.

[12] A. Galanis, D. Štefankovič, and E. Vigoda, *Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models*, Comb. Probab. Comput., 25 (2016), pp. 500–559.

[13] C. D. Godsil, *Matchings and walks in graphs*, J. Graph Theor, 5 (1981), pp. 285–297.

[14] H. Guo, J. Liu, and P. Lu, *Zeros of ferromagnetic 2-spin systems*, in Proceedings of the 31st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2020, pp. 181–192.

[15] H. Guo and P. Lu, *Uniqueness, spatial mixing, and approximation for ferromagnetic 2-spin systems*, ACM Trans. Comput. Theory, 10 (2018).

[16] M. Jerrum and A. Sinclair, *Polynomial-time approximation algorithms for the Ising model*, SIAM J. Comput., 22 (1993), pp. 1087–1116.

[17] M. R. Jerrum, L. G. Valiant, and V. V. Vazirani, *Random generation of combinatorial structures from a uniform distribution*, Theor. Comput. Sci., 43 (1986), pp. 169–188.

[18] F. P. Kelly, *Stochastic models of computer communication systems*, J. R. Stat. Soc., B: Stat. Methodol., 47 (1985), pp. 379–395.

[19] D. A. Levin and Y. Peres, *Markov Chains and Mixing Times*, vol. 107, 2nd ed., American Mathematical Society, 2017.

[20] L. Li, P. Lu, and Y. Yin, *Approximate counting via correlation decay in spin systems*, in Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2012, pp. 922–940.

[21] L. Li, P. Lu, and Y. Yin, *Correlation decay up to uniqueness in spin systems*, in Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2013, pp. 67–84, The latest version is available at arXiv:1111.7064.

[22] L. Li, P. Lu, and Y. Yin, Correlation Decay Up to Uniqueness in Spin Systems, preprint, https://arxiv.org/abs/1111.7064, 2021.

[23] J. Liu , *Approximate Counting, Phase Transitions and Geometry of Polynomials*, Ph.D. thesis, EECS Department, University of California, Berkeley, Aug 2019.

[24] J. Liu, A. Sinclair, and P. Srivastava, *Fisher zeros and correlation decay in the Ising model*, J. Math. Phys., 60 (2019), 103304.

[25] E. Mossel and A. Sly, *Exact thresholds for Ising-Gibbs samplers on general graphs*, Ann. Probab., 41 (2013), pp. 294–328.

[26] V. Patel and G. Regts, *Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials*, SIAM J. Comput., 46 (2017), pp. 1893–1919.

[27] H. Peters and G. Regts, *On a conjecture of Sokal concerning roots of the independence polynomial*, Michigan Math. J., 68 (2019), pp. 33–55.

[28] A. D. Scott and A. D. Sokal, *The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma*, J. Stat. Phys., 118 (2005), pp. 1151–1261.

[29] S. Shao and Y. Sun, *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), 2020, pp. 96:1–15.

[30] A. Sinclair, P. Srivastava, and M. Thurley, *Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs*, J. Stat. Phys., 155 (2014), pp. 666–686.

[31] A. Sly, *Computational transition at the uniqueness threshold*, in Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS), 2010, pp. 287–296.

[32] A. Sly and N. Sun, *The computational hardness of counting in two-spin models on d-regular graphs*, Ann. Probab., 42 (2014), pp. 2383–2416.

[33] D. Štefankovič, S. Vempala, and E. Vigoda, *Adaptive simulated annealing: A near-optimal connection between sampling and counting*, J. ACM, 56 (2009), pp. 1–36.

[34] D. Weitz, *Counting independent sets up to the tree threshold*, in Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC), 2006, pp. 140–149.

[35] J. Zhang, H. Liang, and F. Bai, *Approximating partition functions of the two-state spin system*, Inform. Process. Lett., 111 (2011), pp. 702–710.