

Rapid Mixing of Glauber Dynamics up to Uniqueness via Contraction

Zongchen Chen
 School of Computer Science
 Georgia Institute of Technology
 chenzongchen@gatech.edu

Kuikui Liu
 School of Computer Science & Engineering
 University of Washington
 liukui17@cs.washington.edu

Eric Vigoda
 School of Computer Science
 Georgia Institute of Technology
 vigoda@gatech.edu

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 et al. (2013). Moreover, in the tree non-uniqueness region, Sly (2010) showed that there is no FPRAS to estimate the partition function unless $\text{NP} = \text{RP}$. The algorithmic results follow from the correlation decay approach due to Weitz (2006) or the polynomial interpolation approach developed by Barvinok (2016). However the running time is only polynomial for constant Δ . For the hardcore model, recent work of Anari et al. (2020) 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 importantly 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 or polynomial interpolation approach immediately imply rapid mixing of the Glauber dynamics. Our proof utilizes 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.

Keywords—approximate counting; Glauber dynamics; spectral independence; phase transitions; correlation decay;

I. 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 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: edge weights $\beta, \gamma > 0$ and a 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 $\gamma = 1$ then the only configurations with non-zero 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 $\beta = \gamma$ then 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, when $\beta > 1$ then 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 $\text{poly}(n, \log(1/\epsilon))$. An efficient approximate sampling algorithm implies an FPRAS (fully-polynomial randomized approximation scheme) for the approximate counting problem [9], [22]. Recall, 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 $\text{poly}(n, 1/\epsilon, \log(1/\delta))$, whereas an FPTAS is the deterministic analog (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 $\mu(\sigma_v | \sigma_w = X_t(w) \text{ for all } w \neq v)$. For the case of the hardcore model, then $X_{t+1}(v)$ is set to occupied (i.e., spin 1) with probability $\lambda/(1 + \lambda)$ if no neighbors are currently occupied, and 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 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 , 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 [8] 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/non-uniqueness 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 $h \rightarrow \infty$. 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 $p_{\text{even}} = \lim_{h \rightarrow \infty} p_{2h}$ and $p_{\text{odd}} = \lim_{h \rightarrow \infty} p_{2h+1}$. We say that *tree uniqueness* holds if $p_{\text{even}} = p_{\text{odd}}$ and *tree non-uniqueness* holds if they are not equal. For all $\Delta \geq 3$ there exists a critical fugacity $\lambda_c(\Delta) = (\Delta-1)^{\Delta-1}/(\Delta-2)^\Delta$ [10], where tree uniqueness holds iff $\lambda \leq \lambda_c(\Delta)$.

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 $\lambda < (1 - \delta)\lambda_c(\Delta)$, all graphs of maximum degree Δ , [23] presented an FPTAS for approximating the partition function. On the other side, for all $\delta > 0$, all $\lambda > (1 + \delta)\lambda_c(\Delta)$, [20], [21], [6] proved that, unless $\text{NP} = \text{RP}$, 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, $\lambda < (1 - \delta)\lambda_c$). Weitz’s correlation decay algorithm was extended to the antiferromagnetic Ising model in the tree uniqueness region by Sinclair et al. [19], and to all antiferromagnetic 2-spin systems in the corresponding tree uniqueness region (as we detail below) by Li, Lu, and Yin [12].

An intriguing new algorithmic approach was presented by Barvinok [3] and refined by Patel and Regts [15], 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 for Weitz’s result by Peters and Regts [16], 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 [2]; they proved, 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 (Hardcore model). *Let $\Delta \geq 3$ be an integer and $\delta \in (0, 1)$. 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 correlation decay and polynomial interpolation approach 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. A recent work of Shao and Sun [18] unifies 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 4 and Theorem 5 for a detailed statement. Our proof utilizes several new tools concerning Weitz’s self-

avoiding walk tree, which are detailed in [Section III](#). 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 8](#). 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 [\[23\]](#); see [Lemma 8](#) 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 $\bar{\beta}_c(\Delta) = \frac{1}{\beta_c(\Delta)} = \frac{\Delta}{\Delta-2}$; the first lies in the antiferromagnetic regime, while the second lies in the ferromagnetic regime. If $\beta_c(\Delta) < \beta < \bar{\beta}_c(\Delta)$, 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 [\[8\]](#). Furthermore, Mossel and Sly [\[14\]](#) proved $O(n \log n)$ mixing time of the Glauber dynamics for the ferromagnetic Ising model when $1 \leq \beta < \bar{\beta}_c(\Delta)$. 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, when $\beta \leq \beta_c$ there is an additional uniqueness region for certain values of the external field λ ; this region is covered by [Theorem 3](#).

Theorem 2 (Antiferromagnetic Ising Model). *Let $\Delta \geq 3$ be an integer and $\delta \in (0, 1)$. Assume that $1 > \beta \geq \beta_c(\Delta) + \delta(1 - \beta_c(\Delta))$ 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 the antiferromagnetic case is 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 [\[12\]](#). Roughly speaking, we say uniqueness with gap $\delta \in (0, 1)$ holds on the d -regular tree if for every integer $\ell \geq 1$, 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 $1 \leq d \leq \Delta$; see [Section II](#) for the precise definition.

Both [Theorem 1](#) and [Theorem 2](#) are corollaries of the following general rapid mixing result which holds for gen-

eral antiferromagnetic 2-spin systems in the entire tree uniqueness region.

Theorem 3 (General antiferromagnetic 2-spin system). *Let $\Delta \geq 3$ be an integer and $\delta \in (0, 1)$. 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 [\[7\]](#), [\[18\]](#) for ferromagnetic 2-spin models; see [Section VI](#) for results and the full version [\[5\]](#) of this paper for proofs.

A. 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 $1 \leq i \leq d$ we define T_{v_i} to be the subtree of T rooted at v_i that contains all descendant of v_i . Let $R_r = \mu_T(\sigma_r = 1) / \mu_T(\sigma_r = 0)$ denote the marginal ratio of the root, and $R_{v_i} = \mu_{T_{v_i}}(\sigma_{v_i} = 1) / \mu_{T_{v_i}}(\sigma_{v_i} = 0)$ 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 $F_d : [0, +\infty]^d \rightarrow [0, +\infty]$ 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 interest in the log of marginal ratios. The reason is that we will carefully study the *pairwise influence matrix* \mathcal{I}_G of the Gibbs distribution μ_G , introduced in [\[2\]](#) and defined as for every $r, v \in V$

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

In [\[2\]](#), 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 $\mathcal{I}_G(r \rightarrow v)$ 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 12](#)). Thus, it is more convenient for us to work with the log ratios. To this end, we rewrite the tree recursion as $\log R_r = H_d(\log R_{v_1}, \dots, \log R_{v_d})$ where $H_d : [-\infty, +\infty]^d \rightarrow [-\infty, +\infty]$ 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 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 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 $\Psi : [-\infty, +\infty] \rightarrow (-\infty, +\infty)$ 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 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 $(\tilde{y}_1, \dots, \tilde{y}_d) \in S^d$, we have

$$\|\nabla H_d^\Psi(\tilde{y}_1, \dots, \tilde{y}_d)\|_1 = \sum_{i=1}^d \frac{\psi(y)}{\psi(y_i)} \cdot |h(y_i)| \leq 1 - \alpha$$

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 $y = H_d(y_1, \dots, y_d)$.

- 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 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 III. Note that in both Definition 4 and Theorem 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 [12] 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 that we show

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 the full version [5] of this paper, 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 5 still holds for this larger class of potentials.

We remark that in all previous works of 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, 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 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 $\mathcal{I}_G(r \rightarrow v)$ 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 switching to (F_d, Φ) . Second, the potential function Ψ we will use is obtained from the exact potential Φ in [12], 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 a 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 [12], and also indicates that H_d is a meaningful variant of the tree recursion to consider. To add one more evidence, 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, that just means we can pick Ψ to be the identity function and H_d itself is contracting without any nontrivial potential.

II. 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 to be

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

¹To be more precise, we also multiply a constant factor which only simplifies our calculation and does not matter much; also notice that [12] denotes the potential function by φ and its derivative by $\Phi = \varphi'$.

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 $1 = \lambda_1 \geq \dots \geq \lambda_{|\Omega|} \geq -1$. 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

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

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

$$\begin{aligned} \mathcal{I}_G^{\sigma_\Lambda}(u \rightarrow v) &= \mu_G(\sigma_v = 1 \mid \sigma_u = 1, \sigma_\Lambda) \\ &\quad - \mu_G(\sigma_v = 1 \mid \sigma_u = 0, \sigma_\Lambda). \end{aligned}$$

Write $\mathcal{I}_G^{\sigma_\Lambda}$ for the (pairwise) *influence matrix* whose entries are given by $\mathcal{I}_G^{\sigma_\Lambda}(u \rightarrow v)$.

Weitz's self-avoiding walk tree: Let $G = (V, E)$ be a connected graph and $r \in V$ 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 $r = v_0 - v_1 - \dots - v_\ell$ such that $v_i \neq v_j$ for all $0 \leq i < j \leq \ell$. The SAW tree $T_{\text{SAW}}(G, r)$ is a tree rooted at r , consisting of all self-avoiding walks $r = v_0 - v_1 - \dots - v_\ell$ with $\deg(v_\ell) = 1$, and those appended with one more vertex that closes the cycle (i.e., $r = v_0 - v_1 - \dots - v_\ell - v_i$ 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.

We can define a 2-spin system on $T_{\text{SAW}}(G, r)$ 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 $r = v_0 - v_1 - \dots - v_\ell$ appended with v_i , we fix v_i to be spin 1 if $v_{i+1} < v_\ell$ 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 $T_{\text{SAW}}(G, r)$ 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 $v \in \Lambda$. Note that in general, different copies of v from \mathcal{C}_v can receive different spin assignments. Finally, in the case that 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.

III. PROOF OUTLINE FOR MAIN RESULTS

Step 1 ([2]): Spectral Independence implies rapid mixing.: Our proof builds on [2] 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 $\lambda \leq (1 - \delta)\lambda_c(\Delta)$. One of the key ingredients of their proof is a notion they call spectral independence. [2] shows that the spectral independence property implies rapid mixing. Note that the diagonal entries of $\mathcal{I}_G^{\sigma_\Lambda}$ are 1, as opposed to 0 in the original definition in [2].

Definition 6 (Spectral Independence [2]). We say that the Gibbs distribution μ_G on an n -vertex graph G is $(\eta_0, \dots, \eta_{n-2})$ -*spectrally independent*, if for every $0 \leq k \leq n - 2$, $\Lambda \subseteq V$ of size k and $\sigma_\Lambda \in \{0, 1\}^\Lambda$, one has $\lambda_{\max}(\mathcal{I}_G^{\sigma_\Lambda}) - 1 \leq \eta_k$.

Theorem 7 ([2]). *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 $\|\mathcal{I}_G^{\sigma_\Lambda}\|_1 = \max_{r \in V} \sum_{v \in V} |\mathcal{I}_G^{\sigma_\Lambda}(v \rightarrow r)|$, which corresponds to total influences on a vertex r , and the infinity-norm $\|\mathcal{I}_G^{\sigma_\Lambda}\|_\infty = \max_{r \in V} \sum_{v \in V} |\mathcal{I}_G^{\sigma_\Lambda}(r \rightarrow v)|$, corresponding to total influences of r . In [2] 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. Here 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, much more is true if we look at the influences from r in the self-avoiding tree. We show that for every vertex $v \in V$, the influence $\mathcal{I}_G^{\sigma_\Lambda}(r \rightarrow v)$ in G is preserved in the self-avoiding walk tree $T = T_{\text{SAW}}(G, r)$ rooted at r , in the form of sum of influences $\mathcal{I}_T^{\sigma_\Lambda}(r \rightarrow \hat{v})$ 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 $\lambda = \{\lambda_v : v \in V\}$ 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 $w_G(\sigma | \sigma_\Lambda) = \beta^{m_1(\sigma | \sigma_\Lambda)} \gamma^{m_0(\sigma | \sigma_\Lambda)} \prod_{v \in V \setminus \Lambda} \lambda_v^{\sigma_v}$ where $m_i(\cdot | \sigma_\Lambda)$ is the number of i -edges with at least one endpoint in $V \setminus \Lambda$ for $i = 0, 1$. Furthermore, $Z_G^{\sigma_\Lambda} = \sum_{\sigma \in \{0, 1\}^{V \setminus \Lambda}} w_G(\sigma | \sigma_\Lambda)$ is the partition function conditioned on σ_Λ . We shall view β and γ as some fixed constants and think of λ as $n = |V|$ variables. In this sense, we regard the weights $w_G(\sigma | \sigma_\Lambda)$ as monomials in λ and the partition function $Z_G^{\sigma_\Lambda}$ as a polynomial in λ . Moreover, the marginal ratios $R_G^{\sigma_\Lambda}(v)$ and the influences $\mathcal{I}_G^{\sigma_\Lambda}(r \rightarrow v)$ for $r, v \in V$ are all functions in λ . Our main result is that the partition function of G divides that of $T_{\text{SAW}}(G, r)$ 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 [23], see [Lemma 13](#).

Lemma 8. *Let $G = (V, E)$ be a connected graph, $r \in V$ be a vertex and $\Lambda \subseteq V \setminus \{r\}$ such that $G \setminus \Lambda$ is connected. Let $T = T_{\text{SAW}}(G, r)$ 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 $P_{G,r}^{\sigma_\Lambda} = P_{G,r}^{\sigma_\Lambda}(\lambda)$ independent of λ_r such that*

$$Z_T^{\sigma_\Lambda} = Z_G^{\sigma_\Lambda} \cdot P_{G,r}^{\sigma_\Lambda}. \quad (2)$$

As a corollary, for each vertex $v \in V$,

$$\mathcal{I}_G^{\sigma_\Lambda}(r \rightarrow v) = \sum_{\hat{v} \in \mathcal{C}_v} \mathcal{I}_T^{\sigma_\Lambda}(r \rightarrow \hat{v}), \quad (3)$$

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

Remark 1. We emphasize that for the purposes of bounding the total influence of a vertex in G , only [Eq. \(3\)](#) of [Lemma 8](#) is needed, which can be proved in a purely combinatorial fashion. However, we believe the divisibility property [Eq. \(2\)](#) 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 [Eq. \(2\)](#) has already appeared in [4] for the hardcore model and [13] for the zero-field Ising model in the study of complex roots of the partition function. From [Lemma 8](#), we can get $\sum_{v \in V} |\mathcal{I}_G^{\sigma_\Lambda}(r \rightarrow v)| \leq \sum_{v \in V_T} |\mathcal{I}_T^{\sigma_\Lambda}(r \rightarrow v)|$ 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 us a great tool for computing the (log) ratios of vertices recursively for trees. As we show in [Lemma 12](#), the influence $\mathcal{I}_G^{\sigma_\Lambda}(r \rightarrow v)$ is in fact a version of 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 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 9. *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 $k \geq 1$,*

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

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

[Theorem 5](#) is then proved by combining [Theorem 7](#), [Lemma 8](#) and [Lemma 9](#). We leave its proof to the full version [5] of the paper.

Step 4: Find a good potential.: As our final step, we need to find an (α, c) -potential function as defined in [Definition 4](#). The potential Ψ we choose is exactly the one from [12], adapted to the log marginal ratios and the tree recursion H (see the full version [5] 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 $\gamma \leq 1$, then Ψ is an (α, c) -potential.

Lemma 10. *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 (β, γ, λ) is up-to- Δ unique with gap $\delta \in (0, 1)$. Define the function Ψ implicitly by*

$$\Psi'(y) = \psi(y) = \sqrt{\frac{(1 - \beta\gamma)e^y}{(\beta e^y + 1)(e^y + \gamma)}} = \sqrt{|h(y)|}, \quad (4)$$

$$\Psi(0) = 0.$$

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

We deduce [Theorem 3](#) for the case $\sqrt{\beta\gamma} > \frac{\Delta-2}{\Delta}$ or $\gamma \leq 1$ from [Theorem 5](#) and [Lemma 10](#). The proof of them can be found in the full version [5]. The case that $\sqrt{\beta\gamma} \leq \frac{\Delta-2}{\Delta}$ and $\gamma > 1$ is trickier. As discussed in Section 5 of [12], when $\sqrt{\beta\gamma} \leq \frac{\Delta-2}{\Delta}$ and $\gamma > 1$, the threshold λ_c for the external field is bounded below by some constant for all Δ ; so for some $\lambda > 0$ the spin system is 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 4](#) and [Theorem 5](#) are then modified to a slightly stronger version. The statements and proofs for this case are presented in the full version [5] of the paper.

IV. PRESERVATION OF INFLUENCES FOR SELF-AVOIDING WALK TREES

In this section we show that the self-avoiding walk (SAW) tree, introduced in [23] (see also [17]), maintains all the influence of the root, and thus establishes [Lemma 8](#). To do this, we show that the partition function of G , viewed as a polynomial of the external fields λ , 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 [23], and also all pairwise covariances concerned with v are preserved.

Theorem 11. *Let $G = (V, E)$ be a connected graph, $r \in V$ be a vertex and $\Lambda \subseteq V \setminus \{r\}$ such that $G \setminus \Lambda$ is connected. Let $T = T_{\text{SAW}}(G, r)$ 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 $P_{G,r}^{\sigma_\Lambda} = P_{G,r}^{\sigma_\Lambda}(\lambda)$ 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 2. The proof of [Theorem 11](#) can be adapted to give a purely combinatorial proof of [Eq. \(3\)](#) in [Lemma 8](#). Like in the proof of [23, Theorem 3.1], one can proceed via vertex splitting and telescoping, where instead of telescoping a product of marginal ratios, one instead telescopes a sum of single-vertex influences.

We remark that [4] proved a univariate version of [Theorem 11](#) for the hardcore model, and [13] 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 deduce the preservation of marginal and influences of the root immediately.

Before proving [Theorem 11](#), 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 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 12. *For every graph $G = (V, E)$, $\Lambda \subseteq V$ and $\sigma_\Lambda \in \{0, 1\}^\Lambda$, the following holds:*

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 $u, v \in V$,

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

3) For all $u, v \in V$,

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

Proof: The first two parts are standard. The proofs of them can be found in the full version [5]. For Part 3, we deduce from Part 2 that

$$\begin{aligned} \left(\lambda_v \frac{\partial}{\partial \lambda_v} \right) \log R_G^{\sigma_\Lambda}(u) &= \left(\lambda_v \frac{\partial}{\partial \lambda_v} \right) \log \left(\frac{M_G^{\sigma_\Lambda}(u)}{1 - M_G^{\sigma_\Lambda}(u)} \right) \\ &= \frac{\left(\lambda_v \frac{\partial}{\partial \lambda_v} \right) M_G^{\sigma_\Lambda}(u)}{M_G^{\sigma_\Lambda}(u) (1 - M_G^{\sigma_\Lambda}(u))} \\ &= \frac{K_G^{\sigma_\Lambda}(u, v)}{K_G^{\sigma_\Lambda}(u, u)}. \end{aligned}$$

It remains to show that

$$\mathcal{I}_G^{\sigma_\Lambda}(u \rightarrow v) = \frac{K_G^{\sigma_\Lambda}(u, v)}{K_G^{\sigma_\Lambda}(u, u)},$$

which actually holds for any two binary random variables. To see this, we first compute $K_G^{\sigma_\Lambda}(u, u) \cdot \mathcal{I}_G^{\sigma_\Lambda}(u \rightarrow v)$ by definition:

$$\begin{aligned} &K_G^{\sigma_\Lambda}(u, u) \cdot \mathcal{I}_G^{\sigma_\Lambda}(u \rightarrow v) \\ &= \mu_G(u = 1 \mid \sigma_\Lambda) \cdot \mu_G(u = 0 \mid \sigma_\Lambda) \\ &\quad \cdot [\mu_G(v = 1 \mid u = 1, \sigma_\Lambda) - \mu_G(v = 1 \mid u = 0, \sigma_\Lambda)] \\ &= \mu_G(u = 1, v = 1 \mid \sigma_\Lambda) \cdot \mu_G(u = 0, v = 0 \mid \sigma_\Lambda) \\ &\quad - \mu_G(u = 1, v = 0 \mid \sigma_\Lambda) \cdot \mu_G(u = 0, v = 1 \mid \sigma_\Lambda). \end{aligned}$$

Meanwhile, the covariance can be written as

$$\begin{aligned} K_G^{\sigma_\Lambda}(u, v) &= \mu_G(u = 1, v = 1 \mid \sigma_\Lambda) \\ &\quad - \mu_G(u = 1 \mid \sigma_\Lambda) \cdot \mu_G(v = 1 \mid \sigma_\Lambda) \\ &= \mu_G(u = 1, v = 1 \mid \sigma_\Lambda) \cdot \mu_G(u = 0, v = 0 \mid \sigma_\Lambda) \\ &\quad - \mu_G(u = 1, v = 0 \mid \sigma_\Lambda) \cdot \mu_G(u = 0, v = 1 \mid \sigma_\Lambda). \end{aligned}$$

This shows that $\mathcal{I}_G^{\sigma_\Lambda}(u \rightarrow v) = K_G^{\sigma_\Lambda}(u, v)/K_G^{\sigma_\Lambda}(u, u)$ and thus establishes Part 3. \blacksquare

We deduce [Lemma 8](#) from [Theorem 11](#) and the second item of the following lemma. The proof of [Theorem 11](#) is presented in [Section IV-A](#).

Lemma 13. *Let $G = (V, E)$ be a connected graph, $r \in V$ be a vertex and $\Lambda \subseteq V \setminus \{r\}$ such that $G \setminus \Lambda$ is connected. Let $T = T_{\text{SAW}}(G, r)$ be the self-avoiding walk tree of G rooted at r . Then for every $\sigma_\Lambda \in \{0, 1\}^\Lambda$ we have:*

1) ([23, Theorem 3.1]) Preservation of marginal of the root r :

$$M_G^{\sigma_\Lambda}(r) = M_T^{\sigma_\Lambda}(r), \quad R_G^{\sigma_\Lambda}(r) = R_T^{\sigma_\Lambda}(r);$$

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

$$\begin{aligned} K_G^{\sigma_\Lambda}(r, v) &= \sum_{\hat{v} \in \mathcal{C}_v} K_T^{\sigma_\Lambda}(r, \hat{v}), \\ \mathcal{I}_G^{\sigma_\Lambda}(r \rightarrow v) &= \sum_{\hat{v} \in \mathcal{C}_v} \mathcal{I}_T^{\sigma_\Lambda}(r \rightarrow \hat{v}). \end{aligned}$$

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

Proof: By **Theorem 11**, 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 12** 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 12** 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 12**, we deduce that

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

Finally, we have

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

where the last equality follows as above. \blacksquare

A. Proof of **Theorem 11**

Before presenting our proof, let us first review the notations 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

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

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 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 notations as well.

*Proof of **Theorem 11**:* We will show that there exists a polynomial $P_{G,r}^{\sigma_\Lambda} = P_{G,r}^{\sigma_\Lambda}(\boldsymbol{\lambda})$, independent of λ_r , such that

$$\begin{aligned} Z_T^{\sigma_\Lambda}(r = 1) &= Z_G^{\sigma_\Lambda}(r = 1) \cdot P_{G,r}^{\sigma_\Lambda}, \\ Z_T^{\sigma_\Lambda}(r = 0) &= Z_G^{\sigma_\Lambda}(r = 0) \cdot P_{G,r}^{\sigma_\Lambda}. \end{aligned} \quad (5)$$

The high-level proof idea of **Eq. (5)** is similar to the corresponding result in [23, 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 **Eq. (5)** 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 $\{r_i, d_i\}$ for $1 \leq i \leq d$.

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 an 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 $U_i = \{v_1, \dots, v_d\} \setminus \{v_i\}$. Then, $T = T_{\text{SAW}}(G, r)$ can be generated by the following recursive procedure.

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

- 1) For each i , let $T_i = T_{\text{SAW}}(G_i, v_i)$ plus 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 $1 \leq i \leq d$; output T .

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

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

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_i' ; 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 $Q_{G_i,r}^{\sigma_\Lambda} = Q_{G_i,r}^{\sigma_\Lambda}(\boldsymbol{\lambda})$ by

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

which is independent of λ_r . Then we get

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

Using a similar argument, we also have

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

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

$$\begin{aligned} Z_{T_i}^{\sigma_{\Lambda_i}}(r = 1) &= Z_{G_i}^{\sigma_{\Lambda_i}}(r = 1) \cdot P_{G_i, v_i}^{\sigma_{\Lambda_i}}, \\ Z_{T_i}^{\sigma_{\Lambda_i}}(r = 0) &= Z_{G_i}^{\sigma_{\Lambda_i}}(r = 0) \cdot P_{G_i, v_i}^{\sigma_{\Lambda_i}}; \end{aligned}$$

these polynomials are independent of λ_r since the conditional partition functions for 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

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

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 that $(G \setminus \{r\}) \setminus \Lambda$ 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 component for an integer $k \geq 2$. Let $G'_{(1)}, \dots, G'_{(k)}$ be the subgraphs induced by each component, along 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 $\sigma_{\Lambda_{(j)}}$ 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 $P_{G_{(j)}, r_{(j)}}^{\sigma_{\Lambda_{(j)}}} = P_{G_{(j)}, r_{(j)}}^{\sigma_{\Lambda_{(j)}}}(\boldsymbol{\lambda})$, which is independent of λ_r , such that

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

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

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

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

V. 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., [11], [12], [7]). 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 9](#) in the case that Ψ 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 $\mathcal{I}_T^{\sigma_\Lambda}(r \rightarrow v)$ from the root r to another vertex $v \in V_T$. Notice that if v is disconnected from r when Λ is removed, then $\mathcal{I}_T^{\sigma_\Lambda}(r \rightarrow v) = 0$ 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 descendant 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 interest in 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 $1 \leq d \leq \Delta$ 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 $1 \leq i \leq d \leq \Delta$.

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

Lemma 14. *Let $\Psi : [-\infty, +\infty] \rightarrow (-\infty, +\infty)$ 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 $k \geq 1$,*

$$\begin{aligned} & \sum_{v \in L_r(k)} |\mathcal{I}_T^{\sigma_\Lambda}(r \rightarrow v)| \\ & \leq \Delta_r A_\Psi B_\Psi \left(\max_{1 \leq d < \Delta} \sup_{\mathbf{y} \in S^d} \|\nabla H_d^\Psi(\mathbf{y})\|_1 \right)^{k-1} \end{aligned}$$

where

$$\begin{aligned} A_\Psi &= \max_{u \in L_r(1)} \left\{ \frac{|h(\log R_u)|}{\psi(\log R_u)} \right\}, \\ B_\Psi &= \max_{v \in L_r(k)} \{ \psi(\log R_v) \}. \end{aligned}$$

Before proving [Lemma 14](#), we first present two useful properties of the influences on trees. Firstly, it was shown in [\[2\]](#) that the influences satisfy the following form of chain rule on trees.

Lemma 15 ([\[2, Lemma B.2\]](#)). *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*

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

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

Lemma 16. *Let $v \in V_T$ and u be a child of v in the subtree T_v . Then*

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

Proof: The lemma can be proved through an explicit computation of the influence. Here we present a more delicate proof utilizing [Lemma 12](#), which gives some insights 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 12](#) and the tree recursion imply that

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

where the last equality is because $\mathcal{I}_{T_{u_i}}^{\sigma_\Lambda}(u_i \rightarrow u) = 0$ for $u_i \neq u$ and $\mathcal{I}_{T_u}^{\sigma_\Lambda}(u \rightarrow u) = 1$. \blacksquare

We are now ready to prove [Lemma 14](#).

Proof of Lemma 14: 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_{Δ_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 $\mathcal{I}_T^{\sigma_\Lambda}(r \rightarrow u_i) = 0$ by definition. Then by [Lemma 15](#) and [Lemma 16](#), we get

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

Hence, we obtain that

$$\begin{aligned} \sum_{v \in L_r(k)} |\mathcal{I}_T^{\sigma_A}(r \rightarrow v)| &\leq \Delta_r \cdot \max_{1 \leq i \leq \Delta_r} \left\{ \frac{|h(\log R_{u_i})|}{\psi(\log R_{u_i})} \right\} \quad (6) \\ &\times \max_{1 \leq i \leq \Delta_r} \left\{ \sum_{v \in L_{u_i}(k-1)} \psi(\log R_{u_i}) |\mathcal{I}_T^{\sigma_A}(u_i \rightarrow v)| \right\}. \end{aligned}$$

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

$$\begin{aligned} \sum_{v \in L_u(k)} \psi(\log R_u) |\mathcal{I}_T^{\sigma_A}(u \rightarrow v)| \quad (7) \\ \leq \max_{v \in L_u(k)} \{\psi(\log R_v)\} \cdot \left(\max_{w \in V_{T_u}} \sup_{\tilde{\mathbf{y}} \in S^{d_w}} \|\nabla H_{d_w}^{\Psi}(\tilde{\mathbf{y}})\|_1 \right)^k. \end{aligned}$$

Observe that once we establish Eq. (7), the lemma follows immediately by plugging Eq. (7) into Eq. (6). We will use induction on k to prove Eq. (7). When $k = 0$, if $u \in \Lambda$ is fixed then $L_u(0) = \emptyset$ and there is nothing to show; otherwise, Eq. (7) becomes

$$\psi(\log R_u) |\mathcal{I}_T^{\sigma_A}(u \rightarrow u)| \leq \psi(\log R_u),$$

which holds with equality since $\mathcal{I}_T^{\sigma_A}(u \rightarrow u) = 1$. Now suppose that Eq. (7) holds for some integer $k-1 \geq 0$ (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 $1 \leq d < \Delta$ (if $d = 0$ then $L_u(k) = \emptyset$ and Eq. (7) holds trivially). Again by Lemma 15 and Lemma 16 we have

$$\begin{aligned} \sum_{v \in L_u(k)} \psi(\log R_u) |\mathcal{I}_T^{\sigma_A}(u \rightarrow v)| \\ = \sum_{i=1}^d \psi(\log R_u) |\mathcal{I}_T^{\sigma_A}(u \rightarrow w_i)| \sum_{v \in L_{w_i}(k-1)} |\mathcal{I}_T^{\sigma_A}(w_i \rightarrow v)| \\ = \sum_{i=1}^d \frac{\psi(\log R_u)}{\psi(\log R_{w_i})} |h(\log R_{w_i})| \\ \times \sum_{v \in L_{w_i}(k-1)} \psi(\log R_{w_i}) |\mathcal{I}_T^{\sigma_A}(w_i \rightarrow v)|. \end{aligned}$$

For $U \subseteq V_T$, we let

$$\Xi(U) = \max_{w \in U} \sup_{\tilde{\mathbf{y}} \in S^{d_w}} \|\nabla H_{d_w}^{\Psi}(\tilde{\mathbf{y}})\|_1.$$

Using the induction hypothesis, we get

$$\begin{aligned} \sum_{v \in L_u(k)} \psi(\log R_u) |\mathcal{I}_T^{\sigma_A}(u \rightarrow v)| \\ \leq \max_{v \in L_u(k)} \{\psi(\log R_v)\} \cdot \Xi(V_{T_u} \setminus \{u\})^{k-1} \\ \times \sum_{i=1}^d \frac{\psi(\log R_u)}{\psi(\log R_{w_i})} |h(\log R_{w_i})| \\ \leq \max_{v \in L_u(k)} \{\psi(\log R_v)\} \cdot \Xi(V_{T_u})^k, \end{aligned}$$

where the last inequality follows from that

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

This establishes Eq. (7), and thus proves the lemma. \blacksquare

We then derive Lemma 9 as a corollary.

Proof of Lemma 9: Since Ψ is an (α, c) -potential, the Contraction condition implies that

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

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 $v \in L_r(k)$,

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

Therefore, we get $\Delta_r A_{\Psi} B_{\Psi} \leq c$. The lemma then follows immediately from Lemma 14. \blacksquare

VI. FERROMAGNETIC CASES

In the ferromagnetic case, the best known correlation decay results are given in [7], [18]. Using the potential functions in [7] and [18], we show the following two results, which match the known correlation decay results.

Theorem 17. Fix an integer $\Delta \geq 3$, positive real numbers β, γ, λ 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 λ is arbitrary;
- 2) $\sqrt{\beta\gamma} \geq \frac{\Delta}{\Delta-2}$ and

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

- 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 [18]) 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 3. 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 Eq. (4) in Lemma 10.

For the ferromagnetic case with $\beta = \gamma > 1$ (Ising model), [14] proved a stronger result of $O(n \log n)$ mixing.

The potential function from [7] 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 [7] for unbounded degree graphs.

Theorem 18. 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 given in the full version [5].

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.

REFERENCES

- [1] Vedat Levi Alev and Lap Chi Lau. Improved analysis of higher order random walks and applications. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 1198–1211, 2020.
- [2] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. *arXiv preprint arXiv:2001.00303*, 2020.
- [3] Alexander Barvinok. *Combinatorics and Complexity of Partition Functions*, volume 30. Springer Algorithms and Combinatorics, 2016.
- [4] Ferenc Bencs. On trees with real-rooted independence polynomial. *Discrete Mathematics*, 341(12):3321–3330, 2018.
- [5] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Rapid mixing of Glauber dynamics up to uniqueness via contraction. *Full version, arXiv:2004.09083*, 2020.
- [6] Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models. *Combinatorics, Probability and Computing*, 25(4):500–559, 2016.
- [7] Heng Guo and Pinyan Lu. Uniqueness, spatial mixing, and approximation for ferromagnetic 2-spin systems. *ACM Transactions on Computation Theory*, 10(4), 2018.
- [8] Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM Journal on Computing*, 22(5):1087–1116, 1993.
- [9] Mark R Jerrum, Leslie G Valiant, and Vijay V Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43:169–188, 1986.
- [10] F. P. Kelly. Stochastic models of computer communication systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 47(3):379–395, 1985.
- [11] Liang Li, Pinyan Lu, and Yitong Yin. Approximate counting via correlation decay in spin systems. In *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 922–940, 2012.
- [12] Liang Li, Pinyan Lu, and Yitong Yin. Correlation decay up to uniqueness in spin systems. In *Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 67–84, 2013.
- [13] Jingcheng Liu, Alistair Sinclair, and Piyush Srivastava. Fisher zeros and correlation decay in the Ising model. *Journal of Mathematical Physics*, 60(10):103304, 2019.
- [14] Elchanan Mossel and Allan Sly. Exact thresholds for Ising-Gibbs samplers on general graphs. *Annals of Probability*, 41(1):294–328, 2013.
- [15] Viresh Patel and Guus Regts. Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *SIAM Journal on Computing*, 46:1893–1919, 2017.
- [16] Han Peters and Guus Regts. On a conjecture of Sokal concerning roots of the independence polynomial. *The Michigan Mathematical Journal*, 68(1):33–55, 2019.
- [17] Alexander D. Scott and Alan D. Sokal. The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma. *Journal of Statistical Physics*, 118(5):1151–1261, 2005.
- [18] Shuai Shao and Yuxin 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)*, pages 96:1–15, 2020.
- [19] Alistair Sinclair, Piyush Srivastava, and Marc Thurley. Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs. *Journal of Statistical Physics*, 155(4):666–686, 2014.
- [20] Allan Sly. Computational transition at the uniqueness threshold. In *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 287–296, 2010.
- [21] Allan Sly and Nike Sun. The computational hardness of counting in two-spin models on d -regular graphs. *The Annals of Probability*, 42(6):2383–2416, 2014.
- [22] Daniel Štefankovič, Santosh Vempala, and Eric Vigoda. Adaptive simulated annealing: A near-optimal connection between sampling and counting. *Journal of the ACM*, 56(3):1–36, 2009.
- [23] Dror Weitz. Counting independent sets up to the tree threshold. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC)*, pages 140–149, 2006.