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 $\Delta$ 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 $NP = RP$. 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 $\Delta$. 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 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 $\Delta$ and the statistical physics phase transition on infinite, regular trees of degree $\Delta$ (or up to $\Delta$ 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 $\beta$ controls the (relative) strength of interaction between neighboring 1-spins, $\gamma$ corresponds to neighboring 0-spins, and $\lambda$ is the external field applied to vertices with 1-spins.

Every spin configuration $\sigma \in \{0, 1\}^V$ is assigned a weight $w_G(\sigma) = \beta^{m_1(\sigma)}\lambda^{n_1(\sigma)}$, 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 $n_s(\sigma) = |\{v \in V : \sigma_v = 1\}|$ 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 $\sigma$ is $\lambda^{|\sigma|}$; this example is known as the "hardcore model" where the parameter $\lambda$ 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 $\sigma$ is $w(\sigma) = \beta^m(\sigma)\lambda^n(\sigma)$; this is the classical Ising model where the parameter $\beta$ corresponds to the inverse temperature and $\lambda$ 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 $\pi$ which is within total variation distance $\leq \epsilon$ of the Gibbs distribution $\mu_G$ 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 $\mu(\sigma_v | \sigma_w = X_{t+1}(w))$ 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 $\Delta$-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 $\Delta$-regular tree $T_h$. Let $p_{even} = \lim_{h \to \infty} p_{2h}$, and $p_{odd} = \lim_{h \to \infty} p_{2h+1}$. We say that tree uniqueness holds if $p_{even} = p_{odd}$ 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 $\Delta$ occurs at this same tree critical point. For all constant $\Delta$, all $\delta > 0$, all $\lambda < (1 - \delta)\lambda_c(\Delta)$, all graphs of maximum degree $\Delta$, [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 $\delta$ (recall, $\lambda < (1 - \delta)\lambda_c(\Delta)$). 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 $\Delta$.

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 $\Delta$ and every $0 < \lambda < (1 - \delta)\lambda_c(\Delta)$, the mixing time of the Glauber dynamics for the hardcore model on $G$ with fugacity $\lambda$ 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 $\Delta$; 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 $\Delta$-regular tree are at $\beta_c(\Delta) = \frac{\Delta+2}{\Delta}$ and $\bar{\beta}_c(\Delta) = \frac{1}{\pi(\Delta)} = \frac{\Delta}{2}$. 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 $\lambda$ on the $\Delta$-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 $\lambda$; 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 $1 > \beta > \beta_c(\Delta) + \delta(1 - \beta_c(\Delta))$ and $\lambda > 0$. Then for every $n$-vertex graph $G$ of maximum degree $\Delta$, the mixing time of the Glauber dynamics for the Ising model on $G$ with edge weight $\beta$ and external field $\lambda$ 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 $\Delta$. Hence the appropriate notion is “up-to-$\Delta$ 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 $\ell$ from the root have total “influence” $\leq (1 - \delta)^\ell$ on the marginal of the root. We say up-to-$\Delta$ uniqueness with gap $\delta$ holds if uniqueness with gap $\delta$ 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 general 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 $\beta, \gamma, \lambda$ be reals such that $0 \leq \beta \leq \gamma$ and $\gamma > 0$, $\beta \gamma < 1$ and $\lambda > 0$. Assume that the parameters $(\beta, \gamma, \lambda)$ are up-to-$\Delta$ unique with gap $\delta$. Then for every $n$-vertex graph $G$ of maximum degree $\Delta$, the mixing time of the Glauber dynamics for the antiferromagnetic 2-spin system on $G$ with parameters $(\beta, \gamma, \lambda)$ is $O(n^{2+1.5/\delta})$.

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 $v_1, ..., v_d$. For $1 \leq i \leq \Delta$, 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 $R_{v_1}, ..., R_{v_d}$, due to the independence of $T_{v_i}$’s. More specifically, we can write $R_r = F_d(R_{v_1}, ..., R_{v_d})$ where $F_d : [0, +\infty]^d \to [0, +\infty]$ is a multivariate function such that for $(x_1, ..., x_d) \in [0, \infty]^d$,

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

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 $I_G$ of the Gibbs distribution $\mu_G$, introduced in [2] and defined as for every $r, v \in V$

$$I_G(r \rightarrow v) = \mu_G(\sigma_v = 1 | \sigma_r = 1) - \mu_G(\sigma_v = 1 | \sigma_r = 0).$$

In [2], the authors show that if the maximum eigenvalue of $I_G$ is bounded appropriately, then the Glauber dynamics is rapid mixing. One crucial observation we make in this paper is that the influence $I_G(r \rightarrow v)$ 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 $\log R_v = H_d(\log R_{v_1}, ..., \log R_{v_d})$ where $H_d : [-\infty, +\infty]^d \to [-\infty, +\infty]$ is a function such that for $(y_1, ..., y_d) \in [-\infty, +\infty]^d$,

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

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 \( \frac{\partial}{\partial y_i} H_d(y_1, \ldots, y_d) = h(y_i) \) 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 \( J = \bigcup_{d=0}^{\infty} J_d \) to be the interval containing all log ratios with degree less than \( \Delta \).

**Definition 4 ((\( \alpha, c \))-Potential function).** Let \( \Delta \geq 3 \) be an integer. Let \( \beta, \gamma, \lambda \) be reals such that \( 0 \leq \beta \leq \gamma \), \( \gamma > 0 \) 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 \( \Psi \) is an \( (\alpha, c) \)-potential function with respect to \( \Delta \) and \( (\beta, \gamma, \lambda) \) if it satisfies the following conditions:

1. **(Contraction)** For every integer \( d \) such that \( 1 \leq d < \Delta \) and every \((y_1, \ldots, y_d) \in S^d\), we have
   \[
   \| \nabla H_d^\Psi (y_1, \ldots, y_d) \|_1 = \sum_{i=1}^d \frac{\psi(y_i)}{\psi(y_j)} \cdot |h(y_i)| \leq 1 - \alpha
   \]
   where \( H_d^\Psi = \Psi \circ H_d \circ \Psi^{-1} \), \( y_i = \Psi^{-1}(y_i) \) 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
   \[
   \frac{\psi(y_2)}{\psi(y_1)} \cdot |h(y_1)| \leq c/\Delta
   \]

In the definition of \((\alpha, c)\)-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 \((\alpha, c)\)-potential function.

**Theorem 5.** Let \( \Delta \geq 3 \) be an integer. Let \( \beta, \gamma, \lambda \) be reals such that \( 0 \leq \beta \leq \gamma \), \( \gamma > 0 \) and \( \lambda > 0 \). Suppose that there is an \((\alpha, c)\)-potential with respect to \( \Delta \) and \((\beta, \gamma, \lambda)\) for some \( \alpha \in (0, 1) \) and \( c > 0 \). Then for every \( n \)-vertex graph \( G \) of maximum degree \( \Delta \), the mixing time of the Glauber dynamics for the 2-spin system on \( G \) with parameters \((\beta, \gamma, \lambda)\) is \( O(n^2/\alpha) \).

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 \( \Delta \) and parameters \((\beta, \gamma, \lambda)\) 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 \( \Delta \) and \((\beta, \gamma, \lambda)\). Thus, our mixing time is \( O(n^{2+c/\alpha}) \) with no parameters in the exponent except for \( 1/\alpha \).

In the full version [5] of this paper, we give a slightly more general definition of \((\alpha, c)\)-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 \( \Phi \), a potential function of \( R \). In fact, our results can also be translated into the language of \((F_d, \Phi)\). 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 \( \Phi \), a potential function of \( R \). In fact, our results can also be translated into the language of \((F_d, \Phi)\). To see this, since \( H_d = \log \circ F_d \circ \exp \), in straightforward forward 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 \((F_d, \Psi)\) if and only if the corresponding contraction condition holds for \((F_d, \Phi)\). The Boundedness condition can also be stated equivalently for \((F_d, \Phi)\). Nevertheless, in this paper we choose to work with \((H_d, \Psi)\) for the following two reasons. First, as mentioned earlier, the fact that \( \Omega_C(r \to v) \) is a derivative of \( \log R_t \) 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 \((H_d, \Psi)\) directly rather than switching to \((F_d, \Phi)\). Second, the potential function \( \Psi \) we will use is obtained from the exact potential \( \Phi \) in [12], by the transformation \( \Psi = \Phi \circ \exp \).

It is intriguing to notice that the derivative of this potential is simply \( \psi = \sqrt{|h|} \). Then the Contraction condition has a nice form: \( \sum_{i=1}^d \sqrt{|h(y_i)h(y_j)|} \leq 1 - \alpha \); and the Boundedness condition only involves an upper bound on \( h(y) \). This seems to shed some light on the mysterious potential function \( \Phi \) 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., \( \Delta^2/4 < \sqrt{\beta \gamma} < \Delta^2/2 \)) where the potential \( \Phi = \log \) is picked, that just means we can pick \( \Psi \) 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 \( \Omega \) with stationary distribution \( \mu \). 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

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

To 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 \( \varphi \) 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 $1 = \lambda_1 \geq \cdots \geq \lambda_{|\Omega|} \geq -1$. 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 semi-definite 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 $\beta, \gamma, \lambda$ be real 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 $(\beta, \gamma, \lambda)$ are up-to-$\Delta$ unique with gap $\delta$ 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

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

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

$$T^*_G(u \to v) = \mu_G(\sigma_v = 1 \mid \sigma_u = 1, \sigma_\Lambda) - \mu_G(\sigma_v = 1 \mid \sigma_u = 0, \sigma_\Lambda).$$

Write $T^*_G$ for the (pairwise) influence matrix whose entries are given by $T^*_G(u \to v)$.

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 $r = v_0 - v_1 - \cdots - v_\ell$ 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 $r = v_0 - v_1 - \cdots - v_\ell$ with $\deg(v_\ell) = 1$, and those appended with one more vertex that closes the cycle (i.e., $r = v_0 - v_1 - \cdots - v_\ell - v_1$ 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 $(\beta, \gamma, \lambda)$, 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 $v_\ell$, we fix $v_\ell$ to be spin 1 if $v_{\ell+1} < v_\ell$ with respect to the total ordering on $V$, and spin 0 if $v_{\ell+1} > v_\ell$. For each $v \in V$ we denote the set of all free (unfixed) copies of $v$ in $T_{\text{SAW}}(G,r)$ by $C_v$. For $\Lambda \subseteq V$ and a partial configuration $\sigma_\Lambda \in \{0,1\}^\Lambda$, we define the SAW tree with conditioning $\sigma_\Lambda$ by assigning the spin $\sigma_v$ to every copy $\hat{v}$ of $v$ from $C_v$ and removing all descendants of $\hat{v}$, for each $v \in \Lambda$. Note that in general, different copies of $v$ from $C_v$ can receive different spin assignments. Finally, in the case that every vertex $v$ has a distinct field $\lambda_v$, all copies of $v$ from $C_v$ will have the same field $\lambda_v$ 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 $\Delta$ mixes in $O(n \exp(G(1/\Delta)))$ 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 $T^*_G$ are 1, as opposed to 0 in the original definition in [2].

Definition 6 (Spectral Independence [2]). We say that the Gibbs distribution $\mu_G$ on an $n$-vertex graph $G$ is $(\eta_0, \ldots, \eta_{n-2})$-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

$$\lambda_{\max}(T^*_G(0, \Lambda)) - 1 \leq \eta_k.$$  

Theorem 7 ([2]). If $\mu$ is a $(\eta_0, \ldots, \eta_{n-2})$-spectrally independent distribution, then the Glauber dynamics for sampling from $\mu$ 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 $T^*_G$.

Step 2: Self-avoiding walk trees preserve influences.

From standard linear algebra, we know that the maximum eigenvalue of $T^*_G$ is upper bounded by both the 1-norm $\|T^*_G\|_1 = \max_{v \in V} \sum_{w \in V} |T^*_G(v \to w)|$, which corresponds to total influences on a vertex $r$, and the infinity-norm $\|T^*_G\|_\infty = \max_{r \in V} \sum_{v \in V} |T^*_G(r \to v)|$, corresponding to total influences of $r$. In [2] the authors use $\|T^*_G\|_\infty$ as an upper bound on $\lambda_{\max}(T^*_G)$, 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 $\|T^*_G\|_\infty$ to upper bound $\lambda_{\max}(T^*_G)$ 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 $T^*_G(r \to v)$ 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 $T^*_T(r \to \hat{v})$ over all copies $\hat{v}$ of $v$.

The way we establish this fact is by viewing the partition function as a polynomial in $\lambda$. In fact, it will be useful to consider the more general case with an arbitrary external
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 $(\alpha,c)$-potential function $\Psi$ with respect to $\Delta$ and $(\beta,\gamma,\lambda)$ 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 $k \geq 1$,

$$
\sum_{v \in L_v(k)} |I_{\sigma_\Lambda}^G(r \to v)| \leq c \cdot (1 - \alpha)^{k-1}
$$

where $L_v(k)$ 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 $(\alpha,c)$-potential function as defined in Definition 4. The potential $\Psi$ 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 $(\beta,\gamma,\lambda)$ are up-to-$\Delta$ unique with gap $\delta \in (0,1)$ and either $\sqrt{\Delta} > \frac{\Delta - 2}{\lambda}$ or $\gamma \leq 1$, then $\Psi$ is an $(\alpha,c)$-potential.

**Lemma 10.** Let $\Delta \geq 3$ be an integer. Let $\beta,\gamma,\lambda$ be reals such that $0 \leq \beta \leq \gamma$, $\gamma > 0$, $\beta \gamma < 1$ and $\lambda > 0$. Assume that $(\beta,\gamma,\lambda)$ is up-to-$\Delta$ unique with gap $\delta \in (0,1)$. Define the function $\Psi$ implicitly by

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

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

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

We deduce Theorem 3 for the case $\sqrt{\Delta} > \frac{\Delta - 2}{\lambda}$ 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 $\sqrt{\Delta} \leq \frac{\Delta - 2}{\lambda}$ and $\gamma > 1$ is trickier. As discussed in Section 5 of [12], when $\sqrt{\Delta} \leq \frac{\Delta - 2}{\lambda}$ and $\gamma > 1$, the threshold $\lambda_c$ for the external field is bounded below by some constant for all $\Delta$; so for some $\lambda > 0$ the spin system is in the uniqueness region for arbitrary graphs, even with unbounded degrees (i.e., up-to-$\infty$ 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 $\lambda$, 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 [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 $\sigma_{\Lambda} \in \{0, 1\}^\Lambda$, $Z^{\sigma_{\Lambda}}_G$ divides $Z^{\sigma_{\Lambda}}_T$. More precisely, there exists a polynomial $P^{\sigma_{\Lambda}}_{G,r}$ such that

$$Z^{\sigma_{\Lambda}}_T = Z^{\sigma_{\Lambda}}_G \cdot P^{\sigma_{\Lambda}}_{G,r}.$$ 

Moreover, the polynomial $P^{\sigma_{\Lambda}}_{G,r}$ is independent of $\lambda_r$.

**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 $P^{\sigma_{\Lambda}}_{G,r}$ is independent of the field $\lambda_r$ 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 $M^{\sigma_{\Lambda}}_G(u) = \mu_G(v \upharpoonright \sigma_{\Lambda})$ (henceforth we write $v = i$ for the event $\sigma_v = i$ for convenience), and the covariance of $u$ and $v$ as

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

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 $v \in V$,
   $$\left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) \log Z^{\sigma_{\Lambda}}_G = M^{\sigma_{\Lambda}}_G(v);$$
2. For all $u, v \in V$,
   $$\left(\lambda_u \frac{\partial}{\partial \lambda_u}\right) \left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) \log Z^{\sigma_{\Lambda}}_G = K^{\sigma_{\Lambda}}_G(u, v);$$
3. For all $u, v \in V$,
   $$\left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) \log R^{\sigma_{\Lambda}}_G(u) = I^{\sigma_{\Lambda}}_G(u \to v).$$

**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

$$\left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) \log R^{\sigma_{\Lambda}}_G(u) = \left(\lambda_v \frac{\partial}{\partial \lambda_v}\right) \log \left(\frac{M^{\sigma_{\Lambda}}_G(u)}{1 - M^{\sigma_{\Lambda}}_G(u)}\right) = \frac{M^{\sigma_{\Lambda}}_G(u)}{M^{\sigma_{\Lambda}}_G(u) - 1} = \frac{K^{\sigma_{\Lambda}}_G(u, v)}{K^{\sigma_{\Lambda}}_G(u, v)}.$$ 

It remains to show that

$$I^{\sigma_{\Lambda}}_G(u \to v) = \frac{K^{\sigma_{\Lambda}}_G(u, v)}{K^{\sigma_{\Lambda}}_G(u, u)},$$

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

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

Meanwhile, the covariance can be written as

$$K^{\sigma_{\Lambda}}_G(u, v) = \mu_G(u = 1, v = 1 \mid \sigma_{\Lambda}) \cdot \mu_G(u = 1, v = 0 \mid \sigma_{\Lambda}) \cdot \mu_G(v = 1 \mid \sigma_{\Lambda}) \cdot \mu_G(v = 0 \mid \sigma_{\Lambda}) - \mu_G(u = 1, v = 0 \mid \sigma_{\Lambda}) \cdot \mu_G(u = 0, v = 1 \mid \sigma_{\Lambda}).$$

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

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 $v$:

$$M_G^\sigma(r) = M_T^\sigma(r), \quad R_G^\sigma(r) = R_T^\sigma(r);$$

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

$$K_G^\sigma(r, v) = \sum_{\hat{v} \in C_v} K_T^\sigma(r, \hat{v}),$$

$$I_G^\sigma(r \rightarrow v) = \sum_{\hat{v} \in C_v} I_T^\sigma(r \rightarrow \hat{v}).$$

where $C_v$ is the set of all free copies of $v$ in $T$.

Proof: By Theorem 11, there exists a polynomial $P_{G,r}^\sigma = P_{G,r}^\sigma(\lambda)$ such that $Z_{G,r}^\sigma = Z_{G}^\sigma \cdot P_{G,r}^\sigma$ and $P_{G,r}^\sigma$ is independent of $\lambda_r$. Then it follows from Lemma 12 that

$$M_T^\sigma(r) = \left( \lambda_r \frac{\partial}{\partial \lambda_r} \right) \log Z_T^\sigma = \left( \lambda_r \frac{\partial}{\partial \lambda_r} \right) \log Z_G^\sigma + \log P_{G,r}^\sigma = \left( \lambda_r \frac{\partial}{\partial \lambda_r} \right) M_T^\sigma(r),$$

and therefore $R_T^\sigma(r) = R_G^\sigma(r)$. For the second item, again from Lemma 12 we get

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

Recall that for the spin system on the SAW tree $T$, every free copy $\hat{v}$ of $v$ from $C_v$ has the same external field $\lambda_{\hat{v}} = \lambda_v$. Then, by the chain rule of derivatives and Lemma 12, we deduce that

$$K_G^\sigma(r, v) = \sum_{\hat{v} \in C_v} \left( \lambda_v \frac{\partial}{\partial \lambda_v} \right) M_G^\sigma(r) \frac{\partial \lambda_{\hat{v}}}{\partial \lambda_v} \frac{\lambda_v}{\lambda_{\hat{v}}} = \sum_{\hat{v} \in C_v} K_T^\sigma(r, \hat{v}).$$

Finally, we have

$$I_G^\sigma(r \rightarrow v) = \left( \lambda_v \frac{\partial}{\partial \lambda_v} \right) \log R_G^\sigma(r) = \left( \lambda_v \frac{\partial}{\partial \lambda_v} \right) \log \sum_{\hat{v} \in C_v} I_T^\sigma(r \rightarrow \hat{v}),$$

where the last equality follows as above.

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 $\Lambda = \{ \lambda_v : v \in V \}$. For $\Lambda \subseteq V$ and $\sigma_{\Lambda} \in \{0,1\}^\Lambda$, the weight of $\sigma \in \{0,1\}^{V \setminus \Lambda}$ conditional on $\sigma_{\Lambda}$ is given by

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

where for $i = 0, 1$, $m_i(\cdot | \sigma_{\Lambda})$ 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 $\sigma_{\Lambda}$ is defined as $Z_G^\sigma = \sum_{\sigma \in \{0,1\}^{V \setminus \Lambda}} w_G(\sigma | \sigma_{\Lambda})$. 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 $\sigma_{\Lambda}$ on the SAW tree, we also remove all descendants of $\hat{v} \in C_v$ 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 $Z_G^\sigma(v = i)$ be sum of weights of all configurations with $v = i$. We further extend this notation and write $Z_G^\sigma(U = \sigma_U)$ 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 $P_{G,r}^\sigma = P_{G,r}^\sigma(\lambda)$, independent of $\lambda_r$, such that

$$Z_T^\sigma(r = 1) = Z_G^\sigma(r = 1) \cdot P_{G,r}^\sigma,$$

$$Z_T^\sigma(r = 0) = Z_G^\sigma(r = 0) \cdot P_{G,r}^\sigma. \quad (5)$$

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 $v_1, \ldots, v_d$. Define $G'$ to be the graph obtained by replacing the vertex $r$ with $d$ vertices $r_1, \ldots, r_d$ and then connecting $\{r_i, d_i\}$ for $1 \leq i \leq d$.

Consider first the case where $(G', \{r\}) \setminus \Lambda$ is still connected. For each $i$, let $G_i = G' - r_i$. Define the 2-spin system on $G_i$ with the same parameters $(\beta, \gamma, \lambda)$, plus an additional conditioning that the vertices $r_1, \ldots, r_{i-1}$ are fixed to spin 0 while $r_{i+1}, \ldots, r_d$ are fixed to spin 1; we denote this conditioning by $\sigma_{U_i}$, with $U_i = \{v_1, \ldots, v_d\} \setminus \{v_i\}$. Then, $T = T_{SAW}(G, r)$ can be generated by the following recursive procedure.

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

1) For each $i$, let $T_i = T_{SAW}(G_i, v_i)$ plus the conditioning $\sigma_{U_i};$

2) Let $T = T_{SAW}(G, r)$ be the union of $r$ and $T_1, \ldots, T_d$ 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 $(\beta, \gamma, \lambda)$, with an exception that we let the vertices $r_1, \ldots, r_d$ have no fields (i.e., setting $\lambda_{r_i} = 1$ for $1 \leq i \leq d$ instead of $\lambda_r$). We then observe that

$$Z_G^\sigma(r = 1) = \lambda_r \cdot Z_G^\sigma(r_1 = 1, \ldots, r_d = 1),$$

and the same holds with spin 1 replaced by 0. For $1 \leq i \leq d$, let $\sigma_{A_i}$ denote the union of the conditioning $\sigma_{\Lambda}$ and $\sigma_{U_i},$
where \( \Lambda_i = \Lambda \cup U_i \). Then for every \( 1 \leq i \leq d \) we have
\[
Z_{G_i}^{\sigma_i}(r_1 = 0, \ldots, r_{i-1} = 0, r_i = 1, \ldots, r_d = 1) = \beta \cdot Z_{G_i}^{\sigma_i}(v_i = 1) + Z_{G_i}^{\sigma_i}(v_i = 0).
\]
Notice that both sides are independent of the field \( \lambda_r \); for the left side, all \( r_i \)'s do not have a field for the spin system on \( G^r \); 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 \( Q_{G_i}^{\Lambda} = Q_{G_i,r}^{\Lambda}(\lambda) \) by
\[
Q_{G_i,r}^{\Lambda} = \prod_{i=1}^{d} Z_{G_i}^{\sigma_i}(r_1 = 0, \ldots, r_{i-1} = 0, r_i = 1, \ldots, r_d = 1),
\]
which is independent of \( \lambda_r \). Then we get
\[
Z_{G}^{\sigma}(r = 1) = \lambda_r \prod_{i=1}^{d} Z_{G_i}^{\sigma}(r_1 = 0, \ldots, r_{i-1} = 0, r_i = 1, \ldots, r_d = 1)
= \lambda_r \cdot \prod_{i=1}^{d} \left( \beta \cdot Z_{G_i}^{\sigma}(v_i = 1) + Z_{G_i}^{\sigma}(v_i = 0) \right).
\]
Using a similar argument, we also have
\[
Z_{G}^{\sigma}(r = 0) \cdot Q_{G}^{\Lambda} = \prod_{i=1}^{d} Z_{G_i}^{\sigma}(r_1 = 0, \ldots, r_{i-1} = 0, r_i = 1, \ldots, r_d = 1)
= \prod_{i=1}^{d} \left( Z_{G_i}^{\sigma}(v_i = 1) + \gamma \cdot Z_{G_i}^{\sigma}(v_i = 0) \right).
\]
Since we assume that \((G\setminus \{\{\}) \backslash \Lambda \) is connected, the graph \( G \setminus \Lambda \) is also connected for each \( i \). Then, by the induction hypothesis, for each \( i \) there exists a polynomial \( P_{G_i,v_i} = P_{G_i,v_i}^{\sigma_i}(\lambda) \) such that
\[
Z_{T_i}^{\sigma_i}(r = 1) = Z_{G_i}^{\sigma_i}(r = 1) \cdot P_{G_i,v_i}^{\sigma_i},
Z_{T_i}^{\sigma_i}(r = 0) = Z_{G_i}^{\sigma_i}(r = 0) \cdot P_{G_i,v_i}^{\sigma_i};
\]
these polynomials are independent of \( \lambda_r \) since the conditional partition functions for \( G_i \)'s do not involve \( \lambda_r \). Now if we let
\[
P_{G,r}^{\sigma} = Q_{G,r}^{\Lambda} \cdot \prod_{i=1}^{d} P_{G_i,v_i}^{\sigma_i},
\]
then it follows from the tree recursion that
\[
Z_{G}^{\sigma}(r = 1) = \lambda_r \cdot \prod_{i=1}^{d} (\beta \cdot Z_{T_i}^{\sigma_i}(v_i = 1) + Z_{T_i}^{\sigma_i}(v_i = 0))
= \lambda_r \cdot \prod_{i=1}^{d} \left( \beta \cdot Z_{G_i}^{\sigma}(v_i = 1) P_{G_i,v_i}^{\sigma_i} + Z_{G_i}^{\sigma}(v_i = 0) P_{G_i,v_i}^{\sigma_i} \right)
= Z_{G}^{\sigma}(r = 1) \cdot Q_{G,r}^{\Lambda} \cdot \prod_{i=1}^{d} P_{G_i,v_i}^{\sigma_i}
= Z_{G}^{\sigma}(r = 1) \cdot P_{G,r}^{\sigma}.
\]
The other equality \( Z_{G}^{\sigma}(r = 0) = Z_{G}^{\sigma}(r = 0) \cdot P_{G,r}^{\sigma} \) is established in the same way. This completes the proof for the case that \((G\setminus \{\{\}) \backslash \Lambda \) is connected.

If \((G\setminus \{\{\}) \backslash \Lambda \) has two or more connected components, then we can construct \( T_{\text{SAW}}(G, r) \) by the SAW tree of each component. Recall that \( G^r \) 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 \( G_{(1)}, \ldots, G_{(k)} \) be the subgraphs induced by each component, along with vertices from \( \Lambda \) that are adjacent to \( r \). 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 \( r_{(1)}, \ldots, r_{(k)} \) of \( T_{(1)}, \ldots, T_{(k)} \), the resulting tree is \( T_{\text{SAW}}(G, r) \).

We define the 2-spin system on each \( G_{(j)} \) with the same parameters \((\beta, \gamma, \lambda)\), except that the vertex \( r_{(j)} \) does not have a field (i.e., \( \lambda_{r_{(j)}} = 1 \) instead of \( \lambda_r \)). For \( 1 \leq j \leq k \), let \( \Lambda_{(j)} = \Lambda \cap V(G_{(j)}) \) and \( \sigma_{(j)} = \Lambda_{(j)} \) be the configuration \( \sigma \) restricted on \( \Lambda_{(j)} \). Then \( G_{(j)} \backslash \Lambda_{(j)} \) is connected for every \( j \), and, since \( k \geq 2 \), each \( G_{(j)} \) with conditioning \( \sigma_{(j)} \) has fewer than \( m \) edges. Thus, we can apply the induction hypothesis; namely, for \( 1 \leq j \leq k \) there exists a polynomial \( P_{G_{(j)},r_{(j)}}^{\sigma_{(j)}} = P_{G_{(j)},r_{(j)}}^{\sigma_{(j)}}(\lambda) \), which is independent of \( \lambda_r \), such that
\[
Z_{T_{(j)}}^{\sigma_{(j)}}(r_{(j)} = 1) = Z_{G_{(j)}}^{\sigma_{(j)}}(r_{(j)} = 1) \cdot P_{G_{(j)},r_{(j)}}^{\sigma_{(j)}},
Z_{T_{(j)}}^{\sigma_{(j)}}(r_{(j)} = 0) = Z_{G_{(j)}}^{\sigma_{(j)}}(r_{(j)} = 0) \cdot P_{G_{(j)},r_{(j)}}^{\sigma_{(j)}}.
\]
We define the polynomial \( P_{G,r}^{\sigma} = P_{G,r}^{\sigma}(\lambda) \) to be
\[
P_{G,r}^{\sigma} = \prod_{j=1}^{k} P_{G_{(j)},r_{(j)}}^{\sigma_{(j)}}.
\]
It is then easy to check that
\[
Z_{G}^{\sigma}(r = 1) = \lambda_r \cdot \prod_{j=1}^{k} Z_{G_{(j)}}^{\sigma_{(j)}}(r_{(j)} = 1)
= \lambda_r \cdot \prod_{j=1}^{k} \left( Z_{G_{(j)}}^{\sigma_{(j)}}(r_{(j)} = 1) \cdot P_{G_{(j)},r_{(j)}}^{\sigma_{(j)}} \right)
= Z_{G}^{\sigma}(r = 1) \cdot \prod_{j=1}^{k} P_{G_{(j)},r_{(j)}}^{\sigma_{(j)}} = Z_{G}^{\sigma}(r = 1) \cdot P_{G,r}^{\sigma};
\]
and similarly \( Z_{G}^{\sigma}(r = 0) = Z_{G}^{\sigma}(r = 0) \cdot P_{G,r}^{\sigma} \). The theorem then follows.

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 \( \Psi \), our upper bound is in terms of properties of \( \Psi \), involving
bounds on $\| \nabla H_d^y \|_1$ and $|\psi|$ where $\psi = \Psi'$. We then deduce Lemma 9 in the case that $\Psi$ an $(\alpha, c)$-potential.

Assume that $T = (V_T, E_T)$ is a tree rooted at $r$ of maximum degree at most $\Delta$. 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 $(\beta, \gamma, \lambda)$, conditioned on $\sigma_\Lambda$. We need to bound the influence $\mathcal{I}_T^\Lambda (r \to v)$ from the root $r$ to another vertex $v \in V_T$. Notice that if $v$ is disconnected from $r$ when $\Lambda$ is removed, then $\mathcal{I}_T^\Lambda (r \to v) = 0$ by the Markov property of spin systems. Therefore, we may assume, that, by removing all such vertices, $\Lambda$ contains only leaves of $T$.

For a vertex $v \in V_T$, let $T_v = (V_{T_v}, E_{T_v} \Lambda)$ be the subtree of $T$ rooted at $v$ that contains all descendant of $v$; note that $T_v = T$. 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 $R_v = R_{T_v}^\Lambda (v)$ for simplicity. The log $R_v$’s are related by the tree recursion $H$. If a vertex $v$ has $d$ children, denoted by $u_1, \ldots, u_d$, then the tree recursion is given by

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

where for $1 \leq d \leq \Delta$ and $(y_1, \ldots, y_d) \in [-\infty, \infty]^d$,

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

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

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

and $\frac{\partial}{\partial \lambda} H_d (y_1, \ldots, y_d) = h(y_i)$ 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 $\Psi : [-\infty, +\infty] \to (-\infty, +\infty)$ be a differentiable and increasing (potential) function with image $S = \Psi([-\infty, +\infty])$ and derivative $\psi = \Psi'$. Denote the degree of the root $v$ by $\Delta_r$. Then for every integer $k \geq 1$,

$$\sum_{v \in L_r(k)} |\mathcal{I}_{T_v}^\Lambda (r \to v)| \leq \Delta_r A_y B_y \left( \max_{1 \leq d \leq \Delta} \sup_{y \in S} \| \nabla H_d^y (\tilde{y}) \|_1 \right)^{k-1},$$

where

$$A_y = \max_{u \in L_r(1)} \left\{ \frac{|h(\log R_u)|}{\psi(\log R_u)} \right\},$$

$$B_y = \max_{u \in L_r(k)} \left\{ \psi(\log R_u) \right\}.$$
Hence, we obtain that

$$\sum_{v \in L_u(k)} |T^\alpha_T (r \to v)| \leq \Delta_r \cdot \max_{1 \leq i \leq \Delta_r} \left\{ \frac{h(\log R_w)}{\psi(\log R_w)} \right\}$$

(6)

$$\times \max_{1 \leq i \leq \Delta_r} \left\{ \sum_{v \in L_{w_i}(k-1)} \psi(\log R_w) \cdot |T^{\alpha_T}_T (u \to v)| \right\}.$$ 

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

$$\sum_{v \in L_u(k)} \psi(\log R_w) \cdot |T^{\alpha_T}_T (u \to v)|$$

(7)

$$\leq \max_{v \in L_u(k)} \{\psi(\log R_w)\} \cdot \left( \max_{u \in V_L} \sup_{\psi \in [0,1]} \|\nabla H^\alpha_{d,w}(\psi)\|_1 \right)^k.$$ 

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

$$\psi(\log R_w) \cdot |T^{\alpha_T}_T (u \to v)| \leq \psi(\log R_w),$$

which holds with equality since $T^{\alpha_T}_T (u \to v) = 1$. New 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 $w_1, \ldots, w_d$, 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

$$\sum_{v \in L_u(k)} \psi(\log R_w) \cdot |T^{\alpha_T}_T (u \to v)|$$

$$= \sum_{i=1}^d \psi(\log R_w) \cdot |T^{\alpha_T}_T (u \to w_i)| \sum_{v \in L_{w_i}(k-1)} |T^{\alpha_T}_T (w_i \to v)|$$

$$= \sum_{i=1}^d \psi(\log R_w) \cdot \psi(\log R_w) \cdot |h(\log R_w)|$$

$$\times \sum_{v \in L_{w_i}(k-1)} \psi(\log R_w) \cdot |T^{\alpha_T}_T (w_i \to v)|.$$ 

For $U \subseteq V_T$, we let

$$\Xi(U) = \max_{w \in U} \sup_{\psi \in [0,1]} \|\nabla H^\alpha_{d,w}(\psi)\|_1.$$ 

Using the induction hypothesis, we get

$$\sum_{v \in L_u(k)} \psi(\log R_w) \cdot |T^{\alpha_T}_T (u \to v)|$$

$$\leq \max_{v \in L_u(k)} \{\psi(\log R_w)\} \cdot \Xi(V_T \setminus \{u\})^{k-1}$$

$$\times \sum_{i=1}^d \psi(\log R_w) \cdot |h(\log R_w)|$$

$$\leq \max_{v \in L_u(k)} \{\psi(\log R_w)\} \cdot \Xi(V_T)^k,$$

where the last inequality follows from that

$$\sum_{i=1}^d \psi(\log R_w) \cdot |h(\log R_w)|$$

$$= \sum_{i=1}^d \frac{\partial}{\partial \psi(\log R_w)} H^\alpha_{d,w} (\Psi(\log R_w))$$

$$= \|\nabla H^\alpha_{d,w} (\Psi(\log R_w))\|_1.$$ 

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

We then derive Lemma 9 as a corollary.

Proof of Lemma 9: Since $\Psi$ is an $(\alpha, c)$-potential, the Contraction condition implies that

$$\max_{1 \leq d < \Delta} \sup_{\psi \in [0,1]} \|\nabla H^\alpha_{d,w}(\psi)\|_1 \leq 1 - \alpha.$$ 

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

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

Therefore, we get $\Delta_r, A_c B_q \leq c$. The lemma then follows immediately from Lemma 14.

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 $\beta, \gamma, \lambda$ and $0 < \delta < 1$, and assume $(\beta, \gamma, \lambda)$ satisfies one of the following three conditions:

1) $\frac{\Delta - 2 + \delta}{\Delta - 3} \leq \sqrt{\beta} \gamma \leq \frac{\Delta - 2 - \delta}{\Delta - 3},$ and $\lambda$ is arbitrary;

2) $\lambda \geq \frac{\Delta - 2}{\Delta - 2 - \lambda} \cdot \frac{\max\{1, \beta \Delta - 1\}}{(\lambda - 2)\beta - \Delta};$

3) $\lambda \geq \frac{\Delta - 2}{\Delta - 2 - \lambda} \cdot \frac{(\lambda - 2)\beta - \Delta}{\beta \min(1, \gamma^{\Delta - 1})}.$

Then the identity function $\Psi(y) = y$ (based on the potential given in [18]) is an $(\alpha, c)$-potential function for $\alpha = O(1)$ and $\lambda \leq O(1)$. Furthermore, for every n-vertex graph $G$ of maximum degree at most $\Delta$, the mixing time of the Glauber dynamics for the $2$-spin system on $G$ with parameters $(\beta, \gamma, \lambda)$ is $O(n^{2+c/\delta})$, for a universal constant $c > 0$.

Remark 3. Condition 1 includes both the ferromagnetic case $1 < \sqrt{\beta} \gamma < \frac{\Delta - 2}{\Delta - 3}$ and the antiferromagnetic case $\frac{\Delta - 2}{\Delta - 3} \leq \sqrt{\beta} \gamma < 1.$ Note that in both cases $(\beta, \gamma, \lambda)$ is up-to-$\Delta$ unique with gap $\delta.$ For the antiferromagnetic case, the identity function $\Psi$ is an $(\alpha, c)$-potential with $c \leq 1.5$ and a better contraction rate $\alpha = \delta$, compared with the bound $\alpha \geq \delta/2$ of the potential $\Psi$ 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 $(\alpha, c)$-potential, but $c$ must, unfortunately, depend on $\Delta$. 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 $\beta, \gamma, \lambda$ satisfying $\beta \leq 1 \leq \gamma$, $\sqrt{\beta \gamma} \geq \frac{\Delta}{2} - 2$, and $\lambda < \left( \frac{3}{2} \right)^{\frac{\sqrt{\beta \gamma}}{\Delta - \frac{5}{4}}}$. Then for every $n$-vertex graph $G$ with maximum degree at most $\Delta$, the mixing time of the Glauber dynamics for the ferromagnetic 2-spin system on $G$ with parameters $(\beta, \gamma, \lambda)$ 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**


