Correlation Decay up to Uniqueness in Spin Systems

Liang Li [*] [†]
Peking University
liang.li@pku.edu.cn

Pinyan Lu
Microsoft Research Asia
pinyanl@microsoft.com

Yitong Yin[‡]
Nanjing University
yinyt@nju.edu.cn

Abstract

We give a complete characterization of the two-state anti-ferromagnetic spin systems which exhibit strong spatial mixing on general graphs. We show that a two-state anti-ferromagnetic spin system exhibits strong spatial mixing on all graphs of maximum degree at most Δ if and only if the system has a unique Gibbs measure on infinite regular trees of degree up to Δ, where Δ can be either bounded or unbounded. As a consequence, there exists an FPTAS for the partition function of a two-state anti-ferromagnetic spin system on graphs of maximum degree at most Δ when the uniqueness condition is satisfied on infinite regular trees of degree up to Δ. In particular, an FPTAS exists for arbitrary graphs if the uniqueness is satisfied on all infinite regular trees. This covers as special cases all previous algorithmic results for two-state anti-ferromagnetic systems on general-structure graphs.

Combining with the FPRAS for two-state ferromagnetic spin systems of Jerrum-Sinclair and Goldberg-Jerrum-Paterson, and the very recent hardness results of Sly-Sun and independently of Galanis-Štefankovič-Vigoda, this gives a complete classification, except at the phase transition boundary, of the approximability of all two-state spin systems, on either degree-bounded families of graphs or family of all graphs.

1 Introduction

Spin systems are well studied in the areas of Statistical Physics, Applied Probability and Computer Science as a general framework to capture the essence of how local interactions and constrains affect the macroscopic properties of particle systems. A system is usually described by a graph, with each vertex in one of a fixed number of states called spins, and edges specifying the neighborhood relation of the system.

Let $G(V, E)$ be a graph and q be the number of spin states. A configuration of the system is one of the $q^{|V|}$ possible assignments $\sigma : V \to [q]$. Each configuration σ has an energy $H(\sigma)$ as a sum over all edges and vertices, such that the contribution of an edge $(u, v) \in E$ is determined by a symmetric function of the spin states $\sigma(u)$ and $\sigma(v)$, and the contribution of a vertex $v \in V$ is determined by a function of its spin state $\sigma(v)$. The weight of a configuration σ is $w(\sigma) = \exp(-\frac{H(\sigma)}{T})$, where T is the temperature. We focus on the two-state spin systems. Up to normalization, a two-state spin system is fully captured by three parameters (β, γ, λ), where β and γ determine the symmetric function for edge contribution and λ, also known as the *external field*, determines the function for vertex contribution. The Gibbs measure is a natural probability distribution over all configurations such that the probability of a configuration σ is $\frac{w(\sigma)}{Z}$, where the normalizing factor $Z = \sum w(\sigma)$ is called the partition function. The partition function encodes rich information regarding the macroscopic behavior of the spin system. However, for almost all nontrivial settings it is #P-hard to compute the precise value of partition functions.

One of the most important properties of spin systems is the correlation decay, which says that the correlation between the marginal Gibbs distributions of two vertices decays rapidly with respect to the distance between the two vertices. This property is also called *weak spatial mixing* [25]. Of greater algorithmic significance is the strong spatial mixing, which says that the correlations decay in the presence of arbitrary fixed spins at other vertices. For two-state spin systems, the strong spatial mixing may imply efficient approximation algorithms for the partition function. It is then an important question to characterize the systems which exhibit strong spatial mixing on arbitrary instances of graphs by the parameters of the systems.

A two-state spin system is ferromagnetic if adjacent vertices favor agreement of spins, and is anti-ferromagnetic if otherwise. For all two-state ferromagnetic spin systems, the partition functions can be efficiently approximated [10, 12]. In the anti-ferromagnetic region, the correlation decay plays a central role in the approximability of partition functions. It is believed (see [18]) that for such models the approximability of the partition function is characterized by the uniqueness of Gibbs measure on infinite regular trees, which is equivalent to the weak spatial mixing on the infinite regular trees. This condition is called the *uniqueness condition.*

In a seminal work [25], Weitz shows that for the hardcore model strong spatial mixing is characterized by the uniqueness condition. Restrepo *et al.* [19] gave a new proof of the same result by an elegant message contraction argument (or equivalently the potential method). This is the first time a potential-based analysis is combined with correlation decay techniques to get tight bounds for approximate counting. In a recent work of Sinclair, Srivastava, and Thurley [20] the same tight bound is proved for the anti-ferromagnetic Ising by the potential method. On the hardness side, it is proved for the hardcore model by Sly [22] and Galanis *et al.* [6], and very recently for the general two-state anti-ferromagnetic spin systems by Sly and Sun [23] and independently by Galanis, Štefankovič, and Vigoda [7] that violating the uniqueness condition implies the inapproximability of partition functions. Two questions remain open for our complete understanding of the correlation decay and computation in two-state spin systems: the characterizations of the strong spatial mixing and approximability of general two-state spin systems.

1.1 Our results We characterize the two-state anti-ferromagnetic spin systems which exhibit strong spatial mixing on general degree-bounded graphs or arbitrary graphs by the uniqueness of Gibbs measure on infinite regular trees.

THEOREM 1.1. *For any finite $\Delta \geq 2$ or $\Delta = \infty$, a two-state anti-ferromagnetic spin system exhibits strong spatial mixing on graphs of maximum degree at most Δ if and only if the system exhibits uniqueness on infinite d-regular trees for all $d \leq \Delta$.*

Due to Weitz's self-avoiding walk tree construction [25], the strong spatial mixing of a two-state spin system on degree-bounded graphs immediately implies an FPTAS for the partition function. Indeed, we show an even stronger notion of correlation decay introduced in a previous work [15], namely the computationally ef-

ficient correlation decay, which gives FPTAS not only for the degree-bounded graphs but also for arbitrary graphs with unbounded degrees, when the corresponding uniqueness condition is satisfied.

THEOREM 1.2. *For any finite $\Delta \geq 3$ or $\Delta = \infty$, there exists an FPTAS for the partition function of the two-state anti-ferromagnetic spin system on graphs of maximum degree at most Δ if for all $d \leq \Delta$ the system parameters lie in the interior of the uniqueness region of infinite d-regular tree.*

In the above two theorems, the $\Delta = \infty$ case represents the graphs of unbounded degree.

Due to a very recent hardness results of Sly and Sun [23] for general two-state spin systems and an independent result of Galanis, Štefankovič and Vigoda [7] for a less general setting, violating the uniqueness condition implies inapproximability of the partition function.

THEOREM 1.3. ([23] AND [7]) *For any finite $\Delta \geq 3$ or $\Delta = \infty$, unless $NP = RP$ there does not exist an FPRAS for the partition function of the two-state anti-ferromagnetic spin system on graphs of maximum degree at most Δ if for some $d \leq \Delta$ the system parameters lie in the interior of the non-uniqueness region of infinite d-regular tree.*

The original theorem in [23] holds for d-regular graphs with fixed d, which immediately implies the hardness for degree-bounded graphs or arbitrary graphs. And the hardness condition in both [23] and [7] requires the non-uniqueness on a d-regular tree with $d \geq 3$. But the uniqueness on the infinite 2-regular tree (i.e. the infinite path) always holds for any two-state anti-ferromagnetic spin system, thus the condition in Theorem 1.3 suffices.

For graphs of maximum degree 2 or less, the partition function can be computed exactly in polynomial time. Thus Theorem 1.2 and Theorem 1.3 together with the FPRAS for two-state ferromagnetic spin system [10, 12] give an almost complete (except at the phase transition boundary) classification of the approximability of the partition function of all two-state spin systems, on either degree-bounded families of graphs or family of all graphs.

1.1.1 Regularity and monotonicity In Statistical Physics, the correlation decay is usually studied on regular graphs or even structurally symmetric graphs (e.g. Bethe lattice). Although for hardness results considering only regular graphs will strengthen the lower bounds, from algorithmic perspective it is more

general to consider spin systems on general graphs. The approximation algorithms in [25] for the hardcore model and [20] for the anti-ferromagnetic Ising model are both for general graphs with bounded maximum degrees. As discussed in [20, 23], up to translation of parameters the hardcore and Ising models together are complete for general two-state spin systems on d-regular graphs with fixed d. However, this does not cover the most general case, namely the general two-state spin system on general graphs, of either bounded or unbounded degree. The result in [20](Corollary 1.3) also gave one sufficient condition for correlation decay for general two-state spin system on general graphs with maximum degree Δ. But this condition is not tight. A fundamental reason for this gap is the non-monotonicity of general spin systems.

The well studied hardcore and anti-ferromagnetic Ising models (along with all two-state spin systems with $\beta, \gamma \leq 1$) are both monotone spin systems, in the sense that the uniqueness on infinite d-regular tree implies the uniqueness on all infinite regular trees with smaller degrees. This monotonicity does not necessarily hold in general two-state spin systems.

In [25], Weitz established the following implication in the hardcore model:

THEOREM 2.3 in [25]. *Strong spatial mixing on a d-regular tree implies the strong spatial mixing on all graphs of maximum degree at most d.*

In [25], Weitz also remarked without proof that this implication holds for all two-state spin systems (indeed this is rigorously proved for anti-ferromagnetic Ising model in [20]). With this to be true, devising approximation algorithms for two-state spin systems on degree-bounded graphs is reduced to verifying the strong spatial mixing on d-regular trees. This has been accepted as a fact about the two-state spin systems and has become a building block for approximation algorithms for such systems (see Theorem 2.8 in [20] and the discussions in [21, 22]). It was also raised as a conjecture in [21] whether the claim holds for general multi-state spin systems.

As a byproduct of our analysis (see Section 6), we find that this well-believed implication between the strong spatial mixing on d-regular tree and on graphs of maximum degree at most d holds only for monotone spin systems (including the hardcore and anti-ferromagnetic Ising models). For general two-state spin systems the worst case for uniqueness as well as strong spatial mixing among all degree-bounded graphs is indeed a regular tree, but may no longer be the one of the highest degree. A bright side of this complication is that higher degrees may yield much faster correlation decay, making

possible the FPTAS for graphs with unbounded degrees.

These new phenomena suggest that the general two-state spin systems have much richer structure than the well-studied monotone spin systems such as the hardcore and anti-ferromagnetic Ising models. The former approach via the strong spatial mixing on d-regular tree which succeeds in monotone spin systems on general graphs and general spin systems on regular graphs, meets a barrier when dealing with general spin systems on general graphs. We give a unified approach to the correlation decay in general two-state spin systems, through the strong spatial mixing on arbitrary trees instead of d-regular trees, proved by a potential analysis. The potential functions used in [19] and [20] are suitable only for monotone systems because the potential values depend on vertex degrees. And the potential function used in our previous work [15] can only approach a weaker uniqueness boundary up to continuous relaxation, because the definition of the function involves a continuous version of worst-case degree. In this paper, we devise a unified potential-based analysis which adapts to both the irregularity of the arbitrary tree and the non-monotonicity of general two-state spin system and give tight correlation decay results for all two-state anti-ferromagnetic spin systems on degree-bounded families of graphs and family of all graphs.

1.1.2 Implications of the main results By solving the uniqueness condition we can restate our main results in various threshold forms. Theorem 1.2 covers as special cases all previous algorithmic results for two-state anti-ferromagnetic spin systems on general-structure graphs as well as clears up previously uncovered cases.

In terms of interactions. We can fix the external field λ and discuss the tractable region of (β, γ). Since the roles of β and γ are symmetric, we can further fix one of them and discuss the tractable range of the other. This formulation was used in [10, 15].

Our main result can be restated as follows:

THEOREM 1.4. *For any $\beta \in [0, 1), \lambda > 0, \Delta \geq 3$, there is a critical threshold $\gamma_c(\beta, \lambda, \Delta)$ for the uniqueness on infinite regular trees up to degree Δ such that if $\gamma_c(\beta, \lambda, \Delta) < \gamma < \frac{1}{\beta}$ there is an FPTAS for graphs of maximum degree at most Δ; and in particular, $\gamma_c = \gamma_c(\beta, \lambda, \infty) > 1$ is an absolute constant such that if $\gamma \in (\gamma_c, \frac{1}{\beta})$ there is an FPTAS for arbitrary graphs.*

This covers as special cases all algorithmic results in [10] regarding the anti-ferromagnetic systems and all results in [15], extends the tractable regions in these previous works, and considers the degree-bounded

graphs as well.

In terms of external field. Motivated by the studies of hardcore and anti-ferromagnetic Ising models, we can fix (β, γ) and discuss the tractable range of external field.

Due to the symmetric role of β and γ, we may assume that $\beta \leq \gamma$. Our main results can be restated in specific settings as follows:

- **Hard constraints (when $\beta = 0$):** For any Δ, $\lambda_c(\gamma, \Delta) = \min_{1 < d < \Delta} \frac{\gamma^{d+1} d^d}{(d-1)^{d+1}}$ is a critical threshold for the uniqueness on infinite regular trees up to degree Δ such that if $\lambda < \lambda_c(\gamma, \Delta)$ there exists an FPTAS for graphs of maximum degree at most Δ.

  For $\gamma \leq 1$, the critical threshold equals $\lambda_c(\gamma, \Delta) = \frac{\gamma^\Delta (\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta}$. There is no external field $\lambda > 0$ satisfying uniqueness on infinite regular trees of unbounded degrees. This is consistent with the hardness result for the hardcore model without degree bound [4, 22]. One particularly interesting special case is when $\gamma = 1$, in which case the model is exactly the hardcore model with fugacity parameter λ, and $\lambda_c(1, \Delta) = \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta}$ is the critical fugacity of the uniqueness threshold. This covers the result of [25].

  For $\gamma > 1$, in addition to the results for degree-bounded graphs, there exists an absolute positive constant $\lambda(\gamma) = \min_{d>1} \frac{\gamma^{d+1} d^d}{(d-1)^{d+1}}$ which lower bounds $\lambda(\gamma, \Delta)$ for all Δ, such that if $\lambda < \lambda(\gamma)$ there exists an FPTAS for arbitrary graphs. This is quite different from the case $\gamma \leq 1$ and is not covered by any previous results.

- **Soft constraints (when $\beta > 0$):** For any Δ, $\lambda_c = \lambda_c(\beta, \gamma, \Delta)$ and $\bar{\lambda}_c = \bar{\lambda}_c(\beta, \gamma, \Delta)$ are two critical thresholds for the uniqueness on infinite regular trees up to degree Δ such that if $\lambda \in (0, \lambda_c) \cup (\bar{\lambda}_c, \infty)$ there exists an FPTAS for graphs of maximum degree at most Δ. In particular, if $\sqrt{\beta\gamma} > \frac{\Delta-2}{\Delta}$ the two ranges overlap, thus for any external field $\lambda > 0$ there exists an FPTAS for graphs of maximum degree at most Δ.

  When $\beta = \gamma$, the spin system becomes the anti-ferromagnetic Ising model. And $\lambda_c = 1/\bar{\lambda}_c$, thus if $\lambda \in (0, 1/\bar{\lambda}_c) \cup (\bar{\lambda}_c, \infty)$, i.e. $|\log \lambda| > \log \bar{\lambda}_c$, there exists an FPTAS for graphs of maximum degree at most Δ. This covers the result of anti-ferromagnetic Ising model in [20].

  For unbounded maximum degree Δ, if $\gamma \leq 1$, there is no external field $\lambda > 0$ satisfying uniqueness on

infinite regular trees of unbounded degrees. This is consistent with the hardness results for two-state spin systems in [10]. If $\gamma > 1$, $\lambda_c = \lambda_c(\beta, \gamma)$ is a critical threshold for uniqueness on all infinite regular trees such that if $\lambda < \lambda_c(\beta, \gamma)$ there exists an FPTAS for arbitrary graphs.

1.2 Related works The approximation of partition functions of spin systems has been extensively studied [3–5, 9, 11–13, 16, 24]. In a seminal work [12], Jerrum and Sinclair devised a fully polynomial-time randomized approximation scheme (FPRAS) for the ferromagnetic Ising model. Later in [10], the FPRAS was extended to all two-state ferromagnetic spin systems by translating the parameters to the ferromagnetic Ising model. Also in [10], Goldberg, Jerrum, and Paterson gave an FPRAS and inapproximability results for two-state anti-ferromagnetic spin systems on arbitrary graphs. A gadget based on random regular bipartite graphs was proposed by Dyer, Frieze, and Jerrum in [4] and was also analyzed by Mossel, Weitz, and Wormald in [18] to study the inapproximability on degree-bounded graphs. It is widely believed that the transition of approximability of anti-ferromagnetic spin systems is captured by the phase transition of uniqueness on infinite trees. This was raised openly as a conjecture in [18]. The conjecture was proved by Sly in [22] for the hardcore model. This result was improved by Galanis *et al.* in [6] to a wide range of parameters. Very recently, Sly and Sun [23] proved the hardness of all two-state anti-ferromagnetic spin systems of non-uniqueness on infinite regular trees. A same result for anti-ferromagnetic Ising model without external field was independently proved by Galanis, Štefankovič, and Vigoda in [7].

The correlation decay technique developed independently by Weitz [25] and Bandyopadhyay and Gamarnik [1] is a powerful tool for devising deterministic fully polynomial-time approximation schemes (FPTAS) for partition functions (other important examples include [2, 8]). In [25], Weitz introduced the concept of strong spatial mixing and used it to devise FPTAS for the hardcore model up to the uniqueness threshold. A powerful technique was developed by Restrepo *et al.* in [19] which makes use of the specific structure of graphs for strong spatial mixing. A broader tractable region than the region of uniqueness is achieved on grid lattice by exploiting the structure of the graph. The other most important two-state anti-ferromagnetic spin system, the anti-ferromagnetic Ising model, was studied recently by Sinclair, Srivastava, and Thurley in [20], where a more powerful message-decay method was introduced to analyze the strong spatial mixing and give FPTAS up to uniqueness threshold.

In a previous work [15], we gave an FPTAS for two-state anti-ferromagnetic spin systems without external field on arbitrary graphs with unbounded degrees, up to a continuous relaxation of the uniqueness threshold. The approach used in the current paper was initiated in [15], however the analysis in [15] cannot separate the uniqueness up to certain degree, thus fails in dealing with degree-bounded families of graphs.

2 Definitions and preliminaries

A two-state spin system is described by a graph $G = (V, E)$. A *configuration* of the system is one of the $2^{|V|}$ possible assignments $\sigma : V \to \{0, 1\}$ of states to vertices. We also use two colors *blue* and *green* to denote these two states. The *weight* of a configuration can be described as a product of contributions of individual edges and vertices. Let $A = \begin{bmatrix} A_{0,0} & A_{0,1} \\ A_{1,0} & A_{1,1} \end{bmatrix}$ and $b = (b_0, b_1)$. The weight of a configuration $\sigma : V \to \{0, 1\}$ is given by

$$w(\sigma) = \prod_{(u,v) \in E} A_{\sigma(u),\sigma(v)} \prod_{v \in V} b_{\sigma(v)}.$$

The *Gibbs measure* is a probability distribution over all configurations defined by $\rho(\sigma) = \frac{w(\sigma)}{Z(G)}$. The normalization factor $Z(G) = \sum_\sigma w(\sigma)$ is called the *partition function*.

We can normalize the contributions of a blue-green edge and of a green vertex to be 1. So that $A = \begin{bmatrix} \beta & 1 \\ 1 & \gamma \end{bmatrix}$ for some $\beta, \gamma \geq 0$, and $b = (b_0, b_1) = (\lambda, 1)$ for some $\lambda > 0$. Since the roles of blue and green are symmetric, we can assume that $\beta \leq \gamma$ without loss of generality. The three parameters (β, γ, λ) with $0 \leq \beta \leq \gamma$ and $\lambda > 0$ completely specify a two-state spin system. A two-state spin system with $\beta = \gamma$ is an *Ising model* and a two-state spin system with $\beta = 0, \gamma = 1$ or symmetrically $\beta = 1, \gamma = 0$ is a *hardcore model*.

A two-state spin system is called *anti-ferromagnetic* if adjacent vertices favor disagreeing spins, i.e. if $\beta\gamma < 1$. Without loss of generality, we focus on the cases that $\beta \leq \gamma$.

Definition 2.1. (β, γ, λ) *is* anti-ferromagnetic *if* $0 \leq \beta \leq \gamma$, $\gamma > 0$, $\beta\gamma < 1$, and $\lambda > 0$.

By the symmetry of β and γ and the triviality of the case $\beta = \gamma = 0$, this definition is complete for all nontrivial two-state anti-ferromagnetic systems

2.1 Correlation decay The Gibbs measure defines a marginal distribution of state for each vertex. Let p_v denote the probability of vertex v to be colored blue.

Let $\sigma_\Lambda \in \{0, 1\}^\Lambda$ be a configuration of $\Lambda \subset V$. We call vertices $v \in \Lambda$ *fixed* vertices, and $v \notin \Lambda$ *free* vertices. A σ_Λ is *feasible* if there exists a $\sigma \in \{0, 1\}^E$ with Gibbs measure $\rho(\sigma) > 0$ such that σ is consistent with σ_Λ on Λ. We use $p_v^{\sigma_\Lambda}$ to denote the marginal probability of v to be colored blue conditioning on the configuration of Λ being fixed as σ_Λ.

Definition 2.2. *A spin system on a family of graphs is said to be of* strong spatial mixing *if for any graph* $G = (V, E)$ *in the family, any* $v \in V, \Lambda \subset V$ *and any feasible* $\sigma_\Lambda, \tau_\Lambda \in \{0, 1\}^\Lambda$,

$$|p_v^{\sigma_\Lambda} - p_v^{\tau_\Lambda}| \leq \exp(-\Omega(\mathrm{dist}(v, S))),$$

where $S \subset \Lambda$ *is the subset on which* σ_Λ *and* τ_Λ *differ, and* $\mathrm{dist}(v, S)$ *is the shortest distance from* v *to any vertex in* S.

The *weak spatial mixing* can be defined by measuring the decay with respect to $\mathrm{dist}(v, \Lambda)$ instead of $\mathrm{dist}(v, S)$. The spatial mixing property is also called correlation decay in Statistical Physics.

Let T be a tree rooted by v. Define $R_T^{\sigma_\Lambda} = p_v^{\sigma_\Lambda}/(1 - p_v^{\sigma_\Lambda})$ to be the ratio between the probabilities that the root v is blue and green, when imposing the condition σ_Λ (with the convention that $R_T^{\sigma_\Lambda} = \infty$ when $p_v^{\sigma_\Lambda} = 1$). Suppose that v has d children and T_i is the subtree rooted by the i-th child. Due to the independence of subtrees, we have an easy recursion for calculating $R_T^{\sigma_\Lambda}$:

$$(2.1) \qquad R_T^{\sigma_\Lambda} = \lambda \prod_{i=1}^{d} \frac{\beta R_{T_i}^{\sigma_\Lambda} + 1}{R_{T_i}^{\sigma_\Lambda} + \gamma}.$$

Let $G(V, E)$ be a graph. Similarly define that $R_{G,v}^{\sigma_\Lambda} = p_v^{\sigma_\Lambda}/(1 - p_v^{\sigma_\Lambda})$. In contrast to the case of tree, there is no easy recursion for calculating $R_{G,v}^{\sigma_\Lambda}$ for a general graph G because of the dependencies caused by cycles. In [25], a construction called the self-avoiding walk (SAW) tree was introduced which reduces the computing of marginal distribution in a general graph to that in a tree. Specifically, given a graph $G(V, E)$ and a vertex $v \in V$. The SAW tree $T_{\mathrm{SAW}}(G, v)$ is a tree rooted at v with a new vertex set V_{SAW} (which effectively enumerates all paths originating from v in G and may include fixed leaves). Moreover, any vertex sets $S \subset \Lambda \subset V$ are mapped to respective $S_{\mathrm{SAW}} \subset \Lambda_{\mathrm{SAW}} \subset V_{\mathrm{SAW}}$ and any configuration $\sigma_\Lambda \in \{0, 1\}^\Lambda$ is mapped to a $\sigma_{\Lambda_{\mathrm{SAW}}} \in \{0, 1\}^{\Lambda_{\mathrm{SAW}}}$. We abuse the notation and write $S = S_{\mathrm{SAW}}$ and $\sigma_\Lambda = \sigma_{\Lambda_{\mathrm{SAW}}}$ if no ambiguity is caused. Given a graph $G(V, E)$, $v \in V$ and $S \subset V$, let $\mathrm{dist}_G(v, S)$ be the shortest distance in G from v to any vertex in S.

THEOREM 2.1. (THEOREM 3.1 OF WEITZ [25]) *Let* $G(V, E)$ *be a graph,* $v \in V$, $\sigma_\Lambda \in \{0, 1\}^\Lambda$ *a configuration on* $\Lambda \subset V$, *and* $S \subset V$. *Let* $T = T_{\text{SAW}}(G, v)$. *Then* $R_{G,v}^{\sigma_\Lambda} = R_T^{\sigma_\Lambda}$. *Moreover, it holds that the maximum degree of* T *equals the maximum degree of* G, $\text{dist}_G(v, S) = \text{dist}_T(v, S)$, *and any neighborhood of* v *in* T *can be constructed in time proportional to the size of the neighborhood.*

2.2 The uniqueness condition We consider the uniqueness of Gibbs measure on the infinite $(d+1)$-regular trees $\widehat{\mathbb{T}}^d$, in which the recursion is given by $f_d(x) \triangleq \lambda \left(\frac{\beta x + 1}{x + \gamma} \right)^d$ due to the symmetric structure of $\widehat{\mathbb{T}}^d$.

Let $\hat{x}_d$ be the positive fixed point of $f_d(x)$, that is, $\hat{x}_d = f(\hat{x}_d)$. It is known [14, 17] that the two-state anti-ferromagnetic spin system on $\widehat{\mathbb{T}}^d$ undergoes a phase transition at $|f_d'(\hat{x}_d)| = 1$ with uniqueness when $|f_d'(\hat{x}_d)| \leq 1$. This motivates the following definition.

DEFINITION 2.3. *Let* (β, γ, λ) *be anti-ferromagnetic. Let* $\hat{x}_d$ *be the positive fixed point of function* $f_d(x) = \lambda \left(\frac{\beta x + 1}{x + \gamma} \right)^d$. *We say that* (β, γ, λ) *is* up-to-Δ unique, *if for all integer* $1 \leq d < \Delta$,

$$|f_d'(\hat{x}_d)| = \frac{\lambda d (1 - \beta\gamma)(\beta \hat{x}_d + 1)^{d-1}}{(\hat{x}_d + \gamma)^{d+1}}$$
$$= \frac{d(1 - \beta\gamma)\hat{x}_d}{(\beta \hat{x}_d + 1)(\hat{x}_d + \gamma)}$$
$$< 1.$$

In particular, (β, γ, λ) *is* universally unique *if it is* up-to-∞ unique.

Being up-to-Δ unique is equivalent to that the system is of weak spatial mixing on infinite regular trees up to degree Δ. The uniqueness condition can be described in various threshold forms, which are given in Section 3.

The uniqueness condition is defined by requiring that $|f_d'(\hat{x}_d)| < 1$. The following lemma states that $|f_d'(\hat{x}_d)|$ is bounded by an absolute constant as long as the uniqueness condition holds.

LEMMA 2.1. *Let* (β, γ, λ) *be anti-ferromagnetic. If* (β, γ, λ) *is up-to-Δ unique then there exists an absolute constant* $c < 1$ *which depends only on* β, γ, λ *and* Δ, *such that* $|f_d'(\hat{x}_d)| \leq c$ *for all* $1 \leq d < \Delta$.

Proof. The lemma holds trivially for finite Δ. It then remains to show that in case of universal uniqueness, $|f_d'(\hat{x}_d)|$ cannot be arbitrarily close to 1 as d grows to infinity. If (β, γ, λ) is universally unique, due to Lemma 3.1.2, we must have $\gamma > 1$. For anti-ferromagnetic (β, γ, λ), $\beta \leq \frac{1}{\gamma}$, thus the fixed point $\hat{x}_d = \lambda \left(\frac{\beta \hat{x}_d + 1}{\hat{x}_d + \gamma} \right)^d \leq \frac{\lambda}{\gamma^d}$, therefore $|f_d'(\hat{x}_d)| = \frac{d(1 - \beta\gamma)\hat{x}_d}{(\beta \hat{x}_d + 1)(\hat{x}_d + \gamma)} \leq \frac{d\lambda}{\gamma^d}$. The lemma follows.

3 The Uniqueness Thresholds

The following lemma translates the uniqueness condition into its various threshold forms. The lemma (except for item 2) is not used in the proofs of the main results but is used in the interpretation of the main results and comparisons to the previous results which are mostly stated in threshold forms.

LEMMA 3.1. *Let* (β, γ, λ) *be anti-ferromagnetic.*

1. (β, γ, λ) *is up-to-2 unique.*

2. *If* $\gamma \leq 1$, *then the uniqueness does not hold on infinite d-regular tree for all sufficiently large d.*

3. *If* $\gamma > 1$, *then the uniqueness holds on infinite d-regular tree for all sufficiently large d.*

4. *For any* Δ, *there exists a critical threshold* $\gamma_c = \gamma_c(\beta, \lambda, \Delta)$ *such that* (β, γ, λ) *is up-to-Δ unique if and only if* $\gamma \in (\gamma_c, \frac{1}{\beta})$. *In particular,* $\gamma_c(\beta, \lambda, \infty) > 1$ *and* $\gamma_c(\beta, \lambda, \infty) = \gamma_c(\beta, \lambda, \Delta)$ *for some finite* Δ.

5. *If* $\beta = 0$, *for any* Δ, *there exists a critical threshold* $\lambda_c = \lambda_c(\gamma, \Delta) = \min_{1 < d < \Delta} \frac{\gamma^{d+1} d^d}{(d-1)^{d+1}}$ *such that* (β, γ, λ) *is up-to-Δ unique if and only if* $\lambda \in (0, \lambda_c)$.

6. *If* $\sqrt{\beta\gamma} > \frac{\Delta - 2}{\Delta}$, *then for any external field* λ, (β, γ, λ) *is up-to-Δ unique.*

7. *If* $\beta > 0$, *for any* Δ *that* $\sqrt{\beta\gamma} \leq \frac{\Delta - 2}{\Delta}$, *there exist two critical thresholds* $\lambda_c = \lambda_c(\beta, \gamma, \Delta)$ *and* $\bar{\lambda}_c = \bar{\lambda}_c(\beta, \gamma, \Delta)$ *such that* (β, γ, λ) *is up-to-Δ unique if and only if* $\lambda \in (0, \lambda_c) \cup (\bar{\lambda}_c, \infty)$. *In particular, when* $\beta = \gamma$, *it holds that* $\lambda_c \cdot \bar{\lambda}_c = 1$, *thus* (β, β, λ) *is up-to-Δ unique if and only if* $|\log \lambda| > \log \bar{\lambda}_c$.

8. *If* $\beta > 0$ *and* $\gamma > 1$, *there exists a absolute positive constant* $\lambda_c = \lambda_c(\beta, \gamma)$ *such that* (β, γ, λ) *is universally unique if and only if* $\lambda \in (0, \lambda_c)$.

Proof. Let $f_d(x) = \lambda \left(\frac{\beta x_d + 1}{x_d + \gamma} \right)^d$ and $\hat{x}_d = f_d(\hat{x}_d)$ be the positive fixed point of $f_d(x)$. Then

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

Let (β, γ, λ) be anti-ferromagnetic. That is, $0 \leq \beta \leq \gamma$, $\gamma > 0$, and $\beta\gamma < 1$, thus $\beta < 1$.

1. It is easy to verify that for $\beta\gamma < 1$, $(\beta x+1)(x+\gamma) - (1-\beta\gamma)x > 0$ for any positive $x > 0$. Therefore, when $d = 1$, we have $|f_1(\hat{x}_1)| = \frac{(1-\beta\gamma)\hat{x}_1}{(\beta\hat{x}_1+1)(\hat{x}_1+\gamma)} < 1$, which means that (β, γ, λ) is up-to-2 unique.

2. For all sufficiently large d, it holds that

$$\lambda\beta^d \exp\left(\frac{d}{(1-\beta\gamma)d-3}\right) < \frac{d(1-\beta\gamma)-3}{\beta}, \text{ and}$$

$$\lambda \exp\left(-\frac{\gamma d}{d(1-\beta\gamma)-3}\right) > \frac{\gamma}{d(1-\beta\gamma)-3}.$$

By contradiction, we suppose that $|f_d(\hat{x}_d)| = \frac{d(1-\beta\gamma)\hat{x}_d}{(\beta\hat{x}_d+1)(\hat{x}_d+\gamma)} \leq 1$. Then,

$$1 \geq \frac{d(1-\beta\gamma)\hat{x}_d}{(\beta\hat{x}_d+1)(\hat{x}_d+\gamma)} \geq \frac{d(1-\beta\gamma)}{\beta\hat{x}_d + \frac{\gamma}{\hat{x}_d} + 2}.$$

Case.1: $\hat{x}_d \geq \gamma$. Then $\frac{\gamma}{\hat{x}_d} \leq 1$. Thus,

$$1 \geq \frac{d(1-\beta\gamma)}{\beta\hat{x}_d + \frac{\gamma}{\hat{x}_d} + 2} \geq \frac{d(1-\beta\gamma)}{\beta\hat{x}_d + 3},$$

which implies that $\hat{x}_d \geq \frac{d(1-\beta\gamma)-3}{\beta}$, however we have

$$\hat{x}_d = \lambda\left(\frac{\beta\hat{x}_d+1}{\hat{x}_d+\gamma}\right)^d$$
$$\leq \lambda\left(\frac{\beta\hat{x}_d+1}{\hat{x}_d}\right)^d$$
$$\leq \lambda\left(\beta + \frac{\beta}{d(1-\beta\gamma)-3}\right)^d$$
$$\leq \lambda\beta^d \exp\left(\frac{d}{(1-\beta\gamma)d-3}\right)$$
$$< \frac{d(1-\beta\gamma)-3}{\beta},$$

a contradiction.

Case.2: $\hat{x}_d < \gamma$. Then $\beta\hat{x}_d \leq \beta\gamma < 1$. Thus,

$$1 \geq \frac{d(1-\beta\gamma)}{\beta\hat{x}_d + \frac{\gamma}{\hat{x}_d} + 2} \geq \frac{d(1-\beta\gamma)}{\frac{\gamma}{\hat{x}_d} + 3},$$

which implies that $\hat{x}_d \leq \frac{\gamma}{d(1-\beta\gamma)-3}$, however we

have

$$\hat{x}_d = \lambda\left(\frac{\beta\hat{x}_d+1}{\hat{x}_d+\gamma}\right)^d$$
$$\geq \frac{\lambda}{(\hat{x}_d+1)^d}$$
$$\geq \lambda\left(1 + \frac{\gamma}{d(1-\beta\gamma)-3}\right)^{-d}$$
$$\geq \lambda \cdot \exp\left(-\frac{\gamma d}{d(1-\beta\gamma)-3}\right)$$
$$> \frac{\gamma}{d(1-\beta\gamma)-3},$$

a contradiction.

3. Let $\gamma > 1$. The fixed point $\hat{x}_d = \lambda\left(\frac{\beta\hat{x}_d+1}{\hat{x}_d+\gamma}\right)^d \leq \frac{\lambda}{\gamma^d}$, and

$$|f'_d(\hat{x}_d)| = \frac{d(1-\beta\gamma)\hat{x}_d}{(\beta\hat{x}_d+1)(\hat{x}_d+\gamma)} \leq \frac{d\lambda}{\gamma^d},$$

which is strictly less than 1 for all sufficiently large d. Thus the uniqueness holds on infinite d-regular tree for all sufficiently large d.

4. We first show that there exists a critical threshold $\gamma_c = \gamma_c(\beta, \lambda, \Delta)$ such that (β, γ, λ) is up-to-Δ unique if and only if $\gamma \in (\gamma_c, \frac{1}{\beta})$. It is sufficient to show that if an anti-ferromagnetic (β, γ, λ) is up-to-Δ unique then $(\beta, \gamma', \lambda)$ is up-to-Δ unique for any $\gamma' > \gamma$ and $\beta\gamma' < 1$.

Recall that $\hat{x}_d$ is the positive fixed point of $f_d(x) = \lambda\left(\frac{\beta x+1}{x+\gamma}\right)^d$. Also let $\hat{x}'_d$ denote the positive solution to $x = \lambda\left(\frac{\beta x+1}{x+\gamma'}\right)^d$.

We first show that $\hat{x}'_d < \hat{x}_d$. By contradiction, assume that $\hat{x}'_d \geq \hat{x}_d$. Since for anti-ferromagnetic (β, γ, λ), the function $f_d(x)$ is monotonically decreasing in x, we have

$$\hat{x}_d = \lambda\left(\frac{\beta\hat{x}_d+1}{\hat{x}_d+\gamma}\right)^d$$
$$\geq \lambda\left(\frac{\beta\hat{x}'_d+1}{\hat{x}'_d+\gamma}\right)^d$$
$$> \lambda\left(\frac{\beta\hat{x}'_d+1}{\hat{x}'_d+\gamma'}\right)^d$$
$$= \hat{x}'_d,$$

a contradiction.

Therefore we have $\hat{x}'_d < \hat{x}_d$, which means that

$$\lambda\left(\beta + \frac{(1-\beta\gamma')}{\hat{x}'_d+\gamma'}\right)^d < \lambda\left(\beta + \frac{(1-\beta\gamma)}{\hat{x}_d+\gamma}\right)^d,$$

73

thus

$$\frac{(1-\beta\gamma')}{\hat{x}_d' + \gamma'} < \frac{(1-\beta\gamma)}{\hat{x}_d + \gamma}.$$

For $\hat{x}_d' < \hat{x}_d$, it also holds that

$$\frac{\hat{x}_d'}{\beta\hat{x}_d' + 1} = \frac{1}{\beta + \frac{1}{\hat{x}_d'}} < \frac{1}{\beta + \frac{1}{\hat{x}_d}} = \frac{\hat{x}_d}{\beta\hat{x}_d + 1}.$$

Multiplying the above two inequalities together, we have

$$\frac{d(1-\beta\gamma')\hat{x}_d'}{(\beta\hat{x}_d' + 1)(\hat{x}_d + \gamma')} < \frac{d(1-\beta\gamma)\hat{x}_d}{(\beta\hat{x}_d + 1)(\hat{x}_d + \gamma)}.$$

Note that these are the absolute derivatives at the respective fixed points when the parameters are (β, γ, λ) and $(\beta, \gamma', \lambda)$. Therefore if (β, γ, λ) is up-to-Δ unique then $(\beta, \gamma', \lambda)$ is up-to-Δ unique.

Due to Part 2 of the lemma, if $\gamma \le 1$, $|f_d'(\hat{x}_d)| > 1$ for all sufficiently large d, thus $\gamma_c(\beta, \lambda, \infty) > 1$. And due to Part 3 of the lemma, for any $\gamma \ge \gamma_c(\beta, \lambda, \infty) > 1$, $|f_d'(\hat{x}_d)|$ is arbitrarily close to 0 as d grows to infinity, thus $\gamma_c(\beta, \lambda, \infty) = \gamma_c(\beta, \lambda, \Delta)$ for a finite Δ.

5. When $\beta = 0$, $|f_d'(\hat{x}_d)| = \frac{d(1-\beta\gamma)\hat{x}_d}{(\beta\hat{x}_d + 1)(\hat{x}_d + \gamma)} = \frac{d\hat{x}_d}{\hat{x}_d + \gamma}$, the uniqueness condition $|f_d'(\hat{x}_d)| < 1$ is equivalent to that $\hat{x}_d < \frac{\gamma}{d-1}$ (here we assume $d > 1$ since due to Part 1 of the lemma, for $d = 1$ the uniqueness always holds). Recall that $\hat{x}_d = \lambda\left(\frac{1}{\hat{x}_d + \gamma}\right)^d$. Then $|f_d'(\hat{x}_d)| < 1$ if and only if

$$\lambda = \hat{x}_d(\hat{x}_d + \gamma)^d < \frac{\gamma^{d+1}d^d}{(d-1)^{d+1}}.$$

Let $\lambda_c = \lambda_c(\gamma, \Delta) = \min_{1 < d < \Delta} \frac{\gamma^{d+1}d^d}{(d-1)^{d+1}}$. It holds that (β, γ, λ) is up-to-Δ unique if and only if $\lambda \in (0, \lambda_c)$.

6. We note that $|f_d'(\hat{x}_d)| = \frac{d(1-\beta\gamma)\hat{x}_d}{(\beta\hat{x}_d + 1)(\hat{x}_d + \gamma)}$ is not monotone in $\hat{x}_d$. It achieves its maximum value at $\hat{x}_d = \sqrt{\frac{\beta}{\gamma}}$. Therefore, if for any $1 \le d < \Delta$, $\frac{d(1-\beta\gamma)x}{(\beta x+1)(x+\gamma)} < 1$ for $x = \sqrt{\frac{\beta}{\gamma}}$, then (β, γ, λ) is up-to-Δ unique for any λ. This condition holds when $\sqrt{\beta\gamma} > \frac{d-1}{d+1}$ for all $1 \le d < \Delta$, i.e. when $\sqrt{\beta\gamma} > \frac{\Delta-2}{\Delta}$.

7. The following lemma is needed in the proofs of this part and next part.

LEMMA 3.2. *Given that* $0 < \beta \le \gamma$, $\beta\gamma < 1$, *and* $\sqrt{\beta\gamma} \le \frac{d-1}{d+1}$, *define*

$$x_1(d) = \frac{1}{2\beta}\Big(-1 - \beta\gamma + d(1-\beta\gamma)$$
$$- \sqrt{(-1 - \beta\gamma + d(1-\beta\gamma))^2 - 4\beta\gamma}\Big)$$
$$x_2(d) = \frac{1}{2\beta}\Big(-1 - \beta\gamma + d(1-\beta\gamma)$$
$$+ \sqrt{(-1 - \beta\gamma + d(1-\beta\gamma))^2 - 4\beta\gamma}\Big).$$

which are the two positive solutions to the equation $\frac{d(1-\beta\gamma)x}{(\beta x+1)(x+\gamma)} = 1$.

Let $\lambda_i(d) = x_i(d)\left(\frac{x_i(d)+\gamma}{\beta x_i(d)+1}\right)^d$, $i = 1, 2$.

(a) $x_1(d)$ *is monotonically decreasing in* d; $x_2(d)$ *is monotonically increasing in* d.

(b) $\lambda_2(d)$ *is monotonically increasing in* d *and goes to infinity as* d *grows.*

(c) *If* $\gamma > 1$, $\lambda_1(d)$ *achieves its minimum value at a finite* d.

Proof. Note that $(-1 - \beta\gamma + d(1-\beta\gamma))^2 - 4\beta\gamma \ge 0$ if $\sqrt{\beta\gamma} \le \frac{d-1}{d+1}$, which means that $x_1(d)$ and $x_2(d)$ are well-defined.

(a) It is obvious that $x_2(d)$ is increasing in d and is unbounded as d grows. Note $x_1(d)x_2(d) = \frac{\gamma}{\beta}$, so $x_1(d)$ is decreasing in d and tends to 0 as d grows to infinity.

(b) For $0 \le \beta \le \gamma$, $\beta\gamma < 1$ and $\sqrt{\beta\gamma} \le \frac{d-1}{d+1}$, we have $\beta \le \sqrt{\beta\gamma} \le \frac{d-1}{d+1} < 1$ and $\frac{x+\gamma}{\beta x+1}$ is increasing in x. Moreover, it can be verified that

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

Since $x_2(d)$ is also increasing in d and unbounded as d grows, we can conclude that $\lambda_2(d) = x_2(d)\left(\frac{x_2(d)+\gamma}{\beta x_2(d)+1}\right)^d$ is increasing in d and is unbounded as d grows.

(c) It is sufficient to prove that $\lambda_1(d)$ is unbounded as d grows. When $\gamma > 1$, we have

$$\lambda_1(d) = x_1(d)\left(\frac{x_1(d) + \gamma}{\beta x_1(d) + 1}\right)^d$$
$$\ge x_1(d) \cdot \gamma^d,$$

which can be easily verified to be unbounded as d grows to infinity.

Recall that $\sqrt{\beta\gamma} \leq \frac{\Delta-2}{\Delta}$. Let $\overline{\Delta}$ be the smallest integer d satisfying that $\sqrt{\beta\gamma} \leq \frac{d-1}{d+1}$. Then we have $\sqrt{\beta\gamma} > \frac{d-1}{d+1}$ for all $1 \leq d < \overline{\Delta}$ and $\sqrt{\beta\gamma} \leq \frac{d-1}{d+1}$ for all $\overline{\Delta} \leq d < \Delta$. According to Part 6 of this lemma, for all $\lambda > 0$, (β,γ,λ) is up-to-$\overline{\Delta}$ unique, i.e. $|f_d'(\hat{x}_d)| < 1$ for $1 \leq d < \overline{\Delta}$.

It remains to analyze the $|f_d'(\hat{x}_d)|$ for such d that $\overline{\Delta} \leq d < \Delta$, i.e. $\sqrt{\beta\gamma} \leq \frac{d-1}{d+1}$. By Lemma 3.2, the equation $\frac{d(1-\beta\gamma)x}{(\beta x+1)(x+\gamma)} = 1$ has two positive roots $x_1(d)$ and $x_2(d)$ as given in Lemma 3.2.

It then holds for such d that $|f_d'(\hat{x}_d)| = \frac{d(1-\beta\gamma)\hat{x}_d}{(\beta\hat{x}_d+1)(\hat{x}_d+\gamma)} < 1$ if and only if $\hat{x}_d < x_1(d)$ or $\hat{x}_d > x_2(d)$. Note that $x\left(\frac{x+\gamma}{\beta x+1}\right)^d$ is monotonically increasing in x for any fixed d, thus $\hat{x}_d < x_1(d)$ if and only if

$$\lambda = \hat{x}_d \left(\frac{\hat{x}_d + \gamma}{\beta\hat{x}_d + 1}\right)^d$$
$$< x_1(d)\left(\frac{x_1(d)+\gamma}{\beta x_1(d)+1}\right)^d$$
$$= \lambda_1(d),$$

and $\hat{x}_d > x_2(d)$ if and only if

$$\lambda = \hat{x}_d\left(\frac{\hat{x}_d+\gamma}{\beta\hat{x}_d+1}\right)^d$$
$$> x_2(d)\left(\frac{x_2(d)+\gamma}{\beta x_2(d)+1}\right)^d$$
$$= \lambda_2(d).$$

Therefore, $|f_d'(\hat{x}_d)| < 1$ if and only if $\lambda < \lambda_1(d)$ or $\lambda > \lambda_2(d)$.

For any Δ that $\sqrt{\beta\gamma} \leq \frac{\Delta-2}{\Delta}$, let

$$\lambda_c = \lambda_c(\beta,\gamma,\Delta) = \min_{\overline{\Delta} \leq d < \Delta} \lambda_1(d),$$
$$\bar{\lambda}_c = \bar{\lambda}_c(\beta,\gamma,\Delta) = \max_{\overline{\Delta} \leq d < \Delta} \lambda_2(d).$$

It holds that (β,γ,λ) is up-to-Δ unique if and only if $\lambda \in (0,\lambda_c) \cup (\bar{\lambda}_c, \infty)$.

We then show that when $\beta = \gamma$, it holds that $\lambda_c \cdot \bar{\lambda}_c = 1$. First, it is easy to see that when $\beta = \gamma$, $x_1(d)x_2(d) = \frac{\gamma}{\beta} = 1$, thus $\lambda_1(d) \cdot \lambda_2(d) = x_1(d)x_2(d)\left(\frac{x_1(d)+\gamma}{\beta x_1(d)+1} \cdot \frac{x_2(d)+\gamma}{\beta x_2(d)+1}\right)^d = 1$ because $\beta = \gamma$ and $x_1(d)x_2(d) = 1$. Due to Part 7b of Lemma 3.2, $\lambda_2(d)$ is monotonically increasing thus $\lambda_1(d) = 1/\lambda_2(d)$ is monotonically decreasing. Therefore, $\bar{\lambda}_c = \max_{\overline{\Delta} \leq d < \Delta} \lambda_2(d) = \lambda_2(\Delta-1)$ and $\lambda_c =$

$\min_{\overline{\Delta} \leq d < \Delta} \lambda_1(d) = \lambda_1(\Delta-1)$ and it holds that $\lambda_c \cdot \bar{\lambda}_c = \lambda_1(\Delta-1) \cdot \lambda_2(\Delta-1) = 1$. Then (β,γ,λ) is up-to-Δ unique if and only if $\lambda \in (0, 1/\bar{\lambda}_c) \cup (\bar{\lambda}_c, \infty)$, i.e. $|\log\lambda| > \log\bar{\lambda}_c$.

8. Let $\overline{\Delta}$ be the smallest integer d satisfying that $\sqrt{\beta\gamma} \leq \frac{d-1}{d+1}$. Then by Part 7, for any Δ, (β,γ,λ) is up-to-Δ unique if and only if $\lambda \in (0,\lambda_c) \cup (\bar{\lambda}_c, \infty)$, where $\lambda_c = \min_{\overline{\Delta} \leq d < \Delta} \lambda_1(d)$ and $\bar{\lambda}_c = \max_{\overline{\Delta} \leq d < \Delta} \lambda_2(d)$. By Part 7a and Part 7b of Lemma 3.2, for fixed β,γ, $\lambda_2(d)$ goes to infinity as d grows. Therefore, $\bar{\lambda}_c$ is unbounded for $\Delta = \infty$.

If $\gamma > 1$, by Part 7c of Lemma 3.2, there is an absolute positive constant $\lambda_c = \min_{d \geq \bar{\Delta}} \lambda_1(d)$ such that (β,γ,λ) is universally unique if and only if $\lambda \in (0,\lambda_c) \cup (\bar{\lambda}_c, \infty) = (0,\lambda_c)$.

4 The strong spatial mixing on general graphs

In this section we prove Theorem 1.1. The necessity of the uniqueness condition is trivial since strong spatial mixing on general graphs implies weak spatial mixing on regular trees. It then remains to prove the following theorem.

THEOREM 4.1. *Let (β,γ,λ) be anti-ferromagnetic. For any finite $\Delta \geq 2$ or $\Delta = \infty$, if (β,γ,λ) is up-to-Δ unique, then the two-state spin system of parameters (β,γ,λ) exhibits strong spatial mixing on graphs of maximum degree at most Δ.*

Our approach is to prove the strong spatial mixing on arbitrary trees of maximum degree at most Δ, which by the self-avoiding walk tree construction implies the theorem. This is because for general two-state anti-ferromagnetic spin systems the worst case for strong spatial mixing among all graphs of maximum degree at most d may no longer be the d-regular tree. We will explain this in detail in Section 6.

Given any graph $G(V,E)$ of maximum degree at most Δ, any configuration $\sigma_\Lambda \in \{0,1\}^\Lambda$ on $\Lambda \subset V$ and any $S \subset \Lambda$, fixing an arbitrary vertex $v \in V$, by Theorem 2.1, a self-avoiding walk tree $T = T_{\text{SAW}}(G,v)$ can be constructed such that the maximum degree of T is bounded by Δ, $\text{dist}_G(v,S) = \text{dist}_T(v,S)$ and $R_{G,v}^{\sigma_\Lambda} = R_T^{\sigma_\Lambda}$. Recall that $R_{G,v}^{\sigma_\Lambda} = p_v^{\sigma_\Lambda}/(1-p_v^{\sigma_\Lambda})$ thus $p_v^{\sigma_\Lambda} = R_{G,v}^{\sigma_\Lambda}/(1+R_{G,v}^{\sigma_\Lambda})$. For any σ_Λ and τ_Λ,

$$|p_v^{\sigma_\Lambda} - p_v^{\tau_\Lambda}| = \left|\frac{R_{G,v}^{\sigma_\Lambda}}{1+R_{G,v}^{\sigma_\Lambda}} - \frac{R_{G,v}^{\tau_\Lambda}}{1+R_{G,v}^{\tau_\Lambda}}\right|$$
$$\leq \left|R_{G,v}^{\sigma_\Lambda} - R_{G,v}^{\tau_\Lambda}\right|$$
$$= \left|R_T^{\sigma_\Lambda} - R_T^{\tau_\Lambda}\right|.$$

This motivates the following definition.

DEFINITION 4.1. *Let T be a tree rooted by vertex v, $\tau_\Lambda \in \{0,1\}^\Lambda$ be a configuration on $\Lambda \subset V$ and $S \subseteq \Lambda$ be a vertex set. Define R_v and δ_v as that $R_v \leq R_T^{\tau_\Lambda} \leq R_v + \delta_v$ for all $\sigma_\Lambda \in \{0,1\}^\Lambda$ which differ from τ_Λ only on S.*

It is then sufficient to prove Theorem 4.1 by constructing such R_v and δ_v for $T = T_{\text{SAW}}(G, v)$ and showing that $\delta_v \leq \exp(-\Omega(\text{dist}(v, S)))$.

Let T be a tree rooted by v, who has d children $v_1, \ldots, v_d$, and T_i be the subtree rooted by v_i. It holds that

$$R_T^{\sigma_\Lambda} = f\left(R_{T_1}^{\sigma_\Lambda}, \ldots, R_{T_d}^{\sigma_\Lambda}\right) \triangleq \lambda \prod_{i=1}^d \frac{\beta R_{T_i}^{\sigma_\Lambda} + 1}{R_{T_i}^{\sigma_\Lambda} + \gamma}.$$

The lower and upper bounds R_v and $R_v + \delta_v$ can be recursively constructed as follows. The base cases are: (1) $v \in S$, in which case $R_v = 0$ and $\delta_v = \infty$; and (2) $v \in \Lambda \setminus S$, i.e. v is fixed to be the same color in all σ_Λ, in which case $\delta_v = 0$ and $R_v = \infty$ (or $R_v = 0$) if v is fixed to be blue (or green). For $v \notin \Lambda$, since $f(R_1, \ldots, R_d)$ is monotonically decreasing for antiferromagnetic (β, γ, λ),

$$R_v = f(R_{v_1} + \delta_{v_1}, \ldots, R_{v_d} + \delta_{v_d}),$$
$$R_v + \delta_v = f(R_{v_1}, \ldots, R_{v_d}),$$

where R_{v_i} and $R_{v_i} + \delta_{v_i}$ are the corresponding lower and upper bounds for $R_{T_i}^{\sigma_\Lambda}$, $1 \leq i \leq d$. In particular, when $d = 0$ the empty product equals 1 by convention, thus $R_v = R_v + \delta_v = \lambda$, which is consistent with the case that v is a free vertex having no children.

By the monotonicity of $f(R_1, \ldots, R_d)$, it is easy to check that the R_v and $R_v + \delta_v$ constructed above satisfy the requirement of Definition 4.1. Our goal is to show that δ_v decays exponentially in depth of recursion when the uniqueness holds. A straightforward approach is trying to prove that δ contracts at a constant rate in each recursion step. But this does not have to be true to guarantee the exponential decay. Indeed there are cases that the error does not decay in single steps but decay in a long run. To overcome this, we use a *potential* Φ to amortize the contraction and show that $\delta \cdot \Phi$ contracts at a constant rate. We choose the potential function to be

$$\Phi(R) = \frac{1}{\sqrt{R(\beta R + 1)(R + \gamma)}}.$$

We are analyzing the decay on an arbitrary tree with irregular degrees. In order to adapt this irregularity, the potential function cannot have d as an input, but only caries the information regarding the distribution at the current vertex, yet it has to be able to provide

correct compensation to the step-wise decay at any state of R and for all spin systems satisfying sufficient uniqueness. A heuristic procedure which leads us to this good potential function is discussed in Appendix A.

Let $\varphi(R)$ be a monotone function satisfying that $\varphi'(R) = \Phi(R)$. We define that

$$\epsilon_v \triangleq \varphi(R_v + \delta_v) - \varphi(R_v),$$

and accordingly, $\epsilon_{v_i} \triangleq \varphi(R_{v_i} + \delta_{v_i}) - \varphi(R_{v_i})$, $1 \leq i \leq d$.

We define a function $\alpha(d; x_1, \ldots, x_d)$ as follows:

$$\alpha(d; x_1, \ldots, x_d)$$
$$\triangleq \frac{(1 - \beta\gamma)\left(\lambda \prod_{i=1}^d \frac{\beta x_i + 1}{x_i + \gamma}\right)^{\frac{1}{2}}}{\left(\beta\lambda \prod_{i=1}^d \frac{\beta x_i + 1}{x_i + \gamma} + 1\right)^{\frac{1}{2}} \left(\lambda \prod_{i=1}^d \frac{\beta x_i + 1}{x_i + \gamma} + \gamma\right)^{\frac{1}{2}}}$$
$$\cdot \sum_{i=1}^d \frac{x_i^{\frac{1}{2}}}{(\beta x_i + 1)^{\frac{1}{2}}(x_i + \gamma)^{\frac{1}{2}}}.$$

The following lemma is obtained from applying the Mean Value Theorem. Similar routines were previously used in [15, 19].

LEMMA 4.1. *The followings hold for ϵ_v.*

1. *(relation to δ_v) $\epsilon_v = \delta_v \cdot \Phi(\widetilde{R})$ for some $\widetilde{R} \in [R_v, R_v + \delta_v]$.*

2. *(absolute bound) Assuming that $\gamma > 1$ or the maximum degree of T is bounded by a constant, if $v \notin \Lambda$ then $R_v + \delta_v = O(1)$ and if $v_i \notin S$ for all $1 \leq i \leq d$ then $\epsilon_v = O(1)$.*

3. *(stepwise contraction) There exist $\widetilde{R_i} \in [R_{v_i}, R_{v_i} + \delta_{v_i}]$, $1 \leq i \leq d$, such that*

$$\epsilon_v \leq \alpha(d; \widetilde{R}_1, \ldots, \widetilde{R}_d) \cdot \max_{1 \leq i \leq d}\{\epsilon_{v_i}\}.$$

Proof. 1. Due to the Mean Value Theorem, there exists an $\widetilde{R} \in [R_v, R_v + \delta_v]$ such that

$$\epsilon_v = \varphi(R_v + \delta_v) - \varphi(R_v) = \varphi'(\widetilde{R}) \cdot \delta_v = \Phi(\widetilde{R}) \cdot \delta_v.$$

2. Suppose that each vertex has at most k children. Then due to the assumption either k is bounded or $\gamma > 1$.

If $v \notin \Lambda$, then $\delta_v \leq R_v + \delta_v = f(R_{v_1}, \ldots, R_{v_d}) \leq f(0, \ldots, 0) = \frac{\lambda}{\gamma^d}$, where $0 \leq d \leq k$. If $\gamma > 1$, then $R_v + \delta_v \leq \lambda = O(1)$; and if k is finite, then $R_v + \delta_v \leq \max\{\frac{\lambda}{\gamma^k}, \lambda\} = O(1)$.

Due to the Mean Value Theorem, there exist $\widetilde{R_i} \in [R_{v_i}, R_{v_i} + \delta_{v_i}]$, $1 \le i \le d$, such that

$$
\begin{aligned}
\epsilon_v &= \varphi\left(f(R_{v_1}, \ldots, R_{v_d})\right) \\
&\quad - \varphi\left(f(R_{v_1} + \delta_{v_1}, \ldots, R_{v_d} + \delta_{v_d})\right) \\
&= -\nabla\varphi\left(f(\widetilde{R_1}, \ldots, \widetilde{R_d})\right) \cdot (\delta_{v_1}, \ldots, \delta_{v_d}) \\
&= \frac{(1 - \beta\gamma) \cdot \left(f(\widetilde{R_1}, \ldots, \widetilde{R_d})\right)^{\frac{1}{2}}}{\left(\beta f(\widetilde{R_1}, \ldots, \widetilde{R_d}) + 1\right)^{\frac{1}{2}} \left(f(\widetilde{R_1}, \ldots, \widetilde{R_d}) + \gamma\right)^{\frac{1}{2}}} \\
&\quad \cdot \sum_{i=1}^{d} \frac{\delta_{v_i}}{(\beta\widetilde{R_i} + 1)(\widetilde{R_i} + \gamma)} \\
&\le \sqrt{\frac{f(0, \ldots, 0)}{\gamma}} \cdot \sum_{i=1}^{d} \frac{\delta_{v_i}}{\gamma}.
\end{aligned}
$$

If $v_i \notin S$ for all v_i, then due to the above argument, $\delta_{v_i} \le \frac{\lambda}{\gamma^{d_i}}$ for free v_i, where $0 \le d_i \le k$ is the number of children of v_i; and trivially $\delta_{v_i} = 0$ for fixed v_i. Therefore, $\epsilon_v \le \sqrt{\frac{\lambda}{\gamma^{d+1}}} \sum_{i=1}^{d} \frac{\lambda}{\gamma^{d_i+1}} \le \lambda^{\frac{3}{2}} d\gamma^{-\frac{d+3}{2}} \cdot \max\left\{\frac{1}{\gamma^k}, 1\right\}$. If $\gamma > 1$, then it is easy to see that $d\gamma^{-\frac{d+3}{2}}$ is bounded by a constant, thus it holds that $\epsilon_v = O(1)$; if k is bounded, then $d \le k$ is also bounded, thus clearly $\epsilon_v = O(1)$.

3. We then analyze the stepwise contraction of ϵ_v. Define that $y_v = \varphi(R_v)$ and accordingly $y_{v_i} = \varphi(R_{v_i})$, $1 \le i \le d$. Then $y_v + \epsilon_v = \varphi(R_v + \delta_v)$ and $y_{v_i} + \epsilon_i = \varphi(R_{v_i} + \delta_{v_i})$, $1 \le i \le d$. We have

$$
\begin{aligned}
y_v &= \varphi(f(\varphi^{-1}(y_{v_1} + \epsilon_{v_1}), ..., \varphi^{-1}(y_{v_d} + \epsilon_{v_d}))), \\
y_v + \epsilon_v &= \varphi(f(\varphi^{-1}(y_{v_1}), ..., \varphi^{-1}(y_{v_d}))).
\end{aligned}
$$

Apply the Mean Value Theorem. There exist $\widetilde{y_i} \in [y_{v_i}, y_{v_i} + \epsilon_{v_i}]$ and corresponding $\widetilde{R_i} \in [R_{v_i}, R_{v_i} + \delta_{v_i}]$ satisfying $\widetilde{y_i} = \varphi(\widetilde{R_i})$, $1 \le i \le d$, such that

$$
\begin{aligned}
\epsilon_v &= \varphi(f(\varphi^{-1}(y_{v_1}), ..., \varphi^{-1}(y_{v_d}))) \\
&\quad - \varphi(f(\Phi^{-1}(y_{v_1} + \epsilon_{v_1}), ..., \varphi^{-1}(y_{v_d} + \epsilon_{v_d}))) \\
&= -\nabla\varphi(f(\varphi^{-1}(\widetilde{y_1}), ..., \varphi^{-1}(\widetilde{y_d}))) \cdot (\epsilon_{v_1}, ..., \epsilon_{v_d}) \\
&= -\Phi(f(\widetilde{R_1}, ..., \widetilde{R_d})) \cdot \sum_{i=1}^{d} \frac{\partial f}{\partial R_i} \frac{1}{\Phi(\widetilde{R_i})} \cdot \epsilon_i \\
&= \frac{(1 - \beta\gamma)\left(\lambda \prod_{i=1}^{d} \frac{\beta\widetilde{R_i}+1}{\widetilde{R_i}+\gamma}\right)^{\frac{1}{2}}}{\left(\beta\lambda \prod_{i=1}^{d} \frac{\beta\widetilde{R_i}+1}{\widetilde{R_i}+\gamma} + 1\right)^{\frac{1}{2}} \left(\lambda \prod_{i=1}^{d} \frac{\beta\widetilde{R_i}+1}{\widetilde{R_i}+\gamma} + \gamma\right)^{\frac{1}{2}}} \\
&\quad \cdot \sum_{i=1}^{d} \frac{\epsilon_{v_i}\widetilde{R_i}^{\frac{1}{2}}}{(\beta\widetilde{R_i}+1)^{\frac{1}{2}}(\widetilde{R_i}+\gamma)^{\frac{1}{2}}} \\
&\le \alpha(d; \widetilde{R_1}, \ldots, \widetilde{R_d}) \cdot \max_{1 \le i \le d}\{\epsilon_{v_i}\}.
\end{aligned}
$$

To prove the strong spatial mixing, we first relate δ_v to ϵ_v by Item 1 of Lemma 4.1, and then apply induction on the depth in T, with Item 2 of Lemma 4.1 as basis and Item 3 of Lemma 4.1 as induction step. We then need to bound the contraction rate $\alpha(d; x_1, ..., x_d)$. Note that $\frac{z}{(\beta z+1)(z+\gamma)} \le \left(1 + \sqrt{\beta\gamma}\right)^{-2} \le 1$ for $z \in [0, \infty)$, thus it holds unconditionally for all $x_i \in [0, \infty)$, $1 \le i \le d$, that

$$
\tag{4.2} \alpha(d; x_1, ..., x_d) \le d,
$$

$$
\tag{4.3} \alpha(d; x_1, ..., x_d) \le d \cdot \sqrt{\frac{\lambda}{\gamma^{d+1}}}.
$$

With the uniqueness, the following much tighter contraction bound can be proved.

LEMMA 4.2. *Let (β, γ, λ) be anti-ferromagnetic. If (β, γ, λ) is up-to-Δ unique, then there exists a constant $\alpha < 1$ such that for any integer $1 \le d < \Delta$ and any $x_1, ..., x_d \in [0, +\infty)$, $1 \le i \le d$, it holds that $\alpha(d; x_1, \ldots, x_d) \le \alpha$.*

This lemma is the technical core of our analysis. It crucially relies on the choice of potential function. Theorem 4.1 can be implied by this lemma. The rest of this section is devoted to the proof of this lemma. Given that (β, γ, λ) is up-to-Δ unique, there exists an absolute constant $\alpha < 1$ such that $\alpha(d; x_1, \ldots, x_d) \le \alpha$ for any $1 \le d < \Delta$ and $x_1, \ldots, x_d \ge 0$.

We define the symmetric version of $\alpha(d; x_1, \ldots, x_d)$:

$$
\begin{aligned}
\alpha_d(x) &\triangleq \alpha(d; \underbrace{x, \ldots, x}_{d}) \\
&= d(1 - \beta\gamma) \cdot \sqrt{\frac{x}{(\beta x + 1)(x + \gamma)}} \\
&\quad \cdot \sqrt{\frac{\lambda\left(\frac{\beta x+1}{x+\gamma}\right)^d}{\left(\beta\lambda\left(\frac{\beta x+1}{x+\gamma}\right)^d + 1\right)\left(\lambda\left(\frac{\beta x+1}{x+\gamma}\right)^d + \gamma\right)}}
\end{aligned}
$$

The following lemma shows that the symmetric case dominates the maximum of $\alpha(d; x_1, ..., x_d)$ by using the inequalities of Cauchy-Schwarz and arithmetic and geometric means.

LEMMA 4.3. *Let (β, γ, λ) be anti-ferromagnetic. Then for any integer d and any $x_1, ..., x_d \in [0, +\infty)$, there exists an $\bar{x} \in [0, +\infty)$ such that $\alpha(d; x_1, ..., x_d) \le \alpha(d, \bar{x})$.*

Proof. Let $z_i = \frac{\beta x_i + 1}{x_i + \gamma}$. Then $z_i \in (\beta, \frac{1}{\gamma}]$ and $x_i = \frac{1 - \gamma z_i}{z_i - \beta}$.

Express $\alpha(d; x_1, ..., x_d)$ in terms of z_i:

$$\alpha(d; x_1, ..., x_d) = \frac{\left(\lambda \prod_{i=1}^d z_i\right)^{\frac{1}{2}}}{\left(\beta\lambda \prod_{i=1}^d z_i + 1\right)^{\frac{1}{2}} \left(\lambda \prod_{i=1}^d z_i + \gamma\right)^{\frac{1}{2}}}$$
$$\cdot \sum_{i=1}^d (z_i^{-1} - \gamma)^{\frac{1}{2}} (z_i - \beta)^{\frac{1}{2}}.$$

Due to Cauchy-Schwarz inequality,

$$\sum_{i=1}^d (z_i^{-1} - \gamma)^{\frac{1}{2}} (z_i - \beta)^{\frac{1}{2}}$$
$$\leq d \left(\frac{1}{d} \sum_{i=1}^d (z_i^{-1} - \gamma)(z_i - \beta)\right)^{\frac{1}{2}}$$
$$= d \left(1 + \beta\gamma - \frac{1}{d} \sum_{i=1}^d (z_i \gamma + \beta z_i^{-1})\right)^{\frac{1}{2}}.$$

Due to the inequality of arithmetic and geometric means,

$$d \left(1 + \beta\gamma - \frac{1}{d} \sum_{i=1}^d (z_i \gamma + \beta z_i^{-1})\right)^{\frac{1}{2}}$$
$$\leq d \left(1 + \beta\gamma - \gamma \left(\prod_{i=1}^d z_i\right)^{\frac{1}{d}} - \beta \left(\prod_{i=1}^d z_i\right)^{-\frac{1}{d}}\right)^{\frac{1}{2}}.$$

Let $\bar{z} = \left(\prod_{i=1}^d z_i\right)^{\frac{1}{d}}$. Then combining the above calculations,

$$\alpha(d; x_1, ..., x_d) \leq \frac{(\lambda \bar{z}^d)^{\frac{1}{2}} \cdot d(1 + \beta\gamma - \gamma\bar{z} - \beta\bar{z}^{-1})^{\frac{1}{2}}}{(\beta\lambda\bar{z}^d + 1)^{\frac{1}{2}} (\lambda\bar{z}^d + \gamma)^{\frac{1}{2}}}$$
$$= d \cdot \sqrt{\frac{\lambda\bar{z}^d(\bar{z}^{-1} - \gamma)(\bar{z} - \beta)}{(\beta\lambda\bar{z}^d + 1)(\lambda\bar{z}^d + \gamma)}}.$$

Let $\bar{x}$ be such that $\frac{\beta\bar{x}+1}{\bar{x}+\gamma} = \bar{z}$. Then $\bar{x} \in [0, +\infty)$ and by substituting $\frac{\beta\bar{x}+1}{\bar{x}+\gamma}$ for $\bar{z}$, we have

$$\alpha(d; x_1, ..., x_d)$$
$$\leq d(1 - \beta\gamma) \cdot \sqrt{\frac{\bar{x}}{(\beta\bar{x}+1)(\bar{x}+\gamma)}}$$
$$\cdot \sqrt{\frac{\lambda\left(\frac{\beta\bar{x}+1}{\bar{x}+\gamma}\right)^d}{\left(\beta\lambda\left(\frac{\beta\bar{x}+1}{\bar{x}+\gamma}\right)^d + 1\right)\left(\lambda\left(\frac{\beta\bar{x}+1}{\bar{x}+\gamma}\right)^d + \gamma\right)}}$$
$$= \alpha_d(\bar{x}).$$

LEMMA 4.4. *Let* (β, γ, λ) *be anti-ferromagnetic. If* (β, γ, λ) *is up-to-Δ unique, then there exists a constant* $\alpha < 1$ *such that for any integer* $1 \leq d < \Delta$, *it holds that* $\alpha_d(x) \leq \alpha$ *for all* $x \geq 0$.

Proof. Fix d to be any positive integer. We characterize the x at which $\alpha_d(x)$ achieves its maximum. We denote that $f_d(x) = \lambda \left(\frac{\beta x+1}{x+\gamma}\right)^d$. Taking derivative of $\alpha_d(x)$ with respect to x, we get that

$$\alpha_d'(x) = d(1 - \beta\gamma) \cdot \frac{G'(x)}{2\sqrt{G(x)}},$$

where $G(x) = \frac{x f_d(x)}{(\beta f_d(x)+1)(f_d(x)+\gamma)(\beta x+1)(x+\gamma)}$, whose derivative is

$$G'(x) = \frac{f_d(x) \cdot d(1 - \beta\gamma)x}{(\beta f_d(x) + 1)(f_d(x) + \gamma)(\beta x + 1)^2(x + \gamma)^2}$$
$$\cdot \left(\frac{\gamma - \beta x^2}{d(1 - \beta\gamma)x} - \frac{\gamma - \beta f_d(x)^2}{(\beta f_d(x) + 1)(f_d(x) + \gamma)}\right).$$

As x ranges over $[0, \infty)$, the function $\frac{\gamma - \beta x^2}{d(1-\beta\gamma)x}$ is strictly decreasing in x and ranges from $+\infty$ to $-\infty$, and the function $\frac{\gamma - \beta f_d(x)^2}{(\beta f_d(x)+1)(f_d(x)+\gamma)}$ is strictly increasing in x and has a bounded range. Thus, the equation

$$(4.4) \qquad \frac{\gamma - \beta x^2}{d(1 - \beta\gamma)x} = \frac{\gamma - \beta f_d(x)^2}{(\beta f_d(x) + 1)(f_d(x) + \gamma)}.$$

has unique solution in $(0, \infty)$, denoted by x_d. Moreover, it holds that

$$(4.5) \qquad G'(x) \begin{cases} > 0 & \text{if } 0 \leq x < x_d, \\ = 0 & \text{if } x = x_d, \\ < 0 & \text{if } x > x_d. \end{cases}$$

The same also holds for $\alpha_d'(x)$. Thus, for any fixed d, $\alpha_d(x)$ achieves its maximum when $x = x_d$.

We define that

$$\tilde{\alpha}_d(x) \triangleq \sqrt{d(1 - \beta\gamma) \cdot \frac{(\gamma - \beta x^2)}{(\beta x + 1)(x + \gamma)} \cdot \frac{f_d(x)}{(\gamma - \beta f_d(x)^2)}}.$$

Therefore, for all $x \geq 0$,

$$\alpha_d(x) \leq \alpha_d(x_d)$$
$$= d(1 - \beta\gamma)\sqrt{\frac{x}{(\beta x_d + 1)(x_d + \gamma)}}$$
$$\cdot \sqrt{\frac{f_d(x_d)}{(\beta f_d(x_d) + 1)(f_d(x_d) + \gamma)}}$$
$$(4.6) \qquad = \tilde{\alpha}_d(x_d),$$

where the last equation is obtained by substituting $(\beta f_d(x_d) + 1)(f_d(x_d) + \gamma)$ according to (4.4).

Let $\hat{x}_d$ be the positive fixed point of $f_d(x)$, that is, $\hat{x}_d = f_d(\hat{x}_d)$. With the assumption that (β, γ, λ) is up-to-Δ unique, we claim that for any integral $1 \leq d < \Delta$,

$$(4.7) \qquad \tilde{\alpha}_d(x_d) \leq \tilde{\alpha}_d(\hat{x}_d)$$

Case 1: $\hat{x}_d \leq x_d$. Due to (4.5), we have $G'(\hat{x}_d) \geq 0$. Note that

$$G'(\hat{x}_d) = \frac{d(1 - \beta\gamma)(\gamma - \beta\hat{x}_d^2)\hat{x}_d^2}{(\beta\hat{x}_d + 1)^3(\hat{x}_d + \gamma)^3}$$
$$\cdot \left(\frac{1}{d(1 - \beta\gamma)\hat{x}_d} - \frac{1}{(\beta\hat{x}_d + 1)(\hat{x}_d + \gamma)} \right).$$

Because of the uniqueness, we have $|f_d'(\hat{x}_d)| = \frac{d(1 - \beta\gamma)\hat{x}_d}{(\beta\hat{x}_d + 1)(\hat{x}_d + \gamma)} < 1$, thus $\frac{1}{d(1-\beta\gamma)\hat{x}_d} - \frac{1}{(\beta\hat{x}_d + 1)(\hat{x}_d + \gamma)} > 0$. Therefore $G'(\hat{x}_d) \geq 0$ implies that $\gamma - \beta\hat{x}_d^2 \geq 0$. Since $f_d(x)$ is monotonically decreasing in x and $\hat{x}_d$ is its fixed point, we have

$$\gamma - \beta f_d(x_d)^2 \geq \gamma - \beta f_d(\hat{x}_d)^2$$
$$= \gamma - \beta\hat{x}_d^2$$
$$\geq 0.$$

Since x_d satisfies (4.4), $\gamma - \beta x_d^2$ and $\gamma - \beta f_d(x_d)^2$ must be simultaneously positive or negative, thus it also holds that $\gamma - \beta x_d^2 \geq 0$. Then both $\frac{(\gamma - \beta x^2)}{(\beta x + 1)(x + \gamma)}$ and $\frac{f_d(x)}{(\gamma - \beta f_d(x)^2)}$ are positive and monotonically decreasing in $x \in [\hat{x}_d, x_d]$. Therefore, $\tilde{\alpha}_d(x_d) \leq \tilde{\alpha}_d(\hat{x}_d)$.

Case 2: $\hat{x}_d > x_d$. By the same argument as above, it holds that $\gamma - \beta f_d(\hat{x}_d)^2 = \gamma - \beta\hat{x}_d^2 < 0$, $\gamma - \beta f_d(x_d)^2 < 0$, and $\gamma - \beta x_d^2 < 0$. Thus both $\frac{(\gamma - \beta x^2)}{(\beta x + 1)(x + \gamma)}$ and $\frac{f_d(x)}{(\gamma - \beta f_d(x)^2)}$ are negative and monotonically decreasing in $x \in [x_d, \hat{x}_d]$, hence their product is positive and increasing in $x \in [x_d, \hat{x}_d]$. Therefore, $\tilde{\alpha}_d(x_d) \leq \tilde{\alpha}_d(\hat{x}_d)$.

Combining (4.6) and (4.7), we have

$$\alpha_d(x) \leq \alpha_d(x_d) = \tilde{\alpha}_d(x_d) \leq \tilde{\alpha}_d(\hat{x}_d)$$
$$= \sqrt{\frac{d(1 - \beta\gamma)\hat{x}_d}{(\beta\hat{x}_d + 1)(\hat{x}_d + \gamma)}} = \sqrt{|f_d'(\hat{x}_d)|}.$$

And due to Lemma 2.1, if (β, γ, λ) is up-to-Δ unique then there exists a constant $\alpha < 1$ such that $|f_d'(\hat{x}_d)| < \alpha$ for all integer $1 \leq d < \Delta$. Lemma 4.4 is proved.

Lemma 4.2 is proved by combining Lemma 4.3 and Lemma 4.4.

Proof of Theorem 4.1. Let $T = T_{\text{SAW}}(G, v)$ for a G whose maximum degree is at most Δ. Then the maximum degree of T is at most Δ, thus the root v has at most Δ children in T, and every other vertex in T has less than Δ children. We recursively construct R_u, δ_u and ϵ_u for every subtree in T.

Let $t = \text{dist}(v, S)$. By repeatedly applying Item 3 of Lemma 4.1, without loss of generality, we have a path $u_1 u_2 \cdots u_{t-2}$ in T with $u_1 = v$ such that $\epsilon_{u_j} \leq \alpha(d_j; x_1, \ldots, x_{d_j}) \cdot \epsilon_{u_{j+1}}$ for $j = 1, 2, \ldots, t-3$, where d_j is the number of children of u_j and $x_i \in [0, \infty)$, $1 \leq i \leq d_j$.

Note that $d_1 \leq \Delta$, and $d_j < \Delta$ for all other j. Assume that (β, γ, λ) is up-to-Δ unique. If Δ is bounded, then by Lemma 4.2 there exists a constant $\alpha < 1$, such that $\epsilon_{u_j} \leq \alpha \cdot \epsilon_{u_{j+1}}$ for $2 \leq j \leq t-3$, and $\epsilon_v \leq d_1 \cdot \epsilon_{u_2} \leq \Delta \cdot \epsilon_2$ due to (4.2); and if $\Delta = \infty$, then by Lemma 4.2, $\epsilon_{u_j} \leq \alpha \cdot \epsilon_{u_{j+1}}$ for all $1 \leq j \leq t-3$. In both cases we have $\epsilon_v = O(\alpha^t \cdot \epsilon_{u_{t-2}})$.

Due to Item 1 of Lemma 4.1, $\delta_v = \frac{\epsilon_v}{\Phi(\widetilde{R})} = O\left(\frac{1}{\Phi(\widetilde{R})} \cdot \alpha^t \epsilon_{u_{t-2}}\right)$ for some $\widetilde{R} \in [R_v, R_v + \delta_v]$. We then bound $\Phi(\widetilde{R})$ and $\epsilon_{u_{t-2}}$. Due to Item 2 of Lemma 3.1, the fact that (β, γ, λ) is up-to-Δ unique implies that either Δ is bounded or $\gamma > 1$. Note that v must be free or the theorem is trivial to prove, and none of u_{t-2}'s children is in S because $\text{dist}(v, S) = t$. Thus by Item 2 of Lemma 4.1, $\epsilon_{u_{t-2}} = O(1)$ and $\widetilde{R} \leq R_v + \delta_v = O(1)$, which implies that $\Phi(\widetilde{R}) = \frac{1}{\sqrt{\widetilde{R}(\beta\widetilde{R} + 1)(\widetilde{R} + \gamma)}} = \Omega(1)$.

In conclusion, if (β, γ, λ) is up-to-Δ unique, there exists a constant $\alpha < 1$, such that $\delta_v = O(\alpha^t)$. As discussed in the beginning of section 4, this proves Theorem 4.1.

Strong spatial mixing on regular trees. As a byproduct of our analysis, we prove a strong spatial mixing theorem for regular trees. When the graphs G itself is a regular tree. All vertices (except the root) has the same arity. And all ds (excerpt the one of the root) that appear in the proof are the same and equal that arity. Then the condition that the uniqueness holds on all infinite regular trees of degree up to Δ can be replaced by the uniqueness on infinite Δ-regular tree.

THEOREM 4.2. *For two-state anti-ferromagnetic spin systems, on any infinite Δ-regular tree the uniqueness implies the strong spatial mixing.*

The same result can be obtained by combining the same theorem for the hardcore model [25] and anti-ferromagnetic Ising model [20] and translating the parameters of general two-state anti-ferromagnetic spin systems to these models (as discussed in [20,23]). However, unlike the hardcore and the anti-ferromagnetic

Ising models, for general two-state anti-ferromagnetic spin systems Theorem 4.2 itself is not sufficient to imply the strong spatial mixing on graphs of maximum degree at most Δ. This is discussed in Section 6.

5 Algorithmic implications

In this section we prove Theorem 1.2. That is, if an anti-ferromagnetic (β, γ, λ) is up-to-Δ unique then there exists an FPTAS for the partition function $Z(G)$ for any graph G of maximum degree at most Δ, and in particular the universal uniqueness implies an FPTAS for arbitrary graph G.

It is well-known that $Z(G)$ can be computed from $p_v^{\sigma_\Lambda}$ by the following standard procedure. Let $v_1, \ldots, v_n$ enumerate the vertices in G. For $0 \leq i \leq n$, let σ_i be the configuration fixing the first i vertices $v_1, \ldots, v_i$ as follows: $\sigma_i(v_j) = \sigma_{i-1}(v_j)$ for $1 \leq j \leq i-1$ and $\sigma_i(v_i)$ is fixed so that $p_i \triangleq \Pr[\sigma_i(v_i) \mid \sigma_{i-1}] \geq 1/3$. In particular, $\sigma_n \in \{0,1\}^V$ is a configuration of V. It holds for the Gibbs measure of σ_n that $\rho(\sigma_n) = p_1 p_2 \cdots p_n$ as well as that $\rho(\sigma_n) = \frac{w(\sigma_n)}{Z(G)}$, thus $Z(G) = \frac{w(\sigma_n)}{p_1 p_2 \cdots p_n}$, where the weight $w(\sigma_n) = \prod_{(u,v) \in E} A_{\sigma_n(u), \sigma_n(v)} \prod_{v \in V} b_{\sigma_n(v)}$ can be computed precisely for any particular σ_n in time polynomial in n. Note that p_i equals to either $p_{v_i}^{\sigma_{i-1}}$ or $1 - p_{v_i}^{\sigma_{i-1}}$. Therefore, if $p_v^{\sigma_\Lambda}$ can be approximated within an additive error ϵ in time polynomial in n and $\frac{1}{\epsilon}$, then the configurations σ_i can be efficiently constructed such that all p_i are bounded away from 0, thus the product $p_1 p_2 \cdots p_n$ can be approximated within a factor of $(1 \pm n\epsilon)$ in time polynomial in n and $\frac{1}{\epsilon}$, which implies an FPTAS for $Z(G)$.

Bounded degree graphs. Let G be a graph whose maximum degree is at most Δ and v be any vertex. A self-avoiding walk tree $T = T_{\mathrm{SAW}}(G, v)$ can be constructed so that $R_{G,v}^{\sigma_\Lambda} = R_T^{\sigma_\Lambda}$. We can use the recursive procedure described in Section 4 to compute the upper and lower bounds of $R_T^{\sigma_\Lambda}$, with the setting that for all the vertices more than t steps away from the root v, the trivial bounds $0 \leq R_T^{\sigma_\Lambda} \leq \infty$ is used. Then the proof of Theorem 4.1 shows that the recursive procedure returns R_0 and R_1 such that $R_0 \leq R_T^{\sigma_\Lambda} \leq R_1$, and $R_1 - R_0 = O(\alpha^t)$ for some constant $\alpha < 1$ assuming that (β, γ, λ) is up-to-Δ unique. Note that $R_T^{\sigma_\Lambda} = R_{G,v}^{\sigma_\Lambda} = \frac{p_v^{\sigma_\Lambda}}{1 - p_v^{\sigma_\Lambda}}$. Let $p_0 = \frac{R_0}{R_0+1}$ and $p_1 = \frac{R_1}{R_1+1}$. Then $p_0 \leq p_v^{\sigma_\Lambda} \leq p_1$ and

$$(5.8) \qquad p_1 - p_0 = \frac{R_1}{R_1+1} - \frac{R_0}{R_0+1} \leq R_1 - R_0 = O(\alpha^t).$$

The recursive procedure runs in time $O(\Delta^t)$ since it only needs to construct the first t levels of the self-avoiding

walk tree. If Δ is bounded, by setting $t = \log_\alpha \epsilon$, this gives an algorithm which approximates $p_v^{\sigma_\Lambda}$ within an additive error $O(\epsilon)$ in time polynomial in n and $\frac{1}{\epsilon}$, which implies an FPTAS for $Z(G)$.

Arbitrary graphs. Let G be an arbitrary graph and v be any vertex. Let $T = T_{\mathrm{SAW}}(G, v)$. We use the method of *Computationally Efficient Correlation Decay* introduced in [15] to deal with the unbounded degrees. Intuitively, using this method we observe correlation decay in a refined metric instead of graph distance such that in this new metric a neighborhood of moderate size is sufficient to guarantee desirable correlation decay.

Similarly, we use the recursive procedure described in Section 4 to compute the upper and lower bounds of $R_T^{\sigma_\Lambda}$, but this time the termination condition relies on a new depth defined as follows.

DEFINITION 5.1. *Let T be a rooted tree and $M > 1$ be a constant. For any vertex v in T, define the M-based depth of v, denoted $\ell_M(v)$, as such: $\ell_M(v) = 0$ if v is the root, and $\ell_M(v) = \ell_M(u) + \lceil \log_M(d+1) \rceil$ if v is one of the d children of u.*

Let $M > 1$ to be fixed. Denote by $B(\ell)$ the set of all vertices with M-based depth $< \ell$ along with their children and grandchildren in T. It can be verified by induction that $|B(\ell)| \leq n^2 M^\ell$. The recursion is applied to estimate the $R_T^{\sigma_\Lambda}$ when the current $v \in B(\ell)$ until v is no longer in $B(\ell)$ in which case the trivial bounds $0 \leq R_T^{\sigma_\Lambda} \leq \infty$ is used.

Let ϵ_v be defined as in Section 4. Repeatedly applying Item 3 of Lemma 4.1, without loss of generality, we have a path $u_1 u_2 \cdots u_k$ in T from the root $u_1 = v$ to a u_k with $\ell_M(u_k) \geq \ell$ and $\ell_M(u_{k-1}) < \ell$, such that $\epsilon_{u_j} \leq \alpha(d_j; x_1, \ldots, x_{d_j}) \cdot \epsilon_{u_{j+1}}$ for $j = 1, 2, \ldots, k$, where d_j is the number of children of u_j and $x_i \in [0, \infty)$, $1 \leq i \leq d_j$. The key to overcome the explosion caused by unbounded degrees is to observe that the contraction $\alpha(d; x_1, \ldots, x_d)$ decreases dramatically as the degree d grows.

LEMMA 5.1. *Let (β, γ, λ) be anti-ferromagnetic. If (β, γ, λ) satisfies the universal uniqueness, then there exist constants $\alpha < 1$ and $M > 1$ such that for any integer $d \geq 1$, and any $x_i \in [0, \infty)$, $1 \leq i \leq d$, it holds that $\alpha(d; x_1, ..., x_d) \leq \alpha^{\lceil \log_M(d+1) \rceil}$.*

Proof. Assume the universal uniqueness of (β, γ, λ). Due to Lemma 4.2, there exists a constant $\alpha < 1$ such that $\alpha(d; x_1, \ldots, x_d) \leq \alpha$. By Item 2 of Lemma 3.1, the universal uniqueness implies that $\gamma > 1$, thus there exists a constant $M > 1$ such that $d \cdot \sqrt{\frac{\lambda}{\gamma^{d+1}}} \leq \alpha^{\lceil \log_M(d+1) \rceil}$ for all $d \geq M$. Due to (4.3), it holds

that $\alpha(d; x_1, \ldots, x_d) \leq d \cdot \sqrt{\frac{\lambda}{\gamma^{d+1}}} \leq \alpha^{\lceil \log_M (d+1) \rceil}$ for all $d \geq M$. Note that $\alpha(d; x_1, \ldots, x_d) \leq \alpha$ means that $\alpha(d; x_1, ..., x_d) \leq \alpha^{\lceil \log_M (d+1) \rceil}$ for $d < M$. Therefore, $\alpha(d; x_1, ..., x_d) \leq \alpha^{\lceil \log_M (d+1) \rceil}$ for all d.

By Lemma 5.1, there exist constants $\alpha < 1$ and $M > 1$ such that

$$\begin{aligned}
\epsilon_v &\leq \epsilon_{u_k} \cdot \prod_{j=1}^{k} \alpha^{\lceil \log_M (d_j+1) \rceil} \\
&\leq \epsilon_{u_k} \cdot \alpha^{\sum_{j=1}^{k} \lceil \log_M (d_j+1) \rceil} \\
&= \epsilon_{u_k} \cdot \alpha^{\ell_M(u_k)} \\
&\leq \epsilon_{u_k} \cdot \alpha^{\ell}.
\end{aligned}$$

With the notation used in Section 4, S is the complement of $B(\ell)$. Note that all u_k's children are in $B(\ell)$ thus none of them are in S, and by Item 2 of Lemma 3.1 the universal uniqueness implies that $\gamma > 1$. Thus by Item 2 of Lemma 4.1 it holds that $\epsilon_{u_k} = O(1)$. Therefore, $\epsilon_v \leq \epsilon_{u_k} \cdot \alpha^{\ell} = O(\alpha^{\ell})$.

Let $\delta_v = R_1 - R_0$, where R_0 and R_1 are the bounds returned by the recursive procedure such that $R_0 \leq R_T^{\sigma_\Lambda} \leq R_1$. By the same analysis as in Section 4, $\delta_v = \frac{\epsilon_v}{\Phi(\widetilde{R})} = O(\epsilon_v) = O(\alpha^{\ell})$. Then by (5.8), the marginal probability $p_v^{\sigma_\Lambda}$ is approximated within an additive error of $O(\alpha^{\ell})$. The running time of the recursion is $O(nB(\ell)) = O(n^3 M^{\ell})$. By setting $\ell = \log_\alpha \epsilon$, we have an algorithm which approximates $p_v^{\sigma_\Lambda}$ within an additive error of $O(\epsilon)$ in time polynomial in n and $\frac{1}{\epsilon}$, which implies an FPTAS for $Z(G)$ for arbitrary graph G.

Heterogeneous spin systems. Our analysis in last and this sections actually holds for heterogeneous spin systems which allow that each vertex v has a distinct constant external field $\lambda_v > 0$.

THEOREM 5.1. *For a two-state anti-ferromagnetic heterogeneous spin system with parameters β, γ, and external field λ_v at each vertex v, for any finite $\Delta \geq 2$ or $\Delta = \infty$, if for all v the $(\beta, \gamma, \lambda_v)$ is up-to-Δ unique then the spin system exhibits strong spatial mixing and has FPTAS on graphs of maximum degree at most Δ.*

6 Non-monotonicity of general two-state spin system

In this section, we prove the following theorem, which shows a non-monotone behavior of the strong spatial mixing of general two-state spin systems.

THEOREM 6.1. *There exist two-state anti-ferromagnetic spin systems which exhibit strong*

spatial mixing on infinite d-regular tree but do not exhibit strong spatial mixing on all graphs of maximum degree at most d.

In a seminal work [25], Weitz proved that for the hardcore model the strong spatial mixing on an infinite d-regular tree implies the strong spatial mixing on graphs of maximum degree at most d (Theorem 2.3 in [25]). He further remarked that the same implication holds for all two-state spin systems. That is,

> for any two-state spin system, strong spatial mixing on an infinite d-regular tree implies the strong spatial mixing on graphs of maximum degree at most d.

This claim played important roles in current understanding of correlation decay in two-state spin systems as well as devising FPTAS for such systems. An algorithmic form of this claim was cited in [20] as a theorem for all two-state spin systems (Theorem 2.8 in [20]) and was proved for the anti-ferromagnetic Ising model. It was raised as a conjecture in [21] whether this claim holds for multi-state spin systems.

Here we clarify that this claim holds only for the two-state spin systems under limited settings but does not hold for all general two-state spin systems. This disproves the conjecture in [21] and shows that the common belief that the d-regular tree represents the worst case for strong spatial mixing among all graphs of maximum degree at most d cannot be generalized to general two-state spin systems. We first describe a region that the claim is true.

LEMMA 6.1. *For $0 \leq \beta, \gamma \leq 1$, the strong spatial mixing on infinite d-regular tree implies the strong spatial mixing on trees of maximum degree at most d.*

Proof. Given a rooted tree of maximum degree at most d, for each vertex of k children with $k < d - 1$, we can attach $d-1-k$ dummy children with fixed (distributions of) spin states. This is the method used in [25] and [20]: for the hardcore model the dummy children are fixed to be unoccupied and for the anti-ferromagnetic Ising model the dummy children are of uniform distributions over spin states. In both cases, the dummy children have no effect on their parent. In general, we fix the distribution to be (p_0, p_1) at each dummy child satisfying

$$\begin{aligned}
p_0 + p_1 &= 1, \\
\beta p_0 + p_1 &= p_0 + \gamma p_1.
\end{aligned}$$

When $0 \leq \beta, \gamma \leq 1$, this system has solutions in $0 \leq p_0, p_1 \leq 1$. With the ratio R_i at the i-th child,

$1 \leq i \leq k$, and $R_i = p_0/p_1$ for the dummy children $k < i \leq d-1$, the ratio at the parent is given by the recursion

$$\lambda \prod_{i=1}^{d-1} \frac{\beta R_i + 1}{R_i + \gamma} = \lambda \prod_{i=1}^{k} \frac{\beta R_i + 1}{R_i + \gamma},$$

which is identical to the original quantity.

Due to the self-avoiding walk tree construction (Theorem 2.1), it holds that for $0 \leq \beta, \gamma \leq 1$, strong spatial mixing on infinite d-regular tree implies the strong spatial mixing and the FPTAS on graphs of maximum degree at most d.

For $0 \leq \beta, \gamma \leq 1$, the spin system shows the following monotone property: the uniqueness on infinite d-regular tree implies the uniqueness on all infinite regular trees of smaller degree. This can be verified by the following reasoning: due to Theorem 4.2, on infinite d-regular tree the uniqueness implies the strong spatial mixing, which for $0 \leq \beta, \gamma \leq 1$, implies the strong spatial mixing (including the uniqueness) on all infinite regular trees of smaller degree.

There exist two-state anti-ferromagnetic spin systems which are non-monotone. We can choose anti-ferromagnetic (β, γ, λ) satisfying that $\gamma > 1$ and $\lambda > \lambda_c(\beta, \gamma)$, where $\lambda_c(\beta, \gamma)$ is the critical threshold for universal uniqueness given in Item 8 of Lemma 3.1. According to Item 3 of Lemma 3.1, since $\gamma > 1$ the uniqueness holds on d-regular trees for all sufficiently large d. On the other hand, according to Item 8 of Lemma 3.1, since $\lambda > \lambda_c(\beta, \gamma)$, there exists a finite d' such that the system is non-unique on d'-regular tree.

For such non-monotone systems, due to Theorem 4.2, the uniqueness implies the strong spatial mixing on d-regular tree for sufficiently large d, but the strong spatial mixing does not hold on a regular tree of smaller degree (because of the non-uniqueness on the smaller tree). Therefore the implication between the strong spatial mixing on d-regular tree and on graphs of maximum degree at most d does not hold for general two-state spin systems.

Acknowledgment. We would like to thank Alistair Sinclair and Piyush Srivastava for the helpful comments to an early version of this paper.

References

[1] A. Bandyopadhyay and D. Gamarnik. Counting without sampling: Asymptotics of the log-partition function for certain statistical physics models. *Random Struct. Algorithms*, 33:452–479, 2008.

[2] M. Bayati, D. Gamarnik, D. Katz, C. Nair, and P. Tetali. Simple deterministic approximation algorithms for counting matchings. In *Proceedings of the 39th annual ACM Symposium on Theory of Computing (STOC)*, pages 122–127, 2007.

[3] M. Dyer, M. Jerrum, and E. Vigoda. Rapidly mixing markov chains for dismantleable constraint graphs. In *Proceedings of a DIMACS/DIMATIA Workshop on Graphs, Morphisms and Statistical Physics*, 2001.

[4] M. E. Dyer, A. M. Frieze, and M. Jerrum. On counting independent sets in sparse graphs. *SIAM Journal on Computing*, 31(5):1527–1541, 2002.

[5] M. E. Dyer and C. S. Greenhill. On markov chains for independent sets. *Journal of Algorithms*, 35(1):17–49, 2000.

[6] A. Galanis, Q. Ge, D. Štefankovič, E. Vigoda, and L. Yang. Improved inapproximability results for counting independent sets in the hard-core model. In *Proceedings of the 15th International Workshop on Randomization and Computation (RANDOM)*, pages 567–578, 2011.

[7] A. Galanis, D. Štefankovič, and E. Vigoda. Inapproximability of the partition function for the antiferromagnetic ising and hard-core models. *Arxiv preprint arXiv:1203.2226*, 2012.

[8] D. Gamarnik and D. Katz. Correlation decay and deterministic fptas for counting list-colorings of a graph. In *Proceedings of the 18th annual ACM-SIAM Aymposium on Discrete Algorithms (SODA)*, pages 1245–1254, 2007.

[9] L. A. Goldberg and M. Jerrum. A polynomial-time algorithm for estimating the partition function of the ferromagnetic ising model on a regular matroid. In *Proceedings of the 38th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 521–532, 2011.

[10] L. A. Goldberg, M. Jerrum, and M. Paterson. The computational complexity of two-state spin systems. *Random Structures and Algorithms*, 23(2):133–154, 2003.

[11] M. Jerrum. A very simple algorithm for estimating the number of k-colorings of a low-degree graph. *Random Structures and Algorithms*, 7(2):157–166, 1995.

[12] M. Jerrum and A. Sinclair. Polynomial-time approximation algorithms for the ising model. *SIAM Journal on Computing*, 22(5):1087–1116, 1993.

[13] M. Jerrum, A. Sinclair, and E. Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *Journal of the ACM*, 51:671–697, 2004.

[14] F. Kelly. Stochastic models of computer communication systems. *Journal of the Royal Statistical Society. Series B*, 47(3):379–395, 1985.

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

[16] M. Luby and E. Vigoda. Approximately counting up

to four (extended abstract). In *Proceedings of the twenty-ninth annual ACM Symposium on Theory of Computing (STOC)*, pages 682–687, 1997.

[17] F. Martinelli, A. Sinclair, and D. Weitz. Fast mixing for independent sets, colorings, and other models on trees. *Random Structures and Algorithms*, 31(2):134–172, 2007.

[18] E. Mossel, D. Weitz, and N. Wormald. On the hardness of sampling independent sets beyond the tree threshold. *Probability Theory and Related Fields*, 143:401–439, 2009.

[19] R. Restrepo, J. Shin, P. Tetali, E. Vigoda, and L. Yang. Improved mixing condition on the grid for counting and sampling independent sets. In *Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 140–149, 2011.

[20] A. Sinclair, P. Srivastava, and M. Thurley. Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs. In *Proceedings of the 23rd annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 941–953, 2012.

[21] A. Sly. Uniqueness thresholds on trees versus graphs. *The Annals of Applied Probability*, 18(5):1897–1909, 2008.

[22] A. 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.

[23] A. Sly and N. Sun. The computational hardness of counting in two-spin models on d-regular graphs. To apear in FOCS'12. *Arxiv preprint arXiv:1203.2602*, 2012.

[24] E. Vigoda. Improved bounds for sampling coloring. In *Proceedings of the 40th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 51–60, 1999.

[25] D. Weitz. Counting independent sets up to the tree threshold. In *Proceedings of the thirty-eighth annual ACM Symposium on Theory of Computing (STOC)*, pages 140–149, 2006.

A Heuristics for finding good potential functions

Perhaps the most mysterious step in our proof is the choice of the potential function $\Phi(x) = \frac{1}{\sqrt{x(\beta x+1)(x+\gamma)}}$. As in many cases where potential analysis is applied, there is no standard routine for searching for a suitable potential function. On the other hand, it is quite unlikely that we can just guess such a fairly complicated formula without any hints. Here we present a heuristic approach which leads us to the discovery of a good potential function. This part is not rigorous and logically unnecessary for the soundness of our result. Nevertheless, this heuristic approach is general and interesting enough and may find its applications in other scenarios, thus deserves an exposition here.

The heuristics consists of three steps:

1. Find a necessary condition for the potential function, which is an equation related to the potential function at one point.

2. (heuristic step) Enhance the condition by assuming that the equation holds for the whole range of the variable, which gives us a differential equation.

3. Solve the differential equation and get a potential function.

We first assume that the system is at the boundary of uniqueness. This means that d is the critical degree and $f'(\hat{x}) = -1$, where $f(x) = \lambda \left(\frac{\beta x+1}{x+\gamma} \right)^d$ and $\hat{x} = f(\hat{x})$ is the positive fixed point.

We have the following identities:

$$\hat{x} = \lambda \left(\frac{\beta \hat{x}+1}{\hat{x}+\gamma} \right)^d \text{ and } 1 = \lambda d(1-\beta\gamma)\frac{(\beta \hat{x}+1)^{d-1}}{(\hat{x}+\gamma)^{d+1}},$$

which together implies another identity

$$(A.1) \qquad d(1-\beta\gamma)\hat{x} = (\beta\hat{x}+1)(\hat{x}+\gamma).$$

The goal is to find a potential function such that $\alpha(x) = |f'(x)|\frac{\Phi(f(x))}{\Phi(x)} \leq 1$ for all x. On the other hand, we have $\alpha(\hat{x}) = |f'(\hat{x})|\frac{\Phi(f(\hat{x}))}{\Phi(\hat{x})} = 1 \cdot \frac{\Phi(\hat{x})}{\Phi(\hat{x})} = 1$. So $\alpha(x)$ achieves its maximum when $x = \hat{x}$. As a differentiable function, it holds that $\alpha'(\hat{x}) = 0$, that is

$$\left[f'(x) \cdot \frac{\Phi\left(f(x)\right)}{\Phi(x)} \right]' \Bigg|_{x=\hat{x}} = 0.$$

We have the following calculation:

$$\left[f'(x) \cdot \frac{\Phi(f(x))}{\Phi(x)}\right]'\Bigg|_{x=\hat{x}} = 0$$

$$\Leftrightarrow \ [f'(x)\Phi(f(x))]'\,\Phi(x)\Big|_{x=\hat{x}}$$
$$= f'(x)\Phi(f(x))\Phi'(x)|_{x=\hat{x}}$$

$$\Leftrightarrow \ [f''(\hat{x})\Phi(f(\hat{x})) + f'(\hat{x})\Phi'(f(\hat{x}))f'(\hat{x})]\,\Phi(\hat{x})$$
$$= f'(\hat{x})\Phi(f(\hat{x}))\Phi'(\hat{x})$$

$$\Leftrightarrow \ f''(\hat{x})\Phi(\hat{x}) + \Phi'(\hat{x}) = -\Phi'(\hat{x})$$

$$\Leftrightarrow \ -\frac{f''(\hat{x})}{2} = \frac{\Phi'(\hat{x})}{\Phi(\hat{x})} = (\ln(\Phi(\hat{x})))',$$

where in the above calculation we use the facts that $\hat{x} = f(\hat{x})$ and $f'(\hat{x}) = -1$.

Taking the second derivative of $f(x)$, we have

$$f''(x) = \lambda d(\beta\gamma - 1)\frac{(\beta x + 1)^{d-2}}{(x+\gamma)^{d+2}}$$
$$\cdot ((d-1)\beta(x+\gamma) - (d+1)(\beta x + 1)).$$

This equation already gives an identity for Φ. But it is too complicated. We may further simplify the above formula at the point $x = \hat{x}$.

$$\begin{aligned}
f''(\hat{x}) &= \lambda d(\beta\gamma - 1)\frac{(\beta\hat{x} + 1)^{d-2}}{(\hat{x}+\gamma)^{d+2}}\\
&\quad \cdot ((d-1)\beta(\hat{x}+\gamma) - (d+1)(\beta\hat{x} + 1))\\
&= \frac{(d+1)(\beta\hat{x} + 1) - (d-1)\beta(\hat{x}+\gamma)}{(\beta\hat{x}+1)(\hat{x}+\gamma)}\\
&= \frac{d+1}{\hat{x}+\gamma} - \frac{(d-1)\beta}{\beta\hat{x}+1}\\
&= \frac{d(1-\beta\gamma)}{(\beta\hat{x}+1)(\hat{x}+\gamma)} + \frac{1}{\hat{x}+\gamma} + \frac{\beta}{\beta\hat{x}+1}\\
&= \frac{1}{\hat{x}} + \frac{1}{\hat{x}+\gamma} + \frac{\beta}{\beta\hat{x}+1}.
\end{aligned}$$

So we have

$$\begin{aligned}
(\ln(\Phi(\hat{x})))' &= -\frac{f''(\hat{x})}{2}\\
\text{(A.2)} \qquad &= -\frac{1}{2}\left(\frac{1}{\hat{x}} + \frac{1}{\hat{x}+\gamma} + \frac{\beta}{\beta\hat{x}+1}\right).
\end{aligned}$$

We apply the heuristics and assume that (A.2) simply holds for all x. This gives us a differential equation

$$(\ln(\Phi(x)))' = -\frac{1}{2}\left(\frac{1}{x} + \frac{1}{x+\gamma} + \frac{\beta}{\beta x + 1}\right).$$

The solution of this differential equation is

$$\ln(\Phi(x)) = -\frac{1}{2}\ln(x(x+\gamma)(\beta x + 1)) + C_1,$$

which gives that

$$\Phi(x) = \frac{C_2}{\sqrt{x(\beta x + 1)(x+\gamma)}},$$

where C_1, C_2 are some constants. This gives the potential function we used in the paper.

There are also other equations which hold for the fixed point $\hat{x}$. Choosing which equation for $\hat{x}$ to heuristically extend to all x may affect the potential function we obtained. For example, (A.1) implies that

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

thus we can rewrite $f''(\hat{x})$ as

$$f''(\hat{x}) = \frac{1}{\hat{x}} + \frac{1}{\hat{x}+\gamma} + \frac{\beta}{\beta\hat{x}+1} = \frac{d+1}{d\hat{x}} + \frac{2\beta}{\beta\hat{x}+1}.$$

This gives that

$$(\ln(\Phi(\hat{x})))' = -\frac{d+1}{2d\hat{x}} - \frac{\beta}{\beta\hat{x}+1}.$$

We can also treat this equation as a differential equation for variable x and solving it gives us

$$\Phi(x) = \frac{c}{x^{\frac{d+1}{2d}}(\beta x + 1)},$$

which is the potential function used in [15].