APPROXIMATION ALGORITHMS FOR TWO-STATE ANTI-FERROMAGNETIC SPIN SYSTEMS ON BOUNDED DEGREE GRAPHS

ALISTAIR SINCLAIR, PIYUSH SRIVASTAVA, AND MARC THURLEY

Abstract. In a seminal paper [10], Weitz gave a deterministic fully polynomial approximation scheme for counting exponentially weighted independent sets (which is the same as approximating the partition function of the hard-core model from statistical physics) in graphs of degree at most d, up to the critical activity for the uniqueness of the Gibbs measure on the infinite d-regular tree. More recently Sly [8] (see also [1]) showed that this is optimal in the sense that if there is an FPRAS for the hard-core partition function on graphs of maximum degree d for activities larger than the critical activity on the infinite d-regular tree then NP = RP. In this paper we extend Weitz's approach to derive a deterministic fully polynomial approximation scheme for the partition function of general two-state anti-ferromagnetic spin systems on graphs of maximum degree d, up to the corresponding critical point on the d-regular tree. The main ingredient of our result is a proof that for two-state anti-ferromagnetic spin systems on the d-regular tree, weak spatial mixing implies strong spatial mixing. This in turn uses a message-decay argument which extends a similar approach proposed recently for the hard-core model by Restrepo *et al* [7] to the case of general two-state anti-ferromagnetic spin systems.

1. Introduction

1.1. Background.

Spin systems are a general framework for modeling nearest-neighbor interactions on graphs, and are widely studied in both statistical physics and applied probability. A spin system consists of a large collection of nodes, each of which may be in one of a fixed number of states called *spins*. A neighborhood structure is specified by edges between the nodes. Interactions between neighboring nodes are determined by *edge potentials*, which assign an energy value to each edge based on the spin values of its endpoints. In addition, there are *vertex potentials* which assign an energy value to each node based on the value of its spin. For any configuration σ of spins on the nodes, the energy $H(\sigma)$ is

The first author was supported in part by NSF grant CCF-1016896.

The second author was supported by the Berkeley Fellowship for Graduate Study and NSF grant CCF-1016896. Part of this work was done when this author was a research intern at Microsoft Research India.

The third author was supported in part by a postdoctoral fellowship of the German Academic Exchange Service (DAAD) and by Marie Curie Intra-European Fellowship 271959. Part of this work was done while this author was a postdoctoral scholar at UC Berkeley.

just the sum of its edge and vertex energies. Based on the Gibbs formalism from statistical physics, the probability of finding the system in configuration σ is then proportional to the weight $w(\sigma) = \exp(-H(\sigma))$.

In this paper, we concentrate on *two-state* spin systems, where each vertex can be in one of two states, referred to as "+" and "−". Such a system can be defined by specifying a $(+,+)$ *edge activity* β, a $(-,-)$ edge activity γ, and a *vertex activity* λ, where β, γ and λ are non-negative parameters. For a graph $G(V,E)$, a *configuration* $\sigma : V \mapsto \{+\ ,-\ \}$ is an assignment of $+$ and $-$ spins to the vertices of G. The weight $w(\sigma)$ of the configuration σ is then given by

$$w(\sigma) = \lambda^{m(\sigma)}\beta^{n_+(\sigma)}\gamma^{n_-(\sigma)}, \tag{1}$$

where $m(\sigma)$ denotes the number of vertices assigned spin $-$, and $n_+(\sigma)$ (respectively, $n_-(\sigma)$) denotes the number of edges for which both endpoints are assigned spin $+$ (respectively, $-$). The *partition function* of the model is defined as

$$Z = \sum_{\sigma \in \{+,-\}^V} w(\sigma).$$

The partition function, in addition to being a natural weighted generalization of the notion of counting, is a fundamental quantity in statistical physics. For example, it is the normalizing factor in the Gibbs distribution: the probability of occurrence of configuration σ is given by $\mu(\sigma) = w(\sigma)/Z$. In addition, many other properties of the model can be deduced by studying the partition function [2].

As a simple concrete example of a two-state spin system, consider the setting $\beta = 1$ and $\gamma = 0$ (so that configurations with adjacent "−" spins are assigned weight zero, and thus prohibited), known as the *hard-core* model. The associated Gibbs distribution is a weighted measure on independent sets in the graph G, in which any independent set U has weight $\lambda^{|U|}$. Another important class of examples, known as the *Ising model*[1], is obtained by setting $\beta = \gamma > 0$. There is an important qualitative difference between the Ising model with $\beta = \gamma > 1$ (the *ferromagnetic* case) and with $\beta = \gamma < 1$ (the *anti-ferromagnetic* case). The latter is an example of a "repulsive" model, which means that the edge potentials assign higher weights to edges with different spins at their endpoints, while the ferromagnetic case is "attractive" (higher weights are assigned to edges with the same spin at their endpoints). The parameter λ can be identified with an "external field", i.e., a bias associated with each spin. The case $\lambda = 1$ corresponds to zero field, while $\lambda < 1$ and $\lambda > 1$ correspond to positive and negative fields respectively. More generally, we will refer to any two-state system satisfying $\beta\gamma > 1$ as *ferromagnetic*, and any satisfying $\beta\gamma < 1$ as *anti-ferromagnetic*. Also, a model satisfying $\beta\gamma > 0$ is said to have *soft constraints* (in the sense that no combination of spin values at adjacent vertices is prohibited). In a sense to

[1]The description of the Ising model given here differs slightly from the more popular description in terms of edge and vertex potentials outlined in the first paragraph. However, translating between the two descriptions is easy; see appendix A.

be made precise later (see appendix A), Ising models capture arbitrary two-spin systems with soft constraints, and for this reason we will focus mainly on them.

The theory of spin systems derives in large part from considering the limiting behavior of the Gibbs distribution as the size of the underlying graph goes to infinity. Based on the above formalism for finite graphs, one may define a Gibbs measure μ on an infinite graph $\mathcal{G}$ by requiring that the marginal distribution on any finite subgraph $\mathcal{H}$, given the configuration on $\mathcal{G}\backslash\mathcal{H}$, is given by equation (1). (Here the spins in $\mathcal{G}\backslash\mathcal{H}$ act as a fixed *boundary condition* in (1).) It is a well known result in the statistical physics literature (see, for example, [2]) that at least one such distribution μ can always be defined. However, for certain values of the parameters of the spin system there can be multiple solutions μ, in which case the Gibbs measure is said to be *non-unique.*

We will now look at the phenomenon of non-uniqueness more closely in the special case when the infinite graph $\mathcal{G}$ is a d-ary tree.[2] Further, as already stated above, the anti-ferromagnetic Ising model essentially captures all two-state spin systems on regular graphs, and hence it is sufficient to consider the Ising case. Consider an anti-ferromagnetic Ising model on the d-ary tree with edge activity $\beta(=\gamma)$ and vertex activity λ. It turns out that if $\beta > \frac{d-1}{d+1}$, then the Gibbs measure is unique for all values of λ. In particular, in the zero-field case $\lambda = 1$, the Gibbs measure is unique if and only if $\beta > \frac{d-1}{d+1}$. However, when $\beta \leq \frac{d-1}{d+1}$, the Gibbs measure is no longer unique for all values of the vertex activity λ. In this case, there exists a *critical activity* $\lambda_c(\beta, d) \geq 1$ such that the Gibbs measure is unique if and only if $|\log \lambda| \geq \log \lambda_c(\beta, d)$.

The phenomenon of non-uniqueness of the Gibbs measure can also be described in terms of the more algorithmic notion of *decay of correlations.* We stick to our example of the infinite d-ary tree. Fix a vertex v in the tree, and let S_l be the set of vertices in the tree at distance at least l from v. Let $q_v(l, \sigma)$ be the probability of having spin $+$ at v conditional on the configuration on S_l being σ. It turns out that uniqueness of the Gibbs measure is equivalent to the condition that the inequality

$$|q_v(l, \sigma) - q_v(l, \tau)| \leq \exp(-\Omega(l)) \tag{2}$$

holds for any two configurations σ and τ on S_l.[3] The above condition is referred to in the literature as *weak spatial mixing.*

[2] We remark here that the infinite $(d+1)$-regular tree and the infinite d-ary tree show exactly the same behavior with respect to the uniqueness of the Gibbs measure. This follows immediately from the fact that the $(d+1)$-regular tree can be viewed as a root attached to the roots of $d+1$ infinite d-ary trees. We shall thus move freely between these two objects for ease of exposition throughout the paper.

[3] To be precise, this condition does not hold on the boundary of the uniqueness region, that is, for $|\log \lambda| = \log \lambda_c(\beta, d)$. We will focus mostly on the interior of this region, and by a slight abuse of terminology refer to it as the uniqueness region.

It has been believed for a long time (and proved in various manifestations) that there is an intimate relationship between weak spatial mixing and the running time of algorithms for approximating the associated partition function: roughly speaking, in the uniqueness region (when there is decay of correlations), the system should be amenable to local algorithms and thus be computationally tractable. One of the most spectacular results in this line of work was Weitz's fully polynomial deterministic approximation scheme (FPTAS) for the partition function of the hard-core model, which works on all graphs of degree at most $d+1$ for all activities λ less than the critical activity $\lambda_c(d)$ for the uniqueness of the Gibbs measure on the infinite d-ary tree [10]. This is even more interesting in light of another recent breakthrough due to Sly [8] (see also [1]), who showed that the existence of an FPRAS for the partition function of the hard-core model on graphs of degree at most $d+1$ for any activity larger than $\lambda_c(d)$ would imply that NP = RP. Thus the range of validity of Weitz's algorithm is optimal.

Weitz [10] gave a general two-step framework for designing deterministic algorithms for approximating partition functions of two-state spin systems. To describe this framework, we begin with the observation that in order to give an FPTAS for the partition function, it is sufficient to give an FPTAS for the probability of having spin + at any given vertex v. The first component of the framework is a combinatorial reduction, which shows that the problem of approximating this probability for a general two-state spin system on a graph of maximum degree $d+1$ can be reduced to the problem of approximating the same probability on a related finite subtree of the infinite $(d+1)$-regular tree rooted at v, in which the spins of some of the vertices are fixed to certain values. We emphasize that this is a model-independent reduction, and depends only upon the fact that the number of spin values is two. The associated tree, however, may be exponential in the size of the original graph, and thus one needs to show that it is sufficient to truncate the tree at a depth logarithmic in the size of the graph in order to obtain a good approximation. However, since some of the fixed vertices in the tree might be very close to the root v, it is not possible to argue using weak spatial mixing that a logarithmic depth of recursion suffices for approximating the partition function.

Accordingly, the second component of Weitz's framework is to establish that, for the spin system in question, weak spatial mixing on the infinite d-ary tree (or equivalently, on the infinite $(d+1)$-regular tree) is in fact equivalent to *strong spatial mixing*, which roughly states that the exponential decay of point-to-set correlations guaranteed by weak spatial mixing holds also when the spins at an arbitrary set of vertices are fixed to arbitrarily chosen values (see Section 2 for a precise definition). Weitz [10] established this for the hard core model, using a step-by-step comparison of ratios of occupation probabilities on the standard d-ary tree and on the modified tree with fixed vertices. It was claimed in [10] that such a result holds also for the anti-ferromagnetic Ising model, but to the best of our knowledge no proof of this fact (except in the special zero-field case ($\lambda = 1$); see [7, 11]) has so far been published.

1.2. **Contributions.** In this article, we give a proof for the fact that for the anti-ferromagnetic Ising model with any field, weak spatial mixing implies strong spatial mixing on the d-ary tree, using a message decay argument inspired by Restrepo *et al* [7]. Formally, we have the following theorem:

Theorem 1.1. *For any anti-ferromagnetic two-state spin system with soft constraints on the d-ary tree with $d \geq 2$, weak spatial mixing implies strong spatial mixing.*

Notice that it is easy to see that this holds also for the infinite $(d+1)$-regular tree, since the $(d+1)$-regular tree and the d-ary tree differ only in the degree of the root. Given Weitz's general reduction described above, we obtain as an almost immediate consequence an FPTAS for the partition function of the anti-ferromagnetic Ising model on graphs of maximum degree at most $d+1$ throughout the regime of uniqueness of Gibbs measure on the d-ary tree.

Corollary 1.2. *Let $d \geq 2$. Consider an anti-ferromagnetic Ising model with parameters β and λ. For β and λ in the interior of the uniqueness regime of the d-ary tree, every graph of degree at most $d+1$ exhibits strong spatial mixing. Moreover, for such β and λ, there is a deterministic polynomial time approximation scheme for the partition function of the associated spin system in graphs of degree at most $d+1$.*

By the translation described in appendix A, we can extend this result to general two-state anti-ferromagnetic spin systems. The difference is that the critical activity may now differ for vertices of different degrees. Let $\lambda_c(\beta, d)$ be the critical activity for phase transition of the anti-ferromagnetic Ising model described above (and defined formally in Section 2.2.1). Then, we have the following theorem:

Corollary 1.3. *Let $d \geq 2$. Consider an anti-ferromagnetic two-state spin system with parameters β, γ and λ. Let β' be the edge potential for the equivalent anti-ferromagnetic Ising model. Let $\mathcal{G}$ be the class of graphs with maximum degree $d+1$ in which every vertex v satisfies the condition*

$$|\log \lambda_v| \triangleq |\log \lambda + \frac{\text{degree}\,(v)}{2}(\log \beta - \log \gamma)| > \log \lambda_c(\beta', d).$$

Then every graph in $\mathcal{G}$ exhibits strong spatial mixing, and there is a deterministic polynomial time approximation scheme for the partition function of the associated spin system on graphs in the class $\mathcal{G}$. In particular, the class $\mathcal{G}$ includes all $(d+1)$-regular graphs when β, γ and λ are in the interior of the uniqueness regime of the d-ary tree.

We briefly sketch the approach we use to prove our main technical result, namely that weak spatial mixing implies strong spatial mixing (Theorem 1.1).

The idea is to come up with a "message" (i.e., an invertible function of the probability of a vertex having spin +) such that "disagreements" in the message decay by a constant fraction at each vertex of the tree. The challenge is to ensure that such a message can be designed for all points in the uniqueness region of the d-ary tree. For the special zero-field case of the anti-ferromagnetic Ising model (when $\lambda = 1$), such a message is well known [11]. However, this message does not work up to the threshold for general vertex potentials λ. Restrepo *et al* [7] recently derived a message which works up to the tree uniqueness threshold for the hard-core model. For the general anti-ferromagnetic Ising model, such a message turns out to be more complex than those known for the zero-field case and for the hard-core model. Our message is defined at the beginning of Section 3, and the requisite decay property is established in Section 4.

We conjecture also that our proof of strong spatial mixing based on stepwise decay of messages may lead to further consequences. For example, as shown by Restrepo *et al* [7], the message decay property can be used to extend Weitz's algorithm by exploiting the structure of special classes of graphs to obtain approximation algorithms beyond the tree threshold for those graphs. In addition, our proof demonstrates the versatility of the message approach introduced by Restrepo *et al*.

Remark: After obtaining our message decay proof, we received a sketch of Weitz's original unpublished proof [9]. It is interesting to note that that proof is quite different from ours, and employs a delicate two step analysis of the tree recursion described in Section 2. For reasons described in the previous paragraph, we believe that our message decay proof, in addition to being the first published version of this result, is potentially more robust and flexible than Weitz's approach; for example, it is not clear how to adapt Weitz's analysis to obtain stronger results for special classes of graphs such as lattices, as in [7].

1.3. **Related Work.** Our work is mainly motivated by the deterministic counting algorithm of Weitz [10], which was the first to show an interesting connection between the running time of an algorithm not related to Markov chain Monte Carlo, and the phase transition phenomenon for spin systems. On the complexity side, using the randomized gadget of Mossel, Weitz and Wormald [6], Sly [8] proved that if there is an FPRAS for the partition function of the hard-core model on graphs of degree at most d in the non-uniqueness regime of the d-regular tree, then NP = RP, thus showing that the range of validity of Weitz's algorithm is optimal. Technically Sly's result holds only sufficiently close to the phase transition; this restriction was mostly removed in a recent paper of Galanis *et al* [1]. For the case of unbounded degree graphs, Goldberg, Jerrum and Paterson [4] showed that approximating the partition function for the zero-field case ($\lambda = 1$) is NP-hard in the square $0 \le \beta, \gamma \le 1$.

A related problem is to get exponential lower bounds on the mixing time of any local Markov chain (Glauber dynamics) that samples from the hard-core and anti-ferromagnetic Ising models. Mossel, Weitz and Wormald [6] and Gerschenfeld and Montanari [3] showed that beyond the uniqueness threshold for d-regular trees, Glauber dynamics for these models can take exponential time to mix on d-regular graphs. In general, the running time of Glauber dynamics becomes exponential beyond the reconstruction threshold, and thus these results may be construed as establishing that there exist d-regular graphs on which the reconstruction problem is solvable beyond the uniqueness threshold for the d-regular tree [3].

On the algorithmic side, an analysis of Weitz's algorithm for the zero-field case of the anti-ferromagnetic Ising model appears in [11]. There has been some subsequent progress on the hard-core model on special classes of graphs too: recently, Restrepo *et al* [7] used a message decay proof to get improved strong spatial mixing thresholds on the 2D integer lattice for the hard core model. They achieved this by exploiting the special structure of self-avoiding walk trees obtained when Weitz's reduction is applied to the lattice. The message-decay proof turns out to be crucial in tightening the analysis to obtain strong spatial mixing over a wider range of parameters for these special trees. Much more is known about the ferromagnetic case: Jerrum and Sinclair [5] gave an FPRAS for the Ising model with arbitrary field on graphs of arbitrary degree, while Goldberg, Jerrum and Paterson [4] showed how to extend this to the whole of the ferromagnetic region $\beta\gamma > 1$ with $\lambda = 1$. The latter paper [4] also gave an FPRAS for the partition function on graphs of arbitrary degree for parts of the anti-ferromagnetic region $\beta\gamma < 1$. However, the results of [4] when restricted to bounded degree graphs do not hold throughout the uniqueness region and hence are incomparable to ours.

2. Preliminaries

2.1. **Notation.**

We will mostly follow the notational conventions of [4]. Given a graph $G = (V, E)$, a two-state spin configuration is defined as an assignment $\sigma\colon V \mapsto \{+\,,-\,\}$ of spins to the vertices. Weights for different configurations are computed in terms of the $(+,+)$-edge activity β, the $(-,-)$ edge activity γ and a vertex activity λ, and are given by

$$w(\sigma) = \lambda^{m(\sigma)}\beta^{n_+(\sigma)}\gamma^{n_-(\sigma)}, \tag{3}$$

where given the configuration σ, $m(\sigma)$ denotes the number of vertices assigned spin $-$, and $n_+(\sigma)$ (respectively, $n_-(\sigma)$) denotes the number of edges for which both endpoints are assigned spin $+$ (respectively, $-$). The *partition function* is defined as

$$Z = \sum_{\sigma\in\{-1,1\}^V} w(\sigma).$$

We remark that this representation can be easily translated to the usual description in terms of edge potentials and vertex field: for completeness we give the translation in appendix A.

Definition 2.1. ***(Occupation probability).*** *Given a vertex v in the graph G, the occupation probability p_v is the probability that v is assigned the spin $+$ in a configuration σ sampled according to the weights defined in equation (3).*

2.2. **The Ising model.** The Ising model corresponds to the case $\beta = \gamma$. The model is *ferromagnetic* when $\beta > 1$ and *anti-ferromagnetic* when $\beta < 1$ (the case $\beta = 1$ is trivial). The zero-field case corresponds to $\lambda = 1$, the positive field case to $\lambda < 1$ and the negative field case to $\lambda > 1$. As shown in appendix A, on d-regular graphs the Ising model is equivalent to general two-state spin systems. Thus, in the rest of this paper, we will concentrate mostly on the Ising case. On non-regular graphs the equivalence still holds; however, the vertex activity λ in the Ising model may then be different on different vertices. The adaptation of our results to this setting is described in Corollary 1.3.

2.2.1. *Phase transition.* The anti-ferromagnetic Ising model shows a uniqueness phase transition on the d-ary tree for $d \geq 2$. In particular, one can define a critical activity $\lambda_c(\beta, d)$ as follows:

Definition 2.2. ***(Critical activity).*** *Consider the anti-ferromagnetic Ising model on an infinite d-ary tree with edge activity β and vertex activity λ. If $\beta > \frac{d+1}{d-1}$ then the Gibbs measure is unique for all values of λ. If $\beta \leq \frac{d+1}{d-1}$, then there exists a* critical activity $\lambda_c(\beta, d) \geq 1$ *such that the Gibbs measure is unique if and only if* $|\log \lambda| \geq \log \lambda_c(\beta, d)$.

Using the translation in appendix A, we can now describe the uniqueness region for general two-state anti-ferromagnetic spin systems. For simplicity, we set $\lambda = 1$. The uniqueness region is then sketched in Figure 1. The hyperbola $\beta\gamma = 1$ is the boundary of the anti-ferromagnetic region. Points lying above the hyperbola $\beta\gamma = (\frac{d-1}{d+1})^2$ translate to an Ising model with edge activity $\beta' > \frac{d-1}{d+1}$, and hence lie in the uniqueness region. However, since a critical activity as defined above exists also for edge activities $\beta' \leq \frac{d-1}{d+1}$, some points lying below the above hyperbola are also in the uniqueness region, as Figure 1 shows.

A consequence of uniqueness[4] of the Gibbs measure is *weak spatial mixing*, which captures a weak notion of decay of point to set correlations. Let $p_v(\sigma, S)$ be the probability of occupation of the root ρ of an infinite d-ary tree when the spins of a set S of nodes are fixed according to the configuration σ. Let $\delta(\rho, S)$ denote the distance of ρ from the set S.

Definition 2.3. ***(Weak Spatial Mixing).*** *Given any two-state spin system,* weak spatial mixing *is said to hold if for any set S whose distance $\delta(\rho, S)$ from*

[4] As stated in the introduction, we exclude the boundary of the uniqueness region here.

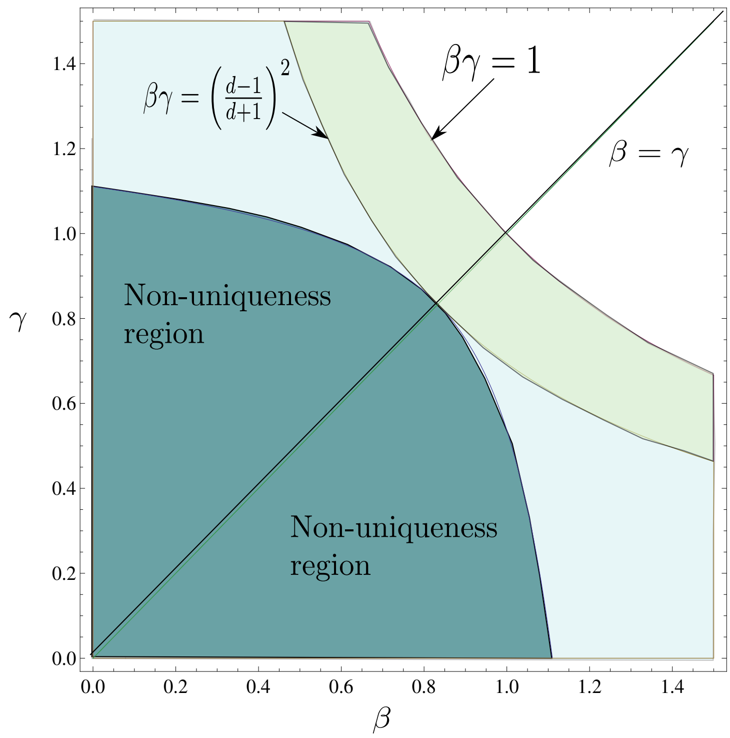

FIGURE 1. The non-uniqueness region (dark shading) for general two-state spin systems, sketched here for the 11-ary tree with $\lambda = 1$. The uniqueness region consists of the other shaded portions.

the root ρ of the tree is finite, and any two configurations σ_1 and σ_2, we have

$$|p_\rho(\sigma_1, S) - p_\rho(\sigma_2, S)| \leq \exp(-\Omega(\delta(\rho, S))).$$

Notice that weak spatial mixing does not guarantee good decay of correlations when the set S contains vertices which are very close to the root ρ, even when σ_1 and σ_2 differ only on vertices which are very far away from ρ. A related but, as the name suggests, stronger notion is that of *strong spatial mixing*, which captures the idea that fixing vertices near the root to the *same* spin should not affect the exponential decay of point-to-set correlations. We note that strong spatial mixing is not in general implied by weak spatial mixing for arbitrary spin systems; see, e.g., [10] for the description of a counterexample involving the ferromagnetic Ising model.

Definition 2.4. ***(Strong Spatial Mixing).*** *Given any two-state spin system,* strong spatial mixing *is said to hold if for any set S whose distance $\delta(\rho, S)$ from the root ρ of the tree is finite, and any two configurations σ_1 and σ_2 which differ only on a set $T \subseteq S$ of vertices, we have*

$$|p_\rho(\sigma_1, S) - p_\rho(\sigma_2, S)| \leq \exp(-\Omega(\delta(\rho, T))).$$

2.2.2. *Phase transition and tree recursions.* It is well known (see, for example, [2]) that the uniqueness condition for two-state spin systems on d-ary trees can be written in terms of the number of fixed points of the recursion for occupation probabilities. Consider a subtree rooted at a vertex v in the d-ary tree, and let $v_i, i = 1, 2, ...d$ be its children. Let p_v be the occupation probability at vertex v and define $R_v = \frac{1-p_v}{p_v}$. One can then write the following recurrence for R_v:

$$R_v = \lambda \prod_{i=1}^{d} \left(\frac{\beta R_{v_i} + 1}{\beta + R_{v_i}} \right). \tag{4}$$

This can easily be converted to a recurrence for occupation probabilities. Define

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

We can then write the recurrence as

$$p_v = F(p_{v_1}, p_{v_2}, \ldots, p_{v_d}) \triangleq \frac{1}{1 + \lambda \prod_{i=1}^{d} h(p_{v_i})}. \tag{5}$$

On the d-ary tree, by symmetry, one can write the recurrence as:

$$p_v = f(p_{v_1}) \triangleq \frac{1}{1 + \lambda h(p_{v_1})^d}. \tag{6}$$

Note that h is an increasing function, and hence F and f are decreasing in each of their arguments.

In terms of the recurrence function f, the condition for uniqueness can be stated as follows:

Theorem 2.5 ([2]). *For given values of β and λ, the infinite d-ary tree has a unique Gibbs measure if and only if the iterated recurrence function $f \circ f$ has a unique fixed point. In particular, the Gibbs measure is unique if and only if there is no x satisfying the conditions:*

$$f(x) = x; \tag{7}$$

$$f'(x) < -1. \tag{8}$$

2.3. **Messages.**

Definition 2.6. (Message). *A* message *is an increasing (and hence invertible on its range) differentiable function $\phi : [0,1] \mapsto \mathbb{R}$.*

Given a recurrence function $f : [0,1] \mapsto \mathbb{R}^+$, and a message ϕ, we denote by f^ϕ the function $\phi \circ f \circ \phi^{-1}$. We use the following fact:

Fact 2.7. *For any message ϕ, the parameters (λ, β) are in the uniqueness region if, and only if, ${f^\phi}'(p^*) \geq -1$ at the unique fixed point p^* of f^ϕ.*

Proof. We use Theorem 2.5. Suppose there exists an x satisfying conditions (7) and (8) for f. It is easy to check that $\phi(x)$ satisfies the same conditions for f^ϕ. Conversely, if x satisfies the conditions for f^ϕ then $\phi^{-1}(x)$ satisfies the same conditions for f. The uniqueness of the fixed point p^* follows since, by construction, f^ϕ is a strictly decreasing function. □

2.4. **Weitz's tree reduction.** Weitz [10] proved the following combinatorial reduction:

Theorem 2.8. *For any two-state spin system, strong spatial mixing on the d-ary tree implies*

i) strong spatial mixing on all graphs of degree at most $d+1$; and
ii) a deterministic fully polynomial approximation scheme for the partition function of the spin system on graphs of maximum degree $d+1$.

3. Messages and contraction on the d-ary tree

In this section, we will prove the main technical ingredient of our result, which is expressed in the following theorem.

Theorem 3.1. *Given d, β and λ, there exists a message ϕ such that the tree recurrence $\phi \circ f \circ \phi^{-1}$ for the quantity $\phi(p_v)$ decays by a constant factor $c(\lambda, d) < 1$ whenever (β, λ) is in the uniqueness region for the d-ary tree.*

Remark: For ease of notation, in the rest of the paper we will prove our results in terms of the uniqueness threshold of the d-ary tree, relating it to algorithms on graphs of degree at most $d+1$. As already noted, the uniqueness thresholds on the $(d+1)$-regular tree and the d-ary tree coincide, and hence our results apply equally to the infinite $(d+1)$-regular tree.

The above theorem shows that the idea of Restrepo *et al* [7] of finding stepwise decaying messages for the hard core model can be extended to arbitrary two-state spin systems. Theorem 1.1 follows as a consequence of the message constructed in Theorem 3.1, and is proved in section 4. Along with Weitz's reduction stated in Theorem 2.8, this immediately implies Corollary 1.2, the FPTAS for the Ising model with arbitrary fields. To derive Corollary 1.3 for general two-spin systems, we will need the translation described in appendix A. Both these latter proofs appear at the end of Section 4.

We begin by setting up the proof of Theorem 3.1, which amounts to showing stepwise decay of f^ϕ in the uniqueness region. In the light of Fact 2.7 the main technical challenge here is to come up with a message ϕ such that the quantity $\left|{f^\phi}'\right|$ is maximized at the unique fixed point of f^ϕ. Let us fix constants

$$A = d(1-\beta^2) + (1-\beta)^2 \text{ and } D = \frac{\sqrt{A+4\beta} - \sqrt{A}}{2\sqrt{A}}.$$

Define

$$\phi(x) = \log\left(\frac{x+D}{1-x+D}\right). \tag{9}$$

Notice that $D > 0$, so this is a continuous real valued function on the interval $[0,1]$. Using this message we will be able to prove the following.

Lemma 3.2. *Consider the anti-ferromagnetic Ising model on a d-ary tree with edge activity β and vertex activity λ. Then, defining $g = f^{\phi}$, $\psi = \phi^{-1}$, $\alpha = \psi(x)$ and $\eta = f(\alpha)$, we have*

$$g''(x) = (\eta-\alpha)g'(x)\psi'(x)\frac{d\beta(1-\beta^2)(2\beta + A\alpha\eta + A(1-\alpha)(1-\eta))}{(\beta+\alpha(1-\alpha)(1-\beta)^2)(\beta+A\eta(1-\eta))(\beta+A\alpha(1-\alpha))}.$$

The proof of Lemma 3.2 is somewhat technical and is given in appendix B.

Before proceeding with the technical development, we pause to give some comments on the design of our message. Notice that the requirement that the derivative of the function $g = f^{\phi}$ should have its maximum magnitude at the unique fixed point of g does not lead to a unique solution for ϕ, and thus we must resort to some educated guesswork for the functional form of ϕ. Our choice is guided by the intuition that, by analogy with the zero field case, where it is well known that the simple message $\phi(x) = \log\left(\frac{x}{1-x}\right)$ is sufficient, a log ratio of probabilities shifted by an additive constant D to account for the field should be appropriate. The choice of D is then determined by the above requirement. The important property of the message that we are able to ensure is that, perhaps surprisingly, it does not depend upon the vertex potential λ, but only upon the edge potential β and the degree d; this is also reflected in the fact that the additive shift D is the same for both probabilities. This property is important in extending our algorithm to the setting of general two-state anti-ferromagnetic spin systems in Corollary 1.3.

Corollary 3.3. *Let $g = f^{\phi}$, with the message ϕ defined above. Then $|g'(x)|$ is maximized at the unique positive fixed point of g.*

Proof. We use the notation established in Lemma 3.2, where we saw that

$$g''(x) = (\eta-\alpha)\underbrace{g'(x)\psi'(x)\frac{d\beta(1-\beta^2)(2\beta + A\alpha\eta + A(1-\alpha)(1-\eta))}{(\beta+\alpha(1-\alpha)(1-\beta)^2)(\beta+A\eta(1-\eta))(\beta+A\alpha(1-\alpha))}}.$$

It is easy to see that the underlined expression above is *negative*: this is because g is a decreasing function, while ψ, being the inverse of the increasing function ϕ, is increasing. Also, we have $0 < \beta < 1$ (for anti-ferromagnetic Ising) and $0 \le \alpha, \eta \le 1$ (since they are probabilities), so that the fraction underlined above is positive.

Let x^* be the unique fixed point of the strictly decreasing function g. From the above discussion, it follows that the sign of $g''(x)$ is the opposite of the sign of $\eta - \alpha = f(\psi(x)) - \psi(x)$. Notice that $\eta - \alpha$ is strictly positive for $x < x^*$

and strictly negative for $x > x^*$. This implies that $g'(x)$ is strictly decreasing for $x < x^*$ and strictly increasing for $x > x^*$. Since g is strictly decreasing this shows that the magnitude of g' is maximized at x^*, which is what we set out to prove. □

Combining Corollary 3.3 with Fact 2.7, we immediately get the following:

Corollary 3.4. *For β and λ in the uniqueness region of the d-ary tree, there exists $c < 1$ such that $|g'(x)| \leq c$ for all x.*

This proves Theorem 3.1.

4. Strong spatial mixing on the d-ary tree

We will now prove Theorems 1.1 and 1.2 with the help of the results obtained in the previous section. Recall that Theorem 1.1 asserts that weak spatial mixing implies strong spatial mixing. To prove this, we consider the vectorized version G of the function $g = f^\phi$ defined in Section 2.3. For $\vec{x} \in \mathbb{R}^d$, define

$$G(x_1, x_2, \ldots, x_d) = \phi\left(\frac{1}{1 + \lambda \prod_{i=1}^d h\left(\psi(x_i)\right)}\right).$$

For this function, it is easy to see that the following condition implies strong spatial mixing on the d-ary tree.

Definition 4.1. ***(Contractive Spatial Mixing).*** *Given the parameters β and λ for the anti-ferromagnetic Ising model on a d-regular tree, contractive spatial mixing holds if there exists a constant $c < 1$ such that*

$$|G(\vec{x}) - G(\vec{y})| \leq c \left|\left|\vec{x} - \vec{y}\right|\right|_\infty, \tag{10}$$

for the vectorized version G of g defined as above with respect to the message ϕ.

To establish this condition, we will rely on the following lemma. As before, we denote ϕ^{-1} by ψ.

Lemma 4.2. *Let $\eta = \psi(G(\vec{x}))$. Let $\bar{\eta}$ be the unique solution of $\psi(g(\bar{\eta})) = \eta$. Then $\left|\left|\nabla\left(G(\vec{x})\right)\right|\right|_1 \leq \left|\left|\nabla\left(G(\bar{\eta}, \bar{\eta}, \ldots, \bar{\eta})\right)\right|\right|_1$.*

Proof. As before, we set $\alpha_i = \psi(x_i)$ for $i = 1, 2, \ldots d$. We then have

$$\eta = \frac{1}{1 + \lambda \prod_{i=1}^d h(\alpha_i)} = \frac{1}{1 + \lambda h(\psi(\bar{\eta}))^d}. \tag{11}$$

With D as defined just before Lemma 3.2, we can now write $\left|\left|\nabla\left(G(\vec{x})\right)\right|\right|_1$ as

$$\left|\left|\nabla\left(G(\vec{x})\right)\right|\right|_1 = \frac{d\eta(1-\eta)(1-\beta^2)}{\beta + A\eta(1-\eta)}\left(1 + (1-\beta^2)\sum_{i=1}^d \frac{\alpha_i(1-\alpha_i)}{\beta + (1-\beta)^2\alpha_i(1-\alpha_i)}\right). \tag{12}$$

For convenience of notation, we define the function $J(x) \triangleq \frac{x(1-x)}{\beta+(1-\beta)^2 x(1-x)}$. Note that maximizing the sum above under the constraint (11) is the same as maximizing $\sum_{i=1}^{d} J(\alpha_i)$ under the constraint that $\prod_{i=1}^{d} h(\alpha_i) = \frac{1-\eta}{\lambda\eta}$. Since h is positive and invertible, it is therefore sufficient to show that the function $K(x) \triangleq J(h^{-1}(e^x))$ is concave in order to show that all α_i's are equal at a maximum. We now show this by direct computation. After differentiating twice and simplifying, we have

$$K''(x) = -\frac{e^{-x}(1+e^{2x})\beta}{(1-\beta^2)^2} < 0.$$

This shows that K is concave. By the discussion above, this implies that the sum in equation (12) is maximized when all α_i's are equal. In conjunction with the condition that $\eta = \frac{1}{1+\prod_{i=1}^{d} h(\alpha_i)}$, this shows that

$$||\nabla\left(G(\vec{x})\right)||_1 \leq ||\nabla\left(G(\bar{\eta},\bar{\eta},\ldots,\bar{\eta})\right)||_1,$$

as claimed. □

Using Corollary 3.3 and the above lemma, we are now ready to prove our main technical result, Theorem 1.1, which says that weak spatial mixing implies strong spatial mixing for general anti-ferromagnetic Ising models.

Proof of Theorem 1.1. Consider a setting of parameters β and λ such that the d-ary tree has weak spatial mixing. Let x^* be the unique fixed point of the function g defined above. By Corollary 3.3, we have $c = |g'(x^*)| < 1$. By Lemma 4.2, this implies that for all $\vec{x}$ in the domain of the function G defined above, $||G(\vec{x})||_1 \leq c$. By the mean value theorem, this implies that

$$||G(\vec{x}) - G(\vec{y})||_1 \leq c\,||\vec{x}-\vec{y}||_\infty,$$

and thus contractive spatial mixing holds. As observed above, this implies strong spatial mixing. □

Combining the above theorem with the general reduction of Weitz [10] stated in Theorem 2.8, we can now prove Corollary 1.2, which asserts the existence of an FPTAS for general anti-ferromagnetic Ising models on bounded-degree graphs up to the uniqueness threshold.

Proof of Corollary 1.2. As observed before, it is sufficient to have an FPTAS for approximating the occupation probability p_ρ of a vertex ρ, under an arbitrary fixing of spin values for an arbitrary subset of vertices, in order to have an FPTAS for the partition function of the associated spin system. Given a vertex ρ in a graph G of maximum degree $(d+1)$, we start by constructing Weitz's self-avoiding walk (SAW) tree rooted at ρ. For non-root vertices in this tree which do not have d children, we can create dummy children (so as to make the arity of the vertex d) all of which independently have occupation probabilities of $1/2$. Is it easy to see that this does not affect the tree recurrence (equation 5). We can now use the message ϕ designed above. Using Lemmas 3.2 and 4.2, we now get strong spatial mixing on the SAW tree. The corollary now follows by using Weitz's reduction (Theorem 2.8). □

Finally, we will see how to use Lemmas 3.2 and 4.2 to prove Corollary 1.3, which extends the FPTAS to general anti-ferromagnetic two-spin systems.

Proof of Corollary 1.3. Given a two-state spin system with parameters β, γ and λ on a graph G of degree at most $d+1$, we can use the translation given in appendix A to come up with an equivalent Ising model given by edge potential $\beta' = \sqrt{\beta\gamma}$ and vertex-dependent potentials $\lambda_v = \lambda(\sqrt{\beta/\gamma})^{d_v}$. Now, as before, in order to estimate the occupation probability p_ρ for a given vertex ρ, we construct the Weitz self-avoiding walk (SAW) tree rooted at ρ, and complete the degree of any vertex in the tree (apart from ρ) which does not have d children by attaching dummy children which are fixed to have occupation probability $\frac{1}{2}$. We now use the message ϕ constructed above for d-ary trees for the parameter β'. By the hypotheses of the corollary, the parameters (β', λ_u) at each vertex u of the SAW tree are in the uniqueness regime of the d-ary tree. Since the message ϕ does not depend upon λ_u, Lemmas 3.2 and 4.2 apply at each vertex u of the tree. Thus, as in the proof of Theorem 1.1, we get contractive spatial mixing and, hence, strong spatial mixing on the SAW tree. Employing Weitz's reduction (Theorem 2.8), this proves the first part of the corollary.

The claim that the class $\mathcal{G}$ in the corollary includes $(d+1)$-regular graphs when β, γ and λ are in the uniqueness region of d-ary tree follows by noticing that in this case the parameters $\lambda' = \lambda_v$ obtained by the translation are the same at each vertex v, and that β' and λ' are in the uniqueness regime of the d-ary tree by the hypotheses of the corollary. Thus, we can now complete the proof for this case in the same manner as above. □

Acknowledgment: We thank Prasad Tetali for providing a manuscript of [7], and Dror Weitz for helpful comments.

References

[1] Galanis, A., Ge, Q., Stefankovic, D., Vigoda, E., and Yang, L. Improved inapproximability results for counting independent sets in the hard-core model. *CoRR abs/1105.5131* (2011).

[2] Georgii, H.-O. *Gibbs Measures and Phase Transitions.* De Gruyter Studies in Mathematics. Walter de Gruyter Inc., October 1988.

[3] Gerschenfeld, A., and Montanari, A. Reconstruction for models on random graphs. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science* (Washington, DC, USA, 2007), IEEE Computer Society, pp. 194–204.

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

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

[6] Mossel, E., Weitz, D., and Wormald, N. On the hardness of sampling independent sets beyond the tree threshold. *Probab. Theory Related Fields 143*, 3-4 (2009), 401–439.

[7] Restrepo, R., Shin, J., Tetali, P., Vigoda, E., and Yang, L. Improved mixing condition on the grid for counting and sampling independent sets. *CoRR abs/1105.0914* (2011).

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

[9] Weitz, D. Private communication. 2011.

[10] Weitz, D. Counting independent sets up to the tree threshold. In *STOC '06: Proceedings of the thirty-eighth annual ACM Symposium on Theory of Computing* (New York, NY, USA, 2006), ACM, pp. 140–149.

[11] Zhang, J., Liang, H., and Bai, F. Approximating partition functions of two-state spin systems. *CoRR abs/0911.5486* (2009).

Appendix

Appendix A. Translation between various descriptions of the Ising model

General two-state spin systems are usually described in terms of (symmetric) energy functions $Q(+\ ,+\ )$, $Q(+\ ,-\ )$ and $Q(-\ ,-\ )$, and an odd vertex field given by $h(+) = -h(-) = h$. For a graph $G = (V, E)$, the partition function of the system is then $Z_2 = \sum_\sigma w_2(\sigma)$ where the sum is over all states of the system $\sigma : V \to \{+, -\}$ and $w_2(\sigma)$ is defined as

$$w_2(\sigma) \triangleq \exp\left(- \sum_{e=\{u,v\}\in E} Q(\sigma(u), \sigma(v)) - \sum_{v\in V} h(\sigma(v))\right). \tag{13}$$

This is in fact equivalent to our formulation of the system given in Section 1.1 (recall the definition of the partition function Z). To see this, define

$$\begin{aligned} \beta &= \exp\left(-Q(+,+) + Q(+,-)\right); \\ \gamma &= \exp\left(-Q(-,-) + Q(+,-)\right); \\ \lambda &= \exp(2h), \end{aligned}$$

which yields

$$Z = Z_2 \exp\left(Q(+,-)|E| + h|V|\right)$$

and, similarly, $w(\sigma) = w_2(\sigma) \exp\left(Q(+,-)|E| + h|V|\right)$ for all σ. We call the above spin systems *soft constraint* systems if β, γ and λ are non-zero, or equivalently, if the energy functions and field are finite for all spin values. As we shall now see, every such soft constraint system can be represented in terms of the Ising model (this translation can also be found, e.g., in [4]). Consider a general two-state spin system with parameters $\beta, \gamma > 0$ and λ. Then the equivalent Ising model has edge activity

$$\beta' = \sqrt{\beta\gamma}, \tag{14}$$

and a degree dependent vertex activity given by

$$\lambda'_v = \lambda \left(\sqrt{\frac{\beta}{\gamma}}\right)^{d_v},$$

where d_v denotes the degree of vertex v. Now, denote the weight of a configuration σ in the Ising model just defined by $w^*(\sigma)$ and its partition function by Z^*. Then one calculates straightforwardly that

$$w(\sigma) = w^*(\sigma) \left(\sqrt{\frac{\gamma}{\beta}}\right)^{|E|}$$

and hence

$$Z = Z^* \left(\sqrt{\frac{\gamma}{\beta}}\right)^{|E|}.$$

Thus we have translated the original spin system with parameters (β, γ, λ) into an Ising model with locally changing field. Note that on regular graphs the resulting field is in fact constant at all vertices. Furthermore, the Ising model is anti-ferromagnetic if and only if $\beta\gamma < 1$. This justifies our use of the term "anti-ferromagnetic" for general spin systems based on the value of $\beta\gamma$.

Appendix B. Proof of Lemma 3.2

In this section, we prove Lemma 3.2. The proof involves a few somewhat lengthy derivative computations, which we isolate in the following lemma.

Lemma B.1. *With the notation used in Lemma 3.2 above, we have*

$$\frac{\phi''(x)}{\phi'(x)} = \frac{A(2x-1)}{\beta + Ax(1-x)} \tag{15}$$

$$\frac{h'(x)}{h(x)} = \frac{1-\beta^2}{\beta + (1-\beta)^2 x(1-x)} \tag{16}$$

$$\frac{h''(x)}{h'(x)} = \frac{2(1-\beta)}{1-(1-\beta)x} \tag{17}$$

$$f'(x) = -df(x)(1-f(x))\frac{h'(x)}{h(x)} \tag{18}$$

$$\frac{f''(x)}{f'(x)} = \frac{f'(x)(1-2f(x))}{f(x)(1-f(x))} + \frac{h''(x)}{h'(x)} - \frac{h'(x)}{h(x)} \tag{19}$$

Proof (sketch). Most of these identities are easily verified by direct computation. In proving equation (15), one needs to keep in mind the definition of the constant D. □

Proof of Lemma 3.2. To ease notation, we will suppress the dependence of the quantities η and α on x. Using the chain rule, we have

$$g'(x) = \frac{\phi'(\eta)}{\phi'(\alpha)} f'(\alpha). \tag{20}$$

Here, we used the fact that since $\psi = \phi^{-1}$, $\psi'(x) = \frac{1}{\phi'(\psi(x))}$. After taking the logarithm, and noticing that the right hand side is more easily expressed as a function of α rather than of x, one can write the second derivative of g as:

$$\frac{1}{\psi'(x)} \frac{g''(x)}{g'(x)} = \frac{\phi''(\eta)}{\phi'(\eta)} \frac{\mathrm{d}\eta}{\mathrm{d}\alpha} - \frac{\phi''(\alpha)}{\phi'(\alpha)} + \frac{f''(\alpha)}{f'(\alpha)}. \tag{21}$$

We now consider each of the terms involved above. Recalling that $\eta = f(\alpha)$, and using equations (18) and (19) to expand the first and last terms in equation

(21) above, we get

$$\frac{1}{\psi'(x)}\frac{g''(x)}{g'(x)} = T_1 - T_2, \tag{22}$$

where T_1 and T_2 are defined as:

$$T_1 \triangleq \frac{h''(\alpha)}{h'(\alpha)} - \frac{h'(\alpha)}{h(\alpha)} - \frac{\phi''(\alpha)}{\phi'(\alpha)}; \tag{23}$$

$$T_2 \triangleq d\frac{h'(\alpha)}{h(\alpha)}\left[\frac{\phi''(\eta)}{\phi'(\eta)}\eta(1-\eta) + 1 - 2\eta\right]. \tag{24}$$

Notice that all terms containing η are isolated in T_2. We now consider each of the terms separately. For T_1, we first have

$$\begin{aligned}\frac{h''(\alpha)}{h'(\alpha)} - \frac{h'(\alpha)}{h(\alpha)} &= \frac{2(1-\beta)}{1-(1-\beta)\alpha} - \frac{1-\beta^2}{\beta+(1-\beta)^2\alpha(1-\alpha)}\\ &= \frac{(1-\beta^2)(2\alpha-1)}{\beta+(1-\beta)^2\alpha(1-\alpha)}.\end{aligned} \tag{25}$$

Here, we used equations (17) and (16) in the first line. Now using equation (15), we have

$$\begin{aligned}T_1 &= \frac{(2\alpha-1)\left((1-\beta)^2\left[\beta+A\alpha(1-\alpha)\right] - A\left[\beta+(1-\beta)^2\alpha(1-\alpha)\right]\right)}{\left(\beta+(1-\beta)^2\alpha(1-\alpha)\right)\left(\beta+A\alpha(1-\alpha)\right)}\\ &= \frac{\beta(2\alpha-1)((1-\beta)^2-A)}{\left(\beta+(1-\beta)^2\alpha(1-\alpha)\right)\left(\beta+A\alpha(1-\alpha)\right)}\\ &= \frac{-d\beta(2\alpha-1)}{(\beta+A\alpha(1-\alpha))}\frac{h'(\alpha)}{h(\alpha)}.\end{aligned} \tag{26}$$

Here, we use $A = d(1-\beta^2) + (1-\beta)^2$, followed by equation (16) in the last line.

We now consider T_2. Again using equation (15), we have

$$\begin{aligned}T_2 &= d\frac{h'(\alpha)}{h(\alpha)}\left[\frac{A(2\eta-1)\eta(1-\eta)}{\beta+A\eta(1-\eta)} - (2\eta-1)\right]\\ &= \frac{-d\beta(2\eta-1)}{\beta+A\eta(1-\eta)}\frac{h'(\alpha)}{h(\alpha)}.\end{aligned} \tag{27}$$

Notice that modulo the $\frac{h'(\alpha)}{h(\alpha)}$ factor, T_1 and T_2 have the same functional form as functions of α and η respectively. In fact, the message ϕ is designed so as to make this possible. We can now substitute these values into equation (22) to get

$$\begin{aligned}
g''(x) &= d\beta g'(x)\psi'(x)\frac{h'(\alpha)}{h(\alpha)}\left[\frac{2\eta-1}{\beta+A\eta(1-\eta)}-\frac{2\alpha-1}{\beta+A\alpha(1-\alpha)}\right]\\
&= d\beta g'(x)\psi'(x)\frac{h'(\alpha)}{h(\alpha)}\frac{(\eta-\alpha)(2\beta+A(\alpha\eta+(1-\alpha)(1-\eta)))}{(\beta+A\alpha(1-\alpha))\,(\beta+A\eta(1-\eta))}\\
&= (\eta-\alpha)g'(x)\psi'(x)\frac{d\beta(1-\beta^2)(2\beta+A\alpha\eta+A(1-\alpha)(1-\eta))}{(\beta+\alpha(1-\alpha)(1-\beta)^2)(\beta+A\eta(1-\eta))(\beta+A\alpha(1-\alpha))}
\end{aligned}$$

where in the last step we used equation (16). □

Alistair Sinclair
Computer Science Division
UC Berkeley
E-mail address: sinclair@cs.berkeley.edu

Piyush Srivastava
Computer Science Division
UC Berkeley
E-mail address: piyushsr@eecs.berkeley.edu

Marc Thurley
Centre de Recerca Matemàtica
Bellaterra, Spain
E-mail address: marc.thurley@googlemail.com