

Fast mixing for independent sets, colorings and other models on trees

Fabio Martinelli[†]

Alistair Sinclair[‡]

Dror Weitz[‡]

Abstract

We study the mixing time of the Glauber dynamics for general spin systems on bounded-degree trees, including the Ising model, the hard-core model (independent sets) and the antiferromagnetic Potts model at zero temperature (colorings). We generalize a framework, developed in our recent paper [18] in the context of the Ising model, for establishing mixing time $O(n \log n)$, which ties this property closely to phase transitions in the underlying model. We use this framework to obtain rapid mixing results for several models over a significantly wider range of parameter values than previously known, including situations in which the mixing time is strongly dependent on the boundary condition.

1 Introduction

1.1 Spin systems on trees

Spin systems capture a wide range of probabilistic models studied in statistical physics, applied probability, artificial intelligence and elsewhere. A (*nearest neighbor*) spin system on a graph $G = (V, E)$ is specified by a finite set S of spin values, a symmetric pair potential $U : S \times S \rightarrow \mathbb{R} \cup \{\infty\}$, and a singleton potential $W : S \rightarrow \mathbb{R}$. A configuration $\sigma \in S^V$ of the system assigns to each vertex (site) $x \in V$ a spin value $\sigma_x \in S$. The probability of finding the system in configuration σ is determined by the *Gibbs distribution*

$$\mu(\sigma) \propto \exp\left[-\left(\sum_{xy \in E} U(\sigma_x, \sigma_y) + \sum_{x \in V} W(\sigma_x)\right)\right].$$

Thus the pair potential specifies the likelihood of seeing a given pair of spins at adjacent sites, while the singleton potential specifies the likelihood of seeing a given

spin. Note that setting $U(s_1, s_2) = \infty$ corresponds to a *hard constraint*, i.e., spin values s_1, s_2 are forbidden to be adjacent. We denote by $\Omega \subseteq S^V$ the set of all *valid* spin configurations, i.e., those for which $\mu(\sigma) > 0$. We will give several concrete examples in a moment.

Usually G is a finite portion of some regular lattice, such as \mathbb{Z}^d . In this paper we concentrate on what is known in statistical physics as the *Bethe lattice* \mathbb{T}^b , i.e., G is a complete rooted tree in which each interior vertex has $b \geq 2$ children.[†] Spin systems on trees are not only a useful simplification of their more classical counterparts on \mathbb{Z}^d , but have recently attracted a lot of attention in their own right as the canonical example of models on “non-amenable” graphs (i.e., those whose boundary is of comparable size to their volume) — see, e.g., [1, 2, 3, 8, 15].

A *boundary condition* corresponds to fixing the spins at the leaves of G to some specified values (e.g., all are set to the same value s). This allows us to formalize the central concept of a *phase transition*. If we let the size of the tree G grow to infinity, the Gibbs distribution tends to a limit known as the (*infinite-volume*) *Gibbs measure*. This limit may or may not depend on the boundary condition: i.e., there may be either a unique Gibbs measure, or multiple Gibbs measures (“phases”) corresponding to different boundary conditions. Informally, the existence of multiple Gibbs measures corresponds to the fact that the spin configuration on the leaves can have a non-zero influence on the spin at the root even as the depth of the tree tends to infinity. A *phase transition* occurs when an infinitesimal change in the potentials leads to a switch from a unique Gibbs measure to multiple Gibbs measures. See, e.g., [7] for more background.

We now illustrate the above ideas with some concrete examples. The following four spin systems are among the most widely studied in the literature, and will serve as the motivating examples in this paper:

[†]Department of Mathematics, University of Roma Tre, Largo San Murialdo 1, 00146 Roma, Italy. Email: martin@mat.uniroma3.it. This work was done while this author was visiting the Departments of EECS and Statistics, University of California, Berkeley, supported in part by a Miller Visiting Professorship.

[‡]Computer Science Division, University of California, Berkeley, CA 94720-1776, U.S.A. Email: {sinclair, dror}@cs.berkeley.edu. Supported in part by NSF Grant CCR-0121555 and DARPA cooperative agreement F30602-00-2-0601.

[†]Strictly speaking, in the Bethe lattice all vertices (including the root) have degree $b + 1$; for convenience we assume that the root has degree b . This difference is purely technical and does not affect our results.

A. The (ferromagnetic) Ising model

There are two spin values $S = \{\pm 1\}$, and the potentials are $U(s_1, s_2) = -\beta s_1 s_2$, $W(s) = 0$, where β is the inverse temperature. Thus $\Omega = \{\pm 1\}^V$. The Gibbs distribution μ assigns higher weight to configurations in which many neighboring spins are aligned with one another. This effect increases with β , so that at high temperatures (low β) the spins behave almost independently, while at low temperatures (high β) large connected regions of equal spins tend to form.

In fact, as is well known [23], a phase transition occurs at the critical value $\beta = \beta_0 = \frac{1}{2} \ln\left(\frac{b+1}{b-1}\right)$; in other words, in the “high temperature” region $\beta \leq \beta_0$ there is a unique Gibbs measure independent of the boundary condition, but as soon as $\beta > \beta_0$ we get (at least) two different Gibbs measures corresponding to the $+1$ and -1 boundary conditions respectively; i.e., if the leaves are $+1$ then the root has probability bounded away from $\frac{1}{2}$ of being $+1$.

B. The hard-core model (independent sets)

This has been used in statistical physics as a model of lattice gases [7], and also in other areas such as the modeling of communication networks [13]. Again there are just two spins $S = \{0, 1\}$, and we refer to a site as *occupied* if it has spin value 1 and *unoccupied* otherwise. The potentials are $U(1, 1) = \infty$, $U(1, 0) = U(0, 0) = 1$, $W(1) = L$ and $W(0) = 0$, where $L \in \mathbb{R}$. The hard constraint here means that no two adjacent sites may be occupied, so Ω can be identified with the set of all *independent sets* in G . Also, the aggregated potential of a valid configuration is proportional to the number of occupied sites. Hence the Gibbs distribution takes the simple form $\mu(\sigma) \propto \lambda^{N(\sigma)}$, where $N(\sigma)$ is the number of occupied sites and the parameter $\lambda = \exp(-L) > 0$, which controls the density of occupation, is referred to as the “activity.”

The hard-core model also undergoes a phase transition at a critical activity $\lambda = \lambda_0 = \frac{b^b}{(b-1)^{b+1}}$ (see, e.g., [24, 13]). For $\lambda \leq \lambda_0$ there is a unique Gibbs measure independent of the boundary condition, while for $\lambda > \lambda_0$ there are (at least) two distinct phases, corresponding to the “odd” and “even” boundary conditions respectively. The even (odd) boundary condition is obtained by making the leaves of the tree all occupied if the depth is even (odd), and all unoccupied otherwise. For $\lambda > \lambda_0$, the probability of occupation of the root in the infinite-volume Gibbs measure differs for odd and even boundary conditions.

C. The (ferromagnetic) Potts model

This model was introduced by Potts [22] as a generalization of the Ising model to more than two spin values; see [28] for a survey. Here $S = \{1, 2, \dots, q\}$ and the po-

tentials are $U(s_1, s_2) = -2\beta\delta_{s_1, s_2}$, $W(s) = 0$. Thus the spin at each site can take one of q possible values, and the aggregated potential of any configuration depends on the number of adjacent pairs of equal spins. Note that the Ising model is the special case $q = 2$. There are no hard constraints, so $\Omega = S^V$.

Qualitatively the behavior of this model is similar to that of the Ising model, though less is known in precise quantitative terms. Again there is a phase transition at a critical $\beta = \beta_0$, which depends on b and q , so that for $\beta > \beta_0$ (and indeed for $\beta \geq \beta_0$ when $q > 2$) there are multiple phases. This value β_0 does not in general have a closed form, but it is known [8] that $\beta_0 < \frac{1}{2} \ln\left(\frac{b+q-1}{b-1}\right)$ for all $q > 2$. (For $q = 2$, this value is exactly β_0 for the Ising model as quoted earlier.)

D. The antiferromagnetic Potts model (colorings)

In this model $S = \{1, 2, \dots, q\}$, and the potentials are $U(s_1, s_2) = 2\beta\delta_{s_1, s_2}$, $W(s) = 0$, where β is the inverse temperature. This is analogous to the Ising and Potts models except that the interactions are *antiferromagnetic*, i.e., neighbors with *unequal* spins are favored. The most interesting case of this model is when $\beta = \infty$ (i.e., zero temperature), which introduces hard constraints. Thus if we think of the q spin values as colors, Ω is the set of *proper colorings* of G , i.e., assignments of colors to vertices so that no two adjacent vertices receive the same color. The Gibbs distribution is uniform over proper colorings. In this model it is q that provides the parameterization. This model has been widely studied not only in statistical physics, but also in computer science because of its connection to graph coloring. See, e.g., [3] for an informative account.

For colorings on the b -ary tree it is well known that, when $q \leq b + 1$, there are multiple Gibbs measures; this follows immediately from the existence of “frozen configurations,” i.e., colorings in which the color of every internal vertex is forced by the colors of the leaves (see, e.g., [3]). Recently it has been proved that, as soon as $q \geq b + 2$, the Gibbs measure is unique [12].

1.2 Glauber dynamics

While classical statistical physics focused on *static* questions about the infinite volume Gibbs measure (such as the existence of a phase transition), the emphasis in recent years has shifted towards the study of the *Glauber dynamics*, a local Markov chain on the set of spin configurations Ω of a finite graph G . For definiteness, we describe the “heat-bath” version of Glauber dynamics: at each step, pick a site x u.a.r., and replace the spin at x by a random spin drawn from the conditional distribution given the spins of its neighbors. It is easy to check that this dynamics converges to μ as its stationary distribution, providing both an algorithm for sampling from

the Gibbs distribution and a plausible model for the actual evolution of the physical system. The key question in the study of the dynamics is to determine the *mixing time*, i.e., the number of steps until the distribution is close to μ , starting from an arbitrary configuration.

Recent developments in statistical physics have revealed a remarkable connection between the mixing time and phase transitions in the special case of the Ising model: for the Ising model on an n -vertex square in \mathbb{Z}^2 (which also has a phase transition at a critical β_0), when $\beta < \beta_0$ the mixing time with arbitrary boundary conditions is $O(n \log n)$ (which is optimal), but as soon as $\beta > \beta_0$ there are boundary conditions for which the mixing time jumps to $\exp(\Omega(\sqrt{n}))$ [17]. This conforms with the intuition that the existence of multiple phases creates a “bottleneck” which dramatically slows down mixing. For the Ising model on trees the situation is even more interesting (see [2, 18]): on an n -vertex b -ary tree, the mixing time remains $O(n \log n)$ for all boundaries,[‡] not only for $\beta \leq \beta_0$ but in fact for all $\beta < \beta_1$, where $\beta_1 = \frac{1}{2} \ln\left(\frac{\sqrt{b+1}}{\sqrt{b-1}}\right) > \beta_0$. As soon as $\beta > \beta_1$, the mixing time for certain boundaries becomes $n^{1+\Omega(1)}$, and the exponent is unbounded as $\beta \rightarrow \infty$. Thus there is again a sharp transition in mixing behavior, but it occurs *inside* the multiple-phase region.[§]

Finally, in the low temperature region $\beta \geq \beta_1$, the mixing time is heavily influenced by the boundary condition: recently, we proved that with the $+1$ (or, symmetrically, -1) boundary condition, the mixing time remains $O(n \log n)$ throughout this region (and hence for *all* values of β) [18]. This formalizes the intuition that the boundary condition breaks the symmetry between the two phases at low temperatures, thus eliminating the bottleneck and enabling rapid mixing.

The above discussion of the Ising model highlights two central issues in the study of the Glauber dynamics for general models. Firstly, for which range of parameter values is the mixing time “optimal” (i.e., $O(n \log n)$) for *arbitrary* boundary conditions? In particular, does this range extend throughout the region of uniqueness of the Gibbs measure, and even beyond (as it does for the Ising model)?[¶] And secondly, for parameter values outside this range, are there natural boundary conditions for which the mixing time is still optimal?

In this paper we present a unified framework for

answering such questions for general spin systems on trees. The framework is adapted and extended from our earlier paper [18] on the Ising model, where we bounded the log-Sobolev constant (and hence the mixing time) in terms of two simple quantities derived from the Gibbs measure, leading to a clean criterion for $O(n \log n)$ mixing time.[†] Using this framework, we are able to substantially extend the range of parameter values for which $O(n \log n)$ mixing time is known for the models described above, and explicitly relate this to properties of the Gibbs measure. In the case of specific boundaries, these are apparently the first results of this kind for any of these models. Beyond providing new results for specific models, our goal is also to demonstrate the wide applicability of our framework.

1.3 Our results

In this subsection we state our results and explain how they relate to previous work.

A. The Ising model

Our results for this model are outlined above, and were derived in our recent paper [18].

B. The hard-core model (independent sets)

The Glauber dynamics for the hard-core model on trees is known to have mixing time polynomial in n at all activities $\lambda > 0$ with free boundaries [11]. Moreover, a rather general result of Luby and Vigoda [14, 26] ensures a mixing time of $O(n \log n)$ when $\lambda < \frac{2}{b-1}$, with arbitrary boundaries. This latter result actually holds for any graph G of maximum degree $b + 1$.

In this paper, we prove results for the hard-core model that mirror those stated above for the Ising model. First, we show that the mixing time is $O(n \log n)$ for all activities $\lambda \leq \lambda_0$ (and indeed beyond), with *arbitrary* boundary conditions. Second, for the *even* (or *odd*) boundary condition, we get the same result for *all* activities λ :

THEOREM 1.1. *For the hard-core model on the n -vertex b -ary tree, the mixing time of the Glauber dynamics is $O(n \log n)$ in both of the following situations:*

- (i) *the boundary condition is arbitrary, and $\lambda \leq \max\left\{\lambda_0, \frac{1}{\sqrt{b-1}}\right\}$;*
- (ii) *the boundary condition is even (or odd), and $\lambda > 0$ is arbitrary.*

[‡]Throughout, a mixing time of $O(n \log n)$ hides constants that depend only on the potentials and on the degree b .

[§]This second critical value β_1 has other interpretations in terms of extremality of the Gibbs measure and the threshold for noisy data transmission on the tree [6].

[¶]Far into the uniqueness region — e.g., for the Ising model at very high temperatures — it is easy to see using the Dobrushin Uniqueness Condition that the mixing time is $O(n \log n)$ (see, e.g., [27]).

[†]A recent paper of Jerrum *et al.* [11] provides alternative tools based on decomposition ideas for bounding the log-Sobolev constant; those tools work in much more general settings, but give substantially weaker bounds than ours for the scenarios discussed in this paper. In particular, it seems unlikely that those methods are sensitive enough to isolate the regime where the mixing time is $O(n \log n)$.

Part (ii) of this theorem is analogous to our earlier result that the mixing time for the Ising model with +1-boundary is $O(n \log n)$ at all temperatures. This is in line with the intuition that the even boundary eliminates the only bottleneck in the dynamics. Part (i) identifies a region in which the mixing time is insensitive to the boundary condition. We would expect this to hold throughout the low-activity region $\lambda \leq \lambda_0$, and indeed, by analogy with the Ising model, also in some intermediate region beyond this. Our bound in part (i) confirms this behavior: note that the quantity $\frac{1}{\sqrt{b-1}}$ exceeds λ_0 for all $b \geq 5$, and indeed for large b it grows as $\Theta(\frac{1}{\sqrt{b}})$ compared to the $\Theta(\frac{1}{b})$ growth of λ_0 . Thus for $b \geq 5$ we establish $O(n \log n)$ mixing time in a region above the critical value λ_0 . To the best of our knowledge this is the first such result. (Note that the result of [14, 26] mentioned earlier establishes $O(n \log n)$ mixing time only for $\lambda < \frac{2}{b-1}$, which is less than λ_0 for all b and so does not even cover the whole uniqueness region.) We also mention that our analysis in this region has consequences for the infinite volume Gibbs measure itself, implying that when $\lambda \leq \frac{1}{\sqrt{b-1}}$ any Gibbs measure that is the limit of finite Gibbs distributions for some fixed boundary configuration is *extremal*, again a new result. (For results on extremality with specific boundary conditions, see [4, 16].) We elaborate on this point in the full version of the paper, where we also derive similar results for an arbitrary system with two spin values.

C. The (ferromagnetic) Potts model

Little is known about the Glauber dynamics for the Potts model on trees, beyond the facts that the mixing time is $O(n \log n)$ for arbitrary boundaries at very high temperatures (by the Dobrushin Uniqueness Condition), and is $\Omega(n^{1+\epsilon})$ for some boundaries at very low temperatures (combining results in [2, 21]). In this paper, we prove:

THEOREM 1.2. *The mixing time of the Glauber dynamics for the Potts model on an n -vertex b -ary tree is $O(n \log n)$ in all of the following situations:*

- (i) *the boundary condition is arbitrary and $\beta < \max \left\{ \beta_0, \frac{1}{2} \ln \left(\frac{\sqrt{b+1}}{\sqrt{b-1}} \right) \right\}$;*
- (ii) *the boundary condition is constant (e.g., all sites on the boundary have spin 1) and β is arbitrary;*
- (iii) *the boundary condition is free (i.e., the boundary spins are unconstrained) and $\beta < \beta_1$, where β_1 is the solution to the equation $\frac{e^{2\beta_1-1}}{e^{2\beta_1+q-1}} \cdot \frac{e^{2\beta_1-1}}{e^{2\beta_1+1}} = \frac{1}{b}$.*

Part (i) of this theorem shows that we get $O(n \log n)$ mixing time for arbitrary boundaries throughout the uniqueness region; also, since $\frac{1}{2} \ln \left(\frac{\sqrt{b+1}}{\sqrt{b-1}} \right) \geq$

$\frac{1}{2} \ln \left(\frac{b+q-1}{b-1} \right) > \beta_0$ when $q \leq 2(\sqrt{b} + 1)$, this result extends into the multiple phase region for many combinations of b and q . Part (ii) of the theorem is an analog of our earlier result that the mixing time of the Ising model with +1-boundaries is $O(n \log n)$ at all temperatures. Part (iii) is of interest for two reasons. First, since $\beta_1 > \beta_0$ always, it exhibits a natural boundary condition under which the mixing time is $O(n \log n)$ beyond the uniqueness region (but not for arbitrary β) for all combinations of b and q . Second, because of an intimate connection between the free boundary case and so-called “reconstruction problems” on trees [20] (in which the edges are noisy channels and the goal is to reconstruct a value transmitted from the root), we obtain an alternative proof of the best known value of the noise parameter under which reconstruction is impossible [21]. As we observe later, a slight strengthening of part (iii) marginally improves on this threshold.

D. The antiferromagnetic Potts model (colorings)

The sharpest result known for the Glauber dynamics on colorings is due to Vigoda [25], who shows that for arbitrary boundary conditions the mixing time is $O(n \log n)$ provided $q > \frac{11}{6}(b+1)$. This result actually holds not only for trees but for any n -vertex graph G of maximum degree $b+1$.[‡] In this paper, we extend this rapid mixing result throughout the uniqueness region, i.e., all the way down to the “critical” value $q = b+2$ for which the Gibbs measure remains unique.

THEOREM 1.3. *The mixing time of the Glauber dynamics for colorings on the n -vertex b -ary tree is $O(n \log n)$ for arbitrary boundary conditions and $q \geq b+2$.*

Note that this result is optimal: when $q = b+1$, it is not too hard to construct boundary conditions under which the Glauber dynamics is not connected.

The remainder of the paper is organized as follows. After some basic definitions in Section 2, in Section 3 we extend the analytic framework from our previous paper to general spin systems on trees, defining the quantities κ and γ and relating them to the mixing time. Then in Sections 4, 5 and 6 we specialize the analysis to the hard-core, colorings and Potts models respectively. Owing to space limitations, some details are left for the full version.

[‡]A recent sequence of papers [5, 19, 9] have reduced the required number of colors further for general graphs, under the assumption that the maximum degree is $\Omega(\log n)$. The current state of the art requires $q \geq (1+\epsilon)(b+1)$, for arbitrarily small $\epsilon > 0$ [10], but these results do not apply in our setting where the degree $b+1$ is fixed.

2 Preliminaries

For $b \geq 2$, let \mathbb{T}^b denote the infinite rooted b -ary tree (in which every vertex has b children). We will be concerned with (complete) finite subtrees T that are initial portions of \mathbb{T}^b , i.e., share the same root; if T has depth m then it has $n = (b^{m+1} - 1)/(b - 1)$ vertices, and its *boundary* ∂T consists of the children (in \mathbb{T}^b) of its leaves, i.e., $|\partial T| = b^{m+1}$. We identify subgraphs of T with their vertex sets, and write $E(A)$ for the edges within a subset A , and ∂A for the boundary of A (i.e., the neighbors of A in $(T \cup \partial T) \setminus A$).

Consider a spin system on T specified by spin values S , pair potential U and singleton potential W as in the Introduction. Let $\tau \in S^{\mathbb{T}^b}$ be a spin configuration on the infinite tree \mathbb{T}^b . We denote by Ω_T^τ the set of configurations $\sigma \in S^{T \cup \partial T}$ that agree with τ on ∂T ; i.e., τ specifies a *boundary condition* on T . The spin at x is denoted σ_x . For any $\eta \in \Omega_T^\tau$ and any subset $A \subseteq T$, the Gibbs distribution on A conditional on the configuration outside A being η is denoted μ_A^η and is defined as follows: if $\sigma \in \Omega$ agrees with η outside A then

$$\mu_A^\eta(\sigma) \propto \exp\left[-\left(\sum_{xy \in E(A \cup \partial A)} U(\sigma_x, \sigma_y) + \sum_{x \in A} W(\sigma_x)\right)\right]; \quad (1)$$

otherwise, $\mu_A^\eta(\sigma) = 0$. In particular, when $A = T$, $\mu_T^\eta = \mu_T^\tau$ is simply the Gibbs distribution on the whole of T with boundary condition τ . We will assume that μ_A^η is well defined (i.e., that the expression in (1) is positive for at least one σ) for every A , τ and η (even when $\mu_T^\tau(\eta) = 0$); we call the spin system *permissive* in this case. (Note that this is an issue only for systems with hard constraints.) All the examples in this paper are clearly permissive. We usually abbreviate Ω_T^τ and μ_T^τ to Ω and μ respectively. When there are hard constraints (i.e., $U(s_1, s_2) = \infty$ for some s_1, s_2) we remove invalid configurations (i.e., those for which $\mu(\sigma) = 0$) from Ω .

The (*heat-bath*) *Glauber dynamics* is the following Markov chain on $\Omega = \Omega_T^\tau$. In configuration $\eta \in \Omega$, transitions are made as follows:

- (i) pick a vertex $x \in T$ u.a.r., and a spin value s chosen from the distribution of the spin at x conditional on the spins of its neighbors (i.e., s has the distribution of the spin at x in $\mu_{\{x\}}^\eta$);
- (ii) go to configuration $\eta^{x,s}$ obtained from η by setting the spin at x to s .

It is a well-known fact (and easily checked) that the Glauber dynamics is ergodic and reversible w.r.t. the Gibbs distribution $\mu = \mu_T^\tau$, and so converges to the

stationary distribution μ . We measure the rate of convergence by the *mixing time*:

$$t_{\text{mix}} = \min\{t : \|P^t(\sigma, \cdot) - \mu\| \leq \frac{1}{2e} \text{ for all } \sigma \in \Omega\}, \quad (2)$$

where $P^t(\sigma, \cdot)$ denotes the distribution of the dynamics after t steps starting from configuration σ , and $\|\cdot\|$ is variation distance. (The constant $\frac{1}{2e}$ in this definition is for algebraic convenience only.)

When we say that the mixing time of the Glauber dynamics is $O(n \log n)$ for some boundary condition τ , we mean that for all finite T , the mixing time for μ_T^τ is $\leq cn \log n$ for a constant c that depends only on b and the potentials.

3 A framework for $O(n \log n)$ mixing time

In [18] we developed a criterion for $O(n \log n)$ mixing time of the Glauber dynamics on trees, and used it to analyze the Ising model both for arbitrary boundary conditions and in the important special case of $+1$ -boundaries. This criterion generalizes immediately to arbitrary spin systems, as we describe in this section along with some useful extensions.

The key ingredients are two quantities, which we call κ and γ , that bound the rate of percolation of disagreements down and up the tree respectively. Both are properties of the collection of Gibbs distributions $\{\mu_T^\tau\}$, where τ is fixed and T ranges over all finite complete subtrees of \mathbb{T}^b . To define κ and γ we need a little notation. For a vertex $x \in T$, write T_x for the (maximal) subtree rooted at x . When x is not the root of T , let $\mu_{T_x}^s$ denote the Gibbs distribution in which the parent of x has its spin fixed to s and the configuration on the bottom boundary of T_x is specified by τ (the global boundary condition on T)[§]. For two distributions μ_1, μ_2 on Ω , $\|\mu_1 - \mu_2\|_x$ denotes the variation distance between the projections of μ_1 and μ_2 onto the spin at x , i.e., $\|\mu_1 - \mu_2\|_x = \frac{1}{2} \sum_{s \in S} |\mu_1(\sigma_x = s) - \mu_2(\sigma_x = s)|$. Recall that $\eta^{x,s}$ is the configuration η with the spin at x set to s .

DEFINITION 3.1. *For a collection of Gibbs distributions $\{\mu_T^\tau\}$ as above, define $\kappa \equiv \kappa(\{\mu_T^\tau\})$ and $\gamma \equiv \gamma(\{\mu_T^\tau\})$ by*

- (i) $\kappa = \sup_T \max_{z, s, s'} \|\mu_{T_z}^s - \mu_{T_z}^{s'}\|_z$;
- (ii) $\gamma = \sup_T \max \|\mu_A^{\eta^{y,s}} - \mu_A^{\eta^{y,s'}}\|_z$, where the maximum is taken over all subsets $A \subset T$, all boundary configurations η , all sites $y \in \partial A$, all neighbors $z \in A$ of y , and all spins $s, s' \in S$.

[§]We do not specify the configuration in the rest of $T \setminus T_x$ as it has no influence on the distribution inside T_x once the spin at the parent of x is fixed.

Remark: Note that κ is the same as γ , except that the maximization is restricted to $A = T_z$ and the boundary vertex y being the parent of z ; hence always $\kappa \leq \gamma$. Since κ involves Gibbs distributions only on maximal subtrees T_z , it may depend on the boundary condition τ at the bottom of the tree. By contrast, γ bounds the worst-case probability of disagreement for an *arbitrary* subset A and arbitrary boundary configuration around A , and hence depends only on the potentials of the system and not on τ . It is the dependence of κ on τ that opens up the possibility of an analysis that is specific to the boundary condition.

The intuition for these definitions comes from the following claim, which relates κ and γ to the rate of disagreement percolation in the tree. For any T and site $x \in T$, write \widehat{T}_x for $T_x \setminus \{x\}$, the subtree T_x excluding its root, and $\mu_{\widehat{T}_x}^s$ for the Gibbs distribution when the spin at x is fixed to s . Also, for $\ell \leq \text{height}(x) + 1$ write $B_{x,\ell}$ for the subtree (or “block”) of height $\ell - 1$ rooted at x (i.e., $B_{x,\ell}$ has ℓ levels). For two configurations $\sigma, \sigma' \in \Omega$, let $|\sigma - \sigma'|_{x,\ell}$ denote the number of sites ℓ levels below x (i.e., on the bottom boundary of $B_{x,\ell}$) at which σ and σ' differ. Note that $|\sigma - \sigma'| \leq b^\ell$.

CLAIM 3.2. *For every $x \in T$ and all $\ell \leq \text{height}(x) + 1$ the following hold:*

- (i) *For all s, s' , there is a coupling $\nu = \nu^{s,s'}$ of $\mu_{\widehat{T}_x}^s$ and $\mu_{\widehat{T}_x}^{s'}$ for which $\mathbb{E}_\nu |\sigma - \sigma'|_{x,\ell} \leq (\kappa b)^\ell$.*
- (ii) *For any $\eta, \eta' \in \Omega$ that have the same spin value at the parent of x , $\|\mu_{B_{x,\ell}}^\eta - \mu_{B_{x,\ell}}^{\eta'}\|_x \leq \gamma^\ell \cdot |\eta - \eta'|_{x,\ell}$.*

The proof of this claim follows from a standard recursive coupling along paths in the tree: see [18, Claim 4.4]. Part (i) shows that κ bounds the probability of a disagreement percolating *down* the tree: i.e., when we fix a disagreement at x and recursively couple the distributions on the children of x , the expected proportion of disagreements after ℓ levels is at most κ^ℓ . Similarly, from part (ii) we see that γ bounds the probability of a disagreement percolating *up* the tree: i.e., when we fix a single disagreement at level ℓ below x , the probability of this disagreement reaching x is at most γ^ℓ .

We now state a theorem that will be our main analytical tool. The theorem gives a sufficient condition for $O(n \log n)$ mixing time in terms of the quantities κ and γ .

THEOREM 3.3. *Consider an arbitrary (permissive) spin system and a boundary condition τ (a configuration on \mathbb{T}^b). If $\kappa \equiv \kappa(\{\mu_T^\tau\})$ and $\gamma \equiv \gamma(\{\mu_T^\tau\})$ satisfy $\max\{\gamma\kappa b, \gamma\} < 1$ then the mixing time of the associated Glauber dynamics is $O(n \log n)$.*

Proof: The proof follows by combining Theorems 3.4 and 5.1 of [18][¶], which together imply that, under the above conditions on κ and γ , the *logarithmic Sobolev constant* of the dynamics is $\Omega(\frac{1}{n})$. By standard facts relating the log-Sobolev constant to the mixing time, this implies that the mixing time is $O(n \log n)$. \square

Theorem 3.3 tells us that, to prove $O(n \log n)$ mixing time, it is enough to estimate the quantities κ and γ for the spin system and boundary condition in question. As we shall see, this can be done using calculations specific to the situation at hand. In [18] we carried out these calculations for the Ising model with various boundary conditions; our goal in this paper is to perform analogous calculations for some other important models, thus demonstrating the utility of the approach.

For some of these other models, we will require two minor but useful generalizations of the above framework, which we now describe. Both generalizations stem from the observation that the role of the definitions of κ and γ is to obtain the bounds on disagreement percolation stated in Claim 3.2. In fact, in Theorem 3.3 we can replace κ and γ by any two values κ' and γ' for which the upper bounds in parts (i) and (ii) of Claim 3.2 are $O((\kappa' b)^\ell)$ and $O(\gamma'^\ell)$ respectively (where the $O(\cdot)$ hides constants independent of ℓ). The arguments leading to Theorem 3.3 are easily seen to hold in this slightly looser setting.

Our first generalization (which will be particularly useful for “non-attractive” systems, including systems with hard constraints) is to consider two levels of the tree at a time, rather than a single level as in Definition 3.1. Accordingly, define

$$\kappa_2 = \sup_T \max_{z, w \prec z, \substack{s_1, s'_1, s_2, s'_2}} \sqrt{\|\mu_{T_z}^{s_1} - \mu_{T_z}^{s'_1}\|_z \cdot \|\mu_{T_w}^{s_2} - \mu_{T_w}^{s'_2}\|_w}, \quad (3)$$

where $w \prec z$ denotes “ w is a child of z .” In fact, we may restrict the maximization to sites z of even (or odd) height. With this definition, it is easy to see that the upper bound in Claim 3.2(i) for the probability of disagreement percolating down the tree can be replaced by $(\kappa_2)^{2(\ell/2-1)} b^\ell = O((\kappa_2 b)^\ell)$. We therefore get the following generalization of Theorem 3.3:

THEOREM 3.3'. *In the setting of Theorem 3.3, if κ_2 and γ satisfy $\max\{\gamma\kappa_2 b, \gamma\} < 1$ then the mixing time of the associated Glauber dynamics is $O(n \log n)$.*

[¶]These theorems are stated for the special case of the Ising model. However, it is easily seen that their proofs make no use of the specific form of the Ising potentials, and thus apply to arbitrary permissive spin systems on trees.

Our second generalization exploits the fact that, when deriving the bound on upward percolation in part (ii) of Claim 3.2, it is enough to control the probability of a disagreement percolating upwards one level from y to z only when z is sufficiently far from the boundary and the root of $B_{x,\ell}$. Let $\hat{\gamma}$ be defined in the same way as γ , but with the maximization restricted to sets A that include the full subtree of depth d rooted at z under the orientation in which y is the parent of z ; here d is an implicit parameter whose value may change from model to model, but will in each case be a constant independent of the size of T . Then the probability of disagreement percolating one level upwards to z , where z is at distance at least d from the boundary and root of $B_{x,\ell}$, is bounded by $\hat{\gamma}$. Thus it is easy to modify the proof of Claim 3.2 so that the factor γ^ℓ in part (ii) is replaced by $\hat{\gamma}^{\ell-2d} = O(\hat{\gamma}^\ell)$. Similarly, we define $\hat{\kappa}$ as before but with the maximization restricted to z that are at distance at least d from the boundary of T . Whenever we use $\hat{\gamma}$, we will also use $\hat{\kappa}$ with the same value of d so that we still have $\hat{\kappa} \leq \hat{\gamma}$. This leads to our second generalization of Theorem 3.3:

THEOREM 3.3''. *In the setting of Theorem 3.3, if $\hat{\kappa}$ and $\hat{\gamma}$ satisfy $\max\{\hat{\gamma}\hat{\kappa}b, \hat{\gamma}\} < 1$ then the mixing time of the associated Glauber dynamics is $O(n \log n)$.*

4 Independent sets (hard-core model)

In this section, we will prove that the mixing time of the Glauber dynamics for sampling independent sets is $O(n \log n)$ in all the situations covered by Theorem 1.1 in the Introduction.

In light of Theorem 3.3, in order to show $O(n \log n)$ mixing time it is enough to bound the quantities κ and γ such that $\max\{\gamma\kappa b, \gamma\} < 1$, where κ and γ can also be replaced by their variants as explained in Section 3. Thus, Theorem 1.1 will follow from:

LEMMA 4.1. *For the independent sets model with activity parameter λ :*

- (i) *For all values of λ , we have $\hat{\gamma} < \frac{\lambda}{1+\lambda}$ (for some choice of the implicit parameter d in $\hat{\gamma}$). In addition, for $\lambda \leq \lambda_0$ and every $\varepsilon > 0$, we have $\hat{\gamma} \leq \frac{1+\varepsilon}{b}$ (again for some $d = d(\varepsilon)$).*
- (ii) *For $\lambda > \lambda_0$, when τ is the 0-boundary condition (i.e., all sites are unoccupied), we have $\kappa_2 \leq \frac{1}{b}$.*

Recall that $\hat{\kappa} \leq \hat{\gamma}$, so from part (i) of this lemma we conclude that $\hat{\gamma}\hat{\kappa}b < 1$ when $(\frac{\lambda}{1+\lambda})^2 \leq \frac{1}{b}$, i.e., when $\lambda \leq \frac{1}{\sqrt{b}-1}$, and also whenever $\lambda \leq \lambda_0$. Since also $\hat{\gamma} < 1$ for all finite λ , part (i) of Theorem 1.1 follows using Theorem 3.3''. This also dispenses with part (ii) of Theorem 1.1 in the region $\lambda \leq \lambda_0$. The remainder

of part (ii) (when $\lambda > \lambda_0$) follows immediately from part (ii) of the above lemma, using Theorem 3.3' and the fact that $\gamma < 1$. (Note that analyzing the 0-boundary for all depths of T handles both odd and even boundary conditions.)

Recall that the calculation of κ and γ involves bounding the effect of changing the spin of a boundary site on an inside neighbor (measured in terms of variation distance). Consider a subset A , an arbitrary boundary configuration η , and sites $y \in \partial A$ and $z \in A$ as in Definition 3.1. We need to compare two distributions, $\mu_A^{\eta^{y,s}}$ and $\mu_A^{\eta^{y,s'}}$. In the first, the site y is occupied (spin value $s = 1$) and in the second y is unoccupied (spin value $s' = 0$). Now, from the definition of the hard-core model, in the first distribution the site z is unoccupied with certainty, and hence the variation distance between the two distributions at z is exactly the probability that z is occupied in the *second* distribution (where y is unoccupied, or equivalently, where the edge connecting y and z is removed). Let p_z stand for this last probability. To summarize, for any subset A , any two adjacent sites $y \in \partial A$ and $z \in A$, and any boundary configuration η ,

$$\|\mu_A^{\eta^{y,1}} - \mu_A^{\eta^{y,0}}\|_z = \mu_A^{\eta^{y,0}}(\sigma_z = 1) \equiv p_z. \quad (4)$$

Our main goal in the rest of this section is to bound the probability of occupation p_z for the relevant values of λ and global boundary condition τ .

Proof of Lemma 4.1 part (i)

We start with the easy observation that, for any subset A , any boundary configuration η and any site $z \in A$, $\mu_A^\eta(\sigma_z = 1) \leq \frac{\lambda}{1+\lambda}$, simply because the r.h.s. is the probability of z being occupied when all its neighbors are unoccupied, and if one of its neighbors is occupied then z is unoccupied with certainty. Using (4), we deduce that $\gamma \leq \frac{\lambda}{1+\lambda}$. We can strengthen this to $\hat{\gamma} < \frac{\lambda}{1+\lambda}$ by noticing that equality is achieved in the above only when all neighbors of z are unoccupied with certainty, which can happen only if all neighbors of z are in ∂A or adjacent to ∂A . So by taking $d = 3$ in the definition of $\hat{\gamma}$ we are done.

We go on to consider the case $\lambda \leq \lambda_0$, i.e., when the infinite-volume Gibbs measure is unique. Here, in order to bound $\hat{\gamma}$, we consider subsets A that include the full subtree $B_{z,d}$ of height d rooted at z , where d is a large enough constant; in fact, w.l.o.g. we may take A to be exactly such a subtree. Recall that we are interested in p_z , the probability that z is occupied when its parent y is unoccupied, or equivalently, when $B_{z,d}$ is disconnected at its root z from the rest of T . We need to show that $p_z \leq \frac{1+\varepsilon}{b}$ for every boundary condition η on the bottom of $B_{z,d}$.

Now since we only consider the case $\lambda \leq \lambda_0$, where by definition the Gibbs measure is unique, if we let the depth d of $B_{z,d}$ grow then the corresponding sequence of probabilities p_z converges to some value p_0 independent of the boundary configuration η . It is therefore enough to show that $p_0 \leq \frac{1}{b}$, since then for large enough d we will have $p_z \leq \frac{1+\varepsilon}{b}$, as required.

The calculation of p_0 can be done by recursively calculating the sequence p_z as was done in, e.g., [13] and as we did for the Ising model in [18]. Let $R_z = \frac{p_z}{1-p_z}$ stand for the ratio of probabilities that the site z is occupied and unoccupied respectively, given that its parent is unoccupied. A simple calculation verifies that $R_z = \lambda \prod_{w \prec z} \left(\frac{1}{1+R_w} \right)$, where R_w is computed w.r.t. $A = B_{w,d-1}$ and the same η . For each λ , we thus define the function

$$J(a) = \lambda \left(\frac{1}{1+a} \right)^b \quad (5)$$

and conclude that $R_0 \equiv \frac{p_0}{1-p_0}$ must be a fixed point of J . Notice that since J is monotonically decreasing and $J(0) = \lambda > 0$, it has a unique fixed point for every λ ; we denote this fixed point by $a_0 = a_0(\lambda)$. Now uniqueness of the Gibbs measure is equivalent to the fixed point a_0 being *attractive* (which corresponds to the existence of the limit p_0). This in turn corresponds to the derivative of J at a_0 being at least -1 ; indeed, when $\lambda = \lambda_0$ is critical the derivative at $a_0 = a_0(\lambda_0)$ is exactly -1 .

Now let $J'(a) = \frac{\partial J(a)}{\partial a}$. We will make use of the fact that $J'(a_0) = J'(R_0) \geq -1$ for $\lambda \leq \lambda_0$ in order to establish that $p_0 \leq \frac{1}{b}$. First note that, by a straightforward calculation,

$$J'(a) = -b \cdot \frac{J(a)}{1+a}. \quad (6)$$

Thus, since R_0 is a fixed point of J , we have $-J'(R_0) = b \cdot \frac{J(R_0)}{1+R_0} = b \cdot \frac{R_0}{1+R_0} = bp_0$. On the other hand, since $-J'(R_0) = -J'(a_0) \leq 1$, we conclude that $p_0 \leq \frac{1}{b}$, as required. This completes the proof of part (i) of Lemma 4.1.

Proof of Lemma 4.1 part (ii)

In this part, the boundary condition τ is set to all-0 and $\lambda > \lambda_0$, meaning that the Gibbs measure is not unique or equivalently, as we saw in the previous part, $J'(a_0) < -1$.

Once again, our aim is to calculate the probabilities of occupation p_z . Here, however, A is a maximal subtree T_z and the bottom boundary condition is τ (rather than arbitrary). We start by noticing that since τ is the all-0 configuration, then $R_z = J^{(\ell)}(0)$, where $J^{(\ell)}$ stands for the ℓ -fold composition of J and ℓ is the distance of z from the bottom boundary of T . However,

this time the sequence R_z does not converge since a_0 is not an attractive fixed point. In fact, the sequence oscillates around a_0 . It hence makes sense to look at two separate sequences, one for even heights and one for odd heights. This leads us to analyze the function $J_2(a) \equiv J^{(2)}(a) \equiv J(J(a))$, which corresponds to jumping two levels at a time. Notice that for a site z at height 2ℓ from the bottom, $R_z = J_2^{(\ell)}(0)$. Let us now describe some properties of the function J_2 (see Fig. 1):

1. J_2 is continuous and increasing on $[0, \infty)$, with $J_2(0) = \lambda/(1+\lambda)^b$ and $\sup_a J_2(a) = \lambda$.
2. a_0 (the unique fixed point of J) is a fixed point of J_2 .
3. If $\lambda \leq \lambda_0$ then a_0 is the unique fixed point of J_2 . If $\lambda > \lambda_0$ then J_2 has three fixed points $a_1 < a_0 < a_2$, where $J(a_1) = a_2$ and $J(a_2) = a_1$.
4. The derivative $J_2'(a) \equiv \frac{\partial J_2(a)}{\partial a}$ is continuous, and for $\lambda > \lambda_0$, $J_2'(a)$ is monotonically increasing on $[0, a_1]$ with $J_2'(a_1) < 1$.

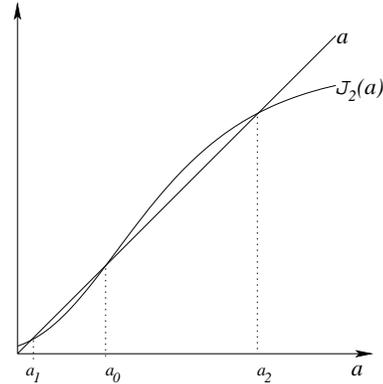


Figure 1: Curve of the function $J_2(a)$, used in the proof of Lemma 4.1, for $\lambda > \lambda_0$. The points a_1, a_0, a_2 are the fixed points of J_2 in increasing order.

It is thus easy to see that, for example, for $\lambda > \lambda_0$, $J_2^{(\ell)}(0)$ converges to a_1 from below, i.e, for all sites z whose height is even, $R_z \leq a_1$. Similarly, if the height of z is odd, $R_z \geq a_2$. We will use the properties of J_2 in order to bound κ_2 by $\frac{1}{b}$. Recall that $\kappa_2^2 = \max_{w \prec z} \|\mu_{T_z}^1 - \mu_{T_z}^0\|_z \cdot \|\mu_{T_w}^1 - \mu_{T_w}^0\|_w = \max_{w \prec z} p_z p_w$, and that it is enough to consider sites z whose height in T is odd. We will show that, for every site z of odd height, and every child w of z , $p_z p_w \leq \frac{1}{b^2}$.

Again we show a connection between the derivative, this time of J_2 , and the relevant probabilities of occupation, i.e., $p_z p_w$. Using (6) to calculate $J_2'(a)$ gives:

$$\begin{aligned}
J'_2(a) &= J'(J(a)) \cdot J'(a) = b \left(\frac{J(J(a))}{1+J(a)} \right) \cdot b \left(\frac{J(a)}{1+a} \right) \\
&= b^2 \cdot \frac{J_2(a)}{a} \cdot \frac{J(a)}{1+J(a)} \cdot \frac{a}{1+a}.
\end{aligned}$$

Recalling that z is the parent of w , together with the fact that since τ is uniformly 0 we have $R_w = R_{w'}$ for any child w' of z , we observe that $R_z = J(R_w)$. Therefore, $J'_2(R_w) = b^2 \cdot \frac{J_2(R_w)}{R_w} \cdot \frac{R_z}{1+R_z} \cdot \frac{R_w}{1+R_w} = b^2 \cdot \frac{J_2(R_w)}{R_w} \cdot p_z p_w$. Now, since z is of odd height then w is of even height and therefore $R_w \leq a_1$. Notice that this means $J_2(R_w) \geq R_w$, because J is continuous with $J(0) > 0$ and a_1 is the smallest fixed point of J . Combining this with the fact that $J'_2(R_w) < 1$ (since $R_w \leq a_1$), we conclude that $p_z p_w \leq \frac{1}{b^2}$. This completes the proof of part (ii) of Lemma 4.1.

5 Colorings (antiferromagnetic Potts model at zero temperature)

In this section we will prove Theorem 1.3: that the mixing time of the Glauber dynamics for sampling proper colorings is $O(n \log n)$ for all boundary conditions provided that the number of colors $q \geq b+2$, i.e., whenever the infinite-volume Gibbs measure is unique. In order to again make use of the machinery from Section 3 involving κ and γ , we prove the following:

LEMMA 5.1. *For the colorings model with q colors, if the infinite-volume Gibbs measure is unique then for every $\varepsilon > 0$ we have $\hat{\gamma} \leq \frac{1+\varepsilon}{q-1-\varepsilon}$ (for a suitable choice $d = d(\varepsilon)$ of the implicit constant in $\hat{\gamma}$).*

Since in [12] it was shown that the Gibbs measure is unique for all $q \geq b+2$, we conclude that for these values of q , $\hat{\gamma} \leq \frac{1+\varepsilon}{q-1-\varepsilon} < \frac{1}{b}$. Theorem 1.3 now follows from Theorem 3.3'' and $\hat{\kappa} \leq \hat{\gamma}$.

Remark: The fact that $\hat{\gamma} < \frac{1}{b}$ means not only that the influence of any boundary configuration on the spin at the root decays with the distance of the boundary from the root (as is already implied by the fact the Gibbs measure is unique), but that it decays *exponentially fast*. This fact is of independent interest, and to the best of our understanding was not obvious from the proof of uniqueness given in [12].

Proof of Lemma 5.1: Recall that in order to bound $\hat{\gamma}$, we need to consider a subset A that includes the full subtree of depth d rooted at z , and bound the variation distance $\max_{s,s'} \|\mu_A^{\eta^{y,s}} - \mu_A^{\eta^{y,s'}}\|_z$ for an arbitrary boundary configuration η , where y is the (unique) neighbor of z in ∂A .

Now it is easy to see that for the colorings model,

$$\|\mu_A^{\eta^{y,s}} - \mu_A^{\eta^{y,s'}}\|_z = \max\{\mu_A^{\eta^{y,s}}(\sigma_z = s'), \mu_A^{\eta^{y,s'}}(\sigma_z = s)\}.$$

This can be shown as follows. Notice that for every color s'' that differs from both s and s' , $\mu_A^{\eta^{y,s}}(\sigma_z = s'') = \frac{1-\mu_A^{\eta^{y,s}}(\sigma_z = s')}{1-\mu_A^{\eta^{y,s'}}(\sigma_z = s)} \cdot \mu_A^{\eta^{y,s'}}(\sigma_z = s'')$. Hence, if $\mu_A^{\eta^{y,s}}(\sigma_z = s') \geq \mu_A^{\eta^{y,s'}}(\sigma_z = s)$, then $\mu_A^{\eta^{y,s}}(\sigma_z = s'') \leq \mu_A^{\eta^{y,s'}}(\sigma_z = s'')$ for every $s'' \neq s'$ and so the event $\mathcal{E} = \{\sigma_z = s'\}$ maximizes the expression $|\mu_A^{\eta^{y,s}}(\mathcal{E}) - \mu_A^{\eta^{y,s'}}(\mathcal{E})|$ over all events \mathcal{E} that only depend on σ_z .

It is now convenient to consider the distribution induced by removing the edge from z to y (i.e., with a “free” condition at y). Call this distribution $\mu_A^{\eta^{y,*}}$. Let $p_z(s) = \mu_A^{\eta^{y,*}}(\sigma_z = s)$, and notice that for the colorings model $\mu_A^{\eta^{y,s}}(\sigma_z = s') = \frac{p_z(s')}{1-p_z(s)}$ simply because $\mu_A^{\eta^{y,s}}(\cdot) = \mu_A^{\eta^{y,*}}(\cdot | \sigma_z \neq s)$ in this model. Thus, $\max_{s,s'} \|\mu_A^{\eta^{y,s}} - \mu_A^{\eta^{y,s'}}\|_z = \max_{s,s'} \mu_A^{\eta^{y,s}}(\sigma_z = s') = \max_{s,s'} \frac{p_z(s')}{1-p_z(s)}$.

So to obtain the claimed bound on $\hat{\gamma}$ we have to show that, for all sets A as above, and all boundary configurations η , $\max_{s,s'} \frac{p_z(s')}{1-p_z(s)} \leq \frac{1+\varepsilon}{q-1-\varepsilon}$. It is at this point that we use the assumption that the Gibbs measure is unique. This means that, if d (the depth of the full subtree contained in A) is large enough, the distribution $p_z(\cdot)$ is arbitrarily close to the uniform distribution, regardless of the boundary configuration η . Thus, for every $\varepsilon > 0$ there exists a (large enough) constant d such that $p_z(s) \leq \frac{1+\varepsilon}{q}$ for all colors s . Hence, $\max_{s,s'} \frac{p_z(s')}{1-p_z(s)} \leq \frac{1+\varepsilon}{q-1-\varepsilon}$, as required. \square

6 Ferromagnetic Potts model

In this section we will prove that, for the Potts model, the mixing time is $O(n \log n)$ in all the situations described in Theorem 1.2. As in the discussion of other models in previous sections, we use the machinery from Section 3. The following lemma sets out the relevant properties of κ and γ ; the proof uses similar ideas to those in the proofs of Lemmas 4.1 and 5.1 and is left for the full version.

LEMMA 6.1. *For the Potts model with q colors at inverse temperature β the following hold:*

- (i) $\gamma \leq \frac{e^{2\beta} - 1}{e^{2\beta} + 1} - \delta$, where $\delta = \delta(b, q, \beta) \geq 0$ with equality if and only if $q = 2$. Furthermore, $\delta(b, q, \beta)$ is increasing in q and decreasing in b and β . [The exact definition of $\delta(b, q, \beta)$ is rather involved and will be given in the full version of the paper].
- (ii) If $\beta < \beta_0$ (i.e., the Gibbs measure is unique) then $\hat{\gamma} < \frac{1}{b}$ (for an appropriate choice of d , the implicit constant in $\hat{\gamma}$).

- (iii) If $\beta \geq \beta_0$ and the boundary condition τ is constant, then $\kappa \leq \frac{1}{b}$.
- (iv) If the boundary condition τ is free, then $\kappa = \frac{e^{2\beta} - 1}{e^{2\beta} + q - 1}$.

Part (i) of Theorem 1.2 follows from parts (i) and (ii) of Lemma 6.1 and the fact that for any boundary condition $\kappa \leq \gamma$. Part (ii) of Theorem 1.2 follows from parts (ii) and (iii) of the lemma and the fact that $\gamma < 1$ (as is apparent from part (i) of the lemma). Finally, part (iii) of Theorem 1.2 follows from parts (i) and (iv) of the lemma. In fact, in light of the bound on γ from part (i), the range of β in part (iii) of Theorem 1.2 can be improved slightly by letting β_1 be the solution to the equation $\frac{e^{2\beta_1} - 1}{e^{2\beta_1} + q - 1} (\frac{e^{2\beta_1} - 1}{e^{2\beta_1} + 1} - \delta) = \frac{1}{b}$, where $\delta = \delta(b, q, \beta)$ is as in part (i) of Lemma 6.1. We note that this modified definition of β_1 is only marginally larger than the original definition of β_1 in Theorem 1.2, and we mention it only in order to show that we can go further than the original threshold, a fact that is interesting due to its implication for the reconstruction problem [20] as mentioned in the Introduction. We elaborate on this point in the full version of the paper.

References

- [1] R.J. BAXTER, *Exactly solved models in statistical mechanics*, Academic Press, London, 1982.
- [2] N. BERGER, C. KENYON, E. MOSSEL and Y. PERES, “Glauber dynamics on trees and hyperbolic graphs,” preprint, 2003. Preliminary version: C. KENYON, E. MOSSEL and Y. PERES, “Glauber dynamics on trees and hyperbolic graphs,” *Proc. 42nd IEEE Symposium on Foundations of Computer Science*, 2001, pp. 568–578.
- [3] G.R. BRIGHTWELL and P. WINKLER, “Random colorings of a Cayley tree,” *Contemporary combinatorics*, Bolyai Soc. Math. Studies **10**, János Bolyai Math. Soc., Budapest, 2002, pp. 247–276.
- [4] G.R. BRIGHTWELL and P. WINKLER, “A second threshold for the hard-core model on a Bethe lattice,” INI Report NI03002-CMP To appear in *Rand. Struc. & Alg.*
- [5] M. DYER and A. FRIEZE, “Randomly colouring graphs with lower bounds on girth and maximum degree,” *Proc. 42nd IEEE Symposium on Foundations of Computer Science*, 2001, pp. 579–587.
- [6] W. EVANS, C. KENYON, Y. PERES and L.J. SCHULMAN, “Broadcasting on trees and the Ising model,” *Ann. Appl. Probab.* **10** (2000), pp. 410–433.
- [7] H.-O. GEORGII, *Gibbs measures and phase transitions*, Walter de Gruyter & Co., Berlin, 1988.
- [8] O. HÄGGSTRÖM, “The random-cluster model on a homogeneous tree,” *Probab. Theory Related Fields* **104** (1996), pp. 231–253.
- [9] T.P. HAYES, “Randomly coloring graphs with girth five,” *Proc. 35th ACM Symposium on Theory of Computing*, 2003, pp. 269–278.
- [10] T.P. HAYES and E. VIGODA, “A non-Markovian coupling for randomly sampling colorings,” to appear in *Proc. 44th IEEE Symp. on Found. of Comp. Sci.*, 2003.
- [11] M. JERRUM, J.-B. SON, P. TETALI and E. VIGODA, “Elementary bounds on Poincaré and log-Sobolev constants for decomposable Markov chains,” INI Report NI03006-CMP To appear in *Ann. Appl. Probab.*
- [12] J. JONASSON “Uniqueness of uniform random colorings of regular trees,” *Statist. Probab. Lett.* **57** (2002), pp. 243–248.
- [13] F.P. KELLY, “Stochastic models of computer communication systems,” *J. Royal Statist. Soc. B* **47** (1985), pp. 379–395.
- [14] M. LUBY and E. VIGODA, “Fast convergence of the Glauber dynamics for sampling independent sets,” *Rand. Struct. & Algo.* **15** (1999), pp. 229–241.
- [15] R. LYONS, Phase transitions on non amenable graphs, *J. Math. Phys.* **41** (2000), pp. 1099–1127.
- [16] J.B. MARTIN, “Reconstruction thresholds on regular trees,” <http://arxiv.org/abs/math.PR/0305400>, to appear in DMTCS volume *Proc. Discrete Random Walks 2003*.
- [17] F. MARTINELLI, “Lectures on Glauber dynamics for discrete spin models,” *Lectures on Probability Theory and Statistics (Saint-Flour, 1997)*, Lecture notes in Mathematics **1717**, pp. 93–191, Springer, Berlin, 1998.
- [18] F. MARTINELLI, A. SINCLAIR and D. WEITZ, “The Ising model on trees: Boundary conditions and mixing time,” Tech. Report UCB//CSD-03-1256, Dept. of EECS, UC Berkeley, July 2003. <http://sunsite.Berkeley.EDU/NCSTRL/> Extended abstract to appear in *Proc. 44th IEEE Symposium on Foundations of Computer Science*, 2003.
- [19] M. MOLLOY, “The Glauber dynamics on colorings of a graph with high girth and maximum degree,” *Proc. 34th ACM Symp. on Th. of Comp.*, 2002, pp. 91–98.
- [20] E. MOSSEL, “Survey: Information flow on trees,” Preprint, October 2002, to appear in DIMACS volume *Graphs, Morphisms and Statistical Physics*.
- [21] E. MOSSEL and Y. PERES, “Information flow on trees,” *Ann. Appl. Probab.* **13** (2003), pp. 817–844.
- [22] R.B. POTTS, “Some generalized order-disorder transformations,” *Proc. Cambridge Philos. Soc.* **48**, pp. 106–109.
- [23] C.J. PRESTON, *Gibbs states on countable sets*, Cambridge University Press, London, 1974.
- [24] F. SPITZER, “Markov random fields on an infinite tree,” *Ann. Probab.* **3** (1975), pp. 387–398.
- [25] E. VIGODA, “Improved bounds for sampling colorings,” *J. Math. Phys.* **41** (2000), pp. 1555–1569.
- [26] E. VIGODA, “A note on the Glauber dynamics for sampling independent sets,” *Electronic Journal of Combinatorics* **8(1)** (2001).
- [27] D. WEITZ “Combinatorial criteria for uniqueness of the Gibbs measure,” preprint, 2003.
- [28] F.Y. WU, “The Potts model,” *Reviews of Modern Physics* **54** (1982), pp. 235–268.