3-2005

Glauber Dynamics on Trees and Hyperbolic Graphs

Noam Berger

Claire Kenyon

Elchanan Mossel
University of Pennsylvania

Yuval Peres

Recommended Citation

Berger, N., Kenyon, C., Mossel, E., & Peres, Y. (2005). Glauber Dynamics on Trees and Hyperbolic Graphs. *Probability Theory and Related Fields, 131* (3), 311-340. http://dx.doi.org/10.1007/s00440-004-0369-4

Glauber Dynamics on Trees and Hyperbolic Graphs

Abstract
We study continuous time Glauber dynamics for random configurations with local constraints (e.g. proper coloring, Ising and Potts models) on finite graphs with n vertices and of bounded degree. We show that the relaxation time (defined as the reciprocal of the spectral gap $|\lambda_1-\lambda_2|$) for the dynamics on trees and on planar hyperbolic graphs, is polynomial in n. For these hyperbolic graphs, this yields a general polynomial sampling algorithm for random configurations. We then show that for general graphs, if the relaxation time τ_2 satisfies $\tau_2=O(1)$, then the correlation coefficient, and the mutual information, between any local function (which depends only on the configuration in a fixed window) and the boundary conditions, decays exponentially in the distance between the window and the boundary. For the Ising model on a regular tree, this condition is sharp.

Disciplines
Statistics and Probability

Glauber Dynamics on Trees and Hyperbolic Graphs

Noam Berger *
University of California, Berkeley
noam@stat.berkeley.edu

Claire Kenyon
LRI, UMR CNRS
Université Paris-Sud, France
kenyon@lix.polytechnique.fr

Elchanan Mossel †
University of California, Berkeley
mossel@stat.berkeley.edu

Yuval Peres‡
University of California, Berkeley
peres@stat.berkeley.edu

February 1, 2008

Abstract

We study continuous time Glauber dynamics for random configurations with local constraints (e.g. proper coloring, Ising and Potts models) on finite graphs with n vertices and of bounded degree. We

*Research supported by Microsoft graduate fellowship.
†Supported by a visiting position at INRIA and a PostDoc at Microsoft research.
‡Research supported by NSF Grants DMS-0104073, CCR-0121555 and a Miller Professorship at UC Berkeley.

show that the relaxation time (defined as the reciprocal of the spectral gap $|\lambda_1 - \lambda_2|$) for the dynamics on trees and on planar hyperbolic graphs, is polynomial in n. For these hyperbolic graphs, this yields a general polynomial sampling algorithm for random configurations. We then show that for general graphs, if the relaxation time τ_2 satisfies $\tau_2 = O(1)$, then the correlation coefficient, and the mutual information, between any local function (which depends only on the configuration in a fixed window) and the boundary conditions, decays exponentially in the distance between the window and the boundary. For the Ising model on a regular tree, this condition is sharp.

1 Introduction

Context

In recent years, Glauber dynamics on the lattice $\mathbf{Z}^d$ was extensively studied. A good account can be found in [25]. In this work, we study this dynamics on other graphs.

The main goal of our work is to determine which geometric properties of the underlying graph are most relevant to the mixing rate of the Glauber dynamics on particle systems.

To define a general particle system [21] on an undirected graph $G = (V, E)$, define a configuration as an element σ of A^V where A is some finite set, and to each edge $(v, w) \in E$, associate a weight function $\alpha_{vw} : A \times A \to \mathbb{R}_+$. The Gibbs distribution assigns every configuration σ probability proportional to $\prod_{\{v,w\}\in E} \alpha_{vw}(\sigma_v, \sigma_w)$. The Ising model (for which $\alpha_{vw}(\sigma_v, \sigma_w) = e^{\beta\sigma_v\sigma_w}$) and the Potts model are examples of such systems; so is the coloring model (for which $\alpha_{vw} = \mathbf{1}_{\sigma_v \neq \sigma_w}$)

On a finite graph, the Heat-Bath Glauber dynamics is a continuous time Markov chain with the generator

$$(\mathcal{L}(f))(\sigma) = \sum_{v\in V} \left(\sum_{a\in A} K[\sigma \to \sigma_v^a] \left(f(\sigma_v^a) - f(\sigma) \right) \right), \tag{1}$$

where σ_v^a is the configuration s.t.

$$\sigma_v^a(w) = \begin{cases} a & \text{if } w = v \\ \sigma(w) & \text{if } w \neq v \end{cases}$$

and

$$K[\sigma \to \sigma_v^a] = \frac{\prod_{w:(w,v)\in E} \alpha_{vw}(a, \sigma_w)}{\sum_{a'\in A} \left(\prod_{w:(w,v)\in E} \alpha_{vw}(a', \sigma_w) \right)}.$$

It is easy to check that this dynamics is reversible with respect to the Gibbs measure. An equivalent representation for the Glauber dynamics, known as the *Graphical representation*, is the following: Each vertex has a rate 1 Poisson clock attached to it. These Poisson clocks are independent of each other. Assume that the clock at v rang at time t and that just before time t the configuration was σ. Then at time t we replace $\sigma(v)$ by a random spin $\sigma'(v)$ chosen according to the Gibbs distribution conditional on the rest of the configuration:

$$\frac{\mathbf{P}[\sigma'(v) = i \mid \sigma]}{\mathbf{P}[\sigma'(v) = j \mid \sigma]} = \prod_{w:\{v,w\}\in E} \frac{\alpha_{vw}(i, \sigma(w))}{\alpha_{vw}(j, \sigma(w))}.$$

We are interested in the rate of convergence of the Glauber dynamics to the stationary distribution. Note that this process mixes $n = |V|$ times faster than the corresponding discrete time process, simply because it performs (on average) n operations per time unit while the discrete time process performs one operation per time unit.

In section 2.1, we describe a connection between the geometry of a graph and the mixing time of Glauber dynamics on it. In particular, we show that for balls in hyperbolic tilings, the Glauber dynamics for the Ising model, the Potts model and proper coloring with $\Delta + 2$ colors (where Δ is the maximal degree), have mixing time polynomial in the volume. An example of such a hyperbolic graph can be obtained from the binary tree by adding horizontal edges across levels; another example is given in Figure 1.

In sections 2.3-4 we study Glauber dynamics for the Ising model on regular trees. For these trees we show that the mixing time is polynomial at all temperatures, and we characterize the range of temperatures for which the spectral gap is bounded away from zero. Thus, the notion that the two sides of the phase transition (high versus low temperatures) should correspond to polynomial versus super-polynomial mixing times for the associated dynamics, fails for the Ising model on trees: here the two sides of the high/intermediate versus low temperature phase transition just correspond

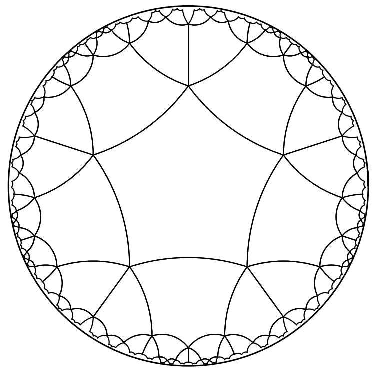

Figure 1: A ball in hyperbolic tiling

to uniformly bounded versus *unbounded* inverse spectral gap. We also exhibit another surprising phenomenon: On infinite regular trees, there is a range of temperatures in which the inverse spectral gap is bounded, even though there are many different Gibbs measures.

In section 5 of the paper we go beyond trees and hyperbolic graphs and study Glauber dynamics for families of finite graphs of bounded degree. We show that if the inverse spectral gap of the Glauber dynamics on the ball centered at ρ stays bounded as the ball grows, then the correlation between the state of a vertex ρ and the states of vertices at distance r from ρ, must decay exponentially in r.

Setup

The graphs. Let $G = (V, E)$ be an infinite graph with maximal degree Δ. Let ρ be a distinguished vertex and denote by $G_r = (V_r, E_r)$ the induced graph on $V_r = \{v \in V : \mathrm{dist}(\rho, v) \le r\}$. Let n_r be the number of vertices in G_r. At some parts of the paper we will focus on the case where $G = T = (V, E)$ is the infinite b-ary tree. In these cases, $T_r^{(b)} = (V_r, E_r)$ will denote the r-level b-ary tree.

The Ising model. In the Ising model on G_r at inverse temperature β, every configuration $\sigma \in \{-1, 1\}^{V_r}$ is assigned probability

$$\mu[\sigma] = Z(\beta)^{-1} \exp\Big(\beta \sum_{\{v,w\} \in E_r} \sigma(v)\sigma(w)\Big)$$

where $Z(\beta)$ is a normalizing constant. When $G_r = T_r^{(b)}$, this measure has the following equivalent definition [8]: Fix $\epsilon = (1 + e^{2\beta})^{-1}$. Pick a random spin ± 1 uniformly for the root of the tree. Scan the tree top-down, assigning vertex v a spin equal to the spin of its parent with probability $1 - \epsilon$ and opposite with probability ϵ.

The Heat-Bath Glauber dynamics for the Ising model chooses the new spin $\sigma'(v)$ in such a way that:

$$\frac{\mathbf{P}[\sigma'(v) = +1 \mid \sigma]}{\mathbf{P}[\sigma'(v) = -1 \mid \sigma]} = \exp\Big(2\beta \sum_{w:\,\{w,v\}\in E_r} \sigma(w)\Big).$$

See [21] or [25] for more background.

Mixing times.

Definition 1.1. *For a reversible continuous time Markov chain, let* $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_k$ *be the eigenvalues of* $-\mathcal{L}$ *where* $\mathcal{L}$ *is the generator. The* **spectral gap** *of the chain is defined as* λ_2*, and the* **relaxation time**, τ_2*, is defined as the inverse of the spectral gap.*

Note that the corresponding discrete time Glauber dynamics has transition matrix $M = \mathbf{I} + \frac{1}{n}\mathcal{L}$, where n is the number of vertices. Moreover, the eigenvalues of M are $1, 1 - \frac{\lambda_2}{n}, 1 - \frac{\lambda_3}{n}, \ldots$ and therefore the spectral gap of the discrete dynamics is the spectral gap of the continuous dynamics divided by n.

Definition 1.2. *For measures* μ *and* ν *on the same discrete space, the* **total-variation distance**, $d_V(\mu,\nu)$*, between* μ *and* ν *is defined as*

$$d_V(\mu,\nu) = \frac{1}{2}\sum_x |\mu(x) - \nu(x)|.$$

Definition 1.3. *Consider an ergodic Markov chain* $\{X_t\}$ *with stationary distribution* π *on a finite state space. Denote by* $\mathbf{P}_x^t$ *the law of* X_t *given* $X_0 = x$*. The* **mixing time** *of the chain,* τ_1*, is defined as*

$$\tau_1 = \inf\{t : \sup_{x,y} d_V(\mathbf{P}_x^t, \mathbf{P}_y^t) \leq e^{-1}\}.$$

For $t \geq \ln(1/\epsilon)\tau_1$, we have

$$\sup_x d_V(\mathbf{P}_x^t, \pi) \leq \sup_{x,y} d_V(\mathbf{P}_x^t, \mathbf{P}_y^t) \leq \epsilon.$$

Using τ_2 one can bound the mixing time τ_1, since every reversible Markov chain with stationary distribution π satisfies (see, e.g., [1]),

$$\tau_2 \leq \tau_1 \leq \tau_2 \left(1 + \log\left((\min_\sigma \pi(\sigma))^{-1}\right)\right). \tag{2}$$

For the Markov chains studied in this paper, this gives $\tau_2 \leq \tau_1 \leq O(n)\tau_2$.

Cut-Width and relaxation time.

Definition 1.4. *The* **cut-width** $\xi(G)$ *of a graph* G *is the smallest integer such that there exists a labeling* $v_1, \dots, v_n$ *of the vertices such that for all* $1 \leq k \leq n$*, the number of edges from* $\{v_1, \dots, v_k\}$ *to* $\{v_{k+1}, \dots, v_n\}$*, is at most* $\xi(G)$*.*

Remark: The *vertex-separation* of a graph G is defined analogously to the cut-width in terms of vertices among $\{v_1, \dots, v_k\}$ that are adjacent to $\{v_{k+1}, \dots, v_n\}$. In [20] it is shown that the vertex-separation of G equals its *path-width*, see [35]. In [19] the cut-width was called the *exposure*.

Generalizing an argument in [25, Theorem 6.4] for $\mathbf{Z}^d$, (see also [15]), we prove:

Proposition 1.1. *Let* G *be a finite graph with* n *vertices and maximal degree* Δ*.*

1. *Consider the Ising model on* G*. The relaxation time of the Glauber dynamics is at most* $ne^{(4\xi(G)+2\Delta)\beta}$*.*

2. *Consider the coloring model on* G*. If the number of colors* q *satisfies* $q \geq \Delta + 2$*, then the relaxation time of the Glauber dynamics is at most* $(\Delta + 1)n(q-1)^{\xi(G)+1}$*.*

Analogous results hold for the independent set and hard core models.

Cut-Width and long-range correlations for hyperbolic graphs. The usefulness of Proposition 1.1 comes about when we bound the relaxation time of certain graphs by estimating their cut-width. The following proposition bounds the cut-width of balls in hyperbolic tilings of the plane. Recall that the **Cheeger constant** of an infinite graph G is

$$c(G) = \inf\left\{ \frac{|\partial A|}{|A|} \,\middle|\, A \subseteq G; 0 < |A| < \infty \right\}, \tag{3}$$

where ∂A is the set of vertices of A which have neighbors in $G \setminus A$.

Proposition 1.2. *For every $c > 0$ and $\Delta < \infty$, there exists a constant $C = C(c, \Delta)$ such that if G is an infinite planar graph with*

- *Cheeger constant at least c,*
- *maximum degree bounded by Δ and*
- *for every r no cycle from G_r separates two vertices of $G \setminus G_r$,*

then $\xi(G_r) \leq C \log n_r$ for all r, where n_r is the number of vertices of G_r.

Combining this with Proposition 1.1 we get that the Glauber dynamics for the Ising models on balls in the hyperbolic tiling has relaxation time polynomial in the volume for every temperature. On the other hand, we have the following proposition:

Proposition 1.3. *Let G be a planar graph with bounded degrees, bounded co-degrees and a positive Cheeger constant. Then there exist $\beta' < \infty$ and $\delta > 0$ such that for all r, all $\beta > \beta'$, and all vertices u, v in G_r, the Ising model on G_r satisfies that $\mathbf{E}[\sigma_u \sigma_v] \geq \delta$. In other words, at low enough temperature there are long-range correlations.*

This shows that for the Ising model on balls of hyperbolic tilings at very low temperature, there are long-range correlations coexisting with polynomial time mixing. While there is no characterization of all planar graphs with positive Cheeger constants, an important family of such graphs is Cayley graphs of nonelementary fuchsian groups which are nonamenable, and hence have a positive Cheeger constant, see [31, 24, 18] for background and [11] for some explicit estimates.

Relaxation time for the Ising model on the tree. The Ising model on the b-ary tree has three different regimes, see [3, 8]. In the high temperature regime, where $1 - 2\epsilon < 1/b$, there is a unique Gibbs measure on the infinite tree, and the expected value of the spin at the root σ_ρ given any boundary conditions $\sigma_{\partial T_r^{(b)}}$ decays exponentially in r. In the intermediate regime, where $1/b < 1 - 2\epsilon < 1/\sqrt{b}$, the exponential decay described above still holds for typical boundary conditions, but not for certain exceptional boundary conditions, such as the all + boundary; consequently, there are infinitely many Gibbs measures on the infinite tree. In the low temperature regime, where $1 - 2\epsilon > 1/\sqrt{b}$, typical boundary conditions impose bias on the expected value of the spin at the root σ_ρ.

Theorem 1.4. *Consider the Ising model on the b-ary tree $T_r = T_r^{(b)}$ with r levels. Let $\epsilon = (1 + e^{2\beta})^{-1}$. The relaxation time τ_2 for Glauber dynamics on $T_r^{(b)}$ can be bounded as follows:*

1. *The relaxation time is polynomial at all temperatures: $\tau_2 = n_r^{O(\log(1/\epsilon))}$. Furthermore, the limit*

$$\lim_{r\to\infty} \frac{\log(\tau_2(T_r^{(b)}, \beta))}{\log(n_r)}$$

exists.

2. *Low temperature regime:*
   (a) *If $1 - 2\epsilon \geq 1/\sqrt{b}$ then $\sup_r \tau_2(T_r) = \infty$. In fact, $\tau_2(T_r) = \Omega(n_r^{\log_b(b(1-2\epsilon)^2)})$ when $1 - 2\epsilon > 1/\sqrt{b}$ and $\tau_2(T_r) = \Omega(\log n_r)$ when $1 - 2\epsilon = 1/\sqrt{b}$.*
   (b) *Moreover, the degree of τ_2 tends to infinity as ϵ tends to zero: $\tau_2(T_t) = n_r^{\Omega(\log(1/\epsilon))}$.*

3. *Intermediate and high temperature regimes:*
   If $1 - 2\epsilon < 1/\sqrt{b}$ then the relaxation time is uniformly bounded: $\tau_2 = O(1)$. Furthermore, this result holds for every external field $\{H(v)\}_{v \in T_r}$.

In particular we obtain from Equation (2) that in the low temperature region $\tau_1 = n_r^{\Theta(\beta)}$, and in the intermediate and high temperature regions $\tau_1 = O(n_r)$.

A recent work by Peres and Winkler [33] compares the mixing times of single site and block dynamics for the heat-bath Glauber dynamics for the Ising model.

They show that if the blocks are of bounded volume, then the same mixing time up to constants is obtained for the single site and block dynamics.

Temp.	$1-2\epsilon$	$\sigma_\rho \vert \sigma_{\partial T} \equiv +$	$I(\sigma_\rho, \sigma_{\partial T})$	τ_2
high	$< 1/2$	unbiased	$\to 0$	$O(1)$
med.	$\in (\frac{1}{2}, \frac{1}{\sqrt{2}})$	biased	$\to 0$	$O(1)$
low	$> \frac{1}{\sqrt{2}}$	biased	$\inf > 0$	$n^{\Omega(1)}$
freeze	$1 - o(1)$	biased	$1 - o(1)$	$n^{\Theta(\beta)}$

Table 1: The Ising model on binary trees. Here the root is denoted ρ, and the vertices at distance r from the root are denoted ∂T.

Combining these results with the path coupling argument of Section 4, it follows that $\tau_1 = O(\log n_r)$ in the intermediate and high temperature regions.

We emphasize that Theorem 1.4 implies that in the intermediate region $1/2 < 1 - 2\epsilon < 1/\sqrt{b}$, the relaxation time is bounded by a constant, yet, in the infinite volume there are infinitely many Gibbs measures. This Theorem is perhaps easiest to appreciate when compared to other results on the Gibbs distribution for the Ising model on binary trees, summarized in Table 1.

The proof of the low temperature result is quite general and applies to other models with "soft" constraints, such as Potts models on the tree .

Spectral gap and correlations. At infinite temperature, where distinct vertices are independent, the Glauber dynamics on a graph of n vertices reduces to an (accelerated by a factor of n) random walk on a discrete n-dimensional cube, where it is well known that the relaxation time is $\Theta(1)$. Our next result shows that at any temperature where such fast relaxation takes place, a strong form of independence holds. This is well known in $\mathbf{Z}^d$, see [25], but our formulation is valid for any graph of bounded degree.

Theorem 1.5. *Denote by σ_r the configuration on all vertices at distance r from ρ. If G has bounded degree and the relaxation time of the Glauber dynamics satisfies $\tau_2(G_r) = O(1)$, then the Gibbs distribution on G_r has the following property. For any fixed finite set of vertices A, there exists $c_A > 0$ such that for r large enough*

$$\mathrm{Cov}[f, g] \le e^{-c_A r} \sqrt{\mathrm{Var}(f)\mathrm{Var}(g)}\,, \tag{4}$$

provided that $f(\sigma)$ depends only on σ_A and $g(\sigma)$ depends only on σ_r. Equivalently, there exists $c'_A > 0$ such that

$$I[\sigma_A, \sigma_r] \le e^{-c'_A r}\,, \tag{5}$$

where I denotes mutual information (see [6].)

This theorem holds in a very general setting which includes Potts models, random colorings, and other local-interaction models.

Our proof of Theorem 1.5 uses "disagreement percolation" and a coupling argument exploited by van den Berg, see [2], to establish uniqueness of Gibbs measures in $\mathbf{Z}^d$; according to F. Martinelli (personal communication) this kind of argument is originally due to B. Zegarlinski. Note however, that Theorem 1.5 holds also when there are multiple Gibbs measures – as the case of the Ising model in the intermediate regime demonstrates. Moreover, combining Theorem 1.5 and Theorem 1.4, one infers that for $1 - 2\epsilon < 1/\sqrt{b}$, we have $\lim_{r\to\infty} I[\sigma_0, \sigma_r] = 0$. This yields another proof of this fact which was proven before in [3, 12, 8].

Plan of the paper

In section 2 we prove Proposition 1.1 via a canonical path argument, and give the resulting polynomial time upper bound of Theorem 1.4 part 1. We also present a more elementary proof of the upper bound on the relaxation time for the tree, which gives sharper exponents and the existence of a limiting exponent; this proof uses Martinelli's block dynamics to show sub-additivity. In section 3 we sketch a proof of Theorem 1.4 part 2a and present a proof of Theorem 1.4 part 2b. These lower bounds are obtained by finding a low conductance "cut" of the configuration space, using global majority of the boundary spins for the former result, and recursive majority for the latter result. In section 4 we establish the high temperature result, using comparison to block dynamics which are analyzed via path-coupling. Finally, in section 5 we prove Theorem 1.5 by a Peierls argument controlling "paths of disagreement" between two coupled dynamics.

Remark: Most of the results proved here were presented (along with proof sketches) in the extended abstract [19]. However, the proofs of our results for hyperbolic graphs (see Section 2.2), which involve some interesting geometry, were not even sketched there. Also, the general polynomial upper bound for trees that we establish in Section 2.3 is a substantial improvement on the results of [19], since it only assumes the dynamics is ergodic and allows for arbitrary hard-core constraints.

2 Polynomial Upper Bounds

2.1 Cut-Width and mixing time

We begin by showing how part 1 of Proposition 1.1 implies the upper bound in part 1 of Theorem 1.4.

Lemma 2.1. *Let $T_r^{(b)}$ be the b-ary tree with r levels. Then,* $\xi(T_r^{(b)}) < (b-1)r+1$.

Proof. Order the vertices using the *Depth first search left to right order* , i.e., use the following labeling for the vertices: The root is labeled $\langle 0,0,\ldots,0\rangle$. The children of the root are labeled $\langle 1,0,\ldots,0\rangle$ through $\langle b,0,\ldots,0\rangle$, and so on, so that the children of $\langle a_1,a_2,\ldots,a_k,0,\ldots,0\rangle$ are $\langle a_1,a_2,\ldots,a_k,1,\ldots,0\rangle$ through $\langle a_1,a_2,\ldots,a_k,b,\ldots,0\rangle$. Then order the vertices lexicographically. Note that in the lexicographic ordering, a vertex always appears before its children. When we enumerated all vertices up to $\langle a_1,a_2,\ldots,a_r\rangle$, the only vertices that were enumerated but whose children were not enumerated are among the set of at most r vertices

$$\{\langle 0,0,\ldots,0\rangle,\langle a_1,0,\ldots,0\rangle,\langle a_1,a_2,\ldots,0\rangle,\ldots,\langle a_1,a_2,\ldots,a_r\rangle\}.$$

Each of these vertices has at most b children, and for all but $\langle a_1,a_2,\ldots,a_r\rangle$ at least one child has already been enumerated. Therefore,

$$\xi(T_r^{(b)}) < (b-1)r+1.$$

□

Corollary 2.2. *1. The relaxation time of the Glauber dynamics for the Ising model on $T_r^{(b)}$ is at most*

$$C(\epsilon)n_r^{1+4(b-1)\log_b\frac{1-\epsilon}{\epsilon}} = n_r^{O(\log(1/\epsilon))}.$$

2. The relaxation time of the Glauber dynamics for the coloring on $T_r^{(b)}$ with $q > b+2$ colors is at most

$$(b+1)n_r^{1+2(b-1)\log_b(q)}$$

Proof. The Corollary follows from Lemma 2.1 and Proposition 1.1. □

The upper bound in part 1 of Theorem 1.4 follows immediately.

Proof of part 1 of Proposition 1.1. The proof follows the lines of the proof given in [25, Theorem 6.4] for the Ising model in $\mathbf{Z}^d$, (see also [15]).

Let Γ be the graph corresponding to the transitions of the Glauber dynamics on the graph G. Between any two configurations σ and η, we define a "canonical path" $\gamma(\sigma,\eta)$ as follows. Fix an order $<$ on the vertices of G which achieves the cut-width. Consider the vertices $v_1 < v_2 < \ldots$ at which $\sigma_v \neq \eta_v$.

We define the k-th configuration $\sigma^{(k)}$ on the path $\gamma(\sigma,\eta)$ by giving spin σ_v to every labeled vertex $v \leq v_k$, spin η_v to every labeled vertex $v > v_k$, and spin $\sigma_v = \eta_v$ for every unlabeled vertex v. Note that $\sigma^{(0)} = \eta$ and $\sigma^{(d(\sigma,\eta))} = \sigma$. Since $\sigma^{(k-1)}$ and $\sigma^{(k)}$ are identical except for the spin of vertex v_k, they are adjacent in Γ. This defines $\gamma(\sigma,\eta)$ (see Figure 2). Note that for every k, there are at most $\xi(G)$ pairs of adjacent vertices (v_i, v_j) such that $i \leq k < j$, hence any configuration on the canonical path between σ and η will have at most $\xi(G)$ edges between spins copied from σ and spins copied from η.

Using canonical paths to bound the mixing rate. For each directed edge $\mathbf{e} = (\omega,\zeta)$ on the configuration graph Γ, we say that $\mathbf{e} \in \gamma(\sigma,\eta)$ if ω and ζ are adjacent configurations in $\gamma(\sigma,\eta)$. Let

$$\rho = \sup_{\mathbf{e}} \sum_{\sigma,\eta:\, \mathbf{e}\in\gamma(\sigma,\eta)} \frac{\mu[\sigma]\mu[\eta]}{Q(\mathbf{e})},$$

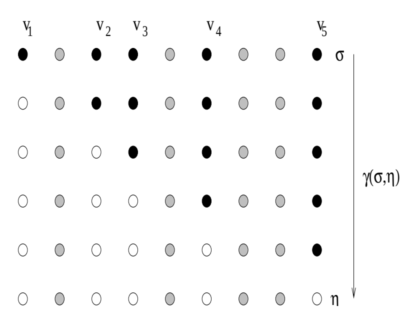

Figure 2: The canonical path from σ to η. The vertices on which σ and η agree are marked in grey; the other vertices are colored black if their spin is chosen according to σ and white if their spin is chosen according to η.

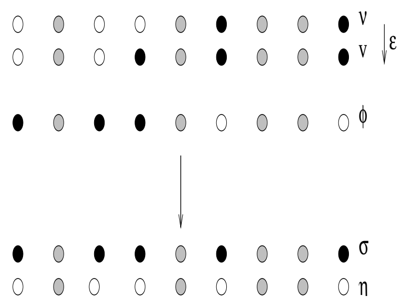

Figure 3: The injection from $(\mathbf{e},\varphi)$ to (σ,η). The vertices on which both endpoints of $\mathbf{e}$ and φ agree are marked in grey; the other vertices are colored black if they precede v_{k_0} and their spin is chosen according to φ, or if they are preceded by v_{k_0} and their spin is chosen according to the endpoints of $\mathbf{e}$; and are colored white otherwise.

where μ is the stationary measure (i.e. the Gibbs distribution), and for any two adjacent configurations ω and ζ, $Q(\mathbf{e}) = Q((\omega,\zeta)) = \mu[\omega]K[\omega \to \zeta]$. If L is the maximal length of a canonical path, then by the argument in [15, 25], the relaxation time of the Markov chain is at most

$$\tau_2 \le L\rho. \tag{6}$$

Since $L \le n$, it follows that $\tau_2 \le n\rho$, thus it only remains to prove an upper bound on ρ.

Analysis of the canonical path. For each directed edge $\mathbf{e}$ in Γ, we define an injection from canonical paths going through $\mathbf{e}$ in the specified direction, to configurations on G. To a canonical path $\gamma(\sigma,\eta)$ going through $\mathbf{e}$, such that $\mathbf{e} = (\sigma^{(k-1)},\sigma^{(k)})$, we associate the configuration φ which has spin η_{v_i} for every v_i s.t. $i \ge k$ and spin σ_{v_i} for every v_i s.t. $i < k$. To verify that this is an injection, note that one can reconstruct σ and η by first identifying the unique k_0 s.t. ω and ζ differ on v_{k_0} and then taking (as in Figure 3)

$$\sigma_{v_k} = \begin{cases} \omega_{v_k} & \omega_{v_k} = \varphi_{v_k} \\ \omega_{v_k} & k \ge k_0 \text{ and } \omega_{v_k} \ne \varphi_{v_k} \\ \varphi_{v_k} & k < k_0 \text{ and } \omega_{v_k} \ne \varphi_{v_k} \end{cases}$$

and

$$\eta_{v_k} = \begin{cases} \omega_{v_k} & \omega_{v_k} = \varphi_{v_k} \\ \varphi_{v_k} & k \ge k_0 \text{ and } \omega_{v_k} \ne \varphi_{v_k} \\ \omega_{v_k} & k < k_0 \text{ and } \omega_{v_k} \ne \varphi_{v_k}. \end{cases}$$

By the property of our labeling,

$$\mu[\sigma]\mu[\eta] \leq \mu[\sigma^{(k-1)}]\mu[\varphi]e^{4\xi(G)\beta}. \tag{7}$$

and $K[\sigma^{(k-1)} \to \sigma^{(k)}] \geq \exp(-2\Delta\beta)$. Now a short calculation concludes the proof:

$$\begin{aligned}
\rho &\leq \sup_e \sum_{\sigma,\eta \text{ s.t. } e\in\gamma(\sigma,\eta)} \frac{\mu[\sigma]\mu[\eta]}{\mu[\sigma^{(k-1)}]K[\sigma^{(k-1)} \to \sigma^{(k)}]} \\
&\leq e^{4\xi(G)\beta} \sup_e \sum_\varphi \frac{\mu[\sigma^{(k-1)}]\mu[\varphi]}{\mu[\sigma^{(k-1)}]K[\sigma^{(k-1)} \to \sigma^{(k)}]} \qquad (8)\\
&\leq e^{4\xi(G)\beta}e^{2\Delta\beta} \sum_\varphi \mu[\varphi] \leq e^{(4\xi(G)+2\Delta)\beta}. \qquad (9)
\end{aligned}$$

The last inequality follows from the fact that the map $\gamma \to \varphi$ is injective and therefore $\sum_\varphi \mu[\varphi] \leq 1$. □

Proof of part 2 of Proposition 1.1. The previous argument does not directly extend to coloring, because the configurations $\sigma^{(k)}$ along the path (as defined above) may not be proper colorings. Assume that $q \geq \Delta + 2$ and let $v_1 < v_2 \cdots < v_n$ be an ordering of the vertices of G which achieves the cut-width. We construct a path $\gamma(\sigma, \eta)$ such that

$$|\gamma(\sigma,\eta)| \leq (\Delta+1)n. \tag{10}$$

Moreover, for all $\tau \in \gamma(\sigma, \eta)$ there exists a k such that

$$\tau_v = \begin{cases} \eta_v & \text{if } v \leq v_k \\ \sigma_v & \text{if } v > v_k \text{ and } v \not\sim \{v,\ldots,v_k\} \end{cases} \tag{11}$$

The way to construct a path $\gamma(\sigma, \eta)$ satisfying (10) and (11) is the following: $\sigma^0 = \sigma$. Given σ^k, we proceed to create σ^{k+1} as follows: Let $i(k) = \inf\{j : \sigma^k_{v_j} \neq \eta_{v_j}\}$. If

$$\rho = \begin{cases} \sigma^k_v & \text{if } v \neq v_{i(k)} \\ \eta_v & \text{if } v = v_{i(k)} \end{cases}$$

is a legal configuration, then $\sigma^{k+1} = \rho$. otherwise, let

$$h(k) = \inf\{j : \sigma^k_{v_j} = \eta_{v_{i(k)}} \text{ and } v_j \sim v_{i(k)}\},$$

and let c be a color that is different from $\eta_{v_{i(k)}}$ and is legal for $v_{h(k)}$ under σ^k. Such a color exists because $q \geq \Delta + 2$. Then, we take

$$\sigma_v^{k+1} = \begin{cases} \sigma_v^k & \text{if } v \neq v_{h(k)} \\ c & \text{if } v = v_{h(k)} \end{cases}$$

It is easy to verify that the path satisfies (10) and (11). Since all legal configurations have the same weight, (7) is replaced by

$$\mu[\sigma]\mu[\eta] = \mu[\sigma^{(k-1)}]\mu[\varphi] \tag{12}$$

On the other hand, the map $\gamma \to \varphi$ is not injective. Instead, by (11), there are at most $(q-1)^{\xi(G)}$ paths which are mapped to the same coloring. We therefore obtain that for the coloring model $\rho \leq n(q-1)^{\xi(G)+1}$ and therefore from (10) and (12),

$$\tau_2 \leq (\Delta + 1)nq(q-1)^{\xi(G)}.$$

□

2.2 Hyperbolic graphs

In this subsection we show that balls in a hyperbolic tiling have logarithmic cut-width. Let $G = (V, E)$ be an infinite planar graph and let $o \in V$. Let G_r be the ball of radius r in G around o, with the induced edges. The following proposition implies Propositions 1.2 and 1.3.

Proposition 2.3. *1. Suppose G has*

- *a positive Cheeger constant,*
- *degrees bounded by Δ,*
- *for all r, no cycle from G_r separates two vertices of $G \setminus G_r$,*

then there exists constants α_1 and α_2 s.t. $\xi(G_r) \leq \alpha_1 \Delta \log(|G_r|) + \alpha_2 \Delta$.

2. Assume that G has bounded degrees, bounded co-degrees, no cycle from G_r separates two vertices of $G \setminus G_r$, and the following weak isoperimetric condition holds:

$$|\partial A| \geq C \log(|A|) \tag{13}$$

for every finite $A \subseteq G$ and for some constant C.

Then there exist $\beta' < \infty$ and $\delta > 0$ s.t. for every $\beta > \beta'$, for every r and for every u, v in G_r, the free Gibbs measure for the Ising model on G_r with inverse temperature β satisfies $\text{cov}(\sigma_u, \sigma_v) \geq \delta$.

Proof of part 1 of Proposition 2.3. Consider a planar embedding of G. Since no cycle from G_r separates two vertices of $G \setminus G_r$, all vertices of $G \setminus G_r$ are in the same face of G_r, and without loss of generality we can assume that it is the infinite face of our chosen embedding of G_r.

Let T be a shortest path tree from o in G_r. In other words, T is a tree such that for every vertex v, the path from o to v in T is a shortest path in G_r. Let $e_1 \in T$ be an edge adjacent to o. We perform a depth-first-search traversal of T, starting from $o = v_0$, traversing e_1 to its end vertex v_1, and continuing in counterclockwise order around T. This defines a linear ordering $v_0 \leq v_1 \leq \cdots \leq v_{n-1}$ of the vertices of G_r.

Consider the induced ordering $w_1 \leq w_2 \leq \cdots \leq w_k$ on the vertices of G_r which are at distance exactly r from o.

Fix $i < j$. We first consider edges between

$$V_{ij} = \{u \in G_r : w_i < u < w_j, u \text{ not ancestor of } w_j \text{ in } T\}$$

and $G_r \setminus V_{ij}$. Note that V_{ij} does not contain any vertex on the paths in T from o to either w_i or w_j. Obviously, there can be edges from V_{ij} to vertices on the paths from o to w_i or w_j. Let $e = \{u, v\}$ be an edge with one endpoint in V_{ij} and the other end in $G \setminus G_r$ but not on $\text{Path}(w_i)$ or $\text{Path}(w_j)$, where $\text{Path}(w_j)$ denote the path in T from o to w_j. Without loss of generality, assume that $w_i < u < w_j < v$. The case where $v < w_i < u < w_j$ is treated similarly.

The path from o to u in T, followed by the edge e, followed by the path from v to o in T, defines a cycle C_e in G_r (see Figure 4). Since $w_i < u < w_j < v$, C_e must enclose exactly one of w_j and w_i. Since the graph is embedded in the plane, the ones among those cycles which enclose w_j must form a nested sequence, and therefore there is an outermost such cycle C_{e^*} with an associated "outermost" edge $e^* = \{u^*, v^*\}$. Similarly, among the edges such that the corresponding cycle encloses w_i, there is an "outermost" edge $f^* = \{x^*, y^*\}$.

There can only be edges from V_{ij} to the vertices enclosed by C_{e^*} or by C_{f^*} (note that this includes the paths from o to w_i and to w_j) . Since all the vertices of $G \setminus G_r$ are in the infinite face of G_r, hence outside C_{e^*}, the set of vertices enclosed by C_{e^*} is the same in G as in G_r. Let A denote the set

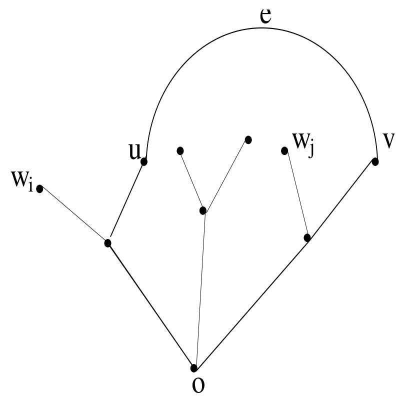

Figure 4: The cycle C_e defined by e.

of vertices enclosed by C_{e^*} (including C_{e^*}). We have: $|\partial A| \leq 2r+1$, hence $|A| \leq (2r+1)/c$, where c is the Cheeger constant of G. Reasoning similarly for C_{f^*}, we obtain that the set of vertices in $G_r \setminus V_{ij}$ adjacent to V_{ij} has size at most $(4r+2)/c$.

Let $B_j = V_{j-1,j}$ for $j \geq 2$, and $B_1 = \{u \in G_r : u < w_1 \text{ or } u > w_k, u \text{ not an ancestor of } w_1 \text{ in } T\}$. Let us bound the cardinality of B_j. As above, we define C_{e^*} and C_{f^*}. Let A denote the union of B_j, of the vertices enclosed by C_{e^*}, and of the vertices enclosed by C_{f^*}.

Since the vertices of B_j are at distance at most $r-1$ from o, they have no neighbors in $G \setminus G_r$. Thus the neighborhood of A in G is such that $\partial A \subset C_{e^*} \cup C_{f^*}$, hence $|\partial A| \leq 4r+2$, and so $|B_j| \leq |A| \leq (4r+2)/c$.

Finally, to compute the cut-width, let $S = \{u : v_0 \leq u \leq v_i\}$, and let j be such that $w_{j-1} < v_i \leq w_j$. We have:

$$V_{1,j-1} \subseteq S \subseteq B_1 \cup V_{1,j-1} \cup B_j \cup \text{Path}(w_1) \cup \text{Path}(w_{j-1}) \cup \text{Path}(w_j).$$

Thus the set of edges between S and $G_r \setminus S$ has size at most $\Delta(4r+2)/c + (|B_1 \cup B_j|)\Delta + (3r+1)\Delta$, which is at most $(3r+1+(12r+6)/c)\Delta$.

Since G has positive Cheeger constant,

$$|G_r| = |G_{r-1} \cup \partial G_{r-1}| \geq |G_{r-1}|(1+c),$$

and so $|G_r| \geq (1+c)^r$, that is, $r \leq \log |G_r| / \log(c+1)$. Hence the set of edges between S and $G \setminus S$ has size at most $(3+12/c)(\Delta/\log(c+1)) \log |G_r| + (1+6/c)\Delta$. This concludes the proof. □

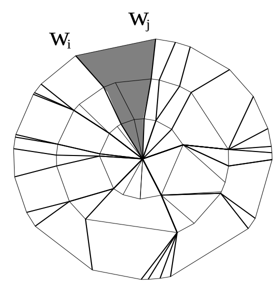

Figure 5: The region between Path(W_i) and Path(W_j)

Proof of part 2 of Proposition 2.3. We use the Random Cluster representation of the Ising model (see, e.g. [9]) and a standard Peierls path-counting argument. For every u and v in G_r, $\text{cov}(\sigma_u, \sigma_v)$ is the probability that u is connected to v in the Random Cluster model. Fix $p < 1$. The exact value of p will be specified later.

Then, β is large enough, i.e., if $(1 - e^{-\beta})/e^{-\beta} > 2p/(1-p)$, then the Random Cluster model dominates percolation with parameter p. So, what we need to show is that for a graph satisfying the requirements of part 2 of the proposition and p high enough, there exists $\delta > 0$ s.t. for every r and every u, v in G_r, we have $\mathbf{P}_p(u \leftrightarrow v) \geq \delta$. By the FKG inequality (see [10]),

$$\mathbf{P}_p(u \leftrightarrow v) \geq \mathbf{P}_p(u \leftrightarrow o)\mathbf{P}_p(v \leftrightarrow o)$$

where o is the center. Therefore we need to show that $\mathbf{P}(v \leftrightarrow o)$ is bounded away from zero. To this end, we will pursue a standard path counting technique: in order for o and v not to be connected, there needs to be a closed path in the dual graph that separates o and v.

Claim 2.4. *There exists $M = M(G)$ s.t. for every r and $v \in G_r$ there are at most M^k paths of length k in the dual graph of G_r that separate o from v.*

By Claim 2.4, if we take $p > 1 - 1/(2M)$ and choose β accordingly then the probability that there exists a closed path in the dual graph that separates o and v is bounded away from 1. □

Proof of Claim 2.4. Here again, we consider an embedding of G_r such that all the vertices of $G \setminus G_r$ lie on the infinite face F of G_r.

Let γ be a shortest path connecting v to o in G_r. Every dual path separating v from o must intersect γ. For an edge e let $\Lambda_k(e)$ be the set of dual paths ψ of length k separating o from v such that ψ intersects e. If $\hat{\Delta}$ is the maximal co-degree in G then $|\Lambda_k(e)| \leq \hat{\Delta}^k$ for every e.

Let $e \in \gamma$ be such that $d(e,o) > \exp(k/C)+k\hat{\Delta}$, $d(e,v) > \exp(k/C)+k\hat{\Delta}$, and $d(e,F) > k\hat{\Delta}$. We will now show that $|\Lambda_k(e)| = 0$. Assume, for a contradiction, that $\psi \in \Lambda_k(e)$. Since ψ has length k and $d(e,F) > \hat{\Delta}k$, ψ does not touch the outer face F, and so the area enclosed by ψ in G_r equals the area enclosed by ψ in G. The dual path ψ encloses either v or o. Without loss of generality, assume that it encloses o. Let e' be the edge of γ closest to o which ψ intersects. Since ψ has length k, we get that $d(e',o) > \exp(k/C)$, and so at least $1+\exp(k/C)$ vertices of γ are enclosed by ψ. By (13) this implies that ψ has length strictly greater than k, a contradiction.

Thus, the total number of paths of length k separating o from v is at most

$$\sum_{e:|\Lambda_k(e)|\neq 0} |\Lambda_k(e)| \leq [2\exp(k/C) + k\hat{\Delta}) + \hat{\Delta}k]\hat{\Delta}^k.$$

□

Remark: An isoperimetric inequality of the type of (13) is necessary. An example where all other conditions of part 2 of Proposition 2.3 are satisfied and yet the conclusion does not hold can be found in Figure 6.

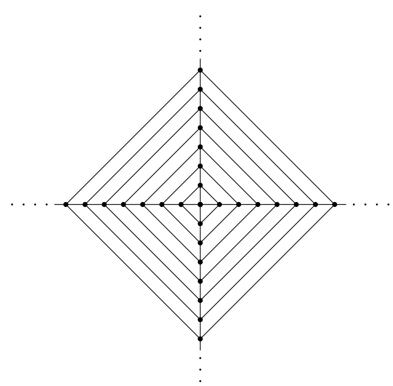

Figure 6: An example of a planar graph s.t. for all temperatures, correlations decay exponentially with distance.

2.3 A polynomial upper bound for trees

In this subsection we give an improved bound on relaxation time for the tree.

Let A be a finite set, and let $\alpha_{vw} : A \times A \to \mathbb{R}_+$ be a weight function. Let G be a graph. Let the Glauber dynamics be as defined above, and let $\mathcal{L} = \mathcal{L}(A, \alpha, G)$ be its generator. We say that the Glauber dynamics on (A, α, G) is ergodic if for every two legal configurations σ_1 and σ_2, we have $(\exp(\mathcal{L}))_{\sigma_1\sigma_2} > 0$. We will prove the following proposition:

Proposition 2.5. *Let $b \geq 2$, and let T denote the infinite b-ary tree, and let T_n be the b-ary tree with n levels. If the Glauber dynamics on (A, α, T_n) is ergodic for every n then*

$$\limsup_{n\to\infty} \frac{1}{n} \log\left(\tau_2\left(\mathcal{L}\left(A, \alpha, T_n\right)\right)\right) < \infty.$$

Conjecture 2.6. *Let $b \geq 2$, T denote the infinite b-ary tree, and let T_n be the b-ary tree with n levels. If the Glauber dynamics on (A, α, T) is ergodic then there exists $0 \leq \tau < \infty$ s.t.*

$$\lim_{n\to\infty} \frac{1}{n} \log\left(\tau_2\left(\mathcal{L}\left(A, \alpha, T_n\right)\right)\right) = \tau. \tag{14}$$

We prove a special case of Conjecture 2.6:

Proposition 2.7. *If the interactions are soft, i.e. $\alpha_{vw}(a, b) > 0$ for all v, w, a and b, then (14) holds.*

The main tool we use for proving Propositions 2.5 and 2.7 is block dynamics (see e.g. [25]). For a spin (or a color) $a \in A$, we denote by $\mathcal{L}(a, \alpha, n)$ the Glauber dynamics on the b-ary tree of depth n, under the interaction matrix α and with the boundary condition that the root has a parent colored a. With a slight abuse of notations, we say that $\tau_2(a, \alpha, n)$ is the relaxation time for $\mathcal{L}(a, \alpha, n)$.

Lemma 2.8. *Let*

$$\hat{\tau}_2(\alpha, n) = \sup_{a\in A} \tau_2(a, \alpha, n)$$

Then, for all m and n,

$$\hat{\tau}_2(\alpha, n + m) \leq \hat{\tau}_2(\alpha, n)\hat{\tau}_2(\alpha, m).$$

Proof. Let $l = n + m$. Partition the tree T_l into disjoint sets $V_1, ..., V_k$ to be specified below. We call $V_1, ..., V_k$ *blocks*, and consider the following block dynamics: Each block V_i has a (rate 1) Poisson clock, and whenever it rings,

V_i updates according to its Gibbs measure determined by the boundary conditions given by the configurations of $T_l^{(b)} - V_i$ and by the external boundary conditions. We denote by $\mathcal{L}^B = \mathcal{L}^B(V_1, ..., V_k)$ the generator for the block dynamics, and let $\mathcal{L}_a^B$ be the generator for the block dynamics with the boundary condition that the parent of the root has color a.

By [25][Proposition 3.4, page 119],

$$\hat{\tau}_2(\alpha, l) \leq \sup_i \hat{\tau}_2(\alpha, V_i) \cdot \sup_{a \in A} \tau_2(\mathcal{L}_a^B)$$

We now define the partition to blocks. For every vertex v up to depth n, the singleton $\{v\}$ is a block, and for every vertex w at depth n, the full subtree of depth m starting at w is a block (see Figure 7). All we need now to finish

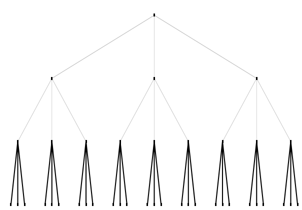

Figure 7: Partition of a tree to blocks

the proof is the following easy claim:

Claim 2.9.

$$\sup_{a \in A} \tau_2(\mathcal{L}_a^B) = \hat{\tau}_2(\alpha, n).$$

Proof. We use the following fact (that could also serve as a definition of the relaxation time). Given the dynamics $\mathcal{L}$ we define the Dirichlet form $\mathcal{E}[g, g] = \frac{1}{2} \sum_{\sigma,\tau} \mu[\sigma] K[\sigma \to \tau](g(\sigma) - g(\tau))^2$. Then

$$\tau_2 = \sup \left\{ \frac{\mu[g^2]}{\mathcal{E}[g, g]} \;:\; \mu[g] = 0 \right\}. \tag{15}$$

Clearly, the expression in (15) evaluated for f and $\mathcal{L}_a^B$ is equal to the one evaluated for g and $\mathcal{L}(a, \alpha, n)$, if

$$g(\eta) = f(\sigma) \text{ for all } \eta \text{ and } \sigma \text{ s.t. } \eta\,|_{T_n} = \sigma. \tag{16}$$

Therefore, we need to show that the maximum in (15) for the dynamics $\mathcal{L}_a^B$ is obtained at a function that satisfies (16). The maximum in (15) is obtained at an eigenfunction of $\mathcal{L}_a^B$. Moreover for every function g, $\mathcal{L}_a^B(g)$ satisfies (16) with some function f. It now follows that the maximum is obtained at a function that satisfies (16). □

□

Proof of Proposition 2.5. From Lemma 2.8 and the sub-additivity lemma, we learn that

$$\limsup_{n\to\infty} \frac{1}{n} \log\left(\hat{\tau}_2\left(\mathcal{L}\left(A, \alpha, T_n\right)\right)\right) < \infty$$

By another application of Matinelli's block dynamics lemma, we get that

$$\tau_2\left(\mathcal{L}\left(A, \alpha, T_n\right)\right) \leq \tau_2\left(\mathcal{L}\left(A, \alpha, T_1\right)\right) \cdot \hat{\tau}_2\left(\mathcal{L}\left(A, \alpha, T_{n-1}\right)\right) \tag{17}$$

and the proposition follows. □

Proof of proposition 2.7. From Lemma 2.8 and the sub-additivity lemma, we learn that there exists $0 \leq \tau < \infty$ s.t.

$$\lim_{n\to\infty} \frac{1}{n} \log\left(\hat{\tau}_2\left(\mathcal{L}\left(A, \alpha, T_n\right)\right)\right) = \tau.$$

For every a, let μ_a be the Gibbs measure for the tree of depth n with the boundary condition that the parent of the root has color a. Note that μ_a is the stationary distribution of $\mathcal{L}(a, \alpha, n)$. Since the interactions are soft, there exists $0 < C < \infty$ s.t. for every a, every n, and every two configurations σ and η on the tree of depth n,

$$\frac{1}{C}\mu(\sigma) \leq \mu_a(\sigma) \leq C\mu(\sigma),$$

and

$$\frac{1}{C}\mathcal{L}(\alpha, n)_{\sigma,\eta} \leq \mathcal{L}(a, \alpha, n)_{\sigma,\eta} \leq C\mathcal{L}(\alpha, n)_{\sigma,\eta}.$$

Therefore, by (15),

$$\frac{1}{C^3}\tau_2\left(\mathcal{L}\left(A, \alpha, T_n\right)\right) \leq \hat{\tau}_2\left(\mathcal{L}\left(A, \alpha, T_n\right)\right) \leq C^3\tau_2\left(\mathcal{L}\left(A, \alpha, T_n\right)\right)$$

and the proposition follows. □

3 Lower Bounds

Proof of Theorem 1.4, part 2a. Theorem 1.4 part 2a is a direct consequence of the extremal characterization of τ_2 given in (15), applied to the particular test function g which sums the spins on the boundary of the tree. It is easy to see that $\mu[g] = 0$ and that

$$\mathcal{E}[g, g] \le \sum_{\sigma,\tau} \mu[\sigma] K[\sigma \to \tau] = O(n_r).$$

We repeat the variance calculation from [8]. When $b(1-2\epsilon)^2 > 1$:

$$\begin{aligned}
\mu[g^2] &= \sum_{w\in\partial\mathbf{T}} \mu[\sigma_w^2] + \sum_{w\in\partial\mathbf{T}} \sum_{\substack{v\in\partial\mathbf{T}\\ v\neq w}} \mu[\sigma_w \sigma_v] \\
&= b^r \cdot \left(1 + \sum_{i=1}^{r} (b-1) b^{i-1} (1-2\epsilon)^{2i}\right) \\
&= b^r \left(1 + \Theta\left(\left(b(1-2\epsilon)^2\right)^r\right)\right) \\
&= \Theta\left(n_r^{1+\log_b(b(1-2\epsilon)^2)}\right).
\end{aligned}$$

It now follows by (15) that if $b(1-2\epsilon)^2 > 1$ then

$$\tau_2 = \Omega\left(n_r^{\log_b(b(1-2\epsilon)^2)}\right),$$

as needed. Repeating the calculation for the case $b(1-2\epsilon)^2 = 1$ yields that

$$\tau_2 = \Omega(\log n_r).$$

The proof follows. □

Remark: Suppose that μ admits a Markovian representation where the conditional distribution of σ_u given its parent σ_v is given by an $|\mathcal{A}| \times |\mathcal{A}|$ mutation matrix P. Let $\lambda_2(P)$ be the second eigen-value of P (in absolute value), and x the corresponding eigen-vector, so that $Px^t = \lambda_2(P)x^t$ and $|x|_2 = 1$.

Let g be the test function $g = c_n x^t$, where $c_n(i)$ is the number of boundary nodes that are labeled by i. It is then easy to see once again that $\mathcal{E}[g, g] =$

$O(n_r)$. Repeating the calculation from [29] it follows that if $b|\lambda_2(P)|^2 > 1$, then

$$\mathbf{Var}[g] = \Theta\left(n_r^{1+\log_b(b|\lambda_2(P)|^2)}\right).$$

Thus in this case,

$$\tau_2 = \Omega\left(n_r^{\log_b(b|\lambda_2(P)|^2)}\right).$$

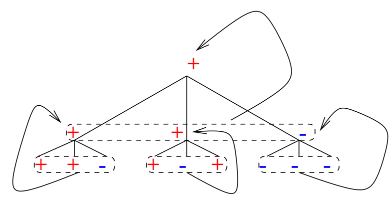

Figure 8: The recursive majority function.

In order to prove the lower bound on the relaxation time for very low temperatures stated in Theorem 1.4 part 2b, we apply (15) to the test function g which is obtained by applying recursive majority to the boundary spins; see [28] for background regarding the recursive-majority function for the Ising model on the tree. For simplicity we consider first the ternary tree T, see Figure 5. Recursive majority is defined on the configuration space as follows. Given a configuration σ, first label each boundary vertex v by its spin σ_v. Next, inductively label each interior vertex w with the label of the majority of the children of w. The value of the recursive majority function g is then the label of the root. We write σ_v for the spin at v and m_v for the recursive majority value at v.

Lemma 3.1. *If u and w are children of the same parent v, then* $\mathbf{P}[m_u \neq m_w] \leq 2\epsilon + 8\epsilon^2$.

Proof:

$$\mathbf{P}[m_u \neq m_w] \leq \mathbf{P}[\sigma_u \neq m_u] + \mathbf{P}[\sigma_w \neq m_w] + \mathbf{P}[\sigma_u \neq \sigma_v] + \mathbf{P}[\sigma_w \neq \sigma_v].$$

We will show that recursive majority is highly correlated with spin, *i.e.* if ϵ is small enough (say $\epsilon < 0.01$), then $\mathbf{P}[m_v \neq \sigma_v] \leq 4\epsilon^2$.

The proof is by induction on the distance ℓ from v to the boundary of the tree. For a vertex v at distance ℓ from the boundary of the tree, write $p_\ell = \mathbf{P}[m_v \neq \sigma_v]$. By definition $p_0 = 0 \leq 4\epsilon^2$.

For the induction step, note that if $\sigma_v \neq m_v$ then one of the following events hold:

- At least 2 of the children of v, have different σ value than that of σ_v, or
- One of the children of v has a spin different from the spin at v, and for some other child w we have $m_w \neq \sigma_w$, or
- For at least 2 of the children of v, we have $\sigma_w \neq m_w$.

Summing up the probabilities of these events, we see that $p_\ell \leq 3\epsilon^2 + 6\epsilon p_{\ell-1} + 3p_{\ell-1}^2$. It follows that $p_\ell \leq 4\epsilon^2$, hence the Lemma. □

Proof of Theorem 1.4 part 2b. Let m be the recursive majority function. Then by symmetry $\mathbf{E}[m] = 0$, and $\mathbf{E}[m^2] = 1$. By plugging m in definition (15), we see that

$$\tau_2 \geq \left(\sum_{\sigma,\tau : m[\sigma]=1, m[\tau]=-1} \mu[\sigma]\mathbf{P}[\sigma \to \tau] \right)^{-1}. \tag{18}$$

Observe that if σ, τ are adjacent configurations (i.e., $\mathbf{P}[\sigma \to \tau] > 0$) such that $m(\sigma) = 1$ and $m(\tau) = -1$, then there is a unique vertex v_r on the boundary of the tree where σ and τ differ. Moreover, if $\rho = v_1, \ldots, v_r$ is the path from ρ to v_r, then for σ we have $m(v_1) = \ldots = m(v_r) = 1$ while for τ we have $m(v_1) = \ldots = m(v_r) = -1$. Writing u_i, w_i for the two siblings of v_i for $2 \leq i \leq k$, we see that for all i, for both σ and τ we have $m(u_i) \neq m(v_i)$. Note that these events are independent for different values of i. We therefore obtain that the probability that $v_1, \ldots, v_r$ is such a path is bounded by $(2\epsilon + 8\epsilon^2)^{r-1}$. Since there are 3^r such paths and since $\mathbf{P}[\sigma \to \tau] \leq 3^{-r}$ we obtain that the right term of (18) is bounded below by

$$(2\epsilon + 8\epsilon^2)^{1-r} \geq n^{\Omega(\beta)}.$$

□

Note that the proof above easily extends to the d-regular tree for $d \geq 3$. A similar proof also applies to the binary tree T, where g is now defined as follows. Look at T_k for even k. For the boundary vertices define $m_v = \sigma_v$. For each vertex v at distance 2 from the boundary, choose three leaves on the boundary below it v_1, v_2, v_3 and let m_v be the majority of the values m_{v_i}. Now continue recursively.

Repeating the above proof, and letting $p_\ell = P[m_v \neq \sigma_v]$ for a vertex at distance 2ℓ, we derive the following recursion: $p_\ell \leq 3(2\epsilon)^2 + 6(2\epsilon)p_{\ell-1} + 3p_{\ell-1}^2$. We then continue in exactly the same way as for the ternary tree.

4 High temperatures

Proof of Theorem 1.4 part 3. Our analysis uses a comparison to block dynamics.

Block dynamics. We view our tree $T = T_r^{(b)}$ as a part of a larger b-ary tree T_* of height $r + 2h$, where the root ρ of T is at level h in T_*. For each vertex v of T_*, consider the subtree of height h rooted at v. A **block** is by definition the intersection of T with such a subtree. Each block has a rate 1 Poisson clock and whenever the clock rings we erase all the spins of vertices belonging to the block, and put new spins in, according to the Gibbs distribution conditional on the spins in the rest of T.

Discrete dynamics: In order to be consistent with [4], we will first analyze the corresponding discrete time dynamics: at each step of the block dynamics, pick a block at random, erase all the spins of vertices belonging to the block, and put new spins in, according to the Gibbs distribution conditional on the spins in the rest of T.

A coupling analysis. We use a weighted Hamming metric on configurations,

$$d(\sigma, \eta) = \sum_v \lambda^{|v|} 1(\sigma_v \neq \eta_v),$$

where $|v|$ denotes the distance from vertex v to the root. Let $\theta = 1 - 2\epsilon$ and $\lambda = 1/\sqrt{b}$. Note that $b\lambda\theta < 1$ and $\theta < \lambda$. Starting from two distinct configurations σ and η, our coupling always picks the same block in σ and in η and chooses the coupling between the two block moves which minimizes $d(\sigma', \eta')$.

We use path-coupling [4], *i.e.*, we will prove that for every pair of configurations which differ by a single spin, applying one step of the block dynamics will reduce the expected distance between the two configurations.

Let v be the single vertex, such that $\sigma_v \neq \eta_v$. Then $d(\sigma, \eta) = \lambda^{|v|}$. Let B denote the chosen block, and σ', η' be the configurations after the move. In order to understand (σ', η'), we will need the following Lemma.

Lemma 4.1. *Let T be a finite tree and let $v \neq w$ be vertices in T. Let $\{\beta_e \geq 0\}_{e \in E(T)}$ be the (ferromagnetic) interactions on T, and let $\{-\infty < H(u) < \infty\}_{u \in V(T)}$ be an external field on the vertices of T. we consider the following conditional Gibbs measures:*
$\mu_{+,H}$: The Gibbs measure with external field H conditioned on $\sigma_v = 1$.
$\mu_{-,H}$: The Gibbs measure with external field H conditioned on $\sigma_v = -1$.
Then, the function $\mu_{+,H}[\sigma_w] - \mu_{+,H}[\sigma_w]$ achieves its maximum at $H \equiv 0$.

Before proving the Lemma, we utilize it to prove Theorem 1.4, part 3. There are four situations to consider.
Case 1. if B contains neither v nor any vertex adjacent to v, then $d(\sigma', \eta') = d(\sigma, \eta)$.
Case 2. If B contains v, then $\sigma' = \eta'$ and $d(\sigma', \eta') = 0 = d(\sigma, \eta) - \lambda^{|v|}$. There are h such blocks, corresponding to the h ancestors of v at $1, 2, \ldots, h$ generations above v. (Note that this holds even when v is the root of T or a leaf of T, because of our definition of blocks).
Case 3. If B is rooted at one of v's children, then the conditional probabilities given the outer boundaries of B are not the same since one block has $+1$ above it and the other block has -1 above it. However both blocks have their leaves adjacent to the same boundary configuration. When considering the process on the block, the influence of the boundary configuration can be counted as altering the external field. Since σ and η have the same external fields and the same boundary configuration on all of the boundary vertices except v, by Lemma 4.1, conditioning on this lower boundary can only reduce $d(\sigma', \eta')$. Therefore, we bound $d(\sigma', \eta')$ by studying the case where one block is conditioned to having a $+1$ adjacent to the root, the other block is conditioned to having a -1 adjacent to the root, and no external field or boundary conditions. Then the block is simply filled in a top-down manner, every edge is faithful (*i.e.* the spin of the current vertex equals the spin of its parent) with probability θ and cuts information (the spin of the current vertex is a new random spin) with probability $1 - \theta$. Coupling these choices for corresponding edges for σ and for η, we see that the distance between σ' and η' will be equal to the weight of the cluster containing v, in expectation $\sum_j \lambda^{|v|+j} b^j \theta^j \leq \lambda^{|v|}/(1 - b\lambda\theta)$. There are b such blocks, corresponding to the b children of v.
Case 4. If B is rooted at v's ancestor exactly $h + 1$ generations above v, then the conditional probabilities are not the same since one block has a leaf v adjacent to a $+1$ and the other block has a leaf adjacent to a -1. There

is exactly one such block. Again we appeal to Lemma 4.1 to show that the expected distance is dominated by the size of the θ cluster of w. The expected weight of v's cluster is bounded by summing over the ancestors w of v:

$$\begin{aligned}\sum_w \theta^{|v|-|w|} \sum_j \lambda^{|w|+j} b^j \theta^j &= \\ &= \frac{\sum_w \lambda^{|w|}\theta^{|v|-|w|}}{1-b\lambda\theta} \\ &= \frac{\lambda^{|v|}}{(1-\theta\lambda^{-1})(1-b\lambda\theta)}.\end{aligned}$$

Overall, the expected change in distance is

$$\mathbf{E}(d(\sigma',\eta') - d(\sigma,\eta)) \le$$
$$\left(\frac{b\lambda^{|v|}}{1-b\lambda\theta} + \frac{\lambda^{|v|}}{(1-\theta\lambda^{-1})(1-b\lambda\theta)} - h\lambda^{|v|}\right)\frac{1}{n+h-1}.$$

If the block height h is a sufficiently large constant, we get that for some positive constant c,

$$\mathbf{E}(d(\sigma',\eta') - d(\sigma,\eta)) \le \frac{-c\lambda^{|v|}}{n} \le \frac{-c}{n} d(\sigma,\eta). \tag{19}$$

Note that $\max d(\sigma,\eta) = \sum_{j\le r} b^j\lambda^j \le \sqrt{n}$. Therefore, by a path-coupling argument (see [4]) we obtain a mixing time of at most $O(n\log n)$ for the blocks dynamics.

Spectral gap of discrete time block dynamics. The $(1-c/n)$ contraction at each step of the coupling implies, by an argument from [5] which we now recall, that the spectral gap of the block dynamics is at least c/n. Indeed, let λ_2 be the second largest eigenvalue in absolute value, and f an eigenvector for λ_2. Let $M = \sup_{\sigma,\eta} |f(\sigma) - f(\eta)|/d(\sigma,\eta)$ and denote by $\mathbf{P}$ the transition operator. Then

$$
\begin{aligned}
|\lambda_2|M &= \sup_{\sigma,\eta} \frac{|\mathbf{P}f(\sigma)-\mathbf{P}f(\eta)|}{d(\sigma,\eta)} \quad \text{since } f \text{ eigenvector for } \lambda_2 \\
&\le \sup_{\sigma,\eta} \sum_{\sigma',\eta'} \mathbf{P}[(\sigma,\eta)\to(\sigma',\eta')] \frac{|f(\sigma')-f(\eta')|}{d(\sigma',\eta')} \frac{d(\sigma',\eta')}{d(\sigma,\eta)} \\
&\le \sup_{\sigma,\eta} \sum_{\sigma',\eta'} \mathbf{P}[(\sigma,\eta)\to(\sigma',\eta')] M \frac{d(\sigma',\eta')}{d(\sigma,\eta)} \\
&= M \sup_{\sigma,\eta} \frac{\mathbf{E}[d(\sigma',\eta')]}{d(\sigma,\eta)} \\
&\le (1-c/n)M \quad \text{by} \quad (19).
\end{aligned}
$$

Thus $|\lambda_2|M \le (1-c/n)M$, whence the (discrete time) block dynamics has relaxation time at most $O(n)$.

Relaxation time for continuous time block dynamics. The continuous time dynamics is n times faster than the discrete time dynamics. This is true because the transition matrix for the discrete dynamics is $M = \mathbf{I} + \frac{1}{n}\mathcal{L}$ where $\mathbf{I}$ is the 2^n-dimensional unit matrix. Therefore

$$\tau_2(\text{block dynamics}) = O(1).$$

Relaxation time for single-site dynamics. Since each block update can be simulated by doing a constant number of single-site updates inside the block, and each tree vertex only belongs to a bounded number of blocks, it follows from proposition 3.4 of [25] that the relaxation time of the single-site Glauber dynamics is also $O(1)$. $\square$

Proof of Lemma 4.1. **Reduction from trees to paths.** We first claim that it suffices to prove the lemma when the tree T consists of a path $v = v_1, \ldots, v_k = w$. (see Figure 9). To see this, let $T_1, T_2, \ldots, T_k$ be the connected components of T when the edges in the path $v_1, v_2, \ldots, v_k$ are erased, s.t. $v_i \in T_i$ for $i = 1, 2, \ldots, k$. Let σ be a configuration on $v_1, \ldots, v_k$, and for a subgraph J let $S(J)$ be the space of configurations on J. The probability of a configuration σ on $v_1, \ldots, v_k$ is

$$\frac{1}{Z} \quad \exp \quad \left(\sum_{i=1}^{k-1} \beta_{\{v_i,v_{i+1}\}} \sigma_{v_i} \sigma_{v_{i+1}} \right) \cdot \prod_{i=1}^{k} \left(\sum_{\tau \in S(T_i - \{v_i\})} \exp\left(\mathcal{H}(\tau \cup \sigma_{v_i})\right) \right)$$

$$= \frac{1}{Z'} \quad \exp \quad \left(\sum_{i=1}^{k-1} \beta_{\{v_i,v_{i+1}\}} \sigma_{v_i} \sigma_{v_{i+1}} + \sum_{i=1}^{k} H'_{v_i} \sigma_{v_i} \right)$$

for some external field $\{H'_u\}$ depending only on $\{H_u\}$ and $\{\beta_e\}$, where Z and Z' are partition functions and $\mathcal{H}(\cdot)$ denotes the Hamiltonian.

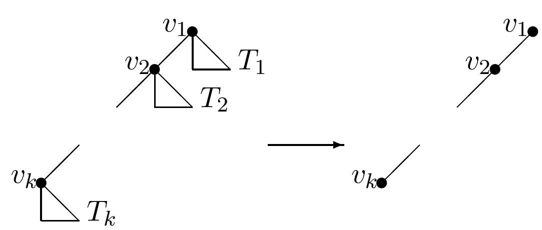

Figure 9: Reduction from trees to paths.

We will now prove the lemma by induction on the length of the path $v_1, \ldots, v_k$.

Paths of length 2. Assume $k = 2$. Writing β for the strength of (v_1, v_2) interaction, H for external field at $w = v_2$. Then,

$$\begin{aligned} \mu_{+,\tau}[\sigma_w] - \mu_{-,\tau}[\sigma_w] &= \frac{e^{\beta+H} - e^{-\beta-H}}{e^{\beta+H} + e^{-\beta-H}} - \frac{e^{-\beta+H} - e^{\beta-H}}{e^{-\beta+H} + e^{\beta-H}} \\ &= \tanh(\beta + H) - \tanh(H - \beta). \end{aligned}$$

It therefore suffices to prove that for $\beta > 0$, the function

$$H \mapsto g(\beta, H) = \tanh(H + \beta) - \tanh(H - \beta)$$

has a unique maximum at $H = 0$. Consider the partial derivative,

$$g_H(\beta, H) = \cosh^{-2}(H + \beta) - \cosh^{-2}(H - \beta). \tag{20}$$

Therefore, if $\beta > 0$ and $H > 0$ then $g_H(\beta, H) < 0$ and if $\beta > 0$ and $H < 0$ then $g_H(\beta, H) > 0$. Thus $H = 0$ is the unique maximum and the claim for $k = 2$ follows.

Induction step. We assume that the claim is true for $k-1$ and prove it for k. We denote $v' = v_{k-1}, \mu'_{+,H} = \mu_H[\cdot|\sigma_{v'} = 1]$ and similarly $\mu'_{-,H}$. Now,

$$\begin{aligned}
&\mu_{+,H}[\sigma_w] - \mu_{-,H}[\sigma_w] \\
&= \big(\mu_{+,H}[\sigma_{v'} = 1]\mu'_{+,H}[\sigma_w] + \mu_{+,H}[\sigma_{v'} = -1]\mu'_{-,H}[\sigma_w]\big) \\
&- \big(\mu_{-,H}[\sigma_{v'} = 1]\mu'_{+,H}[\sigma_w] + \mu_{-,H}[\sigma_{v'} = -1]\mu'_{-,H}[\sigma_w]\big) \\
&= \frac{1}{2}(\mu_{+,H} - \mu_{-,H})[\sigma_{v'}](\mu'_{+,H} - \mu'_{-,H})[\sigma_w]. \qquad (21)
\end{aligned}$$

Since by the induction hypothesis both multipliers in (21) achieve their maximums at $H \equiv 0$, we get that $\mu_{+,H}[\sigma_w] - \mu_{-,H}[\sigma_w]$ also achieves its maximum at $H \equiv 0$. □

5 Proof of Theorem 1.5

Recall that we denoted by σ_r the configuration on all vertices at distance exactly r from ρ. Also recall that μ is the Gibbs measure which is stationary for the Glauber dynamics. We abbreviate $\int f\, d\mu$ as $\mu(f)$.

Mutual information and L^2 estimates. For Markov chains such as $\{\sigma_r\}$, it is generally known [36] that (5) follows from (4), which itself, is a consequence of the following stronger statement:

There exists $c_* > 0$ such that for any vertex set $A \subset G_{r/2}$ and any functions f, g with $\mu(f) = \mu(g) = 0$, we have

$$\mu(fg) \le e^{-c_* r}(\mu(f^2)\mu(g^2))^{1/2}, \qquad (22)$$

provided that $f(\sigma)$ depends only on σ_A and $g(\sigma)$ depends only on σ_r. (22) will follow from a more general proposition below. For a set A of vertices in a graph G we write $\partial_i A$ for the set of vertices v in A for which there exists an edge (v, u) with $u \notin A$.

Proposition 5.1. *Let G be a finite graph, and let A and B be sets of vertices in G. Let d be the distance between A and B and let Δ be the maximum degree in G. For $0 < c < 1$, let*

$$I(c) = c - \log c - 1. \qquad (23)$$

Let c^* be the unique $0 < c < 1$ satisfying $I(c) = \log \Delta$ and for $0 < c < c^*$, let

$$C(c, \Delta) = \big(1 - e^{\log \Delta - I(c)}\big)^{-1/2}. \qquad (24)$$

Further, let λ_2 be the absolute value of the second eigenvalue of the generator of the Glauber dynamics on G, i.e. $\lambda_2 = \frac{1}{\tau_2}$. Let $f = f(\sigma)$ depend only on the values of the configuration in A and $g = g(\sigma)$ depend only on the values of the configuration in B. If $\mu(f) = \mu(g) = 0$, then

$$\mu(fg) \leq \left(e^{-cd\lambda_2} + 2C(c,\Delta)\sqrt{|\partial_i A|\, e^{d(\log\Delta - I(c))}} \right) \|f\|_2 \|g\|_2. \tag{25}$$

In particular (by letting $c = e^{-\log\Delta-\gamma-2}$) for $\gamma \geq 0$,

$$\mu(fg) \leq \left(e^{-d\lambda_2 \exp(-\log\Delta-\gamma-2)} + 4\sqrt{\exp(-(\gamma+1)d)\,|\partial_i A|} \right) \|f\|_2 \|g\|_2. \tag{26}$$

Proof of Theorem 1.5. Note that $|\partial_i A| \leq |A| \leq \Delta^{r/2}$. Therefore, to prove (22) we use (26) with B=$\{v : d(v,o) = r\}$, $d = r/2$ and γ s.t. $e^\gamma > \Delta$. □

Proof of Proposition 5.1. We use a coupling argument. Let μ be the Gibbs measure on G, and let X_0 be chosen according to μ. Let X_t and Y_t be defined as follows: Set $Y_0 = X_0$. For $t > 0$, let X_t and Y_t evolve according to the dynamics with the following graphical representation: Each $v \in G$ has a Poisson clock. Assume the clock at v rang at time t, and let X_{t-} and Y_{t-} be the configurations just before time t. At time t we do the following:

1. If $v \in B$ then X_v updates according to the Gibbs measure, and Y_v does not change.
2. If $v \notin B$ and $X_{t-}(w) = Y_{t-}(w)$ for every neighbor w of v, then both X and Y update according to the Gibbs measure so that $X_t(v) = Y_t(v)$.
3. If $v \notin B$ and there exists a neighbor w of v s.t. $X_{t-}(w) \neq Y_{t-}(w)$ then both X and Y update according to the Gibbs measure, but this time independently of each other.

For a vertex $v \in B$ we define t_v to be the first time the Poisson clock at v rang. For any $v \in G \setminus B$, we define t_v to be the first time the Poisson clock at v rang after $\min_{(w,v)\in E_G} t_w$. Note that $X_t(v) = Y_t(v)$ at any time $t < t_v$, and that t_v depends only on the Poisson clocks, and is independent of the initial configuration X_0. We let $t_A = \min_{v\in A} t_v$.

Let $\mathcal{F}_t$ denote the (σ-algebra of the) Poisson clocks at the vertices up to time t. Let $(P^t f)(\sigma) = \mathbf{E}[f(X_t)|X_0 = \sigma, \mathcal{F}_t]$ and let $(Q^t f)(\sigma) = \mathbf{E}[f(Y_t)|X_0 = \sigma, \mathcal{F}_t]$. Also, let $(\tilde{P}^t f)(\sigma) = \mathbf{E}[f(X_t)|X_0 = \sigma]$ and $(\tilde{Q}^t f)(\sigma) = \mathbf{E}[f(Y_t)|X_0 =$

$\sigma]$. Since for all t the process Y_t is at the stationary distribution and $Y_t|_B = X_0|_B$ for all t, we get

$$\mu[gf] = \mathbf{E}[g(Y_t)f(Y_t)] = \mathbf{E}[g(X_0)f(Y^t)] = \mathbf{E}[g\tilde{Q}^t f]. \tag{27}$$

If $t < t_A$, then clearly $X_t = Y_t$ on A. Therefore, $\|(Q^t f - P^t f)\cdot 1_{t<t_A}\|_2^2 = 0$. On the other hand, $\|Q^t f\|_2 \le \|f\|_2$ $\mathcal{F}_t -$ a.s. and $\|P^t f\|_2 \le \|f\|_2$ $\mathcal{F}_t -$ a.s. This is because the operators $f \to Q^t(f)$ and $f \to P^t(f)$ given $\mathcal{F}_t$ are Markov operators and hence contractions. Therefore

$$\begin{aligned}
\|\tilde{Q}^t f - \tilde{P}^t f\|_2^2 &= \mathbf{E}\left([\tilde{Q}^t f(X_0) - \tilde{P}^t f(X_0)]^2\right) \\
&\le \mathbf{E}\left([Q^t f(X_0) - P^t f(X_0)]^2\right) \\
&= \mathbf{P}(t \le t_A)\int d\mu(\sigma)\mathbf{E}\left([Q^t f(\sigma) - P^t f(\sigma)]^2 | t \le t_A\right) \\
&+ \mathbf{P}(t > t_A)\int d\mu(\sigma)\mathbf{E}\left([Q^t f(\sigma) - P^t f(\sigma)]^2 | t > t_A, X_0 = \sigma\right) \\
&\le 4\mathbf{P}[t_A \le t]\|f\|_2^2
\end{aligned}$$

where the first inequality is because $\tilde{Q}^t f(X_0) - \tilde{Q}^t f(X_0)$ is a conditional expectation of $Q^t f(X_0) - P^t f(X_0)$, and the second inequality is because $\{t > t_A\}$ is $\mathcal{F}_t$-measurable. Therefore, by the Cauchy-Schwartz inequality,

$$\mathbf{E}[(\tilde{Q}^t f - \tilde{P}^t f)g] \le 2\sqrt{\mathbf{P}[t_A \le t]}\|f\|_2\ \|g\|_2\,. \tag{28}$$

Since

$$\mathbf{E}[g\tilde{P}^t f] \le e^{-\lambda_2 t}\|f\|_2\|g\|_2,$$

We infer that from (28) and (27) that

$$\mu[fg] \le \left(e^{-\lambda_2 t} + 2\sqrt{\mathbf{P}[t_A \le t]}\right)\|f\|_2\|g\|_2\,. \tag{29}$$

It remains to bound the two terms in the right-hand side of (29).

For $0 < c < c^*$, we take $t = cd$. We obtain that the first term is $e^{-cd\lambda_2}$, as desired. It remains to bound $\mathbf{P}[t_A \le t]$. We note that $t_A \le t$ only if there is some self-avoiding path (sometimes referred to as “path of disagreement”) between the A and B along which the discrepancy between the two distributions has been conveyed in time less than t.

Time-reversing the process, this means that first-passage-percolation with rate-1 exponential passage times starting at A needs to arrive at distance d

within time cd. There are at most $|\partial_i A|\,\Delta^k$ paths of length k for the first-passage-percolation for each $k \geq d$. Let $\tau(v,w)$ be the time needed to cross the edge (v,w). For each path $v_1, v_2, \ldots, v_k$,

$$\mathbf{P}\left(\tau(v_1,v_2)+\tau(v_2,v_3)+\ldots\tau(v_{k-1},v_k) < cd\right) < e^{-kI(c)}$$

where $I(c) = c - \log c - 1$ is the large deviation rate function for the exponential distribution. Therefore,

$$\begin{aligned}\mathbf{P}(t_A \leq t) &\leq |\partial_i A| \sum_{k=d}^{\infty} \exp\left[k(-I(c') + \log\Delta)\right] \\ &\leq C^2(c,\Delta)\,|\partial_i A|\,e^{d(\log\Delta - I(c))}\end{aligned}$$

□

Plugging this bound into (29), we obtain (25) as needed.

6 Open Problems

In this section we specify some relevant problems that are still open.

Problem 1. *What is the relaxation time $\tau_2(n,b,b^{-1/2})$ of the Glauber dynamics of the Ising model on the b-ary tree of depth n at the critical temperature $1-2\epsilon = \frac{1}{\sqrt{b}}$?*

Using the sum of spins as a test function, we learn that $\Omega(\log n)$ is a lower bound for $\tau_2(n,b,b^{-1/2})$. We conjecture that the relaxation time is of order $\Theta(\log n)$. A weaker conjecture is that

$$\lim_{n\to\infty} \frac{\log(\tau_2(n,b,b^{-1/2}))}{n} = 0.$$

Problem 2. *Fix b, and let*

$$\tau_2(\beta) = \lim_{n\to\infty} \frac{\log(\tau_2(n,b,\beta))}{n}.$$

Theorem 1.4 part 1 tells us that $\tau_2(\beta)$ exists and is finite for all β. Show that $\tau_2(\beta)$ is a monotone function of β. This question is a special case of a more general monotonicity conjecture due to the fourth author, described in [30]. See [30] where a monotonicity result is proven for the Ising model on the cycle.

Problem 3. *For the Ising model (with free boundary conditions and no external field) on a general graph of bounded degree, does the converse of Theorem 1.5 hold, i.e., does uniform exponential decay of point-to-set correlations imply a uniform spectral gap?*
(As pointed out by F. Martinelli (personal communication), the converse fails in certain lattices if plus boundary conditions are allowed).

Problem 4. *Recall the general upper bound* $ne^{(4\xi(G)+2\Delta)\beta}$ *on the relaxation time of Glauber dynamics in terms of cut-width from proposition 1.1. For which graphs does a similar lower bound of the form* $\tau_2 \geq e^{c\xi(G)\beta}$ *(for some constant* $c > 0$*) hold at low temperature?*

Such a lower bound is known to hold for boxes in a Euclidean lattice, our results imply its validity for regular trees, and we can also verify it for expander graphs. A specific class of graphs which could be considered here are the metric balls around a specific vertex in an infinite graph Γ *that has critical probability* $p_c(\Gamma) < 1$ *for bond percolation.*

Remark. After the results presented here were described in the extended abstract [19], striking further results on this topic were obtained by F. Martinelli, A. Sinclair, and D. Weitz [26]. For the Ising model on regular trees, in the temperatures where we show the Glauber dynamics has a uniform spectral gap, they show it satisfies a uniform log-Sobolev inequality; moreover, they study in depth the effects of external fields and boundary conditions.
Acknowledgment. We are grateful to David Aldous, David Levin, Laurent Saloff-Coste and Peter Winkler for useful discussions. We thank Dror Weitz for helpful comments on [19].

References

[1] Aldous, D. and Fill, J. A. (2000) Reversible Markov chains and random walks on graphs, *book in preparation.* Current version available at `www.stat.berkeley.edu/users/aldous/book.html`.

[2] van den Berg, J. (1993) A uniqueness condition for Gibbs measures, with application to the 2-dimensional Ising antiferromagnet. *Comm. Math. Phys.* **152, no. 1**, 161–166.

[3] Bleher, P. M., Ruiz, J. and Zagrebnov V. A. (1995) On the purity of limiting Gibbs state for the Ising model on the Bethe lattice, *J. Stat. Phys* **79**, 473–482.

[4] Bubley, R. and Dyer, M. (1997) Path coupling: a technique for proving rapid mixing in Markov chains. In Proceedings of the 38th Annual Symposium on Foundations of Computer Science (FOCS), 223–231.

[5] Chen, M. F. (1998) Trilogy of couplings and general formulas for lower bound of spectral gap. *Probability towards 2000* Lecture Notes in Statist., **128**, Springer, New York, 123–136.

[6] Cover, T. M. and Thomas, J. A. (1991) *Elements of Information Theory*, Wiley, New York.

[7] Dyer M. and Greenhill C. (2000), 'On Markov chains for independent sets', *J. Algor.* **35,** 17–49.

[8] Evans, W., Kenyon, C., Peres, Y. and Schulman, L. J. (2000) Broadcasting on trees and the Ising Model, *Ann. Appl. Prob.*, **10**, 410–433.

[9] Fortuin C. M., Kasteleyn P. W. (1972) On the random-cluster model. I. Introduction and relation to other models. *Physica* **57**, 536–564.

[10] Fortuin C. M., Kasteleyn P. W. and Ginibre J. (1971) Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.* **22** , 89–103.

[11] Olle Haggstrom, Johan Jonasson and Russell Lyons. Explicit isoperimetric constants and phase transitions in the random-cluster model, Ann. Probab. 30 (2002), 443–473.

[12] Ioffe, D. (1996). A note on the extremality of the disordered state for the Ising model on the Bethe lattice. *Lett. Math. Phys.* **37**, 137–143.

[13] Janson, S., Luczak T. and Ruciński A. (2000) *Random Graphs*, Wiley, New York.

[14] Jerrum, M. (1995) A very simple algorithm for estimating the number of k-colorings of a low-degree graph. *Rand. Struc. Alg.* **7**, 157–165.

[15] Jerrum, M. and Sinclair, A. (1989). Approximating the permanent. *Siam Jour. Comput.* **18**, 1149–1178.

[16] Jerrum, M. and Sinclair, A. (1993). Polynomial time approximation algorithms for the Ising model. *Siam Jour. Comput.* **22**, 1087–1116.

[17] Jerrum, M., Sinclair, A. and Vigoda, E. (2001). A polynomial-time approximation algorithm for the permanent of a matrix with non-negative entries. Proceedings of the 33*rd Annual ACM Symposium on Theory of Computing*, Crete, Greece.

[18] S. Katok (1992) Fuchsian Groups. University of Chicago Press.

[19] Kenyon, C., Mossel, E. and Peres, Y. (2001) Glauber dynamics on trees and hyperbolic graphs. *42nd IEEE Symposium on Foundations of Computer Science (Las Vegas, NV, 2001)*, 568–578, *IEEE Computer Soc., Los Alamitos, CA,*.

[20] Kinnersley, N. G. (1992) The vertex seperation number of a graph equals its path-width. *Infor. Proc. Lett.*, **42**, 345–350.

[21] Liggett, T. (1985) *Interacting particle systems*, Springer, New York.

[22] Luby, M. and Vigoda, E. (1997) Approximately Counting Up To Four, In proceedings of the 29th Annual Symposium on Theory of Computing (STOC),, 682–687.

[23] Luby, M. and Vigoda, E. (1999). Fast Convergence of the Glauber Dynamics for Sampling Independent Sets, Statistical physics methods in discrete probability, combinatorics and theoretical computer science, *Rand. Struc. Alg.* **15**, 229–241.

[24] W. Magnus, Noneuclidean tessellations and their groups, Academic Press, New York and London, 1974.

[25] Martinelli, F. (1998) Lectures on Glauber dynamics for discrete spin models. *Lectures on probability theory and statistics (Saint-Flour, 1997)* 93–191, *Lecture Notes in Math.* **1717**, Springer, Berlin.

[26] Martinelli, F., Sinclair, A. and Weitz, D. (2003) Glauber dynamics on trees: Boundary conditions and mixing time. *Preprint*, available at `http://front.math.ucdavis.edu/math.PR/0307336`

[27] Mossel, E. (2001) Reconstruction on trees: Beating the second eigenvalue, *Ann. Appl. Probab.*, **11 no. 1**,285–300.

[28] Mossel, E. (1998) Recursive reconstruction on periodic trees. *Rand. Struc. Alg.* **13**, 81–97.

[29] Mossel, E. and Peres Y. (2003) Information flow on trees, to appear in *Ann. Appl. Probab.*.

[30] Nacu, S. (2003) Glauber dynamics on the cycle is monotone. *To appear, Probab. Theory Related Fields*

[31] Paterson, A. L. T. (1988) *Amenability.* American Mathematical Soc., Providence.

[32] Propp, J. and Wilson, D. (1996) Exact Sampling with Coupled Markov Chains and Applications to Statistical Mechanics. *Rand. Struc. Alg.* **9**, 223–252.

[33] Peres, Y. and Winkler, P. (2003), in preparation.

[34] Randall, D. and Tetali, P. (2000), Analyzing Glauber dynamics by comparison of Markov chains. *J. of Math. Phys.* **41**, 1598–1615.

[35] N. Robertson and P.D. Seymour (1983), Graph minors. I. Excluding a forest. *J. Comb. Theory Series B* 35, 39–61.

[36] Saloff-Coste, L. (1997) Lectures on finite Markov chains. *Lectures on probability theory and statistics (Saint-Flour, 1996)* 301–413, Lecture Notes in Math. **1665**, Springer, Berlin.

[37] Vigoda, E. (2001). Improved bounds for sampling colorings. Probabilistic techniques in equilibrium and nonequilibrium statistical physics. *J. Math. Phys.* **41**, no. 3, 1555–1569.