The Kesten-Stigum Reconstruction Bound Is Tight for Roughly Symmetric Binary Channels

Christian Borgs, Theory Group, Microsoft Research

Jennifer Chayes, Theory Group, Microsoft Research

Elchanan Mossel*, Dept. of Statistics, U.C. Berkeley

Sebastien Roch†, Dept. of Statistics, U.C. Berkeley

Abstract

We establish the exact threshold for the reconstruction problem for a binary asymmetric channel on the b-ary tree, provided that the asymmetry is sufficiently small. This is the first exact reconstruction threshold obtained in roughly a decade. We discuss the implications of our result for Glauber dynamics, phylogenetic reconstruction, noisy communication and the so-called "replica symmetry breaking" in spin glasses and random satisfiability problems.

1 Introduction

The so-called *reconstruction problem* is a fundamental problem concerning signal decay in noisy communication with duplication, which is intimately related to noisy computation [5, 20], mixing of Glauber dynamics [1, 16] and phylogenetic reconstruction [4, 22]. In the binary case, the problem is defined as follows. Let T_b be a complete b-ary tree and M be a binary (asymmetric) channel, i.e. a 2x2 stochastic matrix. Consider the following Markov process that associates a state in $\{+,-\}$ to each node of T_b: first pick a state at the root in $\{+,-\}$ according to the stationary distribution of M; then move away from the root and iteratively apply the channel M on each edge. The reconstruction problem is the problem of determining the state of the root of T_b, given the state of the Markov chain on level n of the tree, as n goes to $+\infty$. Roughly speaking, the reconstruction problem on (T_b, M) is said to be "solvable" if one can determine the root value significantly better than by guessing according to the stationary distribution (a precise definition is given below).

Solvability of the reconstruction problem is, to some extent, governed by the second eigenvalue θ of the channel M. More precisely, for the binary *symmetric* channel, it was shown in [2] that the reconstruction problem is solvable if and only if $b\theta^2 > 1$. For all other channels, it was also known—and easy to prove—that $b\theta^2 > 1$ implies solvability, but *exact* thresholds for non-solvability were not known before this work. Here we show that the bound $b\theta^2 > 1$—which we refer to as the *Kesten-Stigum bound* [13]—is tight provided that M is close enough to symmetric, i.e., letting

$$M = \frac{1}{2}\left[\begin{pmatrix} 1+\theta & 1-\theta \\ 1-\theta & 1+\theta \end{pmatrix} + \delta \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}\right], \qquad (1)$$

we show that the reconstruction problem for M on T is not solvable if $b\theta^2 \leq 1$ and $|\delta|$ is sufficiently small.

As we explain next, our results have potential applications in noisy computation, mixing, phylogenetics, and random satisfiability problems.

Noisy Communication/Computation. The reconstruction problem is concerned essentially with a trade-off between noise and duplication in a tree communication network. At the root of the tree network a state is chosen. It is then propagated down the tree by applying the channel M at each edge and duplicated at each node towards all outgoing edges. The reconstruction problem is solvable if the signal obtained down the tree has non-trivial correlation with the state at the root.

It is natural to expect that the existence of correlation is a monotone property of the branching ratio and of the noise level. The basic question is: what is the trade-off between the two? The threshold $b\theta^2 = 1$ for the b-ary tree yields a very elegant trade-off. This threshold was previsouly known only for binary symmetric channels and is extended here to roughly symmetric channels. In particular, our results suggest that the $b\theta^2 = 1$ trade-off in the symmetric case is in fact "robust" to a small "systematic bias" in the channel M.

*Supported by an Alfred Sloan fellowship in Matheamtics and by NSF grants DMS-0528488, DMS-0504245 and DMS-0548249 (CAREER). Most of this work was done while visiting Microsoft Research.

†Most of this work was done while visiting Microsoft Research.

The reconstructon problem is also closely related to noisy computation models where each gate independently introduces error, see [27, 6]. The reconstruction problem is not equivalent to noisy computation. However, the two problems are closely related. See [5, 20] for details.

Phylogenetic Reconstruction. Phylogenetic reconstruction is a major task of systematic biology [7]. It was recently shown in [4] that for binary symmetric channels, also called CFN models in evolutionary biology, the sampling efficiency of phylogenetic reconstruction is determined by the reconstruction threshold. Thus, if for all edges of the tree, it holds that $2\theta^2 > 1$ the tree can be recovered efficiently from $O(\log n)$ samples. If $2\theta^2 < 1$, then [22] implies that $n^{\Omega(1)}$ samples are needed. In fact, the proof of the lower bound in [22] implies the lower bound $n^{\Omega(1)}$ whenever the reconstruction problem is "exponentially" unsolvable.

Our results here imply $n^{\Omega(1)}$ lower bounds for phylogenetic reconstruction for asymmetric channels such that $2\theta^2 < 1$ and $|\delta|$ sufficiently small. The details are omitted from this extended abstract. It is natural to conjecture that this is tight and that if $2\theta^2 > 1$ then phylogenetic reconstruction may be achieved with $O(\log n)$ sequences.

Mixing of Markov Chains. One of the main themes at the intersection of statistical physics and theoretical computer science in recent years has been the study of connections between spatial and temporal mixing. It is widely accepted that spatial mixing and temporal mixing of dynamics go hand in hand though this was proven only in restricted settings.

In particular, the spatial mixing condition is usually stated in terms of uniqueness of Gibbs measures. However, as shown in [1] , this spatial condition is too strong. In particular, it is shown in [1] that the spectral gap of continuous-time Glauber dynamics for the Ising model with no external field and no boundary conditions on the b-ary tree is $\Omega(1)$ whenever $b\theta^2 < 1$. This should be compared with the uniqueness condition on the tree given by $b\theta < 1$. In [16] this result is extended to the log Sobolev constant. In [16] it is also shown that for measures on trees, a super-linear decay of point-to-set correlations implies an $\Omega(1)$ spectral gap for the Glauber dynamics with free boundary conditions.

Thus our results not only give the exact threshold for reconstructibility. They also yield an exact threshold for mixing of Glauber dynamics on the tree for Ising models with a small external field. The details are omitted from this extended abstract.

Replica Symmetry Breaking/Random SAT. The replica and cavity methods were invented in theoretical physics to solve Ising spin glass problems on the complete graph—the so-called Sherrington-Kirkpatrick model. Spin-glass problems are random instances of satifiability problems and the main interest in physics is in the distribution of optimal (maximally satisfying) solutions. Spin-glass problems in the complete graph correspond to satisfiability problems where a (weighted) random constraint exists between every pair of variables.

The replica and cavity methods, while not mathematically rigorous, led to numerous predictions on the spin glass and other models on dense graphs, a few of which were proved many years later. These methods were later applied to random satisfiability problems in which every variable appears in expectation in a constant number of randomly chosen clauses. The analysis of these problems, called dilute spin glasses in physics, have led to the best known empirical algorithms for solving random satisfiability problems [18, 24].

A central concept in this theory is the notion of a "glassy phase" of the spin glass measure. In the glassy phase, the distribution on random optimal solutions decomposes into an exponential number of "lumps", where the Hamming distance between every two lumps is $\Omega(n)$. One of the standard techniques in physics for determining the glassy phase is via "replica symmetry breaking". Moreover, there are certain glassy phases for which the replica symmetry breaking is relatively simple—those which are said to have "one-step replica symmetry breaking"; and others in which the replica symmetry breaking is more complicated—those with so-called "full replica symmetry breaking". Rougly speaking, in the "full replica" case each lump decomposes futher into lumps etc., while in the "one-step replica" each lump is "connected" (with respect to the usual Hamming metric).

In a recent paper [19] it is claimed that the parameters for which a "glassy phase occurs" are exactly the same as the parameters for which the reconstruction problem is not solvable. More formally, for determining if the glassy phase occurs for random $(b+1)$-regular graphs and Gibbs measures with some parameters, one needs to check if the reconstruction problem for the b-ary tree and associated parameters is solvable or not.

Furthermore, it is claimed in [19] that the reconstruction problem determines the type of glassy phase as follows. Mezard and Montanari predict that one-step replica symmetry breaking occurs exactly when the Kesten-Stigum bound is not equal to the reconstruction bound; otherwise full replica symmetry breaking occurs.

Thus our results proved here, in conjunction with the theoretical physics predictions of [19], suggest the existence of two types of glassy phases for spin systems on random graphs. It is an interesting challenge to state these predictions in a compact mathematical formulation to prove or disprove them.

1.1 Previous Results

The study of the reconstruction problem began in the seventies [26, 10] where the problem was introduced in terms of the "extremality of the free Gibbs measure" on the tree. (We will not attempt to define this notion here. See the papers [26, 10] for details.). In particular, in [10] it is shown that the reconstruction problem for the binary symmetric channel on the binary tree is solvable when $2\theta^2 > 1$. This in fact follows from a previous work [13] which implies that for any Markov chain M, the reconstruction problem on the b-ary tree is solvable if $b\theta^2 > 1$ where θ is the second largest eigenvalue of M in absolute value.

Proving non-reconstructibility turned out to be harder. While coupling arguments easily yield non-reconstruction, these arguments are typically not tight. A natural way to try to prove non-reconstructibility is to analyze recursions 1) in terms of random variables each of whose values is the expectation of the chain at a vertex, given the state at the leaves of the subtree below it, 2) in terms of ratios of such probabilities, or 3) in terms of log-likelihood ratios of such probabilities. Such recursions were analyzed for a closely related model in [3]. Both the reconstruction model and the model analyzed in [3] deal with the correlation between level n and the root. However, while in the reconstruction problem, the two random variables are generated according to the Markov model on the tree, in [3] the nodes at level n are set to have an i.i.d. distribution and the root has the conditional distribution thus induced.

In spite of this important difference, the two models are closely related. In particular, in [3] it is shown that for the binary tree, the correlation between level n and the root decays if and only if $2\theta^2 \le 1$. Building on the techniques of [3] it was finally shown in [2] that the reconstruction problem for the binary symmetric channel is solvable if and only if $2\theta^2 > 1$. This result was later reproven in various ways [5, 11, 1, 17].

The elegance of the threshold $b\theta^2 = 1$ raised the hope that it is the threshold for reconstruction for general channels. However, previous attempts to generalize any of the proofs to other channels have failed. Moreover in [21] it was shown that for asymmetric binary channels and for symmetric channels on large alphabets the reconstruction problem is solvable in cases where $b\theta^2 < 1$. In fact [21] contains an example of a channel satisfying $\theta = 0$ for which the reconstruction problem is solvable. On the other hand, in [23, 12] it is shown that the threshold $b\theta^2 = 1$ is the threshold for two variants of the reconstruction problem: "census reconstruction" and "robust reconstruction" (see [23, 12] for details).

The results above led some to believe that "reconstruction" unlike its siblings "census reconstruction" and "robust reconstruction" is an extremely sensitive property and that the threshold $b\theta^2 = 1$ is tight only for the binary symmetric channel. This conceptual picture was shaken by recent results in the theoretical physics literature [19] where using variational principles developed in the context of "replica symmetry breaking" it is suggested that the bound $b\theta^2 = 1$ is tight for symmetric channels on 3 and (maybe) 4 letters.

In Theorem 1 (below) we give the first tight threshold for the reconstruction problem for channels other than binary symmetric channels. We show that for asymmetric channels that are close to symmetric, the Kesten-Stigum bound $b\theta^2 = 1$ is tight for reconstruction. Our proof builds on ideas from [3, 2, 5, 25] and is extremely simple. In addition to giving a new result for the asymmetric channel, our proof also provides a very simple proof of the previously known result for the binary *symmetric* channel. See also [11, 5, 17] for other elegant proofs of the result in the symmetric case.

1.2 Definitions and Main Result

Let $T = (V, E, \rho)$ be a tree T with nodes V, edges E and root $\rho \in V$. We direct all edges away from the root, so that if $e = (x, y)$ then x is on the path connecting ρ to y. Let $d(\cdot, \cdot)$ denote the graph-metric distance on T, and $L_n = \{v \in V : d(\rho, v) = n\}$ be the n^{th} level of the tree. For $x \in V$ and $e = (y, z) \in E$, we denote $|x| = d(\rho, x)$, $d(x, (y, z)) = \max\{d(x, y), d(x, z)\}$, and $|e| = d(\rho, e)$. The *b-ary* tree is the infinite rooted tree where each vertex has exactly b children.

A Markov chain on the tree T is a probability measure defined on the state space $\mathcal{C}^V$, where $\mathcal{C}$ is a finite set. Assume first that T is finite and, for each edge e of T, let $M^e = (M^e_{i,j})_{i,j\in\mathcal{C}}$ be a stochastic matrix. In this case the probability measure defined by $(M^e : e \in E)$ on T is given by

$$\overline{\mu}_\ell(\sigma) = \mathbf{1}_{\{\sigma(\rho)=\ell\}} \prod_{(x,y)\in E} M^{(x,y)}_{\sigma(x),\sigma(y)}. \tag{2}$$

In other words, the root state $\sigma(\rho)$ satisfies $\sigma(\rho) = \ell$ and then each vertex iteratively chooses its state from the one of its parent by an application of the Markov transition rule given by M^e (and all such applications are independent). We can define the measure $\overline{\mu}_\ell$ on an infinite tree as well, by Kolmogorov's extension theorem, but we will not need chains on infinite trees in this paper (see [9] for basic properties of Markov chains on trees).

Instead, for an infinite tree T, we let $T_n = (V_n, E_n, \rho)$, where $V_n = \{x \in V : d(x, \rho) \le n\}$, $E_n = \{e \in E : d(e, \rho) \le n\}$ and define $\overline{\mu}^n_\ell$ by (2) for T_n. We are particularly interested in the distribution of the states $\sigma(x)$ for $x \in L_n$, the set of leaves in T_n. This distribution, denoted by μ^n_k, is the projection of $\overline{\mu}^n_k$ on $\mathcal{C}^{L_n}$ given by

$$\mu^n_k(\sigma) = \sum_{\bar{\sigma}:\bar{\sigma}|L_n = \sigma} \overline{\mu}^n_k(\overline{\sigma}) \tag{3}$$

Recall that for distributions μ and ν on the same space Ω the total variation distance between μ and ν is

$$D_V(\mu, \nu) = \frac{1}{2}\sum_{\sigma \in \Omega} |\mu(\sigma) - \nu(\sigma)|. \tag{4}$$

Definition 1 (Reconstructibility) *The reconstruction problem for the infinite tree $\mathcal{T}$ and $(M^e : e \in E)$ is* **solvable** *if there exist $i, j \in \mathcal{C}$ for which*

$$\lim_{n\to\infty} D_V(\mu^n_i, \mu^n_j) > 0. \tag{5}$$

When $M^e = M$ for all e, we say that the reconstruction problem is solvable for T and M.

We will be mostly interested in binary channels, i.e., transition matrices on the state space $\{\pm\}$. In this case, the definition above says that the reconstruction problem is solvable if

$$\lim_{n\to\infty} D_V(\mu^n_+, \mu^n_-) > 0. \tag{6}$$

Our main result is the following:

Theorem 1 (Main Result) *For all $b \ge 2$, there exists a $\delta_0 > 0$ such that for all $|\delta| \le \delta_0$, the reconstruction problem for M on the b-ary tree T_b is not solvable if $b\theta^2 \le 1$.*

See Theorem 2 below for a more general result.

1.3 Proof Idea

Our proof borrows ideas from several previous works, notably [3, 2, 5, 25]. In this section, we give a brief, high-level description of the proof.

First, consider the symmetric channel on the complete binary tree. The fundamental quantity in the proof is the so-called *magnetization of the root* [3] defined as

$$\begin{aligned} X_n \;=\; & \mathbb{P}[\text{root is } + \mid \text{state at level } n \text{ is } \sigma_n] \\ & - \mathbb{P}[\text{root is } - \mid \text{state at level } n \text{ is } \sigma_n]. \end{aligned}$$

Imagine that σ_n is drawn according to the Markov chain on the tree. Then, X_n—as a function of σ_n—is a random variable. Let $\bar{x}_n$ be the second moment of X_n. It is easy to show [5] (see Lemma 3 below) that non-solvability is implied by

$$\limsup_{n\to+\infty} \bar{x}_n = 0.$$

We prove the latter by induction. The proof has two main steps.

1. **Distributional Recursion.** Let X'_{n-1} and X''_{n-1} be two independent copies of X_{n-1}. We show that

$$X_n \overset{d}{=} \frac{\theta(X'_{n-1} + X''_{n-1})}{1 + \theta^2 X'_{n-1} X''_{n-1}}, \tag{7}$$

   where $\overset{d}{=}$ indicates equality in distribution. This follows from the Markov property. See Lemmas 4 and 5.

2. **Moment Recursion.** Expanding (7) and taking expectations, we show that

$$\bar{x}_n \le 2\theta^2 \bar{x}_{n-1}, \tag{8}$$

   for all n. The result follows. See Theorem 3.

In the case of asymmetric channels, we use a weighted version of the magnetization (see (9) below). Correction terms now appear in recursions (7) and (8) which somewhat complicate the analysis. The extra terms can be controlled when δ is small by continuity type arguments (see Proposition 1). Moreover, we prove our result for general trees (rather than complete b-ary trees). This more general result is proved by decomposing (7) and (8) into simple tree operations (see Section 3).

2 Preliminaries and General Result

For convenience, we sometimes write the channel

$$M = \begin{pmatrix} 1-\varepsilon^+ & \varepsilon^+ \\ 1-\varepsilon^- & \varepsilon^- \end{pmatrix}.$$

Note first that the stationary distribution $\pi = (\pi_+, \pi_-)$ of M is given by

$$\pi_+ = \frac{1-\varepsilon^-}{1-\theta} = \frac{1}{2} - \frac{\delta}{2(1-\theta)},$$

and

$$\pi_- = \frac{\varepsilon^+}{1-\theta} = \frac{1}{2} + \frac{\delta}{2(1-\theta)}.$$

In particular, this expression implies that the stationary distribution depends only on the ratio $\delta/(1-\theta)$. Or put differently, each two of the parameters π_+, δ and θ determine the third one uniquely. Note also that

$$\theta = \varepsilon^- - \varepsilon^+, \qquad \pi_- - \pi_+ = \frac{\delta}{1-\theta}.$$

Without loss of generality, we assume throughout that $\pi_- \geq \pi_+$ or equivalently that $\delta \geq 0$. (Note that δ can be made negative by inverting the role of $+$ and $-$.) Below, we will use the notation

$$\pi_{-/+} \equiv \pi_- \pi_+^{-1}, \qquad \Delta \equiv \pi_{-/+} - 1.$$

2.1 General Trees

In this section, we state our Theorem in a more general setting. Namely, we consider general rooted trees where different edges are equipped with different transition matrices—all having the same stationary distribution $\pi = (\pi_+, \pi_-)$. In other words, we consider a general infinite rooted tree $\mathcal{T} = (V, E)$ equipped with a function $\theta : E \to [-1, 1]$ such that the edge e of the tree is equipped with the matrix M^e with $\theta(M^e) = \theta(e)$ and the stationary distribution of M^e is (π_+, π_-).

In this general setting the notion of degree is extended to the notion of *branching number*. In [8], Furstenberg introduced the Hausdorff dimension of a tree. Later, Lyons [14, 15] showed that many probabilistic properties of the tree are determined by this number which he named the branching number. For our purposes it is best to define the branching number via cutsets.

Definition 2 (Cutsets) *A* cutset *S for a tree $\mathcal{T}$ rooted at ρ, is a finite set of vertices separating ρ from ∞. In other words, a finite set S is a cutset if every infinite self avoiding path from ρ intersects S. An* antichain or minimal cutset *is a cutset that does not have any proper subset which is also a cutset.*

Definition 3 (Branching Number) *Consider a rooted tree $\mathcal{T} = (V, E, \rho)$ equipped with an edge function $\theta : E \to [-1, 1]$. For each vertex $v \in V$ we define*

$$\eta(x) = \prod_{e \in \mathtt{path}(\rho, x)} \theta^2(e),$$

where $\mathtt{path}(\rho, x)$ is the set of edges on the unique path between ρ and x in $\mathcal{T}$. The branching number $\mathrm{br}(\mathcal{T}, \theta)$ of $(\mathcal{T}, \theta)$ is defined as

$$\mathrm{br}(\mathcal{T}, \theta) = \inf \left\{ \lambda > 0 : \inf_{\text{cutsets } S} \sum_{x \in S} \eta(x) \lambda^{-|x|} = 0 \right\}.$$

In our main result we show

Theorem 2 (Reconstructibility on General Trees) *Let $0 \leq \theta_0 < 1$. Then there exists $\delta_0 > 0$ such that, for all distributions $\pi = (\pi_+, \pi_-)$ with $\max\{|\delta(\pi, \theta_0)|, |\delta(\pi, -\theta_0)|\} < \delta_0$ and for all trees $(\mathcal{T}, \theta)$ with $\sup_e |\theta(e)| \leq \theta_0$ and $\mathrm{br}(\mathcal{T}, \theta) \leq 1$, the reconstruction problem is not solvable.*

It is easy to see that the conditions of Theorem 2 hold for T_b if $\theta(e) = \theta$ for all e and $b\theta^2 \leq 1$.

2.2 Magnetization

Let T be a finite tree rooted at x with edge function θ. Let σ be the leaf states generated by the Markov chain on (T, θ) with stationary distribution (π_+, π_-). We denote by $\mathbb{P}_T^+, \mathbb{E}_T^+$ (resp. $\mathbb{P}_T^-, \mathbb{E}_T^-$, and $\mathbb{P}_T, \mathbb{E}_T$) the probability/expectation operators with respect to the measure on the leaves of T obtained by conditioning the root to be $+$ (resp. $-$, and stationary). With a slight abuse of notation, we also write $\mathbb{P}_T[+ \mid \sigma]$ for the probability that the state at the the root of T is $+$ given state σ at the leaves. The main random variable we consider is the *weighted magnetization of the root*

$$X = \pi_-^{-1} \left[\pi_- \mathbb{P}_T[+ \mid \sigma] - \pi_+ \mathbb{P}_T[- \mid \sigma] \right]. \tag{9}$$

Note that the weights are chosen to guarantee

$$\mathbb{E}_T[X] = \pi_-^{-1} \left[\pi_- \pi_+ - \pi_+ \pi_- \right] = 0,$$

while the factor π_-^{-1} is such that $|X| \leq 1$ with probability 1.

Note that for any random variable depending only on the leaf states, $f = f(\sigma)$, we have

$$\pi_+ \mathbb{E}_T^+[f] + \pi_- \mathbb{E}_T^-[f] = \mathbb{E}_T[f],$$

so that in particular

$$\pi_+ \mathbb{E}_T^+[X] + \pi_- \mathbb{E}_T^-[X] = \mathbb{E}_T[X] = 0,$$

and

$$\pi_+ \mathbb{E}_T^+[X^2] + \pi_- \mathbb{E}_T^-[X^2] = \mathbb{E}_T[X^2].$$

We define the following analogues of the Edwards-Anderson order parameter for spin glasses on trees rooted at x

$$\bar{x} = \mathbb{E}_T[X^2], \qquad \bar{x}_+ = \mathbb{E}_T^+[X^2], \qquad \bar{x}_- = \mathbb{E}_T^-[X^2].$$

Now suppose $\mathcal{T}$ is an infinite tree rooted at ρ with edge function θ. Let $T_n = (V_n, E_n, x_n)$, where $V_n = \{u \in V : d(u, \rho) \leq n\}$, $E_n = \{e \in E : d(e, \rho) \leq n\}$,

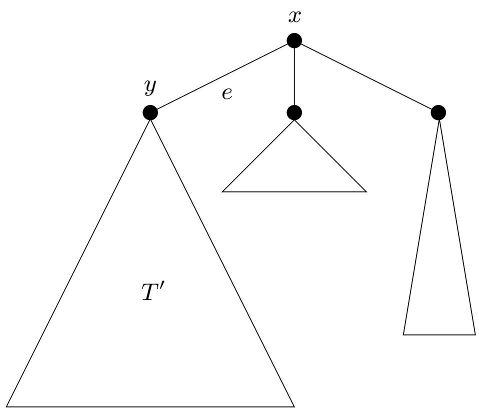

Figure 1. A finite tree T.

and x_n is identified with ρ. It is not hard to see that non-reconstructibility on (T, θ) is equivalent in our notation to

$$\limsup_{n\to\infty} \bar{x}_n = 0.$$

See Lemma 3 below. (Note that the total variation distance is monotone in the cutsets. Therefore the limit goes to 0 with the levels if and only if there exists a sequence of cutsets for which it goes to 0.)

2.3 Expectations

Fix a stationary distribution $\pi = (\pi_+, \pi_-)$. Let $T = (V, E)$ be a finite tree rooted at x with edge function $\{\theta(f), f \in E\}$ and weighted magnetization at the root X. Let y be a child of x and T' be the subtree of T rooted at y. Let Y be the weighted magnetization at the root of T'. See Figure 1. Denote by σ the leaf states of T and let σ' be the restriction of σ to the leaves of T'. Assume the channel on $e = (x, y)$ is given by

$$\begin{aligned} M^e &= \begin{pmatrix} 1-\varepsilon^+ & \varepsilon^+ \\ 1-\varepsilon^- & \varepsilon^- \end{pmatrix} \\ &= \frac{1}{2}\left[\begin{pmatrix} 1+\theta & 1-\theta \\ 1-\theta & 1+\theta \end{pmatrix} + \delta\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}\right]. \end{aligned}$$

We collect in the next lemmas a number of useful identities.

Lemma 1 (Radon-Nikodym Derivative) *The following hold:*

$$\frac{\mathrm{d}\mathbb{P}_T^+}{\mathrm{d}\mathbb{P}_T} = 1 + \pi_{-/+}X, \qquad \frac{\mathrm{d}\mathbb{P}_T^-}{\mathrm{d}\mathbb{P}_T} = 1 - X,$$

$$\mathbb{E}_T^+[X] = \pi_{-/+}\mathbb{E}_T[X^2], \qquad \mathbb{E}_T^-[X] = -\mathbb{E}_T[X^2].$$

Proof: Note that

$$\begin{aligned} X &= \pi_-^{-1}\left[\pi_-\mathbb{P}_T[+\,|\,\sigma] - \pi_+\mathbb{P}_T[-\,|\,\sigma]\right] \\ &= \pi_-^{-1}\left[\mathbb{P}_T[+\,|\,\sigma] - \pi_+\right] \\ &= \pi_{-/+}^{-1}\left[\frac{\mathbb{P}_T[+\,|\,\sigma]}{\pi_+} - 1\right], \end{aligned}$$

so that

$$\frac{\mathrm{d}\mathbb{P}_T^+}{\mathrm{d}\mathbb{P}_T} = \frac{\mathbb{P}_T[+\,|\,\sigma]}{\pi_+} = 1 + \pi_{-/+}X.$$

Likewise,

$$\frac{\mathrm{d}\mathbb{P}_T^-}{\mathrm{d}\mathbb{P}_T} = \frac{\mathbb{P}_T[-\,|\,\sigma]}{\pi_-} = 1 - X.$$

Then, it follows that

$$\mathbb{E}_T^+[X] = \mathbb{E}_T\left[X\left(1 + \pi_{-/+}X\right)\right] = \pi_{-/+}\mathbb{E}_T[X^2],$$

and similarly for $\mathbb{E}_T^-[X]$. ■

Lemma 2 (Child Magnetization) *We have,*

$$\mathbb{E}_T^+[Y] = \theta\mathbb{E}_{T'}^+[Y], \qquad \mathbb{E}_T^-[Y] = \theta\mathbb{E}_{T'}^-[Y].$$

Also,

$$\mathbb{E}_T^+[Y^2] = (1-\theta)\mathbb{E}_{T'}[Y^2] + \theta\mathbb{E}_{T'}^+[Y^2],$$

and

$$\mathbb{E}_T^-[Y^2] = (1-\theta)\mathbb{E}_{T'}[Y^2] + \theta\mathbb{E}_{T'}^-[Y^2].$$

Proof: By the Markov property, we have

$$\begin{aligned} \mathbb{E}_T^+[Y] &= (1-\varepsilon^+)\mathbb{E}_{T'}^+[Y] + \varepsilon^+\mathbb{E}_{T'}^-[Y] \\ &= \left[(1-\varepsilon^+) - \varepsilon^+\frac{\pi_+}{\pi_-}\right]\mathbb{E}_{T'}^+[Y] \\ &= \left[(1-\varepsilon^+) - (1-\varepsilon^-)\right]\mathbb{E}_{T'}^+[Y] \\ &= \theta\mathbb{E}_{T'}^+[Y], \end{aligned}$$

and similarly for $\mathbb{E}_T^-[Y]$.

Also,

$$\begin{aligned} \mathbb{E}_T^+[Y^2] &= (1-\varepsilon^+)\mathbb{E}_{T'}^+[Y^2] + \varepsilon^+\mathbb{E}_{T'}^-[Y^2] \\ &= (1-\varepsilon^+)\mathbb{E}_{T'}^+[Y^2] \\ &\quad + \frac{\varepsilon^+}{\pi_-}(\mathbb{E}_{T'}[Y^2] - \pi_+\mathbb{E}_{T'}^+[Y^2]) \\ &= \theta\mathbb{E}_{T'}^+[Y^2] + (1-\theta)\mathbb{E}_{T'}[Y^2], \end{aligned}$$

where we have used the calculation above. A similar expression holds for $\mathbb{E}_T^-[Y^2]$. ■

2.4 Information-Theoretic Lemma

The following easy lemma implies that, to establish non-solvability, it suffices to show that the second moment of the magnetization goes to 0. See e.g. [5]. We give a proof for completeness.

Lemma 3 *Let T be a finite tree with X the weighted magnetization at the root. Then, it holds that*

$$D_V(\mathbb{P}^+_T, \mathbb{P}^-_T) \leq \frac{1}{2\pi_+}\sqrt{\mathbb{E}_T[X^2]}.$$

Proof: Let L be the leaves of T. Bayes' rule and Cauchy-Schwarz give immediately,

$$\begin{aligned} D_V(\mathbb{P}^+_T, \mathbb{P}^-_T) \\ &= \frac{1}{2}\sum_{\sigma\in\{\pm\}^L}\left|\mathbb{P}^+_T[\sigma] - \mathbb{P}^-_T[\sigma]\right| \\ &= \frac{1}{2}\sum_{\sigma\in\{\pm\}^L}\mathbb{P}_T[\sigma]\left|\frac{\mathbb{P}_T[+\,|\,\sigma]}{\pi_+} - \frac{\mathbb{P}_T[-\,|\,\sigma]}{\pi_-}\right| \\ &= \frac{1}{2\pi_+}\mathbb{E}_T|X| \\ &\leq \frac{1}{2\pi_+}\sqrt{\mathbb{E}_T[X^2]}. \end{aligned}$$

■

3 Tree Operations

To derive moment recursions, the basic graph operation we perform is the following *Add-Merge* operation. Fix a stationary distribution $\pi = (\pi_+, \pi_-)$. Let T' (resp. T'') be a finite tree rooted at y (resp. z) with edge function θ' (resp. θ''), leaf state σ' (resp. σ''), and weighted magnetization at the root Y (resp. Z). Now add an edge $e = (\hat{y}, z)$ with edge value $\theta(e) = \theta$ to T'' to obtain a new tree $\widehat{T}$. Then merge $\widehat{T}$ with T' by identifying $y = \hat{y}$ to obtain a new tree T. To avoid ambiguities, we denote by x the root of T and X the magnetization of the root of T (where we identify the edge function on T with those on T', T'', and e). We let $\sigma = (\sigma', \sigma'')$ be the leaf state of T. See Figure 2. Let also $\widehat{Y}$ be the magnetization of the root on $\widehat{T}$. Assume

$$M^e = \begin{pmatrix} 1-\varepsilon^+ & \varepsilon^+ \\ 1-\varepsilon^- & \varepsilon^- \end{pmatrix}.$$

We first analyze the effect of adding an edge and merging subtrees on the magnetization variable.

Lemma 4 (Adding an Edge) *With the notation above, we have*

$$\widehat{Y} = \theta Z.$$

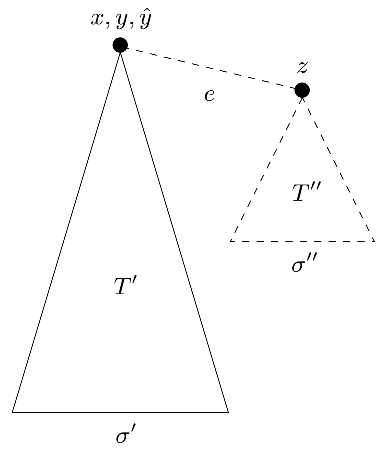

Figure 2. Tree T after the *Add-Merge* of T' and T''. The dashed subtree is $\widehat{T}$.

Proof: Denote

$$\mathcal{F}_\gamma = (1-\varepsilon^\gamma)\frac{\mathbb{P}_{T''}[\sigma''\,|\,+]}{\mathbb{P}_{T''}[\sigma'']} + \varepsilon^\gamma\frac{\mathbb{P}_{T''}[\sigma'\,|\,-]}{\mathbb{P}_{T''}[\sigma'']},$$

for $\gamma = +, -$. By Bayes' rule, the Markov property, and Lemma 1,

$$\begin{aligned} \widehat{Y} &= \pi_+\sum_{\gamma=+,-}\gamma\,\frac{\mathbb{P}_{\widehat{T}}[\gamma\,|\,\sigma'']}{\pi^\gamma} \\ &= \pi_+\sum_{\gamma=+,-}\gamma\,\frac{\mathbb{P}_{\widehat{T}}[\sigma''\,|\,\gamma]}{\mathbb{P}_{\widehat{T}}[\sigma'']} \\ &= \pi_+\sum_{\gamma=+,-}\gamma\,\mathcal{F}_\gamma, \end{aligned}$$

where we have used $\mathbb{P}_{\widehat{T}}[\sigma''] = \mathbb{P}_{T''}[\sigma'']$. We now simplify the expression for $\mathcal{F}_\gamma$. We have

$$\begin{aligned} \mathcal{F}_\gamma &= (1-\varepsilon^\gamma)\left(1+\pi_{-/+}Z\right) + \varepsilon^\gamma(1-Z) \\ &= 1 + \pi_- Z\left[\frac{1-\varepsilon^\gamma}{\pi_+} - \frac{\varepsilon^\gamma}{\pi_-}\right]. \end{aligned}$$

For $\gamma = +$, we get

$$\begin{aligned} \frac{1-\varepsilon^+}{\pi_+} - \frac{\varepsilon^+}{\pi_-} &= (1-\theta)\left[\frac{1-\varepsilon^+}{1-\varepsilon^-} - 1\right] \\ &= (1-\theta)\left[\frac{\varepsilon^- - \varepsilon^+}{1-\varepsilon^-}\right] \\ &= \frac{\theta}{\pi_+}. \end{aligned}$$

A similar calculation for the $-$ case gives for $\gamma = +, -$

$$\mathcal{F}_\gamma = 1 + \gamma\theta\pi_-\pi_\gamma^{-1}Z.$$

Plugging above gives $\widehat{Y} = \theta Z$. ■

Lemma 5 (Merging Subtrees) *With the notation above, we have*

$$X = \frac{Y + \widehat{Y} + \Delta Y\widehat{Y}}{1 + \pi_{-/+}Y\widehat{Y}}.$$

The same expression holds for a general $\widehat{T}$.

Proof: Denote

$$\mathcal{G}_\gamma = 1 + \gamma\pi_-\pi_\gamma^{-1}(Y + \widehat{Y}) + (\pi_-\pi_\gamma^{-1})^2 Y\widehat{Y}.$$

By Bayes' rule, the Markov property, and Lemma 1, we have

$$\begin{aligned} X &= \pi_+ \sum_{\gamma=+,-} \gamma \frac{\mathbb{P}_T[\gamma \mid \sigma]}{\pi^\gamma} \\ &= \pi_+ \sum_{\gamma=+,-} \gamma \frac{\mathbb{P}_T[\sigma \mid \gamma]}{\mathbb{P}_T[\sigma]} \\ &= \pi_+ \frac{\mathbb{P}_{T'}[\sigma']\mathbb{P}_{\widehat{T}}[\sigma'']}{\mathbb{P}_T[\sigma]} \sum_{\gamma=+,-} \gamma \frac{\mathbb{P}_{T'}[\sigma' \mid \gamma]}{\mathbb{P}_{T'}[\sigma']} \frac{\mathbb{P}_{\widehat{T}}[\sigma'' \mid \gamma]}{\mathbb{P}_{\widehat{T}}[\sigma'']} \\ &= \pi_+ \frac{\mathbb{P}_{T'}[\sigma']\mathbb{P}_{\widehat{T}}[\sigma'']}{\mathbb{P}_T[\sigma]} \sum_{\gamma=+,-} \gamma\, \mathcal{G}_\gamma. \end{aligned}$$

Similarly, we have

$$\begin{aligned} &\frac{\mathbb{P}_T[\sigma]}{\mathbb{P}_{T'}[\sigma']\mathbb{P}_{\widehat{T}}[\sigma'']} \\ &\qquad = \frac{1}{\mathbb{P}_{T'}[\sigma']\mathbb{P}_{\widehat{T}}[\sigma'']} \sum_{\gamma=+,-} \pi^\gamma\, \mathbb{P}_T[\sigma \mid \gamma] \\ &\qquad = \sum_{\gamma=+,-} \pi_\gamma\, \mathcal{G}_\gamma. \end{aligned}$$

Note that

$$\sum_{\gamma=+,-} \gamma\, \mathcal{G}_\gamma = \pi_+^{-1}(Y + \widehat{Y}) + \pi_+^{-2}(\pi_- - \pi_+)Y\widehat{Y},$$

where we have used

$$\pi_-^2 - \pi_+^2 = (\pi_- - \pi_+)(\pi_- + \pi_+) = \pi_- - \pi_+.$$

Similarly,

$$\sum_{\gamma=+,-} \pi_\gamma\, \mathcal{G}_\gamma = 1 + \pi_-\pi_+^{-1}Y\widehat{Y}.$$

The result follows. ■

4 Symmetric Channels On Regular Trees

As a warm-up, we start by analyzing the binary symmetric channel on the infinite b-ary tree. Our proof is arguably the simplest proof to date of this result. The same proof structure will be used in the general case. The following theorem is due to [2, 5, 11, 23, 12, 1, 16].

Theorem 3 (Symmetric Channel) *Let M be a transition matrix with $\delta = 0$ and $b\theta^2 \leq 1$. Let $\mathcal{T}$ be the infinite b-ary tree. Then, the reconstruction problem on $(\mathcal{T}, M)$ is not solvable.*

Proof: Consider again the setup of Section 3. Note first that, by Lemma 4, we have $\widehat{Y} = \theta Z$ and therefore

$$\mathbb{E}_{\widehat{T}}[\widehat{Y}^2] = \theta^2 \mathbb{E}_{T''}[Z^2]. \tag{10}$$

In other words, adding an edge to the root of a tree and re-rooting at the new vertex has the effect of multiplying the second moment of the magnetization by θ^2. Now consider the *Add-Merge* operation defined in Section 3. Using the expansion

$$\frac{1}{1+r} = 1 - r + \frac{r^2}{1+r}, \tag{11}$$

the inequality $|X| \leq 1$, and Lemma 5, we get

$$\begin{aligned} X &= Y + \widehat{Y} - Y\widehat{Y}(Y + \widehat{Y}) + Y^2\widehat{Y}^2 X \\ &\leq Y + \widehat{Y} - Y\widehat{Y}(Y + \widehat{Y}) + Y^2\widehat{Y}^2. \end{aligned} \tag{12}$$

Note that from Lemmas 1 and 2, we have

$$\begin{gathered} \mathbb{E}_T^+[X] = \bar{x}, \\ \mathbb{E}_T^+[Y] = \mathbb{E}_T^+[Y^2] = \bar{y}, \quad \mathbb{E}_T^+[\widehat{Y}] = \mathbb{E}_T^+[\widehat{Y}^2] = \theta^2\bar{z}, \end{gathered}$$

where we have used that $\bar{y}_+ = \bar{y}_- = \bar{y}$ and $\bar{z}_+ = \bar{z}_- = \bar{z}$ by symmetry. Taking $\mathbb{E}_T^+$ on both sides of (12), we get

$$\begin{aligned} \bar{x} &\leq \bar{y} + \theta^2\bar{z} - \theta^2\bar{y}\bar{z} - \theta^2\bar{y}\bar{z} + \theta^2\bar{y}\bar{z} \\ &= \bar{y} + \theta^2\bar{z} - \theta^2\bar{y}\bar{z}. \end{aligned}$$

Now, let $T_n = (V_n, E_n, x_n)$ be as in Section 2.2. Repeating the *Add-Merge* operation $(b-1)$ times, we finally have by induction

$$\bar{x}_n \leq b\theta^2\bar{x}_{n-1} - (b-1)\theta^4\bar{x}_{n-1}^2. \tag{13}$$

Indeed, note that for $0 < a < b$,

$$\begin{aligned} &(a\theta^2\bar{x}_{n-1} - (a-1)\theta^4\bar{x}_{n-1}^2) + \theta^2\bar{x}_{n-1} \\ &-\theta^2(a\theta^2\bar{x}_{n-1} - (a-1)\theta^4\bar{x}_{n-1}^2)\bar{x}_{n-1} \\ &\qquad \leq (a+1)\theta^2\bar{x}_{n-1} - a\theta^4\bar{x}_{n-1}^2, \end{aligned}$$

and the first step of the induction is given by (10). This concludes the proof. ■

Remark 1 *Note that equation (13) implies that if $b\theta^2 < 1$ then $\bar{x}_n \leq \exp(-\Omega(n))$, while if $b\theta^2 = 1$ then $\bar{x}_n \leq O(1/n)$.*

5 Roughly Symmetric Channels on General Trees

We now tackle the general case. We start by analyzing the *Add-Merge* operation.

Proposition 1 (Basic Inequality) *Consider the setup of Section 3. Assume $|\theta| < 1$. Then, there is a $\delta_0(|\theta|) > 0$ depending only on $|\theta|$ such that*

$$\bar{x} \leq \bar{y} + \theta^2 \bar{z},$$

whenever δ (on e) is less than $\delta_0(|\theta|)$.

Proof: The proof is similar to that in the symmetric case. By expansion (11), inequality $|X| \leq 1$, and Lemma 5, we have

$$\begin{aligned} X \quad &\leq \quad Y + \widehat{Y} + \Delta Y \widehat{Y} \\ &\quad -\pi_{-/+} Y \widehat{Y}(Y + \widehat{Y} + \Delta Y \widehat{Y}) + \pi_{-/+}^2 Y^2 \widehat{Y}^2. \end{aligned} \tag{14}$$

Let $\rho' = (\bar{y})^{-1}\bar{y}_+$ and $\rho'' = (\bar{z})^{-1}\bar{z}_+$. Then, by Lemmas 1 and 2, we have

$$\mathbb{E}_T^+[X] = \pi_{-/+}\bar{x},$$

$$\mathbb{E}_T^+[Y] = \pi_{-/+}\bar{y}, \quad \mathbb{E}_T^+[Y^2] = \bar{y}\rho',$$

and

$$\begin{aligned} &\mathbb{E}_T^+[\widehat{Y}] = \pi_{-/+}\theta^2\bar{z}, \\ &\mathbb{E}_T^+[\widehat{Y}^2] = \theta^2\bar{z}[(1-\theta) + \theta\rho'']. \end{aligned}$$

Taking $\pi_{-/+}^{-1}\mathbb{E}_T^+$ on both sides of (14), we get

$$\begin{aligned} \bar{x} \quad &\leq \quad \bar{y} + \theta^2\bar{z} + \Delta\pi_{-/+}\theta^2\bar{y}\bar{z} \\ &\quad -\pi_{-/+}\theta^2\bar{y}\bar{z}\rho' - \pi_{-/+}\theta^2\bar{y}\bar{z}[(1-\theta) + \theta\rho''] \\ &\quad -\Delta\theta^2\bar{y}\bar{z}\rho'[(1-\theta) + \theta\rho''] \\ &\quad +\pi_{-/+}\theta^2\bar{y}\bar{z}\rho'[(1-\theta) + \theta\rho''] \\ &\leq \quad \bar{y} + \theta^2\bar{z} - \pi_{-/+}\theta^2\bar{y}\bar{z}[\mathcal{A} - \Delta\mathcal{B}], \end{aligned}$$

where

$$\mathcal{A} = \rho' + (1-\rho')[(1-\theta) + \theta\rho''],$$

and

$$\mathcal{B} = 1 - \pi_{-/+}^{-1}\rho'[(1-\theta) + \theta\rho''].$$

Note that $[(1-\theta)+\theta\rho''] \geq 0$ by Lemma 2. So $\mathcal{B} \leq 1$ and it suffices to have $\mathcal{A} \geq \Delta$. Note also that $\mathcal{A}$ is multilinear in (ρ', ρ''). Therefore, to minimize $\mathcal{A}$, we only need to consider extreme cases in (ρ', ρ''). By

$$\pi_+ y^+ + \pi_- y^- = y$$

it follows that $0 \leq \rho' \leq \pi_+^{-1}$. The same holds for ρ''. At $\rho' = 0$, we have

$$\mathcal{A} = 1 - \theta[1-\rho''] \geq \begin{cases} 1-\theta, & \text{if } \theta \geq 0, \\ 1 - \pi_{-/+}|\theta|, & \text{if } \theta \leq 0, \end{cases}$$

where we have used

$$1 - \pi_+^{-1} = -\pi_{-/+}.$$

At $\rho' = \pi_+^{-1}$, we have

$$\begin{aligned} \mathcal{A} \quad &= \quad \pi_+^{-1} + (1-\pi_+^{-1})[1 - \theta[1-\rho'']] \\ &= \quad 1 + \theta\pi_{-/+}[1-\rho''] \\ &\geq \quad \begin{cases} 1 - \pi_{-/+}^2\theta, & \text{if } \theta \geq 0, \\ 1 - \pi_{-/+}|\theta|, & \text{if } \theta \leq 0. \end{cases} \end{aligned}$$

Since $\pi_{-/+} \geq 1$ by assumption, it follows that

$$\mathcal{A} \geq 1 - \pi_{-/+}^2|\theta|.$$

At $\delta = 0$, this bound is strictly positive and moreover $\Delta = 0$. Therefore, by continuity in δ of Δ and the bound above, the result follows. ■

Proposition 2 (Induction Step) *Let T be a finite tree rooted at x with edge function θ. Let $w_1, \ldots, w_\alpha$ be the children of x in T and denote by e_a the edge connecting x to w_a. Let $\theta_0 = \max\{|\theta(e_1)|, \ldots, |\theta(e_\alpha)|\}$ and assume that on each edge e_a, $\delta \leq \delta_0(\theta_0)$, where δ_0 is defined in Proposition 1. Then*

$$\bar{x} \leq \sum_{a=1}^{\alpha} \theta(e_a)^2 \bar{w}_a.$$

Proof: As noted in the proof of Theorem 3, adding an edge e to the root of a tree and re-rooting at the new vertex has the effect of multiplying the second moment of the magnetization by $\theta^2(e)$. The result follows by applying Proposition 1 $(\alpha - 1)$ times. ■

Proof of Theorem 2: It suffices to show that for all $\varepsilon > 0$ there is an N large enough so that $\bar{x}_n \leq \varepsilon$, $\forall n \geq N$. Fix $\varepsilon > 0$. By definition of the branching number, there exists a cutset S of $\mathcal{T}$ such that

$$\sum_{u \in S} \eta(u) \leq \varepsilon.$$

Assume w.l.o.g. that S is actually an antichain and let N be such that S is in T_N. It is enough to show that

$$\bar{x}_n \leq \sum_{u \in S} \eta(u), \qquad \forall\, n \geq N.$$

Fix $n \geq N$. Applying Proposition 2 repeatedly from the root of T_n down to S, it is clear that

$$\bar{x}_n \leq \sum_{u \in S} \eta(u)\mathbb{E}_{T_n(u)}[U^2] \leq \sum_{u \in S} \eta(u),$$

where $T_n(u)$ is the subtree of T_n rooted at u and U is the magnetization at u on $T_n(u)$ (with $|U| \leq 1$). This concludes the proof. ■

References

[1] N. Berger, C. Kenyon, E. Mossel, and Y. Peres. Glauber dynamics on trees and hyperbolic graphs. *Probab. Theory Related Fields*, 131(3):311–340, 2005. Extended abstract by Kenyon, Mossel and Peres appeared in proceedings of 42nd IEEE Symposium on Foundations of Computer Science (FOCS) 2001, 568–578.

[2] P. M. Bleher, J. Ruiz, and V. A. Zagrebnov. On the purity of the limiting Gibbs state for the Ising model on the Bethe lattice. *J. Statist. Phys.*, 79(1-2):473–482, 1995.

[3] J. T. Chayes, L. Chayes, J. P. Sethna, and D. J. Thouless. A mean field spin glass with short-range interactions. *Comm. Math. Phys.*, 106(1):41–89, 1986.

[4] C. Daskalakis, E. Mossel, and S. Roch. Optimal Phylogenetic Reconstruction. Availible on the Arxiv at math.PR/0509575, Extended abstract to appear at Proceedings of STOC 2006, 2006.

[5] W. S. Evans, C. Kenyon, Y. Y. Peres, and L. J. Schulman. Broadcasting on trees and the Ising model. *Ann. Appl. Probab.*, 10(2):410–433, 2000.

[6] W. S. Evans and L. J. Schulman. Signal propagation and noisy circuits. *IEEE Trans. Inform. Theory*, 45(7):2367–2373, 1999.

[7] J. Felsenstein. *Inferring Phylogenies*. Sinauer, New York, New York, 2004.

[8] H. Furstenberg. Intersections of Cantor sets and transversality of semigroups. In *Problems in analysis (Sympos. Salomon Bochner, Princeton Univ., Princeton, N.J., 1969)*, pages 41–59. Princeton Univ. Press, Princeton, N.J., 1970.

[9] H. O. Georgii. *Gibbs measures and phase transitions*, volume 9 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1988.

[10] Y. Higuchi. Remarks on the limiting Gibbs states on a $(d+1)$-tree. *Publ. Res. Inst. Math. Sci.*, 13(2):335–348, 1977.

[11] D. Ioffe. On the extremality of the disordered state for the Ising model on the Bethe lattice. *Lett. Math. Phys.*, 37(2):137–143, 1996.

[12] S. Janson and E. Mossel. Robust reconstruction on trees is determined by the second eigenvalue. *Ann. Probab.*, 32:2630–2649, 2004.

[13] H. Kesten and B. P. Stigum. Additional limit theorems for indecomposable multidimensional Galton-Watson processes. *Ann. Math. Statist.*, 37:1463–1481, 1966.

[14] R. Lyons. The Ising model and percolation on trees and tree-like graphs. *Comm. Math. Phys.*, 125(2):337–353, 1989.

[15] R. Lyons. Random walks and percolation on trees. *Ann. Probab.*, 18(3):931–958, 1990.

[16] F. Martinelli, A. A. Sinclair, and D. Weitz. Glauber dynamics on trees: boundary conditions and mixing time. *Comm. Math. Phys.*, 250(2):301–334, 2004.

[17] F. Martinelli, A. Sinclair, and D. Weitz. Fast mixing for independent sets, colorings, and other models on trees. In *Proceedings of the 15th ACM-SIAM Symposium on Discrete Algorithms*, pages 449–458, 2004.

[18] M. Mezard and R. Zecchina. Random k-satisfiability: from an analytic solution to an effici ent algorithm. *Phys. Rev. E*, 66, 2002.

[19] M. M. A. Montanari. Reconstruction on Trees and the Spin Glass Transition, year = ”2006”, note = Preprint,.

[20] E. Mossel. Recursive reconstruction on periodic trees. *Random Structures Algorithms*, 13(1):81–97, 1998.

[21] E. Mossel. Reconstruction on trees: beating the second eigenvalue. *Ann. Appl. Probab.*, 11(1):285–300, 2001.

[22] E. Mossel. Survey: Information flow on trees. In J. Nestril and P. Winkler, editors, *Graphs, Morphisms and Statistical Physics. DIMACS series in discrete mathematics and theoretical computer science*, pages 155–170. Amer. Math. Soc., 2004.

[23] E. Mossel and Y. Peres. Information flow on trees. *Ann. Appl. Probab.*, 13(3):817–844, 2003.

[24] M. M. G. Parisi and R. Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, 297, 812, 2002. (Scienceexpress published on-line 27-June-2002; 10.1126/science. 1073287).

[25] R. Pemantle and Y. Peres. The critical Ising model on trees, concave recursions and nonlinear capacity. Avialibe at: arXiv:math.PR/0503137.

[26] F. Spitzer. Markov random fields on an infinite tree. *Ann. Probability*, 3(3):387–398, 1975.

[27] J. von Neumann. Probabilistic logics and the synthesis of reliable organisms from unreliable components. In *Automata studies*, Annals of mathematics studies, no. 34, pages 43–98. Princeton University Press, Princeton, N. J., 1956.