Reconstruction Thresholds on Regular Trees

James B. Martin

Discrete Mathematics and Theoretical Computer Science **AC**, 2003, 191–204

Reconstruction Thresholds on Regular Trees

James B. Martin

CNRS, LIAFA, Univ. Paris 7, case 7014, 2 pl. Jussieu, 75251 Paris 05, France.
martin@liafa.jussieu.fr

We consider the model of *broadcasting on a tree*, with binary state space, on the infinite rooted tree $\mathbb{T}^k$ in which each node has k children. The root of the tree takes a random value 0 or 1, and then each node passes a value independently to each of its children according to a 2×2 transition matrix $\mathbf{P}$. We say that *reconstruction is possible* if the values at the dth level of the tree contain non-vanishing information about the value at the root as $d \to \infty$. Extending a method of Brightwell and Winkler, we obtain new conditions under which reconstruction is impossible, both in the general case and in the special case $p_{11} = 0$. The latter case is closely related to the *hard-core model* from statistical physics; a corollary of our results is that, for the hard-core model on the $(k+1)$-regular tree with activity $\lambda = 1$, the unique simple invariant Gibbs measure is extremal in the set of Gibbs measures, for any $k \geq 2$.

Keywords: broadcasting on a tree, reconstruction, hard-core model, Gibbs measure, extremality

1 Introduction

1.1 Broadcasting on a tree

We consider a model of a broadcasting on the rooted tree $\mathbb{T}^k$, in which every node has k children.

Let $\mathbf{P} = \{p_{ij}, i, j = 0, 1\}$ be a 2×2 stochastic matrix, which we regard as a transition matrix on the set $\{0, 1\}$. Each node $u \in \mathbb{T}^k$ will carry a value $\phi(u) \in \{0, 1\}$, generated as follows. The root takes value 0 with probability $\pi_0 = p_{10}/(p_{01} + p_{10})$ and value 1 with probability $\pi_1 = 1 - \pi_0$. Thereafter the configuration on $\mathbb{T}^k$ is generated recursively; if a node has value $i \in \{0, 1\}$, each of its k children takes the value 0 with probability p_{i0} and the value 1 with probability p_{i1}, all choices being made independently.

We write $\phi = \{\phi(u), u \in \mathbb{T}^k\}$ for a configuration on the whole tree, and denote by μ the probability measure on $\{0, 1\}^{\mathbb{T}^k}$ resulting from this broadcasting construction.

For a node $u \in \mathbb{T}^k$, let $\mathbb{T}^k(u)$ be the subtree consisting of u and all its descendants. By the choice of π_0, we have a translation invariance property for μ; namely that $\mu(\phi(u) = 0) = \pi_0$ for every $u \in \mathbb{T}^k$, and so for any $u, v \in \mathbb{T}^k$, the configurations on $\mathbb{T}^k(u)$ and $\mathbb{T}^k(v)$ have the same distribution, under a natural mapping between the subtrees $\mathbb{T}^k(u)$ and $\mathbb{T}^k(v)$.

We are interested in the following question of *reconstruction*: for $d \geq 1$, how much information about the value at node u is given by the values of the dth generation of its descendants?

Questions of this sort arise in several contexts – for example genetics, communication theory and statistical physics – and have been quite widely studied in the last few years; see Mossel [Mos03] for a survey, and [EKPS00, BRZ95, Iof96, KMP01, Mos01, MP03, BW03, JM03] for a variety of approaches to this sort of model (which can of course be considerably generalised from our particular setting of a binary state space and a regular tree).

The question above can be made precise in several (often equivalent) ways. We use the following formulation.

Let $\mathbb{W}_d(u)$ be the set of descendants of u at distance exactly d from u. For a set $S \subseteq \mathbb{T}^k$, write $\sigma(S)$ for the σ-algebra of events which depend only on the values $\{\phi(u), u \in S\}$.

Define the random variable

$$A(d,u) = \mu\big(\phi(u) = 0 \big| \sigma(\mathbb{W}_d(u))\big),$$

that is, the conditional probability that the value at u is 0, given only the information from the dth generation of its descendants.

From the independence structure given by the broadcasting construction, additional knowledge of any information from nodes beyond the dth generation does not change the conditional distribution of the value of u; that is,

$$A(d,u) = \mu\Big(\phi(u) = 0 \Big| \sigma\big(\bigcup_{d'=d}^{\infty} \mathbb{W}_{d'}(u)\big)\Big).$$

Of course, if $d_1 > d_2$, then

$$\sigma\big(\bigcup_{d'=d_1}^{\infty} \mathbb{W}_{d'}(u)\big) \subseteq \sigma\big(\bigcup_{d'=d_2}^{\infty} \mathbb{W}_{d'}(u)\big),$$

so by the backwards martingale convergence theorem (see e.g. Section 14.4 of [Wil91]), we have that $A(d,u) \to A(u)$ a.s. as $d \to \infty$, where

$$A(u) = \mu(\phi(u) = 0 \big| \mathcal{T}(u));$$

here $\mathcal{T}(u)$ is the *tail* σ-*algebra* of descendants of u, defined by

$$\mathcal{T}(u) = \bigcap_{d=1}^{\infty} \sigma\big(\bigcup_{d'=d}^{\infty} \mathbb{W}_{d'}(u)\big).$$

By the translation invariance property above, the random variable $A(u)$ has the same distribution for all $u \in \mathbb{T}^k$.

Definition: We say that *reconstruction is impossible* (for a given $\mathbf{P}$ and k) if the random variable $A(u)$ is almost surely constant, and otherwise that *reconstruction is possible*.

A complete answer to the question of when reconstruction is possible is currently only known for the case where $\mathbf{P}$ is symmetric. Then let $p_{00} = p_{11} = 1 - \varepsilon$; reconstruction is possible if and only if $k(1-2\varepsilon)^2 > 1$ (see for example [BRZ95, EKPS00, Iof96]).

In general, however, there are gaps between the best known necessary and sufficient conditions for reconstruction to be possible. In this paper we give new conditions on $\mathbf{P}$ under which we show that reconstruction is impossible.

In Proposition 4.1 of [MP03], Mossel and Peres show that reconstruction is impossible whenever

$$\frac{(p_{00} - p_{10})^2}{\min\{p_{00} + p_{10}, p_{01} + p_{11}\}} \leq \frac{1}{k}. \tag{1}$$

We improve the bound to give the following condition:

Theorem 1. *Reconstruction is impossible whenever*

$$\left(\sqrt{p_{00}p_{11}} - \sqrt{p_{01}p_{10}}\right)^2 \leq \frac{1}{k}. \tag{2}$$

A calculation (see Section 4) shows that the LHS of (2) is always less than or equal to that of (1), with equality in the following special cases: (i) $\mathbf{P}$ is symmetric; (ii) $p_{ij} = 0$ for some i, j; (iii) $p_{00} = p_{10}$, $p_{01} = p_{11}$. Note that for symmetric $\mathbf{P}$, (2) becomes the condition that $k(1-2\varepsilon)^2 \leq 1$, and our proof of Theorem 1 gives another proof that reconstruction is impossible under this condition.

We then focus on the special case where $p_{11} = 0$ (of course, the case $p_{00} = 0$ is analogous). This case is closely related to the *hard-core model* from statistical physics, and has been recently studied by Brightwell and Winkler [BW03] and Rozikov and Suhov [RS03]. Certain specific properties in this case allow a more sophisticated argument which gives a much better condition than is obtained by putting $p_{11} = 0$ in Theorem 1.

1.2 Hard-core model

In this section we state our result for the case $p_{11} = 0$ and explain the correspondence with the hard-core model on a regular tree.

Following [BW03], we parametrise $\mathbf{P}$ by the quantity $w > 0$, setting

$$\mathbf{P} = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} = \begin{pmatrix} \frac{1}{1+w} & \frac{w}{1+w} \\ 1 & 0 \end{pmatrix}, \tag{3}$$

or equivalently by the quantity $\lambda = w(1+w)^k > 0$, whose significance we explain later; note that the correspondence between $\lambda > 0$ and $w > 0$ is one-to-one and monotonic.

Let $\lambda_c = \lambda_c(k)$ be the infimum of the set of λ such that reconstruction is possible. If follows from Proposition 12 of [Mos01] that in fact reconstruction is possible for any $\lambda > \lambda_c$ (so that λ_c is also the supremum of the set of λ such that reconstruction is impossible).

Brightwell and Winkler [BW03] show that, as $k \to \infty$,

$$\frac{1+o(1)}{\ln k} \leq \lambda_c(k) \leq (\ln k)^2\big(1+o(1)\big). \tag{4}$$

We improve the lower bound to give the following:

Theorem 2. $\lambda_c(k) > e - 1$ *for all* k.

(For the equivalent threshold value w_c with $w_c(1-w_c)^k = \lambda_c$, one can deduce that $w_c(k) > (\ln k - \ln\ln k)/k$ for all k).

We will now describe the correspondence between the broadcasting model and the hard-core model, and explain (without proofs) the significance of Theorem 2 for the hard-core model on the $(k+1)$-regular tree. For more details on the correspondence between the two models, see also [BW03] and its references.

We denote the $(k+1)$-regular tree by $\tilde{\mathbb{T}}^k$. We can still regard $\tilde{\mathbb{T}}^k$ as a rooted tree, in which the root node has $k+1$ children and every other node has k children. We can then carry out the broadcasting construction on $\tilde{\mathbb{T}}^k$ in exactly the same way as we did on $\mathbb{T}^k$; now the root has $k+1$ rather than k children, but the values at these $k+1$ children are chosen i.i.d. according to the value at the root and the transition matrix $\mathbf{P}$ just as before. We will write $\tilde{\mu}$ for the probability measure on $\{0,1\}^{\tilde{\mathbb{T}}^k}$ resulting from this construction.

The independence structure of the random walk implies that the measure $\tilde{\mu}$ is *simple*, by which we mean that, for any u, the configurations

$$\{\phi(v), v \in C_1(u)\}, \ldots, \{\phi(v), v \in C_{k+1}(u)\}$$

are mutually independent given $\phi(u)$, where the C_i are the connected components of $\tilde{\mathbb{T}}^k \setminus \{u\}$. Although we have defined $\tilde{\mu}$ in an asymmetric way, it's also the case that it is *invariant*, in the sense that it is preserved by any automorphism of $\tilde{\mathbb{T}}^k$. In particular, the choice of the root is not important.

To introduce the hard-core model, we first consider the case of a finite graph with node-set S (and some neighbour relation).

We can identify a configuration $\phi \in \{0,1\}^S$ with the subset $I_\phi := \{u \in S : \phi(u) = 1\}$ of S.

A set $I \subseteq S$ is called an *independent set* if no two neighbours in the graph are both members of I.

The *hard-core measure* on S with *activity* $\lambda > 0$ is the probability measure ν on $\{0,1\}^S$ such that

$$\nu\big(I_\phi \text{ is an independent set }\big) = 1,$$

and such that for an independent set I_0, $\nu(I_\phi = I_0)$ is proportional to $\lambda^{|I_0|}$. Thus in fact

$$\nu(\phi = \phi_0) = Z_\lambda^{-1} \lambda^{|I_{\phi_0}|} 1\big(I_{\phi_0} \text{ is an independent set }\big),$$

where we have the normalising factor

$$Z_\lambda = \sum_{\phi_0 : I_{\phi_0} \text{ is independent}} \lambda^{|I_{\phi_0}|}.$$

When $\lambda = 1$, I_ϕ has the uniform distribution over the set of independent subsets of S.

An equivalent characterisation is that ν is the unique probability measure such that, for any $\phi_0 \in \{0,1\}^S$ and any $u \in S$,

$$\nu\Big(\phi(u) = 1 \Big| \phi(v) = \phi_0(v) \text{ for all } v \neq u\Big) = \frac{\lambda}{1+\lambda} 1\left\{I_{\phi_0} \cup \{u\} \text{ is independent}\right\} \tag{5}$$

The condition (5) makes sense equally when S is infinite, except that (since conditional probabilities are only well defined up to almost sure equality) we should now only demand the condition holds for ν-almost all ϕ_0. Putting $S = \tilde{\mathbb{T}}^k$, we say that a probability measure ν satisfying (5) (for all $u \in \tilde{\mathbb{T}}^k$ and ν-almost all ϕ_0) is a *Gibbs measure* for the hard-core model on $\tilde{\mathbb{T}}^k$ with activity λ.

It is quite straightforward to show that the measure $\tilde{\mu}$ defined above by the broadcasting construction with **P** as in (3) is a Gibbs measure for the hard-core model with activity λ. However, now that the state space is infinite, it's no longer the case that such a measure need be unique. In fact, there is a critical point $\lambda_c' = \lambda_c'(k) = k^k/(k-1)^{(k+1)}$ (identified by Kelly [Kel85]); for $\lambda \leq \lambda_c'$, $\tilde{\mu}$ is the *only* Gibbs measure, whereas for $\lambda > \lambda_c'$, there are others. Nevertheless, for any λ, the measure $\tilde{\mu}$ is the only *simple invariant* Gibbs measure; (this can be deduced, for example, from Theorem 4.1 of [Zac83] – see also Section 5 of that paper for relevant discussion).

The set of Gibbs measures forms a simplex; that is, any mixture of Gibbs measures is also a Gibbs measure, and in particular there is a set of *extremal* Gibbs measures such that every Gibbs measure is expressible in a unique way as a mixture of extremal measures. For $\lambda > \lambda_c'$, we can therefore ask whether the measure $\tilde{\mu}$ is extremal (equivalently, not expressible as a mixture of other Gibbs measures).

It turns out that $\tilde{\mu}$ is extremal at activity λ if and only if reconstruction is impossible for the corresponding broadcasting model on $\mathbb{T}^k$ with transition matrix $\mathbf{P}$. (This is a consequence of the general fact that a Gibbs measure is extremal iff it is trivial on the tail σ-algebra, and of the independence structure given by the broadcasting constructions of μ and $\tilde{\mu}$). Hence the reconstruction threshold λ_c defined after (3) is also the *extremality* threshold for $\tilde{\mu}$; Theorem 2 therefore shows that whenever $\lambda \leq e-1$, the unique simple invariant Gibbs measure $\tilde{\mu}$ for the hard-core model with activity λ is extreme, for any k.

In particular, $\tilde{\mu}$ is extreme in the special case $\lambda = 1$ for any k.

1.3 Outline of proof

Our proof of Theorems 1 and 2 is in the same spirit as the proof by Brightwell and Winkler of the lower bound in (4) for the hard-core model [BW03].

We first develop a coupling between the distributions of the random variable $A(u)$ conditioned on two different events, with certain additional properties beyond those used in [BW03]. We then use this coupling to establish a recursion linking the distribution of $A(u)$ to those of $A(u_1),\ldots,A(u_k)$, where $u_1,\ldots,u_k$ are the children of u. (Of course, we already know from the translation invariance property described in Section 1.1 that all of these distributions are the same). If the recursion relation is contractive in a suitable sense, we obtain that $A(u)$ must be a.s. constant.

In Section 2, we first prove a lemma on conditional probabilities in a more general setting. Specialising to our context, we obtain the existence of a coupling of a pair of random variables A_0, A_1 with the following properties:

(i) The distribution of A_0 is the distribution of $A(u)$ under μ conditioned on the event $\{\phi(u) = 0\}$;

(ii) The distribution of A_1 is the distribution of $A(u)$ under μ conditioned on the event $\{\phi(u) = 1\}$;

(iii) With probability 1, either $A_0 = A_1$ or $A_1 \leq \pi_0 \leq A_0$;

(iv) If $A_0 = A_1$ with probability 1, then both are equal to π_0 with probability 1, and so also $A(u) = \pi_0$ a.s. under μ.

We develop the recursion relations and complete the proofs in Section 3.

The full properties of the coupling are only needed in the hard-core case, where a particular convexity property holds for the recursion relations. The argument in the general case is not as powerful, and rather than all of property (iii) above, we use only that $A_1 \leq A_0$ with probability 1. Restricting the bound in Theorem 1 to the case $p_{11} = 0$ gives a much weaker bound than that in Theorem 2 (in fact, one obtains only the bound $\lambda_c \geq \lambda_c'$ where λ_c' is the threshold for the uniqueness of the Gibbs measure; this bound is obvious in the context of the hard-core model since if the Gibbs measure is unique it is trivially extreme).

2 Conditioned conditional probabilities

We first consider the setting of a general probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $B \in \mathcal{F}$ be an event with probability π_0=$1-\pi_1$, and suppose $0 < \pi_0 < 1$. Write B^C for the complement of B. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$.

We consider the random variable $\mathbb{P}(B\big|\mathcal{G})$ (which is the $\mathcal{G}$-measurable random variable, unique up to almost sure equality, such that for all $D \in \mathcal{G}$

$$\mathbb{P}(D \cap B) = \int_D \mathbb{P}(B\big|\mathcal{G})(\omega)d\mathbb{P}(\omega). \tag{6}$$

See for example Chapter 9 of [Wil91] for background on conditional probabilities).

Lemma 3. *Suppose* $0 \le p_0 \le p_1 \le 1$, *and that* $D \in \mathcal{G}$ *with*

$$\mathbb{P}(B \big| \mathcal{G})(\omega) \in [p_0, p_1] \text{ for all } \omega \in D. \tag{7}$$

Then

$$\frac{\pi_1}{\pi_0}\frac{p_0}{1-p_0}\mathbb{P}(D \big| B^C) \le \mathbb{P}(D \big| B) \le \frac{\pi_1}{\pi_0}\frac{p_1}{1-p_1}\mathbb{P}(D \big| B^C).$$

Proof. From (6) and (7) we have

$$p_0\mathbb{P}(D) \le \mathbb{P}(D \cap B) \le p_1\mathbb{P}(D)$$

and

$$(1-p_1)\mathbb{P}(D) \le \mathbb{P}(D \cap B^C) \le (1-p_0)\mathbb{P}(D).$$

Combining these we get

$$\frac{p_0}{1-p_0}\mathbb{P}(D \cap B^C) \le \mathbb{P}(D \cap B) \le \frac{p_1}{1-p_1}\mathbb{P}(D \cap B^C).$$

Since $\mathbb{P}(D \big| B) = \mathbb{P}(D \cap B)/\pi_0$ and $\mathbb{P}(D \big| B^C) = \mathbb{P}(D \cap B^C)/\pi_1$, the result follows. $\square$

In particular, if J is a subset of the interval $[0, \pi_0)$ then we can set $D = \{\omega : \mathbb{P}(B \big| \mathcal{G})(w) \in J\}$ to obtain

$$\mathbb{P}\Big\{\mathbb{P}(B \big| \mathcal{G}) \in J \big| B\Big\} \le \mathbb{P}\Big\{\mathbb{P}(B \big| \mathcal{G}) \in J \big| B^C\Big\},$$

while if $J \subseteq (\pi, 1]$ then the inequality is reversed. In each case equality holds only if both sides are 0. Hence:

Corollary 4. *There exists a coupling of two random variables* Y_0 *and* Y_1, *such that* Y_0 *has the distribution of* $\mathbb{P}(B \big| \mathcal{G})$ *conditioned on B occurring, such that* Y_1 *has the distribution of* $\mathbb{P}(B \big| \mathcal{G})$ *conditioned on B not occurring, and such that:*

(i) whenever $Y_0 < \pi_0$, *then* $Y_1 = Y_0$, *and*

(ii) whenever $Y_1 > \pi_0$, *then* $Y_1 = Y_0$.

Therefore either $Y_0 = Y_1$ *or* $Y_1 \le \pi_0 \le Y_0$.

Also the distributions of Y_0 *and* Y_1 *are identical iff* $Y_0 = Y_1 = \pi_0$ *with probability 1, or equivalently iff* $\mathbb{P}(B \big| \mathcal{G}) = \pi_0$ *with probability 1.*

Applying this result with $B = \{\phi(u) = 0\}$, with $\mathcal{G} = \mathcal{T}(u)$, with $\mathbb{P} = \mu$ and so with $A(u) = \mathbb{P}(B \big| \mathcal{G})$, we obtain the coupling of A_0, A_1 with the properties claimed in Section 1.3.

3 Recurrences for likelihood ratios

Let $u \in \mathbb{T}^k$ and let $\mathbf{y}$ be a configuration on the set $\mathbb{W}_d(u)$ (the descendants of u at distance exactly d).

For $S \subset \mathbb{T}^k$, write $\phi \restriction_S$ for the configuration ϕ restricted to S.

Define the "likelihood functions"

$$q^{(0)}(d,\mathbf{y}) = \mu\big(\phi \restriction_{\mathbb{W}_d(u)} = \mathbf{y} \big| \phi(u) = 0\big)$$
$$q^{(1)}(d,\mathbf{y}) = \mu\big(\phi \restriction_{\mathbb{W}_d(u)} = \mathbf{y} \big| \phi(u) = 1\big).$$

For $i = 0,1$, the function $q^{(i)}(d,\mathbf{y})$ gives the probability of observing the configuration $\mathbf{y}$ on the set of descendants of u at distance d, given that the value at u itself is i. (Note that because of the translation invariance property noted in Section 1.1, the choice of u is not important).

Define also the "likelihood ratio" function

$$q(d,\mathbf{y}) = \frac{q^{(0)}(d,\mathbf{y})}{q^{(1)}(d,\mathbf{y})}.$$

Let $d \geq 2$ and let the children of u be $u_1,\ldots,u_k$. A configuration $\mathbf{y}$ on $\mathbb{W}_d(u)$ corresponds to a set of configurations $\mathbf{y}_1,\ldots,\mathbf{y}_k$ on $\mathbb{W}_{d-1}(u_1),\ldots,\mathbb{W}_{d-1}(u_k)$. We then have

$$q^{(0)}(d,\mathbf{y}) = \prod_{j=1}^{k}\left[p_{00}q^{(0)}(d-1,\mathbf{y}_j) + p_{01}q^{(1)}(d-1,\mathbf{y}_j)\right]$$
$$q^{(1)}(d,\mathbf{y}) = \prod_{j=1}^{k}\left[p_{10}q^{(0)}(d-1,\mathbf{y}_j) + p_{11}q^{(1)}(d-1,\mathbf{y}_j)\right]$$

and so

$$\begin{aligned} q(d,\mathbf{y}) &= \prod_{j=1}^{k}\left\{\frac{p_{00}q^{(0)}(d-1,\mathbf{y}_j) + p_{01}q^{(1)}(d-1,\mathbf{y}_j)}{p_{10}q^{(0)}(d-1,\mathbf{y}_j) + p_{11}q^{(0)}(d-1,\mathbf{y}_j)}\right\} \\ &= \left(\frac{p_{00}}{p_{10}}\right)^k \prod_{j=1}^{k}\left\{1 + \frac{c_0 - c_1}{q(d-1,\mathbf{y}_j) + c_1}\right\}, \end{aligned} \tag{8}$$

where we define

$$c_0 = \frac{p_{01}}{p_{00}}, \quad c_1 = \frac{p_{11}}{p_{10}}.$$

Define also $a(d,\mathbf{y}) = \mu\left(\phi(u) = 0 \big| \phi \restriction_{\mathbb{W}_d(u)} = \mathbf{y}\right)$. The function a gives the conditional probability that the value at the node u is 0, given a configuration on the set of descendants of u at distance d. We have

$$\begin{aligned} a(d,\mathbf{y}) &= \frac{\pi_0 q^{(0)}(d,\mathbf{y})}{\pi_0 q^{(0)}(d,\mathbf{y}) + \pi_1 q^{(1)}(d,\mathbf{y})} \\ &= \frac{1}{1 + \frac{\pi_1}{\pi_0 q(d,\mathbf{y})}}. \end{aligned} \tag{9}$$

Returning to the random variable $A(d,u) = \mu\left(\phi(u) = 0 \middle| \sigma(\mathbb{W}_d(u))\right)$ defined in Section 1, we have

$$A(d,u) = a\left(d, \phi \restriction_{\mathbb{W}_d(u)}\right),$$

that is, the function $a(d,.)$ applied to the actually observed values of the configuration ϕ on $\mathbb{W}_d(u)$.

Similarly define

$$Q(d,u) = q\left(d, \phi \restriction_{\mathbb{W}_d(u)}\right).$$

From (9), we have

$$A(d,u) = \frac{1}{1 + \frac{\pi_1}{\pi_0 Q(d,u)}}, \quad Q(d,u) = \frac{\pi_1}{\pi_0}\left(\frac{1}{1 - A(d,u)} - 1\right).$$

Recalling $A(d,u) \to A(u)$ a.s., we have $Q(d,u) \to Q(u)$ a.s., where

$$Q(u) = \frac{\pi_1}{\pi_0}\left(\frac{1}{1 - A(u)} - 1\right).$$

From (8) we get

$$Q(d,u) = \left(\frac{p_{00}}{p_{10}}\right)^k \prod_{j=1}^{k}\left(1 + \frac{c_0 - c_1}{Q(d-1,u_j) + c_1}\right),$$

and, taking $d \to \infty$,

$$Q(u) = \left(\frac{p_{00}}{p_{10}}\right)^k \prod_{j=1}^{k}\left(1 + \frac{c_0 - c_1}{Q(u_j) + c_1}\right).$$

Now put

$$L(u,d) = \ln Q(u,d)$$

and

$$\begin{aligned} L(u) &= \ln Q(u) \\ &= \ln\left[\frac{\pi_1}{\pi_0}\left(\frac{1}{1 - A(u)} - 1\right)\right]. \end{aligned}$$

We have $L(u,d) \to L(u)$ a.s. as $d \to \infty$, and

$$L(u) = k \ln\left(\frac{p_{00}}{p_{10}}\right) + \sum_{j=1}^{k} \ln\left(1 + \frac{c_0 - c_1}{\exp\left(L(u_j)\right) + c_1}\right). \tag{10}$$

Since $L(u)$ can be written as a strictly increasing function of $A(u)$, with $A(u) = \pi_0$ corresponding to $L(u) = 0$, we can translate the coupling of A_0, A_1 described in Section 1.3 and proved in Section 2 into a coupling of two random variables L_0, L_1 with the following properties:

(i) The distribution of L_0 is the distribution of $L(u)$ conditioned on the event $\{\phi(u) = 0\}$;

(ii) The distribution of L_1 is the distribution of $L(u)$ conditioned on the event $\{\phi(u) = 1\}$;

(iii) With probability 1, either $L_0 = L_1$ or $L_1 \le 0 \le L_0$;

(iv) If $L_0 = L_1$ with probability 1, then both are equal to 0 with probability 1, and then also $A(u) = \pi_0$ with probability 1.

So to conclude that $A(u)$ is a.s. constant, it's enough to show that $\mathbb{E}(L_0 - L_1) = 0$.

Returning to (10), note that $\exp(L(u_j)) \ge 0$, and so the quantity inside the second logarithm is always at least $\min(c_0/c_1, 1) > 0$. Thus the distribution of $L(u)$ has compact support; hence the same is true for L_0 and L_1, and certainly $\mathbb{E}|L_0| < \infty$, $\mathbb{E}|L_1| < \infty$.

Again let $u_1, \dots, u_k$ be the children of a node u. From the broadcasting construction we get the following information.

Conditional on $\phi(u) = 0$:

> the $\phi(u_j)$, $j = 1, \dots, k$ are i.i.d. taking value 0 w.p. p_{00} and value 1 w.p. p_{01}. Then the $L(u_j)$ are i.i.d., and the distribution of each is a mixture of the distribution of L_0 (with weight p_{00}) and the distribution of L_1 (with weight p_{01}).

Conditional on $\phi(u) = 1$:

> the $\phi(u_j)$, $j = 1, \dots, k$ are i.i.d. taking value 0 w.p. p_{10} and value 1 w.p. p_{11}. Then the $L(u_j)$ are i.i.d., and the distribution of each is a mixture of the distribution of L_0 (with weight p_{10}) and the distribution of L_1 (with weight p_{11}).

Hence from (10),

$$\mathbb{E}L_0 = -k \ln\left(\frac{p_{00}}{p_{10}}\right) + k\left[p_{00}\mathbb{E}\ln\left(1 + \frac{c_0 - c_1}{\exp(L_0) + c_1}\right) + p_{01}\mathbb{E}\ln\left(1 + \frac{c_0 - c_1}{\exp(L_1) + c_1}\right)\right].$$

and

$$\mathbb{E}L_1 = -k \ln\left(\frac{p_{00}}{p_{10}}\right) + k\left[p_{10}\mathbb{E}\ln\left(1 + \frac{c_0 - c_1}{\exp(L_0) + c_1}\right) + p_{11}\mathbb{E}\ln\left(1 + \frac{c_0 - c_1}{\exp(L_1) + c_1}\right)\right].$$

Subtracting and using the fact that $p_{00} - p_{10} = p_{11} - p_{01}$, we obtain

$$\mathbb{E}(L_0 - L_1) = k\mathbb{E}\left[f(L_0) - f(L_1)\right], \tag{11}$$

where

$$\begin{aligned} f(x) &= (p_{11} - p_{01}) \ln\left(1 + \frac{c_0 - c_1}{e^x + c_1}\right) \\ &= \frac{c_1 - c_0}{(1 + c_0)(1 + c_1)} \ln\left(1 + \frac{c_0 - c_1}{e^x + c_1}\right). \end{aligned} \tag{12}$$

3.1 General case

If $c_0 = c_1$, then the function f defined at (12) is constant. In that case, (11) shows that $\mathbb{E}(L_0 = L_1) = 0$, and reconstruction is impossible.

So assume that $c_0 \neq c_1$. Then f is strictly increasing, and one obtains that

$$f'(x) = \frac{(c_1 - c_0)^2}{(1+c_0)(1+c_1)} \frac{e^x}{(e^x + c_0)(e^x + c_1)}. \tag{13}$$

Putting $y = e^x$ and taking the reciprocal, one can find the value of x maximising (13) by finding the value of $y \geq 0$ minimising $(y + c_0)(y + c_1)y^{-1}$. One obtains $y = e^x = (c_0 c_1)^{1/2}$, and so

$$\begin{aligned}
\sup_x f'(x) &= \frac{(c_1 - c_0)^2}{(1+c_0)(1+c_1)} \frac{(c_0 c_1)^{1/2}}{\left((c_0 c_1)^{1/2} + c_0\right)\left((c_0 c_1)^{1/2} + c_1\right)} \\
&= \frac{(c_1 - c_0)^2}{(1+c_0)(1+c_1)} \frac{1}{\left(\sqrt{c_1} + \sqrt{c_0}\right)^2} \\
&= \frac{\left(\sqrt{c_1} - \sqrt{c_0}\right)^2}{(1+c_0)(1+c_1)} \\
&= \left(\sqrt{\frac{1}{1+c_0} \frac{c_1}{1+c_1}} - \sqrt{\frac{c_0}{1+c_0} \frac{1}{1+c_1}} \right)^2 \\
&= \left(\sqrt{p_{00} p_{11}} - \sqrt{p_{01} p_{10}} \right)^2.
\end{aligned} \tag{14}$$

Since we know that $L_1 \leq L_0$ with probability 1, we then have that

$$0 \leq f(L_0) - f(L_1) \leq \sup_x f'(x)(L_0 - L_1)$$

with equality on the RHS iff $L_0 = L_1$, with probability 1. Hence, from (11),

$$\mathbb{E}(L_0 - L_1) \leq k \sup_x f'(x) \mathbb{E}(L_0 - L_1),$$

with equality iff both sides are 0. So to show that $\mathbb{E}(L_0 - L_1) = 0$, and therefore that reconstruction is impossible, it's enough to show that $k \sup f'(x) \leq 1$. Using (14), we see that (2) indeed implies that reconstruction is impossible, and the proof of Theorem 1 is done.

3.2 Hard-core case

Recall that

$$\begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} = \begin{pmatrix} \frac{1}{1+w} & \frac{w}{1+w} \\ 1 & 0 \end{pmatrix};$$

we have also that $\pi_0 = (1+w)/(1+2w)$, $\pi_1 = w/(1+2w)$, and $c_0 = w$, $c_1 = 0$. We have also defined $\lambda = w(1+w)^k$. Equation (12) now becomes

$$f(x) = -\frac{w}{1+w} \ln\left(1 + w e^{-x}\right),$$

and now the function f is concave as well as strictly increasing. Hence in particular, if $x_0 < x_1$ and $y_0 < y_1$ with $x_0 \le y_0$ and $x_1 \le y_1$, then

$$0 \le \frac{f(y_1) - f(y_0)}{y_1 - y_0} \le \frac{f(x_1) - f(x_0)}{x_1 - x_0}. \tag{15}$$

Recurrence (10) now becomes

$$L(u) = -k\ln(1+w) + \sum_{j=1}^{k} \ln\left(1 + we^{-L(u_j)}\right),$$

and so in particular $L(u)$ is always greater than or equal to $-k\ln(1+w)$; the same is therefore true of L_0 and L_1 also. Combining this with property (iii) of the coupling described after (10), we have that, with probability 1, either $L_0 = L_1$ or $-k\ln(1+w) \le L_1 \le 0 \le L_0$. Thus, using (15), we obtain that with probability 1

$$\begin{aligned} 0 \le f(L_0) - f(L_1) &\le (L_0 - L_1)\frac{f(0) - f\big(-k\ln(1+w)\big)}{0 - (-k\ln(1+w))} \\ &= (L_0 - L_1)\frac{w}{1+w}\frac{-\ln(1+w) + \ln(1+\lambda)}{k\ln(1+w)}, \\ &= (L_0 - L_1)\frac{w}{k(1+w)}\left(\frac{\ln(1+\lambda)}{\ln(1+w)} - 1\right), \end{aligned} \tag{16}$$

where we have used

$$f(0) = -w\ln(1+w)/(1+w)$$

and

$$\begin{aligned} f(-k\ln(1+w)) &= -\frac{w}{1+w}\ln\left(1 + we^{k\ln(1+w)}\right) \\ &= -\frac{w}{1+w}\ln\left(1 + w(1+w)^k\right) \\ &= -\frac{w}{1+w}\ln(1+\lambda). \end{aligned}$$

Combining (11) and (16), we get $0 \le \mathbb{E}(L_0 - L_1) \le \rho\left[\mathbb{E}(L_0 - L_1)\right]$, where

$$\rho = \frac{w}{1+w}\left(\frac{\ln(1+\lambda)}{\ln(1+w)} - 1\right).$$

To obtain that $\mathbb{E}(L_0 - L_1) = 0$, and hence that reconstruction is impossible, it's enough that $\rho < 1$. But

$$\begin{aligned} \rho &< \frac{w}{1+w}\frac{\ln(1+\lambda)}{\ln(1+w)} \\ &\le \ln(1+\lambda)\sup_{w>0}\frac{w}{(1+w)\ln(1+w)} \\ &= \ln(1+\lambda), \end{aligned}$$

(since the quantity within the sup is decreasing in w and tends to 1 as $w \downarrow 0$). So certainly if $\lambda \le e-1$, then $\rho < 1$ as desired. Also ρ is continuous as a function of λ (or w), so the threshold value $\lambda_2(k)$ is in fact strictly greater than $e-1$, and the proof of Theorem 2 is done.

4 A calculation

For completeness, we here include details to show that the LHS of (2) is always less than or equal to that of (1). Define $b_0 = \min(p_{00}, p_{11}) \le \max(p_{00}, p_{11}) = b_1$. We obtain

$$\begin{aligned}
\min\{p_{00}+p_{10}, p_{01}+p_{11}\} &= \min\{b_0+1-b_1, b_1+1-b_0\} \\
&= b_0+1-b_1 \\
&= (1-b_0)(1-b_1)+b_0b_1+2b_0(1-b_1) \\
&\le (1-b_0)(1-b_1)+b_0b_1+2\sqrt{b_0b_1}\sqrt{(1-b_1)(1-b_0)} \qquad (17) \\
&= \left[\sqrt{b_0b_1}+\sqrt{(1-b_1)(1-b_0)}\right]^2. \qquad (18)
\end{aligned}$$

(The inequality in (17) follows since $b_0 \le b_1$ and $1-b_1 \le 1-b_0$; equality holds if $b_0 = b_1$ or if one of b_0 and b_1 is 0 or 1).

Then, starting from the LHS of (2) and using (18),

$$\begin{aligned}
\left(\sqrt{p_{00}p_{11}}-\sqrt{p_{01}p_{10}}\right)^2 &= \left[\sqrt{b_0b_1}-\sqrt{(1-b_1)(1-b_0)}\right]^2 \\
&\le \left[\sqrt{b_0b_1}-\sqrt{(1-b_1)(1-b_0)}\right]^2 \frac{\left[\sqrt{b_0b_1}+\sqrt{(1-b_1)(1-b_0)}\right]^2}{\min\{p_{00}+p_{10}, p_{01}+p_{11}\}} \\
&= \frac{[b_0b_1-(1-b_1)(1-b_0)]^2}{\min\{p_{00}+p_{10}, p_{01}+p_{11}\}} \\
&= \frac{(b_0-(1-b_1))^2}{\min\{p_{00}+p_{10}, p_{01}+p_{11}\}} \\
&= \frac{[\pm(p_{00}-p_{10})]^2}{\min\{p_{00}+p_{10}, p_{01}+p_{11}\}},
\end{aligned}$$

which gives the LHS of (1) as required.

Acknowledgements

Thanks to Yuri Suhov for much encouragement throughout the course of this work, to Elchanan Mossel for valuable conversations during the Isaac Newton Institute programme on Computation, Combinatorics and Probability in 2002, and to the referee for helpful comments.

References

[BRZ95] 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.

[BW03] Graham R. Brightwell and Peter Winkler. A second threshold for the hard-core model on a Bethe lattice. Preprint. Currently available from `http://cm.bell-labs.com/cm/ms/who/pw/papers/pubs.html`, 2003.

[EKPS00] William Evans, Claire Kenyon, Yuval Peres, and Leonard J. Schulman. Broadcasting on trees and the Ising model. *Ann. Appl. Probab.*, 10(2):410–433, 2000.

[Iof96] Dmitry Ioffe. On the extremality of the disordered state for the Ising model on the Bethe lattice. *Lett. Math. Phys.*, 37(2):137–143, 1996.

[JM03] Svante Janson and Elchanan Mossel. Robust reconstruction on trees is determined by the second eigenvalue. Preprint. Currently available from `http://www.cs.berkeley.edu/~mossel/`, 2003.

[Kel85] F. P. Kelly. Stochastic models of computer communication systems. *J. Royal Statist. Soc. B*, 47:379–395, 1985.

[KMP01] Claire Kenyon, Elchanan Mossel, and Yuval Peres. Glauber dynamics on trees and hyperbolic graphs. In *42nd IEEE Symposium on Foundations of Computer Science (Las Vegas, NV, 2001)*, pages 568–578. IEEE Computer Soc., Los Alamitos, CA, 2001.

[Mos01] Elchanan Mossel. Reconstruction on trees: beating the second eigenvalue. *Ann. Appl. Probab.*, 11(1):285–300, 2001.

[Mos03] Elchanan Mossel. Survey: information flow on trees. DIMACS volume on Graphs, Morphisms and Statistical Physics. To appear. Currently available from `http://www.cs.berkeley.edu/~mossel/`, 2003.

[MP03] Elchanan Mossel and Yuval Peres. Information flow on trees. Ann. Appl. Probab. To appear, 2003.

[RS03] Utkir A. Rozikov and Yuri M. Suhov. A hard-core model on a Cayley tree: an example of a loss network. Preprint, 2003.

[Wil91] David Williams. *Probability with Martingales*. Cambridge University Press, 1991.

[Zac83] Stan Zachary. Countable state space Markov random fields and Markov chains on trees. *Ann. Probab.*, 11(4):894–903, 1983.