The Annals of Applied Probability
2003, Vol. 13, No. 3, 817–844

INFORMATION FLOW ON TREES

BY ELCHANAN MOSSEL AND YUVAL PERES

University of California, Berkeley

Consider a tree network T, where each edge acts as an independent copy of a given channel M, and information is propagated from the root. For which T and M does the configuration obtained at level n of T typically contain significant information on the root variable? This problem arose independently in biology, information theory and statistical physics.

For all b, we construct a channel for which the variable at the root of the b-ary tree is independent of the configuration at the second level of that tree, yet for sufficiently large $B > b$, the mutual information between the configuration at level n of the B-ary tree and the root variable is bounded away from zero for all n. This construction is related to Reed–Solomon codes.

We improve the upper bounds on information flow for asymmetric binary channels (which correspond to the Ising model with an external field) and for symmetric q-ary channels (which correspond to Potts models).

Let $\lambda_2(M)$ denote the second largest eigenvalue of M, in absolute value. A CLT of Kesten and Stigum implies that if $b|\lambda_2(M)|^2 > 1$, then the *census* of the variables at any level of the b-ary tree, contains significant information on the root variable. We establish a converse: If $b|\lambda_2(M)|^2 < 1$, then the census of the variables at level n of the b-ary tree is asymptotically independent of the root variable. This contrasts with examples where $b|\lambda_2(M)|^2 < 1$, yet the *configuration* at level n is not asymptotically independent of the root variable.

1. Introduction.

Consider a process in which information flows from the root of a tree T to other nodes of T. Each edge of the tree acts as a channel on a finite alphabet $\mathcal{A} = \{1, \dots, k\}$. Denote by $\mathbf{M}_{i,j}$ the transition probability from $i \in \mathcal{A}$ to $j \in \mathcal{A}$, and by M the random function (or channel) which satisfies for all $i \in \mathcal{A}$ and $j \in \mathcal{A}$ that $\mathbf{P}[M(i) = j] = \mathbf{M}_{i,j}$. In other words, $\{M(i)\}_{i \in \mathcal{A}}$ is a collection of random variables satisfying $\mathbf{P}[M(i) = j] = \mathbf{M}_{i,j}$ for all $i, j \in \mathcal{A}$.

Let $\lambda_2(M)$ denote the eigenvalue of $\mathbf{M}$ which has the second largest absolute value [$\lambda_2(M)$ may be negative or nonreal]. At the root ρ one of the symbols of $\mathcal{A}$ is chosen according to some initial distribution. We denote this (random) symbol by σ_ρ. This symbol is then propagated in the tree as follows. For each vertex v having as a parent v', we let $\sigma_v = M_{v',v}(\sigma_{v'})$, where the $\{M_{v',v}\}$ are independent copies of M. Equivalently, for a vertex v, let v' be the parent of v, and let $\Delta(v)$

Received June 2001; revised July 2002.

AMS 2000 subject classifications. Primary 60J80; secondary 60K35, 60F05, 94B99.

Key words and phrases. Information flow, reconstruction problem, tree index Markov chain, Markov random field, census, second eigenvalue, Thompson's principle, Reed–Solomon codes, secret sharing.

be the set of all vertices which are connected to ρ through paths which do not contain v. Then the process satisfies

$$\mathbf{P}\big[\sigma_v = j \,|\, (\sigma_w)_{w \in \Delta(v)}\big] = \mathbf{P}[\sigma_v = j \,|\, \sigma_{v'}] = \mathbf{M}_{\sigma_{v'}, j}.$$

It is natural to study this process in the context of biology, statistical physics and communication theory. See [10, 23] and the references there for more background.

Let $d(\cdot, \cdot)$ denote the graph-metric distance on T, and $L_n = \{v \in V : d(\rho, v) = n\}$ be the nth level of the tree. For $v \in V$ and $e = (v, w) \in E$ we denote $|v| = d(\rho, v)$ and $|e| = \max\{|v|, |w|\}$. We denote by $\sigma_n = (\sigma(v))_{v \in L_n}$ the symbols at the nth level of the tree. We let $c_n = (c_n(1), \dots, c_n(k))$ where

$$c_n(i) = \#\{v \in L_n : \sigma(v) = i\}.$$

In other words, c_n is the *census* of the nth level. Note that $(\sigma_n)_{n=1}^{\infty}$ is a nonhomogeneous Markov chain. If the tree T is spherically symmetric, then $(c_n)_{n=1}^{\infty}$ is a nonhomogeneous Markov chain as well.

For distributions P and Q on the same space Ω, the total variation distance between P and Q is

$$D_V(P, Q) = \tfrac{1}{2} \sum_{\sigma \in \Omega} |P(\sigma) - Q(\sigma)|. \tag{1}$$

DEFINITION 1.1. The reconstruction problem for T and M is *solvable* if there exist $i, j \in \mathcal{A}$ for which

$$\lim_{n \to \infty} D_V(\mathbf{P}_n^i, \mathbf{P}_n^j) > 0, \tag{2}$$

where $\mathbf{P}_n^{\ell}$ denotes the conditional distribution of σ_n given that $\sigma_\rho = \ell$.

DEFINITION 1.2. The reconstruction problem for T and M is *census-solvable* if there exist $i, j \in \mathcal{A}$ for which

$$\lim_{n \to \infty} D_V(\mathbf{P}_n^{(c), i}, \mathbf{P}_n^{(c), j}) > 0, \tag{3}$$

where $\mathbf{P}_n^{(c), \ell}$ denotes the conditional distribution of c_n given that $\sigma_\rho = \ell$.

We note that the total variation distances at (2) and (3) are decreasing in n, and therefore the limits always exist.

It is easy to see that the sequences in (2) and (3) are decreasing. Moreover, assuming that $\mathbf{P}[\sigma_\rho = i] > 0$ for every $i \in \mathcal{A}$, the following conditions are equivalent (see, e.g., [23]).

1. The reconstruction problem for T and M is not solvable.
2. The configurations σ_n are asymptotically independent of σ_ρ; that is, $\lim_{n\to\infty} I(\sigma_\rho, \sigma_n) = 0$, where I is the mutual information operator. Operator I is defined by $I(X, Y) = H(X) + H(Y) - H(X, Y)$ where H is the entropy operator (see [7] for more background).
3. The sequence $(\sigma_n)_{n=1}^\infty$ has a trivial tail σ-field.

Similar conditions apply to census-solvability, where $\{c_n\}$ replace $\{\sigma_n\}$. Note that if the reconstruction problem is census-solvable, it is also solvable.

The reconstruction problem was first studied in statistical physics for the symmetric Ising model on the tree, in the equivalent form as a question regarding the extremality of the free Gibbs measure for this model. See [27, 13, 2, 14–16]. Reference [10] contains extensive background on this problem for the symmetric Ising model. In these papers the reconstruction problem for the symmetric Ising model on trees is analyzed. Writing $\lambda_2(M)$ for the second eigenvalue of $\left(\begin{smallmatrix}1-\varepsilon & \varepsilon\\ \varepsilon & 1-\varepsilon\end{smallmatrix}\right)$ [i.e., $\lambda_2(M) = 1 - 2\varepsilon$], the reconstruction problem is solvable for the b-ary tree $T_b = (V_b, E_b)$ if and only if $b\lambda_2^2(M) > 1$. An analogous threshold for general trees was established in [10].

A few combinatorial questions arise naturally in the context of these tree processes (see also [3–5]). In Section 2 we discuss whether for a channel M there exist any b such that the reconstruction problem is solvable for the infinite b-ary tree T_b and the channel M. In Theorem 2.1 we present a criterion for deciding this problem for a channel M. Using this criterion we construct a channel with the following properties.

THEOREM 1.1. *Write M^h for the hth iteration of the channel M. Then for all $h > 1$ there exists a channel M such that*:

(i) *$M^h(j)$ has the same distribution for all $j \in A$.*

(ii) *For all $h' < h$ there exist i and j such that $M^{h'}(i)$ and $M^{h'}(j)$ have different distributions.*

(iii) *When b is large the reconstruction problem is solvable for the tree T_b and the channel M.*

This is a generalization of an example which appeared in [22] and [23]. In Section 3 we give more delicate constructions which are related to Reed–Solomon codes [25]; see also [26].

THEOREM 1.2. *Let $b > 1$ be an integer and T be the two-level b-ary tree. There exists a channel M such that for any initial distribution, σ_ρ and σ_∂ are independent* (*where σ_ρ is the root label, and σ_∂ is the configuration at the leaves of the two-level b-ary tree*), *yet when B is sufficiently large, the reconstruction problem for the channel M and the B-ary tree T_B is solvable.*

It is tempting to try to find thresholds for the reconstruction problem which depend only on b and $\lambda_2(M)$. For binary symmetric channels the threshold for reconstruction is $b\lambda_2^2(M) = 1$ (this is also the threshold for census reconstruction for general channels, see Theorems 1.4 and 1.5 below).

In [23] it is shown that for some natural generalizations of the binary symmetric channel, the threshold $b\lambda_2^2(M) = 1$ is not the threshold for reconstruction. In particular, it is shown that for the Potts model (the q-ary symmetric channel), where the transition matrix is

$$\mathbf{M} = \begin{pmatrix} 1-(q-1)\delta & \delta & \dots & \delta \\ \delta & 1-(q-1)\delta & \delta & \dots \\ \vdots & \dots & \ddots & \vdots \\ \delta & \dots & \delta & 1-(q-1)\delta \end{pmatrix}, \tag{4}$$

if $b\lambda_2(M) = b(1 - q\delta) > 1$ and q is sufficiently large, then the reconstruction problem is solvable. Similar results hold for asymmetric binary channels. On the other hand, it is well known (see, e.g., [23]), that for both families the reconstruction problem is unsolvable if $0 \le b\lambda_2(M) \le 1$. In Section 4 we improve the bound $b\lambda_2(M) \le 1$ by proving the following.

PROPOSITION 1.3. *Let* $\mathbf{M}$ *be the matrix* (4). *Then the reconstruction problem for* T_b *and* M *is unsolvable when*

$$b\frac{(1-q\delta)^2}{1-(q-2)\delta} \le 1. \tag{5}$$

Similar results hold for asymmetric binary channels. The proofs are based on a reduction to symmetric binary channels on general trees which were analyzed in [10].

In Sections 5, 6 and 7 we consider census reconstruction and show that the threshold for census reconstruction is $b|\lambda_2(M)|^2 = 1$. From the CLT in [18] it follows that if $b|\lambda_2(M)|^2 > 1$, then the reconstruction problem is census-solvable. We extend this result to general trees by proving the theorem.

THEOREM 1.4. *Let* T *be an infinite tree and write* $\mathrm{br}(T)$ *for the branching number of the tree. Let* M *be a channel such that* $\mathrm{br}(T)|\lambda_2(M)|^2 > 1$. *Then the reconstruction problem is solvable for* T *and* M.

The results of [18] also play a crucial role in proving the following.

THEOREM 1.5. *The reconstruction problem for the* b*-ary tree and the channel* M *is not census-solvable if* $b|\lambda_2(M)|^2 < 1$.

We conjecture that the result of Theorem 1.5 should hold also when $b|\lambda_2(M)|^2 = 1$, but we have verified this only for the following channels.

THEOREM 1.6. *Let M be the q-state Potts model, or the asymmetric Ising model, and suppose that $b\lambda_2^2(M) \le 1$; then the reconstruction problem is not census-solvable for the b-ary tree and the channel M.*

In Section 8 we discuss some open problems.

Some of the results in this paper concern general trees. For an infinite tree T many of its probabilistic properties are determined by the branching number $\mathrm{br}(T)$. This is the supremum of the real numbers $\lambda \ge 1$, such that T admits a positive flow from the root to infinity, where on every edge e of T, the flow is bounded by $\lambda^{-|e|}$. Here $|e|$ denotes the number of edges (including e), on the path from e to the root; it is known [20] that $\mathrm{br}(T)^{-1}$ is the critical probability for Bernoulli percolation on T. See [20] and [10] for equivalent definitions of $\mathrm{br}(T)$ in terms of percolation, cutset sums and electrical conductance. We note that for the regular tree T_b we have $\mathrm{br}(T_b) = b$.

2. Distinguishing states by tree networks. Let M be a channel on an alphabet $\mathcal{A}$ of size k with $\lambda_2(M) = 0$. Looking at the Jordan form of $\mathbf{M}$ we see that $\mathrm{rank}(\mathbf{M}^k) = 1$, and therefore $M^k(i)$ has the same distribution for all $i \in \mathcal{A}$. Moreover, by Theorem 1.5 it follows that for all b, the reconstruction problem is not *census* solvable for the channel M and the tree T_b. Does this mean that the reconstruction problem is unsolvable for the channel M and T_b? In this section we answer this question.

DEFINITION 2.1. Let M be a channel on $\mathcal{A}$ and consider the minimal equivalence relation $\sim$ on $\mathcal{A}$ which satisfies: If i and j satisfy for all ℓ,

$$\sum_{\ell' \sim \ell} \mathbf{M}_{i,\ell'} = \sum_{\ell' \sim \ell} \mathbf{M}_{j,\ell'}, \tag{6}$$

then $i \sim j$. If $i \sim j$, we say that i and j are *indistinguishable*. Otherwise, we say that i and j are *distinguishable*.

This equivalence relation is obtained by:

1. First identifying $i, j \in \mathcal{A}$ whose one-step transitions are the same.
2. Then looking at the chain where such i, j are lumped and repeating the first step.

THEOREM 2.1. *If the states i and j are indistinguishable, then there exists N such that for all $n \ge N$ and for all infinite rooted trees T,*

$$D_V(\mathbf{P}_n^i, \mathbf{P}_n^j) = 0. \tag{7}$$

On the other hand, for every channel M, there exists b, such that for the tree T_b,

$$\inf_{n \ge 1, i \nsim j} D_V(\mathbf{P}_n^i, \mathbf{P}_n^j) > 0. \tag{8}$$

EXAMPLE 2.2. *Let $\{Z_i\}_{i=1}^{\infty}$ be an i.i.d. sequence satisfying $\mathbf{P}[Z_i = 0] = 1 - \mathbf{P}[Z_i = 1] = p$, where $0 < p < 1$. Let $h \geq 1$ and consider the channel M defined by the Markov chain $Y_i = (Z_{i+1}, \dots, Z_{i+h})$. Thus M has state space $\{0, 1\}^h$, with the product probability measure, $(p\delta_0 + (1-p)\delta_1)^{\otimes^h}$, as stationary measure. It is easily seen that all the states of M are indistinguishable. Therefore by Theorem* 2.1, *there exists N such that for all trees and all $n \geq N$, it holds that $D_V(\mathbf{P}_n^i, \mathbf{P}_n^j) = 0$ (one can take $N = h$).*

PROOF OF THEOREM 1.1. Let $\{Z_i\}_{i=1}^{\infty}$ be as in Example 2.2. Set

$$Y_i = \max\{0 \leq j \leq h : Z_{i+1} = \cdots = Z_{i+j} = 1\}$$

($Y_i = 0$ if $Z_{i+1} = 0$). It is easily seen that Y_i defines a channel M on the space $\{0, \dots, h\}$. Moreover, it is clear that for all ℓ, the variables $\{Z_i\}_{i \geq \ell+h+1}$ and $\{Y_i\}_{i \leq \ell}$ are independent. Therefore $\{Y_i\}_{i \geq \ell+h}$ and $\{Y_i\}_{i \leq \ell}$ are independent. It follows that the variables $M^h(j)$ have the same distribution for all j. Thus $\operatorname{rank}(\mathbf{M}^h) = 1$, and $\lambda_2(M) = 0$. Writing $\mathbf{M}$,

$$\mathbf{M} = \begin{pmatrix} p & p(1-p) & p(1-p)^2 & \dots & p(1-p)^h \\ 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & p & 1-p \end{pmatrix},$$

we see that if $h > 1$, then all the states of M are distinguishable. Therefore by Theorem 2.1, when b is sufficiently large we have for the tree T_b that $\inf_{n \geq 1, i, j} D_V(\mathbf{P}_n^i, \mathbf{P}_n^j) > 0$. Note that the channel above is obtained by lumping the channel of Example 2.2. This is a generalization of a channel appearing in [22]; see also [19]. □

REMARKS.

1. Theorem 2.1 implies that if for a channel M all the states are indistinguishable, then $\lambda_2(M) = 0$.
2. Suppose that the channel M is reversible and has $\lambda_2(M) = 0$. Then looking at the diagonal form of $\mathbf{M}$, it follows that $\operatorname{rank}(\mathbf{M}) = 1$. In particular, all the states of M are indistinguishable.

PROOF OF THEOREM 2.1. The first claim is easy. We define a new equivalence relation between states in $\mathcal{A}$. We let i and j be equivalent if there exists an N such that for all trees and all $n \geq N$ we have $D_V(\mathbf{P}_n^i, \mathbf{P}_n^j) = 0$. It is clear that this is indeed an equivalence relation and that it satisfies (6). The proof of the first claim follows.

For the proof of the second claim let $\mathcal{A}'$ be the set of equivalence classes for $\sim$. Let M' be the channel on $\mathcal{A}'$ defined as follows. For $i', j' \in \mathcal{A}'$ we choose $i \in i'$ and define

$$\mathbf{P}[M'(i') = j'] = \sum_{j \in j'} \mathbf{P}[M(i) = j]. \tag{9}$$

It is clear that (9) does not depend on the choice of i. Moreover, it is clear that for any tree T, we may couple the tree process for M and the M' tree process for M' in such a way that for all v in T we have that $\sigma_v \in \sigma'_v$. In particular, it follows that for all trees and all $i \in i'$ and $j \in j'$,

$$D_V(\mathbf{P}_n^{i'}, \mathbf{P}_n^{j'}) \le D_V(\mathbf{P}_n^{i}, \mathbf{P}_n^{j}).$$

Therefore, in order to prove the second claim it suffices to show that there exists a b such that for the tree T_b and the channel M' we have

$$\inf_{i', j' \in \mathcal{A}', n} D_V(\mathbf{P}_n^{i'}, \mathbf{P}_n^{j'}) > 0. \tag{10}$$

The crucial property of $\mathbf{M}'$ we will exploit is that for all $i \neq j \in \mathcal{A}'$ we have $(\mathbf{M}'_{i,\ell}) \neq (\mathbf{M}'_{j,\ell})$. Thus in order to simplify the presentation of the proof, we will assume that the channel $\mathbf{M}$ on the alphabet $\mathcal{A}$ satisfies $(\mathbf{M}_{i,\ell})_{\ell=1}^{|\mathcal{A}|} \neq (\mathbf{M}_{j,\ell})_{\ell=1}^{|\mathcal{A}|}$ for all $i \neq j \in \mathcal{A}$.

We are going to show that for every $\varepsilon > 0$, there exists $b = b(\varepsilon)$, and a recursive algorithm, which given the symbols at the nth level of T_b, reconstructs the symbol at the root with probability at least $1 - \varepsilon$ (uniformly for all initial distributions of the root). This implies (see [23]) that (10) holds.

We let $\delta = \min_{i,j} |(\mathbf{M}_{i,\ell})_\ell - (\mathbf{M}_{j,\ell})_\ell|_\infty$, and assume that $\varepsilon < \delta/8$. By standard results in the theory of large deviations (see, e.g., [8]), there exist positive constants C_1 and C_2 such that for all i, if X_i is the census of b independent trials with distribution $(\mathbf{M}_{i,\ell})_\ell$, then

$$\mathbf{P}\bigl[|X_i - b(\mathbf{M}_{i,\ell})_\ell|_\infty > b\delta/8\bigr] < C_1 \exp(-C_2 b\delta^2). \tag{11}$$

Similarly, if X is the sum of b i.i.d. $\{0, 1\}$-valued variables, each of which has the value 1 with probability ε and the value 0 with probability $1 - \varepsilon$, then

$$\mathbf{P}[|X - \varepsilon b| > b\delta/8] \le C_1 \exp(-C_2 b\delta^2). \tag{12}$$

We now choose b such that $C_1 \exp(-C_2 b\delta^2) < \varepsilon/2$, and apply the following recursive algorithm in order to reconstruct the symbol at the root of the tree.

1. For each vertex v, construct a census vector X_v, of the reconstructed values for the children of that vertex.
2. Reconstruct at v the value i, where i minimizes the distance $|X_v - b(\mathbf{M}_{i,\ell})_\ell|_\infty$.

It now follows by induction that the probability of correct reconstruction is at least $1 - \varepsilon$. Indeed if this is true for all the children of v, then by (12), with probability at least $1-\varepsilon/2$, for at least $(1-\delta/4)b$ of the children we reconstructed the correct value. So if Y_v is the census of the original labels at the children of v, then $\mathbf{P}[|X_v - Y_v|_\infty \geq b\delta/4] < \varepsilon/2$. Letting i be the original label at v, we obtain by (11),

$$\begin{aligned}\mathbf{P}[|X_v - b\mathbf{M}_{i,\ell}|_\infty \geq b\delta/2] &\leq \mathbf{P}[|X_v - Y_v|_\infty \geq b\delta/4]\\ &\quad + \mathbf{P}[|Y_v - b\mathbf{M}_{i,\ell}|_\infty \geq b\delta/4] < 2\varepsilon/2 = \varepsilon.\end{aligned}$$

Thus, with probability at least $1-\varepsilon$, the vector X_v satisfies $|X_v - b\mathbf{M}_{i,\ell}|_\infty < b\delta/2$, where i is the symbol at v. It now follows that we reconstruct the correct value with probability at least $1-\varepsilon$. □

REMARK. Note that we have shown that when b is sufficiently large, it is possible to reconstruct the equivalence class of a state with probability close to 1.

3. The distinguishing power of tree networks.

Let M be a channel on $\mathcal{A}$ and let T be a tree which consists of two nodes, the root ρ and an additional vertex v (the 1-ary one-level tree). Suppose that the label at the root is chosen in such a way that $\mathbf{P}[\sigma_\rho = i] > 0$ for all $i \in \mathcal{A}$, and $I(\sigma_\rho, \sigma_{\partial T}) = 0$, where ∂T are the vertices in the boundary of the tree T (in this case v), and I is the mutual information operator. The assumption $I(\sigma_\rho, \sigma_{\partial T}) = 0$ is equivalent to rank($\mathbf{M}$) $= 1$. Therefore, all the states of M are indistinguishable and the reconstruction problem for T_b and the channel M is unsolvable for all b.

On the other hand, let T be the two-level 1-ary tree. In Theorem 1.1 we constructed a channel M such that for the tree T we have $I(\sigma_\rho, \sigma_{\partial T}) = 0$, yet the reconstruction problem for T_b is solvable when b is sufficiently large.

It is natural to ask a similar question for other finite trees T: Suppose that for any initial distribution the channel M satisfies $I(\sigma_\rho, \sigma_{\partial T}) = 0$; does this imply that the reconstruction problem for T_b is unsolvable for all b?

If the trees T and T' have the same inner nodes, but $\partial T \subset \partial T'$, then by the "data processing lemma" (see, e.g., [7]), $I(\sigma_\rho, \sigma_{\partial T}) \leq I(\sigma_\rho, \sigma_{\partial T'})$. In particular, if there exists a vertex v in ∂T which is at distance 1 from the root and $I(\sigma_\rho, \sigma_{\partial T}) = 0$, then $I(\sigma_\rho, \sigma_v) = 0$ and therefore the reconstruction problem is unsolvable for T_b for all b.

Below we show that trees having boundary vertices at distance 1 from the root are the only trees for which the 0-information condition implies nonreconstruction for all b. The following construction exploits properties of Reed–Solomon codes ([25], see also [26]).

THEOREM 1.2. *Let $b > 1$ be an integer and T be the two-level b-ary tree. There exists a channel M such that for any initial distribution $I(\sigma_\rho, \sigma_{\partial T}) = 0$, yet when B is sufficiently large, the reconstruction problem for the channel M and T_B is solvable.*

Now given any finite rooted tree T having no boundary vertices at distance 1 from the root ρ, let b be the maximal degree in T, and let M be the channel constructed at Theorem 1.2 for b. If T' is the two-level b-ary tree, then by the "data processing lemma," $I(\sigma_\rho, \sigma_{\partial T}) \le I(\sigma_\rho, \sigma_{\partial T'}) = 0$, yet the reconstruction problem is solvable for T_B when B is sufficiently large.

The construction in the theorem is via polynomials over finite fields.

DEFINITION 3.1. Let $\mathcal{F}$ be a finite field with $q \ge b+2$ elements. Let $x_1, \dots, x_{b+1}$ be a fixed set of nonzero elements of $\mathcal{F}$. We define a channel on the state space

$$\mathcal{F}^b[x] = \{f(x) : f(x) \in \mathcal{F}[x],\ \deg f \le b\}.$$

Given f, take I to be a uniform variable in the set $\{1, \dots, b+1\}$, then take $M(f)$ to be $g \in \mathcal{F}^b[x]$ which satisfy $g(0) = f(x_I)$ with probability

$$\big|\{g \in \mathcal{F}^b[x] : g(0) = f(x_I)\}\big|^{-1} = q^{-b}.$$

In other words,

$$\mathbf{P}[M(f) = g] = \frac{|\{i : f(x_i) = g(0)\}|}{(b+1)q^b}.$$

PROPOSITION 3.1. *Let $\{x_1, \dots, x_{b+1}\}$ be a set of elements of $\mathcal{F}$. There is a linear bijection from $\mathcal{F}^b[x]$ to $\mathcal{F}^{b+1}$ defined by*

$$f \to \big(f(x_1), \dots, f(x_{b+1})\big). \tag{13}$$

The inverse map is defined by the interpolation polynomial,

$$(y_1, \dots, y_{b+1}) \to f(x) = \sum_{i=1}^{b+1} \frac{\prod_{j \ne i}(x - x_j)}{\prod_{j \ne i}(x_i - x_j)} y_i. \tag{14}$$

LEMMA 3.2. *Let T be the two-level b-ary tree. Then for any initial distribution, $I(\sigma_\rho, \sigma_{\partial T}) = 0$.*

PROOF. We show that for all $f \in \mathcal{F}^b[x]$ and $h = (h_{i,j})_{i,j=1}^b \in (\mathcal{F}^b[x])^{b^2}$ it holds that

$$\mathbf{P}[\sigma_{\partial T} = h | \sigma_\rho = f]$$

is independent of f. This implies the claim of the lemma. We denote by $(v_i)_{i=1}^b$ the children of ρ and by $(w_{i,j})_{j=1}^b$ the children of v_i. We then have

$$\begin{aligned} &\mathbf{P}[\sigma_{\partial T} = h | \sigma_\rho = f] \\ &= \sum_{a_1, \dots, a_b \in F} \mathbf{P}[\forall i, \sigma_{v_i}(0) = a_i | \sigma_\rho = f]\mathbf{P}[\sigma_{\partial T} = h | \forall i, \sigma_{v_i}(0) = a_i], \end{aligned} \tag{15}$$

$$= \sum_{a_1,\dots,a_b \in F} \mathbf{P}[\forall\, i, \sigma_{v_i}(0) = a_i | \sigma_\rho = f] \prod_{i=1}^{b} \mathbf{P}[\forall\, j, \sigma_{w_{i,j}} = h_{i,j} | \sigma_{v_i}(0) = a_i]$$

$$= q^{-b^2} \sum_{a_1,\dots,a_b \in F} \mathbf{P}[\forall\, i, \sigma_{v_i}(0) = a_i | \sigma_\rho = f]$$

$$\times \prod_{i=1}^{b} \mathbf{P}[\forall\, j, \sigma_{w_{i,j}}(0) = h_{i,j}(0) | \sigma_{v_i}(0) = a_i].$$

The first equality in (15) follows from the fact that h is independent of f given $\{\sigma_{v_i}(0)\}_{i=1}^{b}$. The last equality follows as $h_{i,j} = h_{i,j}(0) + x\hat{h}_{i,j}(x)$, where $\hat{h}_{i,j}(x)$ is independent of v_i.

Note that Proposition 3.1 implies that if $z_1, \dots, z_j$ are distinct nonzero elements and $1 \le j \le b$, then the uniform distribution measure on $\mathcal{F}^b[x]$ satisfies, for all $y_1, \dots, y_j, a \in \mathcal{F}$, that

$$\mathbf{P}[\forall\, 1 \le i \le j,\ f(z_i) = y_i | f(0) = a] = q^{-j}.$$

Therefore if $z_1, \dots, z_b$ is a sequence of nonzero elements (not necessarily distinct), then

$$\mathbf{P}[\forall\, 1 \le i \le b, f(z_i) = y_i | f(0) = a]$$

$$= \begin{cases} 0, & \exists\, i, j, \text{ s.t. } z_i = z_j \text{ and } y_i \ne y_j, \\ q^{-|\{z_1,\dots,z_b\}|}, & \text{otherwise,} \end{cases} \tag{16}$$

does not depend on a. In particular, it follows that if $z_1, \dots, z_b$ are i.i.d. uniform variables in $\{x_1, \dots, x_{b+1}\}$, then $\mathbf{P}[\forall\, 1 \le i \le j,\ f(z_i) = y_i | f(0) = a]$ does not depend on a. Thus the probability at the right-hand side of (15),

$$\mathbf{P}[\forall\, j, \sigma_{w_{i,j}}(0) = h_{i,j}(0) | \sigma_{v_i}(0) = a_i],$$

does not depend on a_i. It follows that (15) may be written as

$$cq^{-b^2} \sum_{a_1,\dots,a_b \in F} \mathbf{P}[\forall\, i, \sigma_{v_i}(0) = a_i | \sigma_\rho = f] = cq^{-b^2},$$

where $c = c(h)$ is independent of f as needed. □

LEMMA 3.3. *Let $\{x_1, \dots, x_{b+1}\}$ be a fixed set of nonzero elements in $\mathcal{F}$. For an element $f \in \mathcal{F}^b[x]$ satisfying $f(x_k) = y_k$ for $1 \le k \le b+1$ and a permutation $\pi \in S_{b+1}$ we define f_π to be the element of $\mathcal{F}^b[x]$ which satisfies*

$$f_\pi(x_{\pi(k)}) = y_k. \tag{17}$$

Moreover, for $a, b \in \mathcal{F}$ define

$$n_{a,b} = \left|\{(f, \pi) \in \mathcal{F}^b[x] \times S_{b+1} : f(0) = a,\ f_\pi(0) = b\}\right|. \tag{18}$$

We then have for all $a \neq b$, $c \neq d$ that

$$n_{a,b} = n_{c,d}, \qquad n_{a,a} = n_{c,c}. \tag{19}$$

Moreover, for all $a \neq b$,

$$n_{a,a} > n_{a,b}. \tag{20}$$

PROOF. We note that if $c \in \mathcal{F}$ and $f \in \mathcal{F}^b[x]$, then $(f + c)_\pi = f_\pi + c$. Similarly, if $0 \neq d \in \mathcal{F}$ then $(df)_\pi = df_\pi$. Thus for all a, b, we have $n_{a+c,b+c} = n_{a,b}$ and $n_{da,db} = n_{a,b}$. In particular, $n_{0,0} = n_{a,a}$ for all a, and if $a \neq b$ and $c \neq d$, then $n_{a,b} = n_{0,b-a} = n_{0,d-c} = n_{c,d}$. Relations (19) now follow.

By (19), in order to prove (20), it suffices to show that if $a \neq 0$, then $n_{0,0} > n_{0,a}$. Looking at (14) one sees that the elements g of $\mathcal{F}^b[x]$ which satisfy $g(0) = a$ are exactly those elements which satisfy the equation

$$\sum_{i=1}^{b+1} c_i g(x_i) = a,$$

where c_i are some constants such that $c_i \neq 0$ for all i (c_i are functions of $x_1, \ldots, x_{b+1}$).

Thus the elements g of $\mathcal{F}^b[x]$ which satisfy $g(0) = 0$ and $g_\pi(0) = a$ are exactly those elements of $\mathcal{F}^b[x]$ which satisfy the equations

$$\sum_{i=1}^{b+1} c_i g(x_i) = 0 \tag{21}$$

and

$$\sum_{i=1}^{b+1} c_{\pi(i)} g(x_i) = a. \tag{22}$$

Note that by Proposition 3.1 the number of solutions of these equations, $s_\pi(0, a)$ in $\mathcal{F}^b[x]$ is the same as the number of solutions of these equations, as linear equations over $\mathcal{F}$ in the variables $g(x_1), \ldots, g(x_{b+1})$. It now follows that for every permutation $\pi \neq 1$ and $a \neq 0$,

$$s_\pi(0, a) \leq s_\pi(0, 0). \tag{23}$$

Moreover, when $\pi = 1$,

$$0 = s_\pi(0, a) < s_\pi(0, 0) = q^b. \tag{24}$$

Now,

$$n_{0,a} = \sum_{\pi \in S_{b+1}} s_\pi(0, a) < \sum_{\pi \in S_{b+1}} s_\pi(0, 0) = n_{0,0},$$

as needed. □

LEMMA 3.4. *Let $f, g \in \mathcal{F}^b[x]$. Then f and g are indistinguishable in the sense of Definition* 2.1 *if and only if $f = g_\pi$ for some $\pi \in S_{b+1}$.*

PROOF. We will write $f \sim_1 g$ to denote that f and g are indistinguishable, and $f \sim_2 g$ when there exists $\pi \in S_{b+1}$ such that $f = g_\pi$. It is clear that if $f \sim_2 g$, then $f \sim_1 g$.

On the other hand, suppose that $f \nsim_2 g$. We will show that there exist an $h \in \mathcal{F}^b[x]$ such that

$$\sum_{h' \sim_2 h} \mathbf{P}[M(f) = h'] \neq \sum_{h' \sim_2 h} \mathbf{P}[M(g) = h']. \tag{25}$$

This would imply that indeed $\sim_1 = \sim_2$ (see Definition 2.1). We may write (25) equivalently,

$$\sum_{\pi \in S_{b+1}} \mathbf{P}[M(f) = h_\pi] \neq \sum_{\pi \in S_{b+1}} \mathbf{P}[M(g) = h_\pi]. \tag{26}$$

Writing $n_f(a) = |\{1 \le j \le b+1 : f(x_j) = a\}|$, we may write the left-hand side of (26) as

$$\frac{1}{(b+1)q^b} \sum_{\pi \in S_{b+1}} n_f(h_\pi(0)). \tag{27}$$

Summing over all $h \in \mathcal{F}^b[x]$ with $h(0) = a$ in (27) we obtain

$$\frac{n_{a,a} n_f(a) + \sum_{b \neq a} n_{a,b} n_f(b)}{(b+1)q^b} = \frac{(n_{a,a} - n_{a,a+1}) n_f(a) + (b+1) n_{a,a+1}}{(b+1)q^b} \tag{28}$$

[by (19) we may write $n_{a,a+1}$ for $n_{a,b}$ when $a \neq b$]. Since $f \nsim_2 g$ there exists $a \in \mathcal{F}$ such that $n_f(a) > n_g(a)$. Thus,

$$\begin{aligned}
\sum_{h : h(0) = a, \pi \in S_{b+1}} \mathbf{P}[M(f) = h_\pi] &= \frac{(n_{a,a} - n_{a,a+1}) n_f(a) + (b+1) n_{a,a+1}}{(b+1)q^b} \\
&> \frac{(n_{a,a} - n_{a,a+1}) n_g(a) + (b+1) n_{a,a+1}}{(b+1)q^b} \qquad (29) \\
&= \sum_{h : h(0) = a, \pi \in S_{b+1}} \mathbf{P}[M(g) = h_\pi],
\end{aligned}$$

where we have used the fact that by (20) $n_{a,a} > n_{a,a+1}$. It now follows that there exists an h for which (26) holds. □

PROOF OF THEOREM 1.2. Fix b and let T be the two-level b-ary tree. Let M be the channel defined at Definition 3.1. Proposition 3.1 implies that for any initial distribution $I(\sigma_\rho, \sigma_{\partial T}) = 0$.

On the other hand, by Lemma 3.4, the channel M satisfies that f and g are indistinguishable for M if and only if $g = f_\pi$ for some π. It follows by Theorem 2.1 that when B is sufficiently large the reconstruction problem for T_B is solvable (moreover, when B is sufficiently large we may reconstruct $\{f_\pi\}_{\pi \in S_{b+1}}$ the equivalence class of f, with probability close to 1). □

4. Improved bounds for Potts models. In this section we will focus on the following two families of channels:

1. The asymmetric binary channels. These channels have the state space $\{0, 1\}$ and the matrices
$$\mathbf{M} = \begin{pmatrix} 1-\delta_1 & \delta_1 \\ 1-\delta_2 & \delta_2 \end{pmatrix}, \tag{30}$$
with $\lambda_2(M) = \delta_2 - \delta_1$.
2. The symmetric channels on q symbols. These have the state space $\{1, \dots, q\}$ and the matrices
$$\mathbf{M} = \begin{pmatrix} 1-(q-1)\delta & \delta & \dots & \delta \\ \delta & 1-(q-1)\delta & \delta & \dots \\ \vdots & \dots & \ddots & \vdots \\ \delta & \dots & \delta & 1-(q-1)\delta \end{pmatrix}, \tag{31}$$
with $\lambda_2(M) = 1 - q\delta$.

The reconstruction problem is unsolvable for (30) when $|b\lambda_2(M)| \le 1$ and unsolvable for (31) when $0 \le b\lambda_2(M) \le 1$ (see Propositions 3 and 4 in [23]). On the other hand, Proposition 5.1 implies that when $b\lambda_2^2(M) > 1$ the reconstruction problem is solvable for these channels. Moreover, the main results of [23] state that the reconstruction problem for (30) is solvable when $b\lambda_2(M) > 1$ and δ_1 is sufficiently small. Similarly, the reconstruction problem for (31) is solvable when $b\lambda_2(M) > 1$ and q is sufficiently large. In this section we improve the existing bounds for nonreconstruction by showing the following.

PROPOSITION 4.1. *Let M be defined by the transition matrix* (30). *Then the reconstruction problem for T_b is unsolvable when*
$$b\frac{(\delta_2 - \delta_1)^2}{\min\{\delta_1 + \delta_2, 2 - \delta_1 - \delta_2\}} \le 1. \tag{32}$$
Similarly, the reconstruction problem for a general tree T is unsolvable if
$$\mathrm{br}(T)\frac{(\delta_2 - \delta_1)^2}{\min\{\delta_1 + \delta_2, 2 - \delta_1 - \delta_2\}} < 1. \tag{33}$$

PROPOSITION 4.2. *Let M be the channel* (31). *Then the reconstruction problem for T_b is unsolvable when*

$$b\frac{(1-q\delta)^2}{1-(q-2)\delta}\le 1. \tag{34}$$

Similarly, the reconstruction problem for a general tree T is unsolvable if

$$\operatorname{br}(T)\frac{(1-q\delta)^2}{1-(q-2)\delta}<1. \tag{35}$$

Note that the claim of Proposition 1.3 is contained in Proposition 4.2. Propositions 4.2 and 4.1 follow from the following theorem.

THEOREM 4.3. *Let M be a channel and let i, j be two states such that for all $k\notin\{i,j\}$ it holds that $\mathbf{M}_{i,k}=\mathbf{M}_{j,k}$, and there exist $0\le\varepsilon\le 1$ and $\alpha\ge 0$, $\beta\ge 0$, $\gamma\ge 0$ such that*

$$\begin{pmatrix}\mathbf{M}_{i,i} & \mathbf{M}_{i,j}\\ \mathbf{M}_{j,i} & \mathbf{M}_{j,j}\end{pmatrix}=\alpha\begin{pmatrix}1-\varepsilon & \varepsilon\\ \varepsilon & 1-\varepsilon\end{pmatrix}+\begin{pmatrix}\beta & \gamma\\ \beta & \gamma\end{pmatrix}. \tag{36}$$

If T_b is the b-ary tree and $b\alpha(1-2\varepsilon)^2\le 1$, then $\lim_{n\to\infty}D_V(\mathbf{P}_n^i,\mathbf{P}_n^j)=0$. Similarly, if T is a general tree with $\operatorname{br}(T)\alpha(1-2\varepsilon)^2<1$, then $\lim_{n\to\infty}D_V(\mathbf{P}_n^i,\mathbf{P}_n^j)=0$.

PROOF OF PROPOSITION 4.1. If $\delta_1+\delta_2\le 1$ write

$$\begin{pmatrix}1-\delta_1 & \delta_1\\ 1-\delta_2 & \delta_2\end{pmatrix}=(\delta_1+\delta_2)\begin{pmatrix}\delta_2/(\delta_1+\delta_2) & \delta_1/(\delta_1+\delta_2)\\ \delta_1/(\delta_1+\delta_2) & \delta_2/(\delta_1+\delta_2)\end{pmatrix}+\begin{pmatrix}1-\delta_1-\delta_2 & 0\\ 1-\delta_1-\delta_2 & 0\end{pmatrix}$$

and the proposition follows from Theorem 4.3. Otherwise, we let $\delta_1'=1-\delta_1$, $\delta_2'=1-\delta_2$ and use the decomposition

$$\begin{aligned}\begin{pmatrix}1-\delta_1 & \delta_1\\ 1-\delta_2 & \delta_2\end{pmatrix}&=\begin{pmatrix}\delta_1' & 1-\delta_1'\\ \delta_2' & 1-\delta_2'\end{pmatrix}\\ &=(\delta_1'+\delta_2')\begin{pmatrix}\delta_1'/(\delta_1'+\delta_2') & \delta_2'/(\delta_1'+\delta_2')\\ \delta_2'/(\delta_1'+\delta_2') & \delta_1'/(\delta_1'+\delta_2')\end{pmatrix}\\ &\quad+\begin{pmatrix}0 & 1-\delta_1'-\delta_2'\\ 0 & 1-\delta_1'-\delta_2'\end{pmatrix}.\end{aligned}$$

□

PROOF OF PROPOSITION 4.2. For all $i\ne j$ we have

$$\begin{aligned}&\begin{pmatrix}\mathbf{M}_{i,i} & \mathbf{M}_{i,j}\\ \mathbf{M}_{j,i} & \mathbf{M}_{j,j}\end{pmatrix}\\ &\quad=(1-(q-2)\delta)\\ &\qquad\times\begin{pmatrix}(1-(q-1)\delta)/(1-(q-2)\delta) & \delta/(1-(q-2)\delta)\\ \delta/(1-(q-2)\delta) & (1-(q-1)\delta)/(1-(q-2)\delta)\end{pmatrix}\end{aligned}$$

and the proposition follows from Theorem 4.3. □

Below we use the "data processing lemma" which implies that if X, Y and Z are random variables such that X and Z are independent given Y, then $I(X, Z) \leq \min\{I(X, Y), I(Y, Z)\}$ (see, e.g., [7] for an exact formulation and for a proof).

PROOF OF THEOREM 4.3. Assume that the label of the root is chosen to be i or j with probability $1/2$ each and let X denote this symbol. For a set W, we denote by $\widetilde{Y}_W$ the symbols at the vertices of W, and $\widetilde{Y}_n = \widetilde{Y}_{L_n}$. We will show that

$$\lim_{n\to\infty} I(X, \widetilde{Y}_n) = 0. \tag{37}$$

Equation (37) is equivalent to $\lim_{n\to\infty} D_V(\mathbf{P}^i_n, \mathbf{P}^j_n) = 0$ (see, e.g., [23]).

We will split the proof into three main steps.

Step 1 [10]. If M is the binary symmetric channel,

$$\mathbf{M} = \begin{pmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{pmatrix}, \tag{38}$$

then for a general tree T and a set W,

$$I(X, \widetilde{Y}_W) \leq \sum_{w\in W} (1-2\varepsilon)^{2|w|}. \tag{39}$$

This is Theorem 1.3 in [10].

Step 2. We will show that it suffices to prove the theorem assuming that

$$\begin{pmatrix} \mathbf{M}_{i,i} & \mathbf{M}_{i,j} \\ \mathbf{M}_{j,i} & \mathbf{M}_{j,j} \end{pmatrix} = \alpha \begin{pmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{pmatrix} \tag{40}$$

instead of (36). Indeed, assume that the theorem is true under the condition (40). Now let M be a channel which satisfies (36). Consider the following auxiliary channel N. The channel has the state space $\mathcal{A}' = \mathcal{A} \cup \{i^*, j^*\}$ and the following transition matrix $\mathbf{N}$:

1. For all $\ell \notin \{i, j, i^*, j^*\}$ and all $\ell' \in \mathcal{A}$, set $\mathbf{N}_{\ell,\ell'} = \mathbf{M}_{\ell,\ell'}$.
2. $$\begin{pmatrix} \mathbf{N}_{i,i} & \mathbf{N}_{i,j} \\ \mathbf{N}_{j,i} & \mathbf{N}_{j,j} \end{pmatrix} = \alpha \begin{pmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{pmatrix}.$$
3. $$\begin{pmatrix} \mathbf{N}_{i,i^*} & \mathbf{N}_{i,j^*} \\ \mathbf{N}_{j,i^*} & \mathbf{N}_{j,j^*} \end{pmatrix} = \begin{pmatrix} \beta & \gamma \\ \beta & \gamma \end{pmatrix}.$$
4. For all ℓ, set $\mathbf{N}_{i^*,\ell} = \mathbf{N}_{i,\ell}$ and $\mathbf{N}_{j^*,\ell} = \mathbf{N}_{j,\ell}$.

It is clear that if the original channel M satisfies the conditions of Theorem 4.3, then the channel N satisfies these conditions with (40) replacing (36). By our assumption this implies that $\lim_{n\to\infty} I(X, \widehat{Y}_n) = 0$, where $\widehat{Y}_n$ is the labeling of level n for N.

Recalling that $\widetilde{Y}_n$ is the labeling of level n for M, we note that $\widetilde{Y}_n$ may be obtained from $\widehat{Y}_n$ by replacing each occurrence of i^* by i and each occurrence of j^* by j. By the data processing lemma it now follows that

$$I(X,\widetilde{Y}_n) \le I(X,\widehat{Y}_n) \to 0,$$

as needed.

Step 3. We prove the theorem assuming that (40) holds. We begin by introducing two random variables Z and Y. Let $\mathtt{path}(v)$ be the path from ρ to v. We let

$$Z = \bigl\{(v,\sigma_v) : \exists\, w \in \mathtt{path}(v), \sigma_w \notin \{i,j\}\bigr\}, \tag{41}$$

and given a set W, let

$$Y_W = \bigl\{(v,\sigma_v) : v \in W, \forall\, w \in \mathtt{path}(v), \sigma_w \in \{i,j\}\bigr\}. \tag{42}$$

We denote Y_{L_n} by Y_n. Roughly speaking, Y_n contains all information on the $i \to j$ and $j \to i$ process from the root to level n; Z contains all information which is independent of this process.

By the data processing lemma,

$$I(X,\widetilde{Y}_n) \le I\bigl(X,(Y_n,Z)\bigr). \tag{43}$$

Since for all $\ell \notin \{i,j\}$ it holds that $\mathbf{M}_{i,\ell} = \mathbf{M}_{j,\ell}$, it follows that Z is independent of X. We therefore obtain

$$\begin{aligned} I(X,(Y_n,Z)) &= H(X) + H(Y_n,Z) - H(X,Y_n,Z) \\ &= H(X|Z) + H(Y_n|Z) - H(X,Y_n|Z) \\ &= \mathbf{E}_z I(X,Y_n|Z=z). \end{aligned} \tag{44}$$

We will show that for almost all z,

$$\lim_{n\to\infty} I(X,Y_n|Z=z) = 0. \tag{45}$$

Since $0 \le I(X,Y_n|Z=z) \le 1$ for all z, this would imply the theorem by (44) and (43). By the data processing lemma, in order to prove (45), it suffices to prove that for a.e. z there exist cutsets W_n for which

$$\lim_{n\to\infty} I(X,Y_{W_n}|Z=z) = 0. \tag{46}$$

[Recall that W is a cutset if W intersects every infinite path emanating from ρ. If W_n is a cutset and for all $v \in W_n$ we have $\ell \le |v| \le \ell'$, then by the data processing lemma it follows that

$$I(X,Y_{\ell'}) \le I(X,Y_{W_n}) \le I(X,Y_\ell),$$

so (45) is equivalent to (46).]

The key observation is noting that given $Z = z$ we have a broadcast process with the binary symmetric channel (38) on the tree

$$T_z = \{v : \forall\, w \in \mathtt{path}(v), \sigma_w \in \{i, j\}\}.$$

Therefore by Step 1, in order to prove (46) it suffices to show that for a.e. z, there exist cutsets W_n such that

$$\lim_{n\to\infty} \sum_{w\in W_n} (1-2\varepsilon)^{2|w|} = 0. \tag{47}$$

However, for a general tree T, we have $\mathrm{br}(T_z) = \alpha\,\mathrm{br}(T)$ for a.e. z. So if $\alpha\,\mathrm{br}(T)(1-2\varepsilon)^2 < 1$, then (47) holds for a.e. z for appropriate W_n's (depending on z). Equation (47) holds for T_b when $b\alpha(1-2\varepsilon)^2 = 1$ by [21], Theorem 3. □

5. $\mathbf{br}(T)|\lambda_2|^2 > 1$ implies reconstruction.

The following proposition is a consequence of the results of [18].

PROPOSITION 5.1. *Suppose that* $b|\lambda_2(M)|^2 > 1$, *then the reconstruction problem is census-solvable.*

In this section we prove an extension of the proposition to general trees. The proof is based on a second-moment argument generalizing further an argument for the symmetric Ising model for regular trees in [13], which was generalized to the symmetric Ising model on general trees in [10].

THEOREM 1.4. *Let T be an infinite tree and M a channel such that* $\mathrm{br}(T)|\lambda_2(M)|^2 > 1$, *then the reconstruction problem for M on T is solvable.*

In order to prove this theorem we are going to bound from below $D_{\chi^2}(\mathbf{P}_n^{(c),i}, \mathbf{P}_n^{(c),j})$ where

$$D_{\chi^2}(P, Q) = \frac{1}{2}\sqrt{\sum_\sigma \frac{2(P(\sigma)-Q(\sigma))^2}{P(\sigma)+Q(\sigma)}} \tag{48}$$

is the χ^2 distance.

LEMMA 5.2.

$$D_{\chi^2}^2 \le D_V \le D_{\chi^2}. \tag{49}$$

PROOF. The first inequality in (49) follows when we use the estimate

$$\frac{2|P(\sigma)-Q(\sigma)|}{P(\sigma)+Q(\sigma)} \le 2,$$

while the second follows by the Cauchy–Schwarz inequality when we write

$$\frac{1}{2}|P(\sigma) - Q(\sigma)| = \frac{|P(\sigma) - Q(\sigma)|}{\sqrt{2(P(\sigma) + Q(\sigma))}} \sqrt{\frac{P(\sigma) + Q(\sigma)}{2}}. \qquad \square$$

We prove Theorem 1.4 by constructing linear estimators of the root variable for finite trees and then evaluating the first and second moments of these estimators.

We use the notion of *flows* and view the tree as an *electrical network*. We refer the reader to [9] and [10] for definitions and more background. We say that a set of vertices W is an *anti-chain* if no vertex in W is a descendant of another.

LEMMA 5.3. *Abbreviate $\lambda = \lambda_2(M)$, and take $v \neq 0$, s.t. $\mathbf{M}v^t = \lambda v^t$, with $|v|_2 = 1$. Let T be a finite tree, and consider the tree as an electrical network where the edge e is assigned the resistance*

$$R(e) = (1 - |\lambda|^2)|\lambda|^{-2|e|}.$$

Let W be an antichain in T, and μ a unit flow from ρ to W. For a vertex x, we let c_x be the ith unit vector, $c_x = e_i$, if the label of x is i, that is, if $\sigma_x = i$. Consider the estimator

$$S_\mu = \sum_{x \in W} \frac{\mu(x) c_x v^t}{\lambda^{|x|}}. \tag{50}$$

Then for all $\ell \in \mathcal{A}$,

$$\mathbf{E}^\ell[S_\mu] = e_\ell v^t, \tag{51}$$

and there exists a constant $0 < c(M) < 1$, which depends on M only such that

$$\begin{aligned} c(M)\big(1 + \mathcal{R}_{\text{eff}}(\rho \leftrightarrow W)\big) &\le \min_\mu \min_l \mathbf{E}^\ell[|S_\mu|^2] \le \min_\mu \max_l \mathbf{E}^\ell[|S_\mu|^2] \\ &\le \big(1 + \mathcal{R}_{\text{eff}}(\rho \leftrightarrow W)\big). \end{aligned} \tag{52}$$

PROOF. If x is at level n, then

$$\mathbf{E}^\ell[c_x] = e_\ell \mathbf{M}^n. \tag{53}$$

It follows from (53) that for every vector w,

$$\mathbf{E}^\ell[c_x w^t] = e_\ell \mathbf{M}^n w^t. \tag{54}$$

Since v is an eigenvalue of $\mathbf{M}$ which corresponds to $\lambda = \lambda_2(M)$, we obtain

$$\mathbf{E}^\ell[c_x v^t] = \lambda^n e_\ell v^t. \tag{55}$$

Now (51) follows by linearity.

Recall that $\mathtt{path}(x)$ is the path from ρ to x [more generally let $\mathtt{path}(x, y)$ be the path from x to y]. We let $x \wedge y$, be the *meeting point* of x and y, that is, the vertex farthest from the root ρ on $\mathtt{path}(x) \cap \mathtt{path}(y)$. We have

$$\begin{aligned} |\mathbf{E}^{\ell}[(c_x v^t)\overline{(c_y v^t)}]| &= \left|\sum_i \mathbf{P}^{\ell}[c_{x\wedge y} = e_i]\mathbf{E}[|(c_x v^t)\overline{(c_y v^t)}|\,|c_{x\wedge y} = e_i]\right| \\ &\le \sum_i \mathbf{P}^{\ell}[c_{x\wedge y} = e_i]|e_i v^t|^2|\lambda|^{d(x,y)} \le |\lambda|^{d(x,y)}. \end{aligned} \tag{56}$$

Therefore,

$$\mathbf{E}^{\ell}[|S_\mu|^2] \le \sum_{x,y} \frac{\mu(x)\mu(y)}{|\lambda|^{|x|}|\lambda|^{|y|}}|\mathbf{E}^{\ell}[(c_x v^t)\overline{(c_y v^t)}]| \le \sum_{x,y} \frac{\mu(x)\mu(y)}{|\lambda|^{2|x\wedge y|}}. \tag{57}$$

Since

$$\frac{1}{|\lambda|^{2|u|}} = 1 + \sum_{e \in \mathtt{path}(u)} R(e),$$

if follows that for all ℓ,

$$\begin{aligned} \mathbf{E}^{\ell}[|S_\mu|^2] &\le \left(1 + \sum_e R(e) \sum_{x,y \in W} \mathbf{1}_{e \in \mathtt{path}(x\wedge y)}\mu(x)\mu(y)\right) \\ &= \left(1 + \sum_e R(e)\mu^2(e)\right). \end{aligned} \tag{58}$$

When we take the minimum in (58) we obtain,

$$\min_\mu \max_\ell \mathbf{E}^{\ell}[|S_\mu|^2] \le \min_\mu \left(1 + \sum_e R(e)\mu^2(e)\right) = \big(1 + \mathscr{R}_{\mathrm{eff}}(\rho \leftrightarrow W)\big), \tag{59}$$

where the equality in (59) is Thompson's principle (see [9]). Equation (59) is the right-hand side inequality in (52). We omit the proof of the other inequality in (52) as it is similar and it is not used in the proof of Theorem 1.4. □

PROOF OF THEOREM 1.4. Since $|\lambda|^2 \operatorname{br}(T) > 1$, it follows that

$$\mathscr{R}_{\mathrm{eff}}(\rho \leftrightarrow \infty) = \sup_n \mathscr{R}_{\mathrm{eff}}(\rho \leftrightarrow L_n) < \infty. \tag{60}$$

We consider linear estimators as in Lemma 5.3, and note that it cannot be that v is a multiple of $(1, \dots, 1)$, since by the Perron–Frobenius theorem the eigenvector which corresponds to the eigenvalue 1 is the unique eigenvector which has only positive entries. In particular,

$$\widetilde{C}(M) = \max_{i,j} |e_i v^t - e_j v^t| > 0. \tag{61}$$

We fix a level n, take i and j such that we obtain the maximum in (61), and consider a linear estimator as in Lemma 5.3 for $W = L_n$, such that the upper bound in (52) holds. We obtain that

$$\left|\sum_\sigma S_\mu(\sigma)(\mathbf{P}_n^i[\sigma] - \mathbf{P}_n^j[\sigma])\right| = \widetilde{C}(M). \tag{62}$$

On the other hand, by the Cauchy–Schwarz inequality we obtain that

$$\begin{aligned}&\left(\sum_\sigma S_\mu(\sigma)(\mathbf{P}_n^i[\sigma] - \mathbf{P}_n^j[\sigma])\right)^2\\ &\quad\le \sum_\sigma |S_\mu(\sigma)|^2(\mathbf{P}_n^i[\sigma] + \mathbf{P}_n^j[\sigma]) \sum_\sigma \frac{(\mathbf{P}_n^i[\sigma] - \mathbf{P}_n^j[\sigma])^2}{\mathbf{P}_n^i[\sigma] + \mathbf{P}_n^j[\sigma]} \qquad (63)\\ &\quad\le 4\big(1 + \mathcal{R}_{\mathrm{eff}}(\rho \leftrightarrow L_n)\big) D_{\chi^2}^2(\mathbf{P}_n^i, \mathbf{P}_n^j).\end{aligned}$$

Combining (62) and (63) we obtain that, for all n,

$$D_{\chi^2}^2(\mathbf{P}_n^i, \mathbf{P}_n^j) \ge \frac{\widetilde{C}(M)^2}{4(1 + \mathcal{R}_{\mathrm{eff}}(\rho \leftrightarrow L_n))}. \tag{64}$$

So the theorem follows by Lemma 5.2. □

PROOF OF PROPOSITION 5.1. For the b-ary tree T_b we have $\mathrm{br}(T_b) = b$. Moreover, the flow μ which minimizes $\sum_e R(e)\mu^2(e)$ in (59) is the flow which satisfies $\mu(e) = b^{-|e|}$. This implies that we may take S_n to be

$$S_n = \frac{c_n v^t}{b^n \lambda^n}.$$

In particular S_n is a function of c_n. Now, arguing as in the proof of Theorem 1.4, we see that there exist i, j for which $D_V(\mathbf{P}_n^{(c),i}, \mathbf{P}_n^{(c),j})$ is bounded away from 0 for all n. □

6. Census reconstruction fails when $b|\lambda_2|^2 < 1$. In this section we prove nonreconstruction when $b|\lambda_2|^2 < 1$. The proof relies on the Kesten–Stigum CLT [18]. Let M be a channel and consider the process for M on the b-ary tree. From the Kesten–Stigum limit theorem [17], it follows that a.s. $\lim_{n\to\infty} c_n^j/b^n = \pi$, where π is the stationary distribution of M. In the Kesten–Stigum CLT [18], the asymptotic behavior of c_n^j is studied by proving limit theorems for $c_n^j v^t$ for v's which satisfy $\pi v^t = 0$. The following is an immediate consequence of Theorem 2.3 of [18].

KESTEN–STIGUM CLT [18]. *Let M be a transition matrix such that $b|\lambda_2(M)|^2 < 1$ and let π be the stationary distribution for M; that is, the nor-*

malized left eigenvector for the eigenvalue 1. *Then for any vector v which is orthogonal to π, and for all $j \in \mathcal{A}$,*

$$\frac{c_n^j v^t}{b^{n/2}} \to \mathcal{N}(0, \sigma), \tag{65}$$

where $\mathcal{N}$ denotes a normal random variable and σ does not depend on j. Moreover, the convergence rate may be bounded in terms of $|v|_2$ only.

An immediate consequence is the following.

PROPOSITION 6.1. *For all $i, j \geq 1$, all $\varepsilon > 0$ and all n, one may couple c_n^j and c_n^i on a space with probability measure $\mathbf{P}_n$ in such a way that*

$$\lim_{n\to\infty} \mathbf{P}_n[|c_n^j - c_n^i| > \varepsilon b^{n/2}] = 0. \tag{66}$$

PROOF. If follows from the Kesten–Stigum CLT that

$$\frac{c_n^\ell - b^n\pi}{b^{n/2}} \to \mathcal{N},$$

where $\mathcal{N}$ is a normal vector which does not depend on ℓ. The claim follows, as for all $\delta > 0$, for sufficiently large n, we can couple both c_n^i and c_n^j with the same normal variable $b^n\pi + b^{n/2}\mathcal{N}$ in such a way that

$$\mathbf{P}[|c_n^\ell - b^n\pi - b^{n/2}\mathcal{N}| > \varepsilon b^{n/2}] < \delta,$$

for both $\ell = i$ and $\ell = j$. □

THEOREM 1.5. *If $b|\lambda_2(M)|^2 < 1$, then the reconstruction problem is not census-solvable for the b-ary tree and the channel M.*

The proof is easier when all the entries of $\mathbf{M}$ are strictly positive, in which case the following lemma is trivial.

LEMMA 6.2. *Let M be irreducible and aperiodic, then there exists an h such that, for all i and j, all $n \geq h$ and all v,*

$$\mathbf{P}[c_n^i = v] > 0 \qquad \textit{iff}\ \mathbf{P}[c_n^j = v] > 0. \tag{67}$$

PROOF. Since $(\mathbf{M}_{i,j})_{i,j=1}^k$ is irreducible and aperiodic it follows that there exists an ℓ such that the all the entries of matrix $\mathbf{M}^\ell$ are strictly positive. Thus, for every two states i and j we may construct an ℓ level tree such that the root of the tree is labeled by i and all of the leaves are labeled by j. Let $h = \ell(k+4)$ and $n \geq h$. Suppose that $\mathbf{P}[c_n^i = v] > 0$. Therefore there exists a labeling of the tree of n levels which has i at the root and σ at level n, and such the census of σ is v.

We will prove the lemma by constructing a labeling where the root is labeled by j, level n is labeled by τ, and the census of τ is v.

We denote by $x_1, \dots, x_{b^{n-\ell}}$ the vertices at level $n-\ell$. We define $c(x_t)$ to be the census of the subtree rooted at x_t for the labeling σ. For a nonnegative vector $u = (u_1, \dots, u_k)$ which satisfies $\sum u_i = b^\ell$, define $c(\sigma, u) = |\{t : c(x_t) = u\}|$. Note that if $c(\sigma, u) > 0$, then there exists a labeling of the ℓ level tree, denoted $\tau(u)$, such that the labeling of the root is $\sigma_\rho(u)$ and $\tau(u)$ has u as its census.

Since

$$v - \sum_u \bigl(c(\sigma, u) \mod b^\ell\bigr)u = 0 \qquad \mod b^\ell,$$

it follows that

$$v = b^\ell w + \sum_u \bigl(c(\sigma, u) \mod b^\ell\bigr)u,$$

for some integer-valued vector w.

Note that if for root value j, we could label $\sum_u (c(\sigma, u) \mod b^\ell)$ of the vertices of level $n-\ell$ by labels $\sigma_\rho(u)$ with multiplicity $c(\sigma, u) \mod b^\ell$, then we are done. Since using these vertices, it is possible to construct $\sum_u (c(\sigma, u) \mod b^\ell)u$ of the census. Now using the other vertices at level $n-\ell$ it is possible to build the $b^\ell w$ part of the census (whatever labels these vertices have). Note that

$$\sum_u \bigl(c(\sigma, u) \mod b^\ell\bigr) \le b^\ell \left| \left(u : \forall i, u_i \ge 0, \sum_{i=1}^k u_i = b^\ell \right) \right| \le b^\ell b^{k\ell} = b^{(k+1)\ell}.$$

Therefore by assigning $b^{(k+1)\ell}$ of the vertices at level $n - 2\ell$ the task of producing the prescribed labels at level $n-\ell$ we obtain the required result. □

PROOF OF THEOREM 1.5. We take h such that (67) holds for $m \ge h$. Fixing the level n, we may write

$$c^j_{n+h} = \sum_{i=1}^k S^i\bigl(c^j_n(i)\bigr), \tag{68}$$

where S^i is a random walk on $\mathbf{Z}^k$ which satisfies

$$\mathbf{P}[S^i(t+1) = S^i(t) + v] = \mathbf{P}[c^i_h = v].$$

By (67), all the S^i are random walks with the same support.

The Kesten–Stigum limit theorem [17] implies that $\lim_{n\to\infty} c^j_n / b^n = \pi$, and therefore with probability going to 1 as $n \to \infty$, for all i and j it holds that $c^j_n(i) > 0.5 b^n \min_\ell \pi_\ell$. By the local CLT it now follows that if $|c^j_n - c^i_n| < \varepsilon\sqrt{b^n}$, then it is possible to couple c^j_{n+h} and c^i_{n+h} in such a way that $\mathbf{P}[c^j_{n+h} \ne c^i_{n+h}] \le f(\varepsilon, n)$, where $f(\varepsilon, n) \to 0$ as $\varepsilon \to 0$ and $n \to \infty$. Now the result follows from Proposition 6.1. □

7. Census reconstruction for Potts models fails at criticality. In Section 6 we presented a proof that $b|\lambda_2|^2 < 1$ implies that the reconstruction problem is not census-solvable. We believe that this result is also valid when $b|\lambda_2|^2 = 1$. In this section we prove that this is the case when M is the q-state Potts model (31) or the asymmetric Ising model (30).

THEOREM 1.6. *Let M be the q-state Potts model* (31), *or the asymmetric Ising model* (30), *and suppose that $b\lambda_2^2(M) \le 1$, then the reconstruction problem is unsolvable.*

PROOF. We prove the theorem for the three-state Potts model, the general proof being similar. It is convenient to denote the states of the channel by -1, 0 and 1. We show that for all $i, j \in \{-1, 0, 1\}$ there exists a coupling such that $\mathbf{P}[c_n^i = c_n^j] \to 1$ as $n \to \infty$. By symmetry, it suffices to find such a coupling for c_n^1 and c_n^{-1}. We denote by σ^+ the coloring with 1 at the root and by σ^- the coloring with -1 at the root.

During all the steps of the coupling we require that for all v we have $\sigma_v^- = 0$ iff $\sigma_v^+ = 0$. The main step in proving the existence of the required coupling is given in the following lemma.

LEMMA 7.1. *There exists $p^* > 0$ such that given $(\sigma_v)_{|v|\le n}$ and $(\tau_v)_{|v|\le n}$ which are coupled in such a way that $\sigma_v = 0$ iff $\tau_v = 0$, there exists $N \ge n$ and a coupling procedure for $(\sigma_v)_{|v|\le N}$ and $(\tau_v)_{|v|\le N}$ such that $\sigma_v = 0$ iff $\tau_v = 0$ and such that $\mathbf{P}[c_N^\sigma = c_N^\tau] \ge p^*$* (*where c_N^σ and c_N^τ are the σ and τ level N census vectors*, *respectively*).

We apply Lemma 7.1 for $n_1 = 1$, and $\sigma_\rho = 1, \tau_\rho = -1$. We obtain $n_2 > n_1$ s.t. $\mathbf{P}[c_{n_1}^\sigma = c_{n_1}^\tau] \ge p^*$. On the event that $c_{n_2}^\sigma = c_{n_2}^\tau$, it is easy to couple σ and τ in such a way that for all $n \ge n_2$, $c_n^\sigma = c_n^\tau$. When $c_{n_2}^\sigma \ne c_{n_2}^\tau$, we apply Lemma 7.1 to obtain $n_3 > n_2$, s.t. $\mathbf{P}[c_{n_3}^\sigma = c_{n_3}^\tau | c_{n_2}^\sigma \ne c_{n_2}^\tau] \ge p^*$, so $\mathbf{P}[c_{n_3}^\sigma = c_{n_3}^\tau] \ge p^* + (1 - p^*)p^*$. Iterating the above argument k times we obtain that for $n \ge n_k$, the coupling probability satisfies $\mathbf{P}[c_n^1 = c_n^{-1}] \ge 1 - (1 - p^*)^{k-1}$, as needed. □

PROOF OF LEMMA 7.1. Recall that the Potts model has matrix (4), where $\lambda_2(M) = 1 - 3\delta$. We let $\theta = \lambda_2(M) = 1 - 3\delta$. At level n there are two types of vertices: those for which $\sigma_v = \tau_v$, and those for which $\sigma_v = -\tau_v \in \{-1, 1\}$. For all the vertices w whose ancestors v at level n satisfy $\sigma_v = \tau_v$, we let $\sigma_w = \tau_w$.

Denote the vertices v at level n which satisfy $\sigma_v = -\tau_v \in \{-1, 1\}$ by $v_1, \dots, v_\ell$. Let T be the graph which consists of the subtrees rooted at $v_1, \dots, v_\ell$.

We think of T as drawn in the plane, and denote the vertices of level m of this graph by $v_1^m, \dots, v_{\ell b^{n-m}}^m$. We will slightly abuse the notation by redefining

$$\begin{aligned} c_m^\sigma(i) &= |\{j : 1 \le j \le \ell b^{n-m},\ \sigma(v_j^m) = i\}|, \\ c_m^\tau(i) &= |\{j : 1 \le j \le \ell b^{n-m},\ \tau(v_j^m) = i\}|. \end{aligned} \tag{69}$$

Note that $c^\sigma_m = c^\tau_m$ in the old definition iff it does in the new one.

We define $X^\sigma_m = c^\sigma_m(1) - c^\sigma_m(-1)$ and $X^\tau_m = c^\tau_m(1) - c^\tau_m(-1)$. Note that if a coupling of σ and τ satisfies $\sigma_v = 0$ iff $\tau_v = 0$, then we have $c^\sigma_m = c^\tau_m$ iff $X^\sigma_m = X^\tau_m$.

We define the coloring of T as a dynamic process starting at the root, running level by level from left (v^m_1) to right ($v^m_{\ell b^{n-m}}$). Writing v' for the parent of v, we let

$$\begin{aligned} S^\sigma(v^m_i) &= (b\theta)^{-n} X^\sigma_n + \sum_{n \le k < m} (b\theta)^{-k} \sum_{j \le \ell b^{k-n}} \bigl(\sigma(v^k_j) - \theta\sigma(v^{k'}_j)\bigr) \\ &\quad + (b\theta)^{-m} \sum_{j \le i} \bigl(\sigma(v^m_j) - \theta\sigma(v^{m'}_j)\bigr), \end{aligned} \tag{70}$$

where we formally define $\sigma_{\rho'} = 0$. We define S^τ similarly. Note that all the terms in (70) except the first one have mean zero, and therefore both S^σ and S^τ are martingales. Also,

$$\begin{aligned} S^\sigma(v^m_{\ell b^{m-n}}) &= (b\theta)^{-m} X^\sigma_m, \\ S^\tau(v^m_{\ell b^{m-n}}) &= (b\theta)^{-m} X^\tau_m. \end{aligned} \tag{71}$$

The coupling consists of a few steps, during all of which we require that for all v we have $\sigma_v = 0$ iff $\tau_v = 0$.

1. Reflection until we reach N such that

$$-\sqrt{b^N} C_1 \le X^\sigma_N = -X^\tau_N \le C_1 \sqrt{b^N}. \tag{72}$$

We label the vertices level by level, left to right using the rule $\sigma_v = -\tau_v$, until we reach v^N_i with $|S^\sigma(v^N_i)| \le (1+\theta)(b\theta)^{-N}$ and N large. From Lemma 7.2 below, it follows that such a vertex exists a.s. Now $X^\sigma_N - (b\theta)^N S^\sigma(v^N_i)$ is a sum of $\ell b^{N-n} - i$ independent variables with values $-1, 0, 1$ and expected value 0. It therefore follows by the CLT that with probability going to 1 as $C_1 \to \infty$ we have $|X^\sigma_N| \le C_1 \sqrt{b^N}$. Therefore, by continuing the reflection $\sigma(v^N_j) = -\tau(v^N_j)$ for $i < j \le \ell b^{N-n}$, we obtain (72). We let $Y^\sigma_{N+1}(i,j)$ be the number of vertices in the σ labeling at level $N+1$ of type j and having a type i parent. All these variables are binomial and

$$\mathbf{E}\bigl[Y^{\sigma,\tau}_{N+1}(i,j) | c^{\sigma,\tau}_N\bigr] = \bigl(b(1-\theta)/3 + \delta_{i,j} b\theta\bigr) c^{\sigma,\tau}_N(i). \tag{73}$$

2. Labeling the 0's of level $N+1$. We label all vertices of level $N+1$ with $\sigma(v^{N+1}_j) = \sigma(v^{N+1}_j) = 0$ independently with probability

$$\mathbf{P}\bigl[\sigma(v^{N+1}_j) = 0 | \sigma\bigl((v^{N+1}_j)'\bigr)\bigr] = \mathbf{P}\bigl[\tau(v^{N+1}_j) = 0 | \tau\bigl((v^{N+1}_j)'\bigr)\bigr].$$

We thus achieve that $c^\sigma_{N+1}(0) = c^\tau_{N+1}(0)$.

By the CLT it follows that with probability going to 1 as $C_3 \to \infty$,

$$\bigl|Y^{\sigma,\tau}_{N+1}(i,0) - \mathbf{E}\bigl[Y^{\sigma,\tau}_{N+1}(i,0) | c^{\sigma,\tau}_N\bigr]\bigr| \le C_3 \sqrt{b^{N+1}},$$

and therefore for $j \neq 0$,

$$|\mathbf{E}[Y^{\sigma,\tau}_{N+1}(i,j)|c^{\sigma,\tau}_N, Y^{\sigma,\tau}_{N+1}(i,0)] - \mathbf{E}[Y^{\sigma,\tau}_{N+1}(i,j)|c^{\sigma,\tau}_N]| \leq C_3\sqrt{b^{N+1}}. \tag{74}$$

3. Labeling some of the ± of level $N+1$. Let i_0 be the maximizer of $c^{\sigma}_N(i)$. For $i \neq i_0$ we label all vertices v at level $N+1$ which have a parent of type i and satisfy $\sigma(v) \neq 0$ by ± 1 using the reflecting coupling $\sigma(v) = -\tau(v)$. By the CLT it follows that with probability going to 1 as $C_4 \to \infty$ we have for $i \neq i_0$ and all j,

$$|Y^{\sigma,\tau}_{N+1}(i,j) - \mathbf{E}[Y^{\sigma,\tau}_{N+1}(i,j)|c^{\sigma,\tau}_N, Y^{\sigma,\tau}_{N+1}(i,0)]| \leq C_4\sqrt{b^{N+1}}. \tag{75}$$

4. Local CLT for the final coupling. We now want to label the remaining vertices with ± 1 to achieve $X^{\sigma}_{N+1} = X^{\tau}_{N+1}$. We note that $Z^{\sigma,\tau} = Y^{\sigma,\tau}(i_0,1) - Y^{\sigma,\tau}(i_0,-1)$ are sums of i.i.d. ± 1 random variables and therefore satisfy the local CLT. Moreover, by (74), it follows that with high probability each of the Z's is a sum of at least $\ell b^{N+1-n}(1-\theta)/6$ such variables. In order that $X^{\sigma}_{N+1} = X^{\tau}_{N+1}$ we need that $\sum_i Y^{\sigma}(i,1) - Y^{\sigma}(i,-1) = \sum_i Y^{\tau}(i,1) - Y^{\tau}(i,-1)$. By (73)–(75) it follows that for $i \neq i_0$ and with probability going to 1 as $C_5 \to \infty$,

$$|(Y^{\sigma,\tau}(i,1) - Y^{\sigma,\tau}(i,-1)) - b\theta c^{\sigma,\tau}_N(i)| \leq C_5\sqrt{b^{N+1}}. \tag{76}$$

Similarly, by (73) and (74),

$$\begin{aligned}&|\mathbf{E}[Y^{\sigma,\tau}(i_0,1) - Y^{\sigma,\tau}(i_0,-1)|c^{\sigma,\tau}_N, Y^{\sigma,\tau}_{N+1}(i,0)] - b\theta c^{\sigma,\tau}_N(i_0)|\\ &\quad \leq C_5\sqrt{b^{N+1}}.\end{aligned} \tag{77}$$

Now by (76), (77), (72) and the fact that until now we used reflection, it follows that $Z^{\sigma} + W^{\sigma} = Z^{\tau} + W^{\tau}$ where $W^{\sigma,\tau}$ are random variables which satisfy $W^{\sigma} = -W^{\tau}$ and with probability going to 1 as $C_6 \to \infty$, $|W^{\sigma} - W^{\tau} - (\mathbf{E}[Z^{\sigma}] - \mathbf{E}[Z^{\tau}])| \leq C_6\sqrt{b^{N+1}}$. By the local CLT, it follows that with probability $p^* > 0$ we may couple Z^{σ} and Z^{τ} in order to achieve $X^{\sigma}_N = X^{\tau}_N$. □

LEMMA 7.2. *S^{σ} changes sign infinitely often a.s.*

PROOF. We prove that X^{σ}_n changes sign infinitely often a.s. By the Borel–Cantelli lemma it suffices to prove that for all n and all $\sigma(v)_{|v|\leq n}$ there exists $m \geq n$ such that $\mathbf{P}[X^{\sigma}_m > 0|\sigma(v)_{|v|\leq n}] \geq 1/4$. However, by the Kesten–Stigum theorem [18], there exists some $a_m \to \infty$ such that given $\sigma(v)_{|v|\leq n}$, X_m/a_m converges to a nondegenerate normal random variable with expected value 0. So when m is large, $\mathbf{P}[X^{\sigma}_m > 0|\sigma(v)_{|v|\leq n}] \geq 1/4$ as needed. □

8. Unsolved problems.

8.1. *Critical values.* Except for symmetric binary channels, there are no interesting families of channels for which the critical value for reconstruction is known.

PROBLEM 1. *For the three-state Potts model on the b-ary tree* (31), *find the critical value for reconstruction.*

The fact that there exists a critical δ_c s.t. the reconstruction problem is solvable when $\delta < \delta_c$ and unsolvable when $\delta > \delta_c$ follows from Proposition 12 in [23].

The same question applies to asymmetric binary channels, and to proper colorings (which correspond to the zero temperature anti-ferromagnetic Potts model). Proper colorings of trees were studied in [5]. The corresponding transition matrix is

$$\mathbf{M} = \frac{1}{q-1}\begin{pmatrix} 0 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots \\ \vdots & \dots & \ddots & \vdots \\ 1 & \dots & 1 & 0 \end{pmatrix}. \tag{78}$$

A simple coupling argument (see [5]) shows that if $b \le q-1$, then the reconstruction problem is unsolvable for colorings of T_b. On the other hand, applying standard coupon-collector estimates recursively (similarly to Theorem 2.1), it is easy to see that if $b \ge (1+\varepsilon)q\log q$ and q is large, then the reconstruction problem is solvable. By Proposition 5.1 and Theorem 1.5, the census reconstruction problem is solvable if $b > (q-1)^2$, and unsolvable if $b < (q-1)^2$.

PROBLEM 2. *For fixed q, for which b is the reconstruction problem solvable for the channel* (78)?

8.2. *Soft inputs*: *robust phase transitions.* We would like to mention briefly the notion of "robust" phase transition which first appeared in [24]. Consider the usual reconstruction problem, but suppose that the data at the boundary is given with some additional noise. The proofs that if $b\lambda_2^2(M) > 1$, then the reconstruction problem is (census) solvable are immune to this noise. However, it may be the case that the threshold for the reconstruction problem changes when the inputs are replaced by noisy inputs. Indeed, we suspect that adding this additional noise (assuming it is fixed but sufficiently strong) will shift the phase transition to the point $b\lambda_2^2(M) = 1$. A similar phenomenon was proved in [24] for the phase transition of uniqueness of Gibbs measures. For n and m, we denote by $\sigma_{n,m}$ the configuration which is obtained from σ_n by applying the random function M^m independently to each of the symbols in σ_n. We denote by $\mathbf{P}_{n,m}^{\ell}$ the conditional distribution on $\sigma_{n,m}$ given that $\sigma_\rho = \ell$. We then have following.

CONJECTURE 1. *For all M and b, such that $b\lambda_2^2(M) < 1$, there exists m such that*

$$\sup_{i,j} \lim_{n\to\infty} D_V(\mathbf{P}^i_{n,m}, \mathbf{P}^j_{n,m}) = 0.$$

8.3. *Monotonicity.* For Potts models (4), it is easy to see that if the reconstruction problem is solvable for q and δ and $q' < q$, then the reconstruction problem is also solvable for q' and δ. Here is a sketch of the argument. Suppose that the reconstruction problem is not solvable for q' on the b-ary tree. We construct a coupling of measures with different root values for q on the b-ary tree. Consider the measure with 1 at the root, and the measure with 2 at the root. The two measures can be coupled in such a way that all transitions to one of the $q - q'$ states are the same for both measures. Thus in order to couple these measure, it suffices to couple two q' measures on a subtree of the b-ary tree. This can be done with high probability for the full b-ary tree, and therefore (using projection) for any subtree of the b-ary tree. Now the claim follows.

We expect that for fixed $\lambda = \lambda_2(M)$, reconstruction is easier when q is larger.

CONJECTURE 2. *Consider two symmetric channels M_1 and M_2 [as in (4)] on q_1 and q_2 symbols, respectively, where $q_1 < q_2$. If $\lambda_2(M_1) = \lambda_2(M_2)$ and the reconstruction problem is solvable for M_1, then it is also solvable for M_2.*

This is obvious when q_2 is a multiple of q_1. Using the reconstruction criteria for the binary symmetric channel on 2 symbols, it is easy to prove the conjecture when $q_1 = 2$.

Acknowledgments. Most of this work was done when the first author was visiting the Algo group in INRIA Rocquencourt and when the second author was at Hebrew University of Jerusalem. Parts of this research were conducted in the Theory group at Microsoft research. We thank Olle Häggström, Claire Kenyon, László Lovász, Jeff Steif and Peter Winkler for helpful discussions. Thanks to Noam Berger, Alan Hammond, Olle Häggström, Svante Janson, Serban Nacu, Gabor Pete and an anonymous referee for their comments on previous drafts of this paper.

REFERENCES

[1] ATHREYA, K. B. and NEY, P. E. (1972). *Branching Processes*. Springer, New York.

[2] 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. Statist. Phys.* **79** 473–482.

[3] BRIGHTWELL, G. and WINKLER, P. (1999). Graph homomorphisms and phase transitions. *J. Combin. Theory Ser. B* **77** 415–435.

[4] BRIGHTWELL, G. and WINKLER, P. (2000). Gibbs measures and dismantlable graphs, *J. Combin. Theory Ser. B* **78** 141–169.

[5] BRIGHTWELL, G. and WINKLER, P. (2001). Random colorings of a Cayley tree. Unpublished manuscript.

[6] CAVENDER, J. (1978). Taxonomy with confidence. *Math. Biosci.* **40** 271–280.

[7] COVER, T. M. and THOMAS, J. A. (1991). *Elements of Information Theory*. Wiley, New York.

[8] DEMBO, A. and ZEITOUNI, O. (1997). *Large Deviations, Techniques and Applications*. Springer, New York.

[9] DOYLE, P. G. and SNELL, E. J. (1984). *Random Walks and Electrical Networks.* Math. Assoc. Amer., Washington, DC.

[10] EVANS, W., KENYON, C., PERES, Y. and SCHULMAN, L. J. (2000). Broadcasting on trees and the Ising model. *Ann. Appl. Probab.* **10** 410–433.

[11] FITCH, W. M. (1971). Toward defining the course of evolution: Minimum change for a specific tree topology. *Systematic Zoology* **20** 406–416.

[12] HAJEK, B. and WELLER, T. (1991). On the maximum tolerable noise for reliable computation by formulas. *IEEE Trans. Inform. Theory* **37** 388–391.

[13] HIGUCHI, Y. (1977). Remarks on the limiting Gibbs state on a $(d+1)$-tree. *Publ. RIMS Kyoto Univ.* **13** 335–348.

[14] 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.

[15] IOFFE, D. (1996). A note on the extremality of the disordered state for the Ising model on the Bethe lattice. In *Trees* (B. Chauvin, S. Cohen and A. Rouault, eds.). Birkhäuser, Boston.

[16] KENYON, C., MOSSEL, E. and PERES, Y. (2001). Glauber dynamics on trees and hyperbolic graphs (extended abstract). In *Proccedings of FOCS 2001*. To appear.

[17] KESTEN, H. and STIGUM, B. P. (1966). Limit theorems for decomposable multidimensional Galton–Watson processes. *J. Math. Anal. Appl.* **17** 309–338.

[18] KESTEN, H. and STIGUM, B. P. (1966). Additional limit theorem for indecomposable multidimensional Galton–Watson processes. *Ann. Math. Statist.* **37** 1463–1481.

[19] LOVÁSZ, L. and WINKLER, P. (1998). Mixing times. *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.* **41** 85–133.

[20] LYONS, R. (1990). Random walks and percolation on trees. *Ann. Probab.* **18** 931–958.

[21] LYONS, R. and PEMANTLE, R. (1992). Random walk in a random environment and first-passage percolation on trees. *Ann. Probab.* **20** 125–136.

[22] MOSSEL, E. (1998). Recursive reconstruction on periodic trees. *Random Structures Algorithms* **13** 81–97.

[23] MOSSEL, E. (2001). Reconstruction on trees: Beating the second eigenvalue. *Ann. Appl. Probab.* **11** 285–300.

[24] PEMANTLE, R. and STEIF, J. E. (1999). Robust phase transitions for Heisenberg and other models on general trees. *Ann. Probab.* **27** 876–912.

[25] REED, I. S. and SOLOMON, G. (1954). A class of multiple-error-correcting codes and the decoding scheme. *IEEE Trans. Inform. Theory* **4** 38–49.

[26] SHAMIR, A. (1979). How to share a secret? *Comm. ACM* **22** 612–613.

[27] SPITZER, F. (1975). Markov random fields on an infinite tree. *Ann. Probab.* **3** 387–394.

[28] STEEL, M. (1989). Distributions in bicolored evolutionary trees. Ph.D. thesis, Massey Univ., Palmerston North, New Zealand.

COMPUTER SCIENCE DIVISION
UNIVERSITY OF CALIFORNIA
SODA HALL
BERKELEY, CALIFORNIA 94720
E-MAIL: mossel@stat.berkeley.edu

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
367 EVANS HALL
BERKELEY, CALIFORNIA 94720-3860
E-MAIL: peres@stat.berkeley.edu