Communications in
**Mathematical
Physics**

Reconstruction of Random Colourings

Allan Sly⋆

Statistics Department, University of California, Berkeley, CA 94720, USA.
E-mail: sly@stat.berkeley.edu

Abstract: Reconstruction problems have been studied in a number of contexts including
biology, information theory and statistical physics. We consider the reconstruction prob-
lem for random k-colourings on the Δ-ary tree for large k. Bhatnagar et al. [2] showed
non-reconstruction when $\Delta \leq \frac{1}{2} k \log k - o(k \log k)$. We tighten this result and show non-
reconstruction when $\Delta \leq k[\log k + \log \log k + 1 - \log 2 - o(1)]$, which is very close to the
best known bound establishing reconstruction which is $\Delta \geq k[\log k + \log \log k + 1 + o(1)]$.

1. Introduction

Determining the reconstruction threshold of a Markov random field has been of interest
in a number of areas including biology, information theory and statistical physics. Recon-
struction thresholds on trees are believed to determine the dynamical phase transitions
in many constraint satisfaction problems including random K-SAT and random colour-
ings on random graphs. It is thought that at this point the space of solutions splits into
exponentially many clusters. The properties of the space of solutions of these problems
are of interest to physicists, probabilists and theoretical computer scientists.

It is known [18,20,21] that reconstruction holds when the number of colours sat-
isfies $k[\log k + \log \log k + 1 + o(1)] \leq \Delta$. This bound is given by the analysis of a
naive reconstruction algorithm which reconstructs the root only when it is known with
absolute certainty given the leaves. The problem of finding good bounds when non-
reconstruction holds is more difficult, it requires showing that the spins on the root
and the leaves are asymptotically independent. The best previous rigorous result was
that $\Delta \leq \frac{1}{2} k \log k - o(k \log k)$ implies non-reconstruction [2]. We improve this to
$\Delta \leq k[\log k + \log \log k + 1 - \log 2 - o(1)]$. Even at a heuristic level no non-reconstruc-
tion bound as good as ours was known.

1.1. Definitions. We begin by giving a general description of broadcast models on trees and the reconstruction problem. The broadcast model on a tree T is a model in which information is sent from the root ρ across the edges, which act as noisy channels, to the leaves of T. For some given finite set of characters $\mathcal{C}$ a configuration on T is an element of $\mathcal{C}^T$, that is an assignment of a character $\mathcal{C}$ to each vertex. The broadcast model is a probability distribution on configurations defined as follows. Some $|\mathcal{C}| \times |\mathcal{C}|$ probability transition matrix M is chosen as the noisy channel on each edge. The spin σ_ρ is chosen from $\mathcal{C}$ according to some initial distribution and is then propagated along the edges of the tree according to the transition matrix M. That is if vertex u is the parent of v in the tree then the spin at v is defined according to the probabilities

$$P(\sigma_v = j | \sigma_u = i) = M_{i,j}.$$

We will focus on the colouring model with $|\mathcal{C}| = k$ which is given by the transition matrix $M_{i,j} = \frac{1_{i \neq j}}{k-1}$.

Broadcast models and in particular colourings can also be considered as Gibbs measures on trees. Given a finite set of colours k and a graph $T = (V, E)$, a k-colouring is an assignment of a colour to each vertex so that adjacent vertices have different colours. The random k-colouring model is then the uniform probability distribution on valid k-colourings of the graph. It is a Gibbs measure or Markov random field on the space of configurations $\sigma \in \{1, \ldots, k\}^V$ given by

$$P(\sigma) = \frac{1}{Z} \prod_{(u,v) \in E} 1_{\sigma_u \neq \sigma_v},$$

where Z is a normalizing constant given by the number of colourings of T. On an infinite tree more than one Gibbs measure may existi; the broadcast colouring model corresponds to the free Gibbs measure.

We will restrict our attention to Δ-ary trees, that is the infinite rooted tree where every vertex has Δ offspring. Let $L(n)$ denote the spins at distance n from the root.

Definition 1. *We say that a model is **reconstructible** on a tree T if,*

$$\limsup_n \sum_{i,L} \left| P(\sigma_\rho = i, L(n) = L) - P(\sigma_\rho = i) P(L(n) = L) \right| > 0,$$

*where the sum is over all $i \in \mathcal{C}$ and all configurations L on the vertices at distance n from the root. When the limsup is 0 we will say the model has **non-reconstruction** on T.*

Non-reconstruction is equivalent to the mutual information between $\sigma_\rho = L(0)$ and $L(n)$ going to 0 as n goes to infinity and also to $\{L(n)\}_{n=1}^\infty$ having a non-trivial tail sigma-field. More equivalent formulations are given in [17] Prop. 2.1. As increasing Δ only increases the information on the root, we can define $\Delta^*(k)$ to be the reconstruction threshold, that is the smallest Δ such that reconstruction holds on the Δ-ary tree.

In contrast to reconstruction consider the uniqueness property of a model.

Definition 2. *We say that a model has* uniqueness *on a tree T if*

$$\limsup_n \sup_{L,L'} \sum_{i \in \mathcal{C}} \left| (P(\sigma_\rho = i | L(n) = L) - P(\sigma_\rho = i | L(n) = L')) \right| > 0,$$

where the supremum is over all configurations L, L' on the vertices at distance n from the root.

   Reconstruction implies non-uniqueness and is a strictly stronger condition. Essentially uniqueness says that there is some configuration on the leaves which provides information on the root while reconstruction says that a typical configuration on the leaves provides information on the root.

1.2. Background. For some parameterized collection of models the key question in studying reconstruction is finding which models have reconstruction, which typically involves finding a threshold. This problem naturally arises in biology, information theory and statistical physics and involves the trade off between increasing numbers of leaves with increasingly noisy information as the distance from the root to the leaves increases. The simplest collection of model is the binary symmetric channel which is defined on two characters with

$$M = \begin{pmatrix} 1 - \epsilon & \epsilon \\ \epsilon & 1 - \epsilon \end{pmatrix}$$

for $0 < \epsilon < \frac{1}{2}$, which corresponds to the ferromagnetic Ising model on the tree with no external field. It was shown in [3 and 7] that this channel has reconstruction if and only if $\Delta(1 - 2\epsilon)^2 > 1$.

   The broadcast model is a natural model for the evolution of characters of DNA. In phylogenetic reconstruction the goal is to reconstruct the ancestral tree of a collection of species given their genetic data. Daskalakis, Mossel and Roch [5, 16] proved the conjecture of Mike Steel that the number of samples required for phylogenetic reconstruction undergoes a phase transition at the reconstruction threshold for the binary symmetric channel.

   Exact reconstruction thresholds have only been calculated in the binary symmetric model and binary asymmetric models with sufficiently small asymmetry [4]. In both these cases the threshold corresponds to the Kesten-Stigum bound [10]. The Kesten-Stigum bound shows that reconstruction holds whenever $\Delta\lambda_2(M)^2 > 1$, where $\lambda_2(M)$ denotes the second largest eigenvalue of M. In fact when $\Delta\lambda_2(M)^2 > 1$, it is possible to asymptotically reconstruct the root from just knowing the number of times each character appears on the leaves (census reconstruction) without using the information on their positions on the leaves. Mossel [15, 17] showed that the Kesten-Stigum bound is not the bound for reconstruction in the binary-asymmetric model with sufficiently large asymmetry or in the Potts model with sufficiently many characters.

   It was shown in [9] that k-colourings have uniqueness on Δ-ary trees if and only if $k \geq \Delta + 2$ which therefore also establishes non-reconstruction in this regime. Exactly finding the threshold for reconstruction is difficult so most attention has been focused on finding its asymptotics as the number of colours and the degree goes to infinity. Recently [2] greatly improved this bound showing that $\Delta^*(k) \geq (\frac{1}{2} + o(1))k \log k$. On the other hand [18] showed that when $\Delta \geq (1 + o(1))k \log k$ then with high probability in k the spin of the root is exactly determined by the leaves and so reconstruction is possible. With a more detailed analysis this argument can be improved to show reconstruction when $k[\log k + \log \log k + 1 + o(1)] \leq \Delta$, as was shown in [20, 21].

   This is a large improvement on the Kesten-Stigum bound which implies reconstruction when $\Delta > (k - 1)^2$. In related work Mezard and Montanari [14] found a variational principle which establishes bounds on reconstruction for colourings but which is asymptotically weaker than Lemma 7. Our results establish extremely tight bounds on $\Delta^*(k)$ with the upper and lower bounds differing by just $(\log 2 + o(1))k$ rather than $\frac{1}{2}k \log k$ previously.

Theorem 1. *The k-colouring model has reconstruction threshold $\Delta^*(k)$ satisfying,*

$$\Delta^*(k) \leq k[\log k + \log \log k + 1 + o(1)]$$

and

$$\Delta^*(k) \geq k[\log k + \log \log k + 1 - \log(2) - o(1)].$$

1.3. Applications to Statistical Physics. The reconstruction threshold on trees is believed to play a critical role in the dynamical phase transitions in certain glassy systems given by random constraint satisfaction problems. Important examples include random K-SAT and random colourings on random graphs. We will briefly describe what is conjectured by physicists about such systems [11,21], generally without rigorous proof, and why understanding the reconstruction threshold for colourings plays an important role in such systems.

The Erdős-Rényi random graph $G(n, p)$ is a random graph on n vertices where every pair of vertices is connected with probability p. To maintain constant average degree Δ we let $p = \Delta/n$. The k-colouring model on $G(n, \Delta/n)$ or random Δ-regular graphs undergoes several phase transitions as Δ grows. If we consider the space of solutions to the random colouring model where two colourings are adjacent if they differ at at most $o(n)$ vertices, then for the smallest values of Δ the space of solutions forms a large connected component. Above the clustering transition Δ_d the space of solutions breaks into exponentially many disconnected clusters and has no giant component with a constant fraction of the probability. This replica symmetry breaking transition is believed [11,12] to occur at $\Delta_d = k[\log k + \log \log k + \alpha + o(1)]$. In a recent remarkable result [1] rigorously proved that when $(1 + o(1))k \log k \leq \Delta \leq (2 - o(1))k \log k$, then the space of solutions indeed breaks into exponentially many small clusters. A second transition occurs when most clusters have frozen spins, that is vertices which have the same colour in every colouring in the cluster. This phase transition is believed to occur at $\Delta_r = k[\log k + \log \log k + 1 + o(1)]$ [20,21] and is the best upper bound known for Δ_d. Two more transitions are believed to occur: condensation where the size of the clusters is given by a Poisson-Dirichlet process, and the colouring threshold beyond which no more colourings are possible. These transitions are conjectured to occur at $\Delta_c = 2k \log k - \log k - 2 \log 2 + o(1)$ and $\Delta_s = 2k \log k - \log k - 1 + o(1)$ respectively [21]. Similar results are also expected to hold for K-SAT and other random constraint satisfiability problems [11].

Both random regular and Erdős-Rényi random graphs are locally tree-like. Asymptotically in a random regular graph the neigbourhood of a random vertex is a regular tree and for Erdős-Rényi random graphs it is a Galton-Watson branching process tree with Poisson offspring distribution. It is conjectured [11] that the reconstruction threshold on the corresponding tree is exactly the clustering threshold Δ_d on the random graph. As such, rigorous estimates of the reconstruction problem can be seen as part of a larger program of understanding glassy phases in constraint satisfaction problems.

The clustering threshold is also believed to play an important role in the efficiency of MCMC algorithms for finding and sampling from colourings of the graphs. MCMC algorithms are believed to be efficient up to the clustering threshold but experience an exponential slowdown beyond it [11]. This is to be expected since a local MCMC algorithm cannot move between clusters each of which has exponentially small probability. Rigorous proofs of rapid mixing of MCMC algorithms, such as the Glauber dynamics, fall a long way behind. For random regular graphs, results of [6] imply rapid mixing when

$k \geq 1.49\Delta$, well below the reconstruction threshold and even the uniqueness threshold. Even less is known for Erdős-Rényi random graphs as almost all MCMC results are given in terms of the maximum degree which in this case grows with n. Polynomial time mixing of the Glauber dynamics has been shown [19] for a constant number of colours in terms of Δ.

1.4. Open Problem. If the probability that the leaves uniquely determine the spin at the root does not go to 0 as n goes to infinity then the model has reconstruction. It is natural to ask is this a necessary condition for reconstruction. When $k = 5$ and $\Delta = 14$ it was shown in [14] using a variational principle that reconstruction holds but the probability that the leaves fix the root goes to 0. However, this is the only case in which the variational principle gives an upper bound on the number of colours required for reconstruction which is better than the bound of the leaves fixing the root. It remains open to determine if for large numbers of colours/high degree if this is exactly the reconstruction threshold. Numerical results of [21] suggest this is in fact not the case and there are two separate thresholds. Answering this question would be of significant interest.

2. Proofs

We introduce the notation we use in the proofs. We denote the colours by $\mathcal{C} = \{1, \ldots, k\}$ and let T be the Δ-ary tree rooted at ρ. Let $u_1, \ldots, u_\Delta$ be the children of ρ and let T_j denote the subtree of descendants of u_j. Let $P(\sigma)$ denote the free measure on colourings on the Δ-ary tree. Let $L(n)$ denote the spins at distance n from ρ and let $L_j(n)$ denote the spins on level n in the subtree T_j. We let E^i and P^i denote the expectation and probability with respect to the measure conditioned to have i at the root. For a random variable U, a function of σ, we will let $\mathcal{L}(U)$ denote the law of U and $\mathcal{L}^i(U)$ denote its conditional law with respect to the measure conditioned to have i at the root.

For a configuration L on the spins at distance n from ρ define the deterministic function f_n as

$$f_n(i, L) = P(\sigma_\rho = i | L(n) = L).$$

By the recursive nature of the tree we also have that

$$f_n(i, L) = P(\sigma_{u_j} = i | L_j(n) = L).$$

Now define $X_i(n) = X_i$ by

$$X_i(n) = f_n(i, L(n)).$$

These random variables are a deterministic function of the random configuration $L(n)$ of the leaves which gives the marginal probability that the root is in state i. By symmetry the X_i are exchangable. Now we define two distributions

$$X^+ = X^+(n) = \mathcal{L}^1 f_n(1, L(n)),$$

and

$$X^- = X^-(n) = \mathcal{L}^2 f_n(1, L(n)).$$

We will establish non-reconstruction by showing that the distributions X^+ and X^- both converge to $\frac{1}{k}$ as n goes to infinity. By symmetry we have

$$\mathcal{L}^{i_1}\left(f_n(i_2, L(n))\right) \overset{d}{=} \begin{cases} X^+ & i_1 = i_2, \\ X^- & \text{otherwise,} \end{cases}$$

and the set $\{f_n(i, L(n)) : 2 \leq i \leq k\}$ is conditionally exchangeable when conditioned on the event $\sigma_\rho = 1$. Moreover, they are conditionally exchangeable given $\sigma_\rho = 1$ and the value of $f_n(1, L(n))$. Now define

$$Y_{ij} = Y_{ij}(n) = f_n(i, L_j(n)).$$

This is equal to the probability that $\sigma_{u_j} = i$, given the random configuration $L_j(n)$ on the spins on level n in the subtree T_j. The following proposition follows immediately from the symmetries of the model.

Proposition 1. *The Y_{ij} satisfy the following properties:*

– *The random vectors $Y_j = (Y_{1j}, \ldots, Y_{qj})$ are conditionally independent given σ_ρ for $j = 1, \ldots, d$.*
– *Conditional on σ_{u_j} the random variable $Y_{\sigma_{u_j}j}$ is equal in distribution to $X^+(n)$ while for $i \neq \sigma_{u_j}$ the random variables Y_{ij} are equal in distribution to $X^-(n)$.*
– *Further, for fixed j, given σ_{u_j} and $Y_{\sigma_{u_j}j}$ the random variables $\{Y_{ij}\}_{i \neq \sigma_{u_j}}$ are conditionally exchangeable over $i \neq \sigma_{u_j}$.*

We make use of these symmetries to simplify the anaylsis. Given the standard Gibbs measure recursions on trees we have that

$$f_{n+1}(1, L(n+1)) = \frac{\prod_{j=1}^{\Delta}(1 - f_n(1, L_j(n)))}{\sum_{i=1}^{k} \prod_{j=1}^{\Delta}(1 - f_n(i, L_j(n)))}$$

and so

$$X_1(n+1) = \frac{Z_1}{\sum_{i=1}^{k} Z_i},$$

where

$$Z_i = \prod_{j=1}^{\Delta}(1 - Y_{ij}).$$

We let x_n and z_n denote $E^1 X_1(n) = E X^+(n)$ and $E^1(X_1(n) - \frac{1}{k})^2 = E(X^+(n) - \frac{1}{k})^2$ respectively. These quantities, in particular x_n, play a major role in our analysis. The following lemma, which can be viewed as the analogue of Lemma 1 of [4], allows us to relate the first and second moments of X^+.

Lemma 1. *We have that*

$$x_n = E X^+ = E^1 \sum_{i=1}^{k} X_i(n)^2 = E \sum_{i=1}^{k} (X_i(n))^2,$$

and

$$x_n - \frac{1}{k} = EX^+ - \frac{1}{k} = E\sum_{i=1}^{k}\left(X_i(n) - \frac{1}{k}\right)^2 \geq E\left(X^+ - \frac{1}{k}\right)^2 = z_n.$$

Proof. From the definition of conditional probabilities and of f_n and the fact that $P(\sigma_\rho = 1) = \frac{1}{k}$ we have that

$$
\begin{aligned}
E^1 f_n(1, L(n)) &= \sum_L f_n(1, L)P(L(n) = L|\sigma_\rho = 1) \\
&= \sum_L \frac{P(L(n) = L, \sigma_\rho = 1)}{P(\sigma_\rho = 1)} f_n(1, L) \\
&= k\sum_L P(L(n) = L)f_n(1, L)^2 \\
&= kE(X_1(n))^2 \\
&= E\sum_{i=1}^{k}(X_i(n))^2.
\end{aligned}
$$

By symmetry for any $i_1, i_2 \in \mathcal{C}$,

$$E^{i_1}\sum_{i=1}^{k}(X_i(n))^2 = E^{i_2}\sum_{i=1}^{k}(X_i(n))^2,$$

and so

$$E\sum_{i=1}^{k}(X_i(n))^2 = \frac{1}{k}\sum_{i'=1}^{k}E^{i'}\sum_{i=1}^{k}(X_i(n))^2 = E^1\sum_{i=1}^{k}(X_i(n))^2.$$

Finally we have that

$$E\sum_{i=1}^{k}(X_i(n) - \frac{1}{k})^2 = E\sum_{i=1}^{k}(X_i(n))^2 - \frac{2}{k}E\sum_{i=1}^{k}X_i(n) + k\frac{1}{k^2} = EX^+ - \frac{1}{k},$$

which completes the proof.

Corollary 1. *We have that $x_n \geq \frac{1}{k}$ and that*

$$\lim_n x_n = \frac{1}{k}$$

implies non-reconstruction.

Proof. We have that $x_n \geq z_n + \frac{1}{k} \geq \frac{1}{k}$. If x_n converges to $\frac{1}{k}$ then

$$\sum_{i=1}^{k}E\left(X_i(n) - \frac{1}{k}\right)^2 \to 0$$

which implies non-reconstruction.

2.1. Non-reconstruction. Our analysis is split into two phases, the first when x_n is close to 1 and the second when x_n is close to $\frac{1}{k}$.

Lemma 2. *Suppose that $\beta < 1 - \log 2$. Then for sufficiently large k if $\Delta < k[\log k + \log \log k + \beta]$ then*

$$\limsup_n x_n \leq \frac{2}{k}.$$

Proof. We fix the colour of the root to be 1 and let $\mathcal{F}$ denote the sigma-algebra generated by $\{\sigma_{u_j} : 1 \leq j \leq \Delta\}$, the colours of the neighbours of the root. For $1 \leq i \leq k$ let $b_i = \#\{j : \sigma_{u_j} = i\}$, the number of times each colour appears amongst the neighbours of the root. Of course $b_1 = 0$ since the neighbours of the root cannot be 1. For $1 \leq i \leq k$ define

$$U_i = \prod_{1 \leq j \leq \Delta : \sigma_{u_j} = i} (1 - Y_{ij}).$$

Note that with this definition $U_1 = 1$. We will use the symmetries and exchangeability of the model to reduce the problem to considering a random variable only involving the U_i. Conditional on $\mathcal{F}$, the U_i are independent and are distributed as the product of b_i independent copies of $(1 - X^+(n))$ and $0 \leq U_i \leq 1$ for all i. Fix an ℓ with $2 \leq \ell \leq k$. Let W_1 and W_ℓ be defined by

$$W_1 = \prod_{1 \leq j \leq \Delta : \sigma_{u_j} \neq \ell} (1 - Y_{1j}), \quad W_\ell = \prod_{1 \leq j \leq \Delta : \sigma_{u_j} \neq \ell} (1 - Y_{\ell j})$$

so $Z_\ell = W_\ell U_\ell$. Note that for $j \in \{1 \leq j \leq \Delta : \sigma_{u_j} \neq \ell\}$ we have that $\sigma_{u_j} \notin \{1, \ell\}$, since of none of the σ_{u_j} are 1. So by Proposition 1, conditional on $\mathcal{F}$ and $\sigma_{u_j} \notin \{1, \ell\}$, we have that Y_{1j} and $Y_{\ell j}$ are conditionally exchangeable and so W_1 and W_ℓ are conditionally exchangeable. We will analyse the effect of swapping W_1 with W_ℓ. Recall that $Z_i = \prod_{j=1}^{\Delta}(1 - Y_{ij})$ so define

$$\widetilde{Z}_\ell = W_1 U_\ell = W_1 \prod_{1 \leq j \leq \Delta : \sigma_{u_j} = \ell} (1 - Y_{\ell j}),$$

and

$$\widetilde{Z}_1 = W_\ell \prod_{1 \leq j \leq \Delta : \sigma_{u_j} = \ell} (1 - Y_{1j}),$$

and for $i \notin \{1, \ell\}$,

$$\widetilde{Z}_i = Z_i.$$

Proposition 1 noted that $Y_j = \{Y_{1j}, \ldots, Y_{kj}\}$ are conditionally independent given $\mathcal{F}$ and for each j given σ_{u_j} and $Y_{\sigma_{u_j} j}$ the random variables $\{Y_{ij} : i \neq \sigma_{u_j}\}$ are conditionally exchangeable. It follows that

$$(W_1, W_\ell, Z_1, \ldots, Z_k, U_1 \ldots, U_k, \sigma_1, \ldots, \sigma_\Delta)$$
$$\stackrel{d}{=} \left(W_\ell, W_1, \widetilde{Z}_1, \ldots, \widetilde{Z}_k, U_1 \ldots, U_k, \sigma_1, \ldots, \sigma_\Delta \right), \tag{1}$$

where we denote equality as in distributions of random vectors since this just swaps Y_{1j}'s
with $Y_{\ell j}$'s which are conditionally exchangeable given all the other random variables.
Since $0 \le U_\ell \le 1$, and $\sum_{i=2}^{k} Z_i - \sum_{i=2}^{k} \widetilde{Z}_i = Z_\ell - \widetilde{Z}_\ell = U_\ell(W_\ell - W_1)$ it follows that
$(W_1 - W_\ell)$ has the same sign as

$$\left(W_1 + \sum_{i=2}^{k} Z_i \right) - \left(W_\ell + \sum_{i=2}^{k} \widetilde{Z}_i \right) = (W_1 - W_\ell)(1 - U_\ell)$$

and so

$$\frac{1}{W_1 + \sum_{i=2}^{k} Z_i} - \frac{1}{W_\ell + \sum_{i=2}^{k} \widetilde{Z}_i}$$

has the opposite sign as $W_1 - W_\ell$. Applying the equality in distribution of Eq. (1) we
have that

$$E^1 \left[\left. \frac{W_\ell - W_1}{W_1 + \sum_{i=2}^{k} Z_i} \right| \mathcal{F}, \{U_i\} \right]$$

$$= \frac{1}{2} E^1 \left[\left. \frac{W_\ell - W_1}{W_1 + \sum_{i=2}^{k} Z_i} + \frac{W_1 - W_\ell}{W_\ell + \sum_{i=2}^{k} \widetilde{Z}_i} \right| \mathcal{F}, \{U_i\} \right]$$

$$= \frac{1}{2} E^1 \left[\left. (W_\ell - W_1) \left(\frac{1}{W_1 + \sum_{i=2}^{k} Z_i} - \frac{1}{W_\ell + \sum_{i=2}^{k} \widetilde{Z}_i} \right) \right| \mathcal{F}, \{U_i\} \right]$$

$$\ge 0,$$

where the first equality follows using equality in distributions of the random vectors and
the inequality follows from the two terms of the product having the same sign. Since
$0 \le Z_1 \le W_1 \le 1$ we have that,

$$E^1 \left[\left. \frac{Z_1}{Z_1 + \sum_{i=2}^{k} Z_i} \right| \mathcal{F}, \{U_i\} \right] \le E^1 \left[\left. \frac{W_1}{W_1 + \sum_{i=2}^{k} Z_i} \right| \mathcal{F}, \{U_i\} \right]$$

$$\le E^1 \left[\left. \frac{W_\ell}{W_1 + \sum_{i=2}^{k} Z_i} \right| \mathcal{F}, \{U_i\} \right]$$

$$\le E^1 \left[\left. \frac{W_\ell}{Z_1 + \sum_{i=2}^{k} Z_i} \right| \mathcal{F}, \{U_i\} \right],$$

and so since $Z_\ell = U_\ell W_\ell$ and we are conditioning on U_ℓ,

$$E^1 \left[\left. \frac{Z_1 U_\ell}{Z_1 + \sum_{i=2}^{k} Z_i} \right| \mathcal{F}, \{U_i\} \right] \le E^1 \left[\left. \frac{Z_\ell}{Z_1 + \sum_{i=2}^{k} Z_i} \right| \mathcal{F}, \{U_i\} \right].$$

Recall that $\ell \ge 2$ is arbitrary so the previous equation holds for all $2 \le \ell \le k$ simulta-
neously. Summing over all values of ℓ we get that,

$$E^1 \left[\left. \frac{Z_1 \left(1 + \sum_{l=2}^{k} U_l \right)}{\sum_{i=1}^{k} Z_i} \right| \mathcal{F}, \{U_i\} \right] \le \sum_{l=1}^{k} E^1 \left[\left. \frac{Z_l}{\sum_{i=1}^{k} Z_i} \right| \mathcal{F}, \{U_i\} \right] = 1,$$

and hence since we are conditioning on the U_i,

$$E^1\left[X_1(n+1)|\,\mathcal{F},\{U_i\}\right] = E^1\left[\frac{Z_1}{\sum_{i=1}^{k} Z_i}\bigg|\,\mathcal{F},\{U_i\}\right] \le \frac{1}{1+\sum_{i=2}^{k} U_i}.$$

We now estimate the expected value of the right-hand side of the previous equation. Using the fact that $\frac{1}{1+x} = \int_0^1 s^x ds$ we have that

$$\frac{1}{1+\sum_{i=2}^{k} U_i} = \int_0^1 s^{\sum_{i=2}^{k} U_i} ds.$$

As s^u is convex as a function of u we have that $s^u \le s^0(1-u) + s^1 u$ when $0 \le u \le 1$ and so since $0 \le U_i \le 1$ we have that $E^1 s^{U_i} \le (1 - E^1 U_i) + s E^1 U_i = 1 - (1-s)E^1 U_i$. Since it is conditional on $\mathcal{F}$ the U_i are independent and are distributed as the product of b_i independent copies of $(1 - X^+(n))$ we have that,

$$E^1\left[X_1(n+1)|\,\mathcal{F}\right] \le \int_0^1 \prod_{i=2}^{k} (1 - (1-s)E^1[U_i|\mathcal{F}])ds$$

$$= \int_0^1 \prod_{i=2}^{k} (1 - (1-s)(1-x_n)^{b_i})ds.$$

Now the colours σ_{u_j} are chosen independently and uniformly from the set $\{2, \ldots, k\}$ so $(b_2, \ldots, b_k)$ has a multinominal distribution. Let $\beta < \beta^* < 1 - \log 2$ and let $\widetilde{b}_i$ be iid random variables distributed as Poisson(D), where $D = \log k + \log\log k + \beta^*$. By Lemma 4 we can couple the b's and $\widetilde{b}$'s so that $(b_2, \ldots, b_k) \le (\widetilde{b}_2, \ldots, \widetilde{b}_k)$ whenever $\sum_{j=2}^{k} \widetilde{b}_j \ge \Delta$. It follows that

$$x_{n+1} = E^1 X_1(n+1)$$

$$\le E^1 \mathbf{1}_{\{\sum_{j=2}^{k} \widetilde{b}_j < \Delta\}} + \int_0^1 E^1 \prod_{i=2}^{k} (1 - (1-s)(1-x_n)^{\widetilde{b}_i})ds$$

$$\le p + \int_0^1 (1 - (1-s)\exp(-x_n D))^{k-1} ds$$

$$\le p + \int_0^1 \exp\left(-(1-s)(k-1)\exp(-x_n D)\right) ds$$

$$= p + \frac{1 - \exp\left(-(k-1)\exp(-x_n D)\right)}{(k-1)\exp(-x_n D)},$$

where $p = P(\text{Poisson}((k-1)D) < \Delta)$. Now $p = \exp\left(-\Omega\left(\frac{k}{\sqrt{\Delta}}\right)\right) = o(k^{-1})$ and the function

$$g(y) = p + \frac{1 - \exp\left(-(k-1)\exp(-yD)\right)}{(k-1)\exp(-yD)}$$

is increasing in y so the result follows by Lemma 3.

Lemma 3. *Let $y_0, y_1, \ldots$ be a sequence of positive real numbers such that $y_0 = 1$ and $y_{n+1} = g(y_n)$, where $g(y_n) = p + \frac{1-\exp(-(k-1)\exp(-y_n D))}{(k-1)\exp(-y_n D)}$, $D = \log k + \log \log k + \beta^*$, $\beta^* < 1 - \log 2$ and $p = o(k^{-1})$. Then for large enough k,*

$$\limsup_n y_n < \frac{2}{k}.$$

Proof. Since $\frac{d}{dx} \frac{1-e^{-x}}{x}\Big|_{x=0} = -\frac{1}{2}$ we can find $\epsilon, \delta > 0$ such that when $0 < x < \delta$, then

$$\frac{1-e^{-x}}{x} < 1 - \left(\frac{1}{2} - \epsilon\right) x.$$

Assuming our choice of ϵ is sufficiently small we can also choose $r' > r > 0$ such that $(\frac{1}{2} - \epsilon)e^{-\beta^*} > e^{-1}(1+r')$. Now for large enough k, $(k-1)\exp(-D) = \frac{(k-1)e^{-\beta^*}}{k \log k} < \delta$ and so using the fact that $r < r'$ and $p = o(k^{-1})$,

$$y_1 = g(1) \leq p + 1 - \left(\frac{1}{2} - \epsilon\right) \frac{(k-1)e^{-\beta^*}}{k \log k} \leq 1 + p - \frac{(1+r)e^{-1}}{\log k} \leq 1 - \frac{e^{-1}}{\log k},$$

provided k is sufficiently large. Now since g is a continuous increasing function and $y_1 < y_0$ it follows that the sequence y_i is decreasing. Suppose that $(k-1)\exp(-y_i D) < \delta$. Then

$$y_{i+1} \leq p + 1 - \left(\frac{1}{2} - \epsilon\right)(k-1)\exp(-y_i D),$$

and so for k sufficiently large

$$\begin{aligned}
1 - y_{i+1} &\geq \left(\frac{1}{2} - \epsilon\right)(k-1)\exp(-y_i D) - p \\
&\geq \left(\frac{1}{2} - \epsilon\right)\frac{(k-1)e^{-\beta^*}}{k \log k} \exp((1-y_i)\log k) - p \\
&\geq \frac{(1+r')e^{-1}}{\log k} \exp((1-y_i)\log k) - p \\
&\geq (1+r')(1-y_i) - p \\
&\geq (1+r)(1-y_i),
\end{aligned}$$

where the second to last inequality uses the fact that $e^x \geq ex$ and the final inequality uses the fact that $1 - y_i \geq \frac{e^{-1}}{\log k}$, while $p = o(k^{-1})$. It follows that y_i decreases until for some i, $(k-1)\exp(-y_i D) \geq \delta$. Now let $\frac{1-e^{-\delta}}{\delta} = \alpha' < \alpha'' < \alpha < 1$ for some α. When k is large enough then

$$y_{i+1} \leq p + \frac{1-e^{-\delta}}{\delta} \leq \alpha''.$$

Then for k large enough, $\exp(-y_{i+1}D) \geq \exp(-\alpha''D) \geq \exp(-\alpha \log k) = k^{-\alpha}$. It follows that

$$y_{i+2} \leq p + \frac{1}{(k-1)\exp(-y_{i+1}D)} \leq 2k^{\alpha-1}.$$

Finally we have $\exp(-y_{i+2}D) \geq \exp(-2k^{\alpha-1}D) \geq \frac{2}{3}$ and so

$$y_{i+3} \leq p + \frac{1}{(k-1)\exp(-y_{i+2}D)} < 2k^{-1}$$

when k is large enough, which completes the proof.

In the preceding lemma we note that the requirement that $\beta^* < 1 - \log 2$ comes from the fact that $x < \frac{1}{2}e^{x-\beta^*}$ for all x when $\beta^* < 1 - \log 2$.

Lemma 4. *Suppose that* $(b_1, \ldots, b_k)$ *has the multinominal distribution* $M\left(n, \left(\frac{1}{k}, \frac{1}{k}, \ldots \frac{1}{k}\right)\right)$. *Let* $\widetilde{b}_j$ *be iid random variables distributed as Poisson(D). We can couple the b's and $\widetilde{b}$'s so that* $(b_1, \ldots, b_k) \leq (\widetilde{b}_1, \ldots, \widetilde{b}_k)$ *(respectively $\geq$) whenever* $\sum_{j=1}^{k} \widetilde{b}_j \geq n$ *(respectively $\leq$).*

Proof. Since the $\widetilde{b}_j$ are independent and Poisson, conditional on the sum $N = \sum_{j=1}^{k} \widetilde{b}_j$, the distribution of $(\widetilde{b}_1, \ldots, \widetilde{b}_k)$ is multinominal $M\left(N, \left(\frac{1}{k}, \frac{1}{k}, \ldots \frac{1}{k}\right)\right)$ (see [13] Prop. 6.2.1). Now if $n \leq m$ then two multinomial distributions A and B distributed as $M\left(n, \left(\frac{1}{k}, \frac{1}{k}, \ldots \frac{1}{k}\right)\right)$ and $M\left(m, \left(\frac{1}{k}, \frac{1}{k}, \ldots \frac{1}{k}\right)\right)$ respectively can be trivially coupled so that $A \leq B$, which completes the proof.

Janson and Mossel [8] studied "robust reconstruction", the question of when reconstruction is possible from a very noisy copy of the leaves. They found that the threshold for robust reconstruction is exactly the Kesten-Stigum bound. Lemma 2 establishes that the leaves provide very little information about the spin at a vertex a long distance from the leaves. So as information over long distances is very noisy the results of [8] suggest that reconstruction would only be possible after the Kesten-Stigum bound whereas, in our context, Δ is much less than $\lambda_2(M)^{-2}$. As such, only crude bounds are needed to establish the following lemma.

Lemma 5. *For sufficiently large k if $\Delta \leq 2k \log k$ and if $x_n \leq \frac{2}{k}$ then*

$$x_{n+1} - \frac{1}{k} \leq \frac{1}{2}\left(x_n - \frac{1}{k}\right).$$

Proof. Using the identity

$$\frac{1}{s+r} = \frac{1}{s} - \frac{r}{s^2} + \frac{r^2}{s^2}\frac{1}{s+r}$$

and taking $s = E^1 \sum_{i=1}^{k} Z_i$ and $r = \sum_{i=1}^{k}(Z_i - E^1 Z_i)$ we have that

$$x_{n+1} - \frac{1}{k} = E^1 \frac{Z_1 - \frac{1}{k}\sum_{i=1}^{k} Z_i}{\sum_{i=1}^{k} Z_i}$$

$$= E^1 \frac{Z_1 - \frac{1}{k}\sum_{i=1}^{k} Z_i}{E^1 \sum_{i=1}^{k} Z_i} - E^1 \frac{\left(Z_1 - \frac{1}{k}\sum_{i=1}^{k} Z_i\right)\left(\sum_{i=1}^{k}(Z_i - E^1 Z_i)\right)}{\left(E^1 \sum_{i=1}^{k} Z_i\right)^2}$$

$$+ E^1 \frac{Z_1 - \frac{1}{k}\sum_{i=1}^{k} Z_i}{\sum_{i=1}^{k} Z_i} \frac{\left(\sum_{i=1}^{k}(Z_i - E^1 Z_i)\right)^2}{\left(E^1 \sum_{i=1}^{k} Z_i\right)^2}.$$

Now by Lemma 6,

$$
\frac{E^1\left(Z_1 - \frac{1}{k}\sum_{i=1}^{k} Z_i\right)}{E^1 \sum_{i=1}^{k} Z_i} \leq \frac{\frac{k-1}{k}\left(1 + \frac{2\Delta}{k}\left(x_n - \frac{1}{k}\right)\right) - \frac{k-1}{k}\left(1 - \frac{2\Delta}{k^2}\left(x_n - \frac{1}{k}\right)\right)}{1 + (k-1)\left(1 - \frac{2\Delta}{k^2}\left(x_n - \frac{1}{k}\right)\right)}
$$

$$
\leq \frac{3\Delta}{k^2}\left(x_n - \frac{1}{k}\right). \tag{2}
$$

Using the inequality $\frac{1}{2}(a^2 + b^2) \geq ab$ we have that

$$
-\left(Z_1 - \frac{1}{k}\sum_{i=1}^{k} Z_i\right)\left(\sum_{i=1}^{k} Z_i - E^1 Z_i\right)
$$

$$
= -\left(\left(Z_1 - E^1 Z_1\right) + \left(E^1 Z_1 - \frac{1}{k}\sum_{i=1}^{k} E^1 Z_i\right) - \frac{1}{k}\left(\sum_{i=1}^{k}(Z_i - E^1 Z_i)\right)\right)
$$

$$
\cdot\left(\sum_{i=1}^{k}(Z_i - E^1 Z_i)\right)
$$

$$
\leq \frac{1}{2}\left|Z_1 - E^1 Z_1\right|^2 + \left(\frac{1}{2} + \frac{1}{k}\right)\left|\sum_{i=1}^{k}(Z_i - E^1 Z_i)\right|^2
$$

$$
-\left(E^1 Z_1 - E^1 \frac{1}{k}\sum_{i=1}^{k} Z_i\right)\left(\sum_{i=1}^{k}(Z_i - E^1 Z_i)\right)
$$

so by Lemma 6 we have that,

$$
E^1\left[-\left(Z_1 - \frac{1}{k}\sum_{i=1}^{k} Z_i\right)\left(\sum_{i=1}^{k}(Z_i - E^1 Z_i)\right)\right]
$$

$$
\leq \left(\frac{k-1}{k}\right)^{2\Delta}\left(x_n - \frac{1}{k}\right)\left[\frac{4\Delta}{k} + 4\Delta\right]
$$

and

$$
E^1\left[-\frac{\left(Z_1 - \frac{1}{k}\sum_{i=1}^{k} Z_i\right)\left(\sum_{i=1}^{k}(Z_i - E^1 Z_i)\right)}{\left(E^1 \sum_{i=1}^{k} Z_i\right)^2}\right]
$$

$$
\leq \frac{\left(x_n - \frac{1}{k}\right)\left[\frac{4\Delta}{k} + 4\Delta\right]}{\left(1 + (k-1)\left(1 - \frac{2\Delta}{k^2}\left(x_n - \frac{1}{k}\right)\right)\right)^2}
$$

$$
\leq \frac{5\Delta}{k^2}\left(x_n - \frac{1}{k}\right). \tag{3}
$$

Finally since $0 \leq \frac{Z_1}{\sum_{i=1}^k Z_i} \leq 1$ we have that $\left| \frac{Z_1 - \frac{1}{k} \sum_{i=1}^k Z_i}{\sum_{i=1}^k Z_i} \right| \leq 1$, and so

$$E^1 \frac{Z_1 - \frac{1}{k} \sum_{i=1}^k Z_i}{\sum_{i=1}^k Z_i} \frac{\left(\sum_{i=1}^k (Z_i - E^1 Z_i) \right)^2}{\left(E^1 \sum_{i=1}^k Z_i \right)^2} \leq E^1 \frac{\left(\sum_{i=1}^k (Z_i - E^1 Z_i) \right)^2}{\left(E^1 \sum_{i=1}^k Z_i \right)^2}$$

$$\leq \frac{5\Delta}{k^2} \left(x_n - \frac{1}{k} \right). \tag{4}$$

Combining Eqs. (2), (3) and (4) we have that

$$x_{n+1} - \frac{1}{k} \leq \frac{13\Delta}{k^2} \left(x_n - \frac{1}{k} \right) \leq \frac{1}{2} \left(x_n - \frac{1}{k} \right) \tag{5}$$

for large enough k, which completes the result.

Lemma 6. *For sufficiently large k if $\Delta \leq 2k \log k$ and if $x_n \leq \frac{2}{k}$ then the following all hold:*

$$\left(\frac{k-1}{k} \right)^\Delta \leq E^1 Z_1 \leq \left(\frac{k-1}{k} \right)^\Delta \left(1 + \frac{2\Delta}{k} \left(x_n - \frac{1}{k} \right) \right), \tag{6}$$

and for $i \neq 1$,

$$\left(\frac{k-1}{k} \right)^\Delta \left(1 - \frac{2\Delta}{k^2} \left(x_n - \frac{1}{k} \right) \right) \leq E^1 Z_i \leq \left(\frac{k-1}{k} \right)^\Delta, \tag{7}$$

$$\mathrm{Var}^1 Z_1 \leq \left(\frac{k-1}{k} \right)^{2\Delta} \frac{4\Delta}{k} \left(x_n - \frac{1}{k} \right), \tag{8}$$

$$\mathrm{Var}^1 \left(\sum_{i=1}^k Z_i \right) \leq \left(\frac{k-1}{k} \right)^{2\Delta} 4\Delta \left(x_n - \frac{1}{k} \right). \tag{9}$$

Proof. From Eq. (15) of Lemma 9 we have that

$$E^1 Z_1 = \left(\frac{k-1}{k} + \frac{1}{k-1} \left(x_n - \frac{1}{k} \right) \right)^\Delta,$$

and since by Corollary 1, $x_n \geq \frac{1}{k}$ we have that

$$E^1 Z_1 \geq \left(\frac{k-1}{k} \right)^\Delta.$$

Then since $\exp(x) = 1 + x + O(x^2)$ and $\frac{k\Delta}{(k-1)^2} \left(x_n - \frac{1}{k} \right)$ is small for large k,

$$E^1 Z_1 \leq \left(\frac{k-1}{k} \right)^\Delta \exp \left(\frac{k\Delta}{(k-1)^2} \left(x_n - \frac{1}{k} \right) \right)$$

$$\leq \left(\frac{k-1}{k} \right)^\Delta \left(1 + \frac{2\Delta}{k} \left(x_n - \frac{1}{k} \right) \right),$$

which establishes Eq. (6). Equations (7), (8) and (9) are established similarly using identities from Lemma 9.

2.2. Reconstruction. An upper bound on the reconstruction threshold $\Delta^*(k)$ is found by estimating the probability that the colour of the root is uniquely determined by the colours at the leaves. This method was described in [18] and used to a higher level of precision in [21,20]. We restate the result and give a full proof for completeness.

Lemma 7. *Suppose that $\beta > 1$. Then for sufficiently large k if $\Delta > k[\log k + \log \log k + \beta]$ then the colour of the root is uniquely determined by the colours at the leaves with probability at least $1 - \frac{1}{\log k}$, that is*

$$\inf_n P(X^+(n) = 1) > 1 - \frac{1}{\log k}.$$

Proof. Let p_n be the probability that the leaves at distance n determine the spin at the root, that is $p_n = P^1(X_1(n) = 1)$. We will show that when k is large then $\liminf_n p_n$ is close to 1.

Suppose we fix the colour of the root to be 1 and let $\mathcal{F}$ denote the sigma-algebra generated by $\{\sigma_{u_j} : 1 \leq j \leq \Delta\}$, the colours of the neighbours of the root. For $2 \leq i \leq k$ let $b_i = \#\{j : \sigma_{u_j} = i\}$, the number of times each colour appears in the neighbours of the root. Now each colour σ_{u_j} is chosen uniformly from the set $\{2, \ldots, k\}$ so $(b_2, \ldots, b_k)$ has a multinominal distribution. Let $\beta > \beta^* > 1$ and let $\widetilde{b}_i$ be iid random variables distributed as Poisson(D), where $D = \log k + \log \log k + \beta^*$. By Lemma 4 we can couple the b's and $\widetilde{b}$'s so that $(b_2, \ldots, b_k) \geq (\widetilde{b}_2, \ldots, \widetilde{b}_k)$ whenever $\sum_{i=2}^k \widetilde{b}_j \leq \sum_{i=2}^k b_j = \Delta$. If for each colour $2 \leq i \leq k$ there is some vertex u_j such that the states of the leaves, $L_j(n)$ fix the colour of u_j to be i, then the leaves $L(n + 1)$ fix the colour of ρ to be 1. Conditional on $\mathcal{F}$ the probability that there is such a vertex u_j for a given colour i is at least $1 - (1 - p_n)^{b_i}$. Moreover these are conditionally independent of $\mathcal{F}$ so it follows that

$$p_{n+1} \geq \prod_{i=2}^k E^1 \left[1 - (1 - p_n)^{b_i} | \mathcal{F} \right]$$

$$\geq \prod_{i=2}^k E^1 \left[1 - (1 - p_n)^{\widetilde{b}_i} \right] - s$$

$$= (1 - \exp(-p_n D))^{k-1} - s,$$

where $s = P(\text{Poisson}((k-1)D) > \Delta) = o(k^{-1})$. Now

$$f(x) = (1 - \exp(-xD))^{k-1} - s$$

is an increasing function in x and hence when k is large enough

$$f\left(1 - \frac{1}{\log k}\right) = \left(1 - \exp\left(-(1 - \frac{1}{\log k})(\log k + \log \log k + \beta^*)\right)\right)^{k-1} - s$$

$$> 1 - \frac{1}{\log k},$$

and since $p_0 = 1$,

$$\inf_n p_n \geq 1 - \frac{1}{\log k},$$

which completes the proof.

2.3. Main Theorem.

*Proof (*Theorem 1*).* Combining Lemmas 2 and 5 establishes non-reconstruction when $\Delta \leq k[\log k + \log \log k + 1 - \log(2) - o(1)]$. Lemma 7 shows that the root can be reconstructed correctly with probability at least $1 - \frac{1}{\log k}$, which establishes reconstruction when $\Delta \geq k[\log k + \log \log k + 1 + o(1)]$.

Remarks. For large k the Poisson(Δ) distribution is concentrated around Δ with standard deviation $O(\sqrt{\Delta})$ which is significantly smaller than the error bounds in Theorem 1. With some minor modifications the bounds for Δ-ary trees can be extended to Galton-Watson branching processes with offspring distribution Poisson(Δ). The reconstruction of Galton-Watson branching processes with offspring distribution Poisson(Δ) is of interest because, as noted before, it is believed to be related to the clustering phase transition for colourings on Erdős-Rényi random graphs.

To be more specific, for the proof of non-reconstruction we can again bound $x_n = E^1 X_1(n)$, where the expected value is taken over all possible trees. In Lemma 2 we repeat the same bounds on x_n, the only difference being Δ is now random, which does not affect the results for large k. Then similar estimates can be made in Lemma 5 provided $\frac{\Delta}{k} \left(x_n - \frac{1}{k} \right)$ is very small. As Δ is concentrated around its expected value the probability of this not holding is very small and this can be used to complete the proof of non-reconstruction.

When $\beta > \beta^* > 1$, with probability going to 1 as k goes to infinity, the Galton-Watson branching process contains a subgraph which is a $(k[\log k + \log \log k + \beta^*])$-ary tree rooted at ρ. Reconstruction then follows from Lemma 7.

A. Appendix

In this appendix we calculate identities which are used in the proof of Lemma 6. Observe that since $E X^+(n) + (k-1) E X^- = 1$ we have that $E X^+ - \frac{1}{k} = -(k-1)(E X^- - \frac{1}{k})$. We will show that the means and variances of the Y_{ij} and Z_i can all be calculated in terms of x_n and z_n.

Lemma 8. *We have the identities*

$$E^1 Y_{1j} = \frac{1}{k} - \frac{1}{k-1} \left(x_n - \frac{1}{k} \right), \tag{10}$$

$$E^1 Y_{1j}^2 = \frac{1}{k^2} + \frac{k-2}{k(k-1)} \left(x_n - \frac{1}{k} \right) - \frac{1}{k-1} z_n. \tag{11}$$

For $2 \leq i \leq k$,

$$E^1 Y_{ij} = \frac{1}{k} + \frac{1}{(k-1)^2} \left(x_n - \frac{1}{k} \right) \tag{12}$$

and

$$E^1 Y_{ij}^2 = \frac{1}{k^2} + \frac{k^2 - 2k + 2}{k(k-1)^2}\left(x_n - \frac{1}{k}\right) + \frac{1}{(k-1)^2} z_n. \tag{13}$$

For any $1 \le i_1 < i_2 \le k$,

$$Cov^1(Y_{i_1 j}, Y_{i_2 j}) \le 0. \tag{14}$$

Proof. When the root is conditioned to be 1, $\sigma_{u_j} \ne 1$ and so Y_{1j} is distributed as X^- and we have that

$$E^1 Y_{1j} = EX^- = \frac{1}{k} - \frac{1}{k-1}\left[EX^+ - \frac{1}{k}\right] = \frac{1}{k} - \frac{1}{k-1}\left(x_n - \frac{1}{k}\right),$$

and

$$E^1 Y_{1j}^2 = E(X^-)^2$$

$$= \frac{1}{k-1}\left[E\sum_{i=1}^{k}(X_i)^2 - E(X^+)^2\right]$$

$$= \frac{1}{k-1}[EX^+ - E(X^+)^2]$$

$$= \frac{1}{k-1}\left[\frac{k-2}{k}E\left(X^+ - \frac{1}{k}\right) - E\left(X^+ - \frac{1}{k}\right)^2 + \frac{k-1}{k^2}\right]$$

$$= \frac{1}{k^2} + \frac{k-2}{k(k-1)}\left(x_n - \frac{1}{k}\right) - \frac{1}{k-1} z_n,$$

where the third equality follows from Lemma 1. For $2 \le i \le k$ we have that

$$E^1 Y_{ij} = \frac{1}{k-1}[1 - E^1 Y_{1j}] = \frac{1}{k-1}\left[1 - \frac{1}{k} + \frac{1}{k-1}\left[x_n - \frac{1}{k}\right]\right]$$

$$= \frac{1}{k} + \frac{1}{(k-1)^2}\left(x_n - \frac{1}{k}\right),$$

and again using Lemma 1,

$$E^1 Y_{ij}^2 = \frac{1}{k-1}\left[E^1 \sum_{i=1}^{k}(X_i)^2 - E^1 Y_{1j}^2\right]$$

$$= \frac{1}{k^2} + \frac{k^2 - 2k + 2}{k(k-1)^2}\left(x_n - \frac{1}{k}\right) + \frac{1}{(k-1)^2} z_n.$$

Also for $2 \le i \le k$,

$$E^1 Y_{1j} Y_{ij} = \frac{1}{k-1}\sum_{i'=2}^{k} E^1 Y_{1j} Y_{i'j}$$

$$= \frac{1}{k-1} E^1 Y_{1j}(1 - Y_{1j})$$

$$\le \frac{1}{k-1} E^1 Y_{1j} E(1 - Y_{1j})$$

$$= E^1 Y_{1j} E^1 Y_{ij},$$

so $\mathrm{Cov}^1(Y_{1j}, Y_{ij}) \leq 0$. Finally for $2 \leq i_1 < i_2 \leq k$,

$$\mathrm{Var}^1(1 - Y_{1j}) = \sum_{i=2}^{k} \mathrm{Var}^1(Y_{ij}) + (k-1)(k-2)\mathrm{Cov}^1(Y_{i_1j}, Y_{i_2j}),$$

and so

$$\begin{aligned}\mathrm{Cov}^1(Y_{i_1j}, Y_{i_2j}) &= \mathrm{Var}^1(1 - Y_{1j}) - \sum_{i=2}^{k} \mathrm{Var}^1(Y_{ij}) \\ &= \mathrm{Var}(X^-) - ((k-2)\mathrm{Var}(X^-) + \mathrm{Var}(X^+)) \\ &\leq 0,\end{aligned}$$

so $\mathrm{Cov}^1(Y_{i_1j}, Y_{i_2j}) \leq 0$.

Using Lemma 8 we can calculate the means and covariances of the Z_j.

Lemma 9. *We have the following results:*

$$E^1 Z_1 = \left(\frac{k-1}{k} + \frac{1}{k-1}\left(x_n - \frac{1}{k}\right)\right)^\Delta, \tag{15}$$

$$E^1 Z_1^2 = \left(\left(\frac{k-1}{k}\right)^2 + \frac{3k-2}{k(k-1)}\left(x_n - \frac{1}{k}\right) - \frac{1}{k-1}z_n\right)^\Delta. \tag{16}$$

For each $2 \leq i \leq k$ then

$$E^1 Z_i = \left(\frac{k-1}{k} - \frac{1}{(k-1)^2}\left(x_n - \frac{1}{k}\right)\right)^\Delta \tag{17}$$

and

$$E^1 Z_i^2 = \left(\left(\frac{k-1}{k}\right)^2 + \frac{k^2-4k+2}{k(k-1)^2}\left(x_n - \frac{1}{k}\right) + \frac{1}{(k-1)^2}z_n\right)^\Delta. \tag{18}$$

For any $1 \leq i_1 < i_2 \leq k$,

$$\mathrm{Cov}^1(Z_{i_1j}, Z_{i_2j}) \leq 0. \tag{19}$$

Proof. By Eq. (10) we have that

$$\begin{aligned}E^1 Z_1 &= E\prod_{j=1}^{\Delta}(1 - Y_{1j}) \\ &= \left(1 - \left(\frac{1}{k} - \frac{1}{k-1}\left(x_n - \frac{1}{k}\right)\right)\right)^\Delta \\ &= \left(\frac{k-1}{k} + \frac{1}{k-1}\left(x_n - \frac{1}{k}\right)\right)^\Delta,\end{aligned}$$

which establishes Eq. (15). Equations (16), (17) and (18) follow similarly. Using Eq. (14) we have that for $1 \le i_1 < i_2 \le k$,

$$E^1 Z_{i_1} Z_{i_2} = E^1 \prod_{j=1}^{\Delta} (1 - Y_{i_1 j})(1 - Y_{i_2 j})$$

$$\le \prod_{j=1}^{\Delta} E^1 (1 - Y_{i_1 j}) E(1 - Y_{i_2 j})$$

$$= E^1 Z_{i_1} E^1 Z_{i_2},$$

which establishes Eq. (19).

References

1. Achlioptas, D., Coja-Oghlan, A.: *Algorithmic barriers from phase transition*. http://front.math.ucdavis.edu/0803.2122, 2008
2. Bhatnagar, N., Vera, J., Vigoda, E.: *Reconstruction for colorings on trees*. http://front.math.ucdavis.edu/0711.3664, 2007
3. Bleher, P.M., Ruiz, J., Zagrebnov, V.A.: On the purity of limiting Gibbs state for the Ising model on the Bethe lattice. J. Stat. Phys. **79**, 473–482 (1995)
4. Borgs, C., Chayes, J.T., Mossel, E., Roch, S.: The Kesten-Stigum reconstruction bound is tight for roughly symmetric binary channels. In: *FOCS* 2006, Los Alamitos, CA: IEEE Computer Society, 2006, pp. 518–530
5. Daskalakis, C., Mossel, E., Roch, S.: Optimal phylogenetic reconstruction. In: *STOC'06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, New York: ACM, 2006, pp. 159–168
6. Dyer, M., Frieze, A., Hayes, T.P., Vigoda, E.: Randomly coloring constant degree graphs. In: *FOCS '04: Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, Washington, DC: IEEE Computer Society, 2004, pp. 582–589
7. Evans, W., Kenyon, C., Peres, Y., Schulman, L.J.: Broadcasting on trees and the Ising model. Ann. Appl. Probab. **10**(2), 410–433 (2000)
8. Janson, S., Mossel, E.: Robust reconstruction on trees is determined by the second eigenvalue. Ann. Probab. **32**(3B), 2630–2649 (2004)
9. Jonasson, J.: Uniqueness of uniform random colorings of regular trees. Stat. Prob. Lett. **57**, 243–248 (2002)
10. Kesten, H., Stigum, B.P.: Additional limit theorems for indecomposable multidimensional Galton-Watson processes. Ann. Math. Stat. **37**, 1463–1481 (1966)
11. Krząkała, F., Montanari, A., Ricci-Tersenghi, F., Semerjian, G., Zdeborova, L.: Gibbs states and the set of solutions of random constraint satisfaction problems. Proc. Nat. Acad. Sci. **104**, 10318–10323 (2007)
12. Krząkała, F., Pagnani, A., Weigt, M.: Threshold values, stability analysis, and high-q asymptotics for the coloring problem on random graphs. Phys. Rev. E **70**(4), 046705 (2004)
13. Lange, K.: *Applied probability*. Springer Texts in Statistics. New York: Springer-Verlag, 2003
14. Mézard, M., Montanari, A.: Reconstruction on trees and spin glass transition. J. Stat. Phys. **124**(6), 1317–1350 (2006)
15. Mossel, E.: Reconstruction on trees: beating the second eigenvalue. Ann. Appl. Prob. **11**(1), 285–300 (2001)
16. Mossel, E.: Phase transitions in phylogeny. Trans. Amer. Math. Soc. **356**(6), 2379–2404 (electronic), (2004)
17. Mossel, E.: Survey: information flow on trees. In: *Graphs, morphisms and statistical physics*, Volume **63** of DIMACS Ser. Discrete Math. Theoret. Comput. Sci., Providence, RI: Amer. Math. Soc., 2004, pp. 155–170
18. Mossel, E., Peres, Y.: Information flow on trees. Ann. Appl. Probab. **13**, 817–844 (2003)
19. Mossel, E., Sly, A.: *Gibbs rapidly samples colorings of G(n,d/n)*. http://arxiv.org/abs/0707.3241V2[math.PR], 2007
20. Semerjian, G.: On the freezing of variables in random constraint satisfaction problems. J. Stat. Phys. **130**, 251 (2008)
21. Zdeborová, L., Krząkała, F.: Phase transitions in the coloring of random graphs. Phys. Rev. E **76**, 031131 (2007)