The Tightness of the Kesten-Stigum Reconstruction Bound for a Symmetric Model With Multiple Mutations

Wenjian Liu

City University of New York

Frontier Probability Days 2018

Joint work with Dr. Sreenivasa Rao Jammalamadaka and Dr. Ning Ning
Department of Statistics and Applied Probability, UCSB

Outline

Outline

Motivations

INFLUENCE MAXIMIZATION: Propagation of Opinions

- Rapid and global growth of online social networks with millions of Internet users
- How does a persons opinion change the opinions of other people in the network?

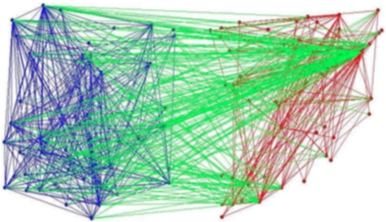
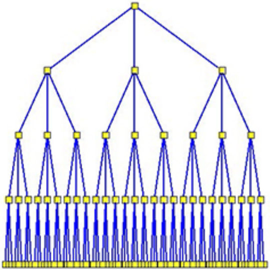


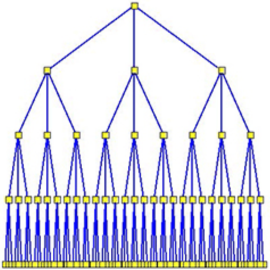
Fig. 1. The US Senator network. Links indicate cosponsorship of bills by senators, while the link colors indicate party affiliations.

Preliminaries

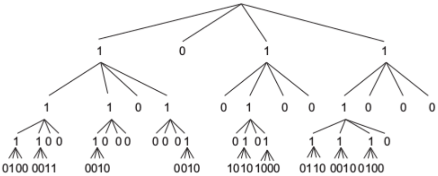


- We focus on d-ary trees $\mathbb{T} = (\mathbb{V}, \mathbb{E}, \rho)$, i.e. the infinite rooted tree where every vertex has exactly d offspring, with nodes $\mathbb{V}$, edges $\mathbb{E}$ and root $\rho \in \mathbb{V}$

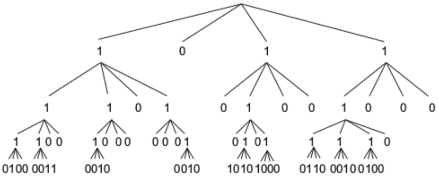
Preliminaries



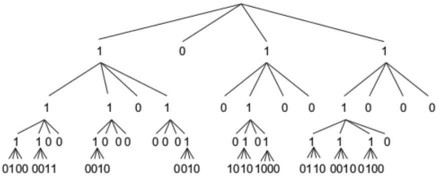
- We focus on d-ary trees $\mathbb{T} = (\mathbb{V}, \mathbb{E}, \rho)$, i.e. the infinite rooted tree where every vertex has exactly d offspring, with nodes $\mathbb{V}$, edges $\mathbb{E}$ and root $\rho \in \mathbb{V}$
- Each edge of the tree acts as a channel on a finite characters set $\mathcal{C}$, whose elements are configurations on $\mathbb{T}$, denoted by σ.



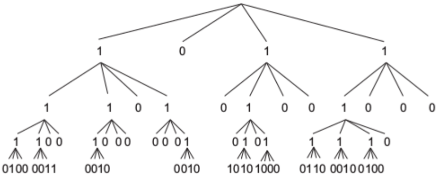
- The probability transition matrix $\mathbf{M} = (M_{ij})$ is chosen as the noisy communication channel on each edge.



- The probability transition matrix $\mathbf{M} = (M_{ij})$ is chosen as the noisy communication channel on each edge.
- The state of the root ρ, denoted by σ_ρ, is chosen according to an initial distribution π on $\mathcal{C}$.



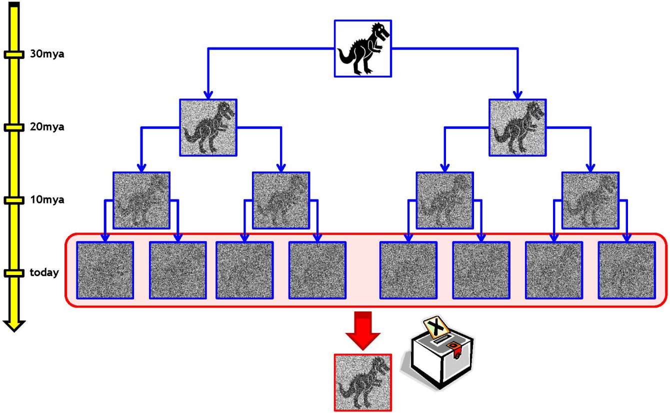
- The probability transition matrix $\mathbf{M} = (M_{ij})$ is chosen as the noisy communication channel on each edge.
- The state of the root ρ, denoted by σ_ρ, is chosen according to an initial distribution π on $\mathcal{C}$.
- Let $\sigma(n)$ denote the spins at distance n from the root and let $\sigma^i(n)$ denote $\sigma(n)$ conditioned on $\sigma_\rho = i$.



- The probability transition matrix $\mathbf{M} = (M_{ij})$ is chosen as the noisy communication channel on each edge.
- The state of the root ρ, denoted by σ_ρ, is chosen according to an initial distribution π on $\mathcal{C}$.
- Let $\sigma(n)$ denote the spins at distance n from the root and let $\sigma^i(n)$ denote $\sigma(n)$ conditioned on $\sigma_\rho = i$.
- This is a discrete, irreducible, aperiodic, and reversible Markov chain.

Reconstruction Question

Consider all the symbols received at the vertices of the n^{th} generation. Does this configuration contain a non-vanishing information on the letter transmitted by the root, as n goes to ∞?



Reconstruction Definition

Definition

The reconstruction problem for the infinite tree $\mathbb{T}$ is *solvable* if for some $i, j \in \mathcal{C}$,

$$\limsup_{n\to\infty} d_{TV}(\sigma^i(n), \sigma^j(n)) > 0$$

where d_{TV} is the total variation distance, i.e.

$$d_{TV}(\sigma^i(n), \sigma^j(n)) = \sup_A |\mathbb{P}(\sigma(n) = A \mid \sigma_\rho = i) - \mathbb{P}(\sigma(n) = A \mid \sigma_\rho = j)|$$

When the $\limsup$ is 0 we will say the model has *non-reconstruction* on $\mathbb{T}$.

Application

- Phylogenetic reconstruction is a major task of systematic biology, which is to construct the ancestry tree of a collection of species, given the information of present species.
- The corresponding reconstruction threshold answers the question whether the ancestral DNA information can be reconstructed from a known phylogenetic tree.
- This threshold is also crucial to determine the number of samples required, in the sense that, only enumerations of each type of spin at the leaves are collected, regardless of their positions on the leaves.

Applications

- Reconstruction threshold on trees plays an important role in the dynamic phase transitions in certain glassy systems subject to random constraints.
- For random colorings on the Erdös–Rényi random graph with average connectivity d, Achlioptas and Coja-Oghlan [2008] proved that there is a phase transition, from the situation that most of the mass is contained in one giant component, to the case that the space of solutions breaks into exponentially many smaller clusters.
- This phase transition has been proved corresponding to known bounds on the reconstruction threshold for proper colorings on trees, see e.g. Mossel and Peres [2003], Semerjian [2008] and Sly [2009].

Outline

Probability Transition Matrix (Classical Models)

$$\textbf{Symmetric Ising: } \mathbf{M} = \frac{1}{2}\begin{pmatrix} 1+\theta & 1-\theta \\ 1-\theta & 1+\theta \end{pmatrix}$$

$$\textbf{Asymmetric Ising: } \mathbf{M} = \frac{1}{2}\left[\begin{pmatrix} 1+\theta & 1-\theta \\ 1-\theta & 1+\theta \end{pmatrix} + \Delta \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}\right]$$

$$\textbf{Potts: } \mathbf{M} = \begin{pmatrix} 1-p & \frac{p}{q-1} & \cdots & \frac{p}{q-1} \\ \frac{p}{q-1} & 1-p & \cdots & \frac{p}{q-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{p}{q-1} & \frac{p}{q-1} & \cdots & 1-p \end{pmatrix}_{q \times q}$$

Background

- Kesten-Stigum bound: the reconstruction problem is solvable if $d|\lambda|^2 > 1$.

Background

- Kesten-Stigum bound: the reconstruction problem is solvable if $d|\lambda|^2 > 1$.
- What about $d|\lambda|^2 < 1$?

Background

- Kesten-Stigum bound: the reconstruction problem is solvable if $d|\lambda|^2 > 1$.
- What about $d|\lambda|^2 < 1$?
- For the binary symmetric channel, the reconstruction problem is solvable if and only if $d\lambda^2 > 1$ (Bleher et al. [1995]).

Background

- Kesten-Stigum bound: the reconstruction problem is solvable if $d|\lambda|^2 > 1$.
- What about $d|\lambda|^2 < 1$?
- For the binary symmetric channel, the reconstruction problem is solvable if and only if $d\lambda^2 > 1$ (Bleher et al. [1995]).
- The KS 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 (Mossel [2001, 2004]).

Background

- Kesten-Stigum bound: the reconstruction problem is solvable if $d|\lambda|^2 > 1$.
- What about $d|\lambda|^2 < 1$?
- For the binary symmetric channel, the reconstruction problem is solvable if and only if $d\lambda^2 > 1$ (Bleher et al. [1995]).
- The KS 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 (Mossel [2001, 2004]).
- The first exact reconstruction threshold in roughly a decade, was obtained by Borgs et al. [2006].

Background

- Kesten-Stigum bound: the reconstruction problem is solvable if $d|\lambda|^2 > 1$.
- What about $d|\lambda|^2 < 1$?
- For the binary symmetric channel, the reconstruction problem is solvable if and only if $d\lambda^2 > 1$ (Bleher et al. [1995]).
- The KS 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 (Mossel [2001, 2004]).
- The first exact reconstruction threshold in roughly a decade, was obtained by Borgs et al. [2006].
- The KS bound on the Potts models was completely investigated by Sly [2011].

Background

- Kesten-Stigum bound: the reconstruction problem is solvable if $d|\lambda|^2 > 1$.
- What about $d|\lambda|^2 < 1$?
- For the binary symmetric channel, the reconstruction problem is solvable if and only if $d\lambda^2 > 1$ (Bleher et al. [1995]).
- The KS 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 (Mossel [2001, 2004]).
- The first exact reconstruction threshold in roughly a decade, was obtained by Borgs et al. [2006].
- The KS bound on the Potts models was completely investigated by Sly [2011].
- The KS bound on the Asymmetric Ising model was analyzed by Liu and Ning [2016, 2017].

Outline

Probability Transition Matrix (Motivation)

- In computational biology, the broadcast model is the main model for the evolution of base pairs of DNA.

  The K80 model (Kimura, 1980), as one of the most common models of DNA evolution, distinguishes between transitions ($A \leftrightarrow G$, i.e. from purine to purine, or $C \leftrightarrow T$, i.e. from pyrimidine to pyrimidine) and transversions (from purine to pyrimidine or vice versa).

Probability Transition Matrix

Consider a characters set $\mathcal{C} = \{1, \ldots, q\} \cup \{q+1, \ldots, 2q\}$, with $q \geq 2$, and the state of the root ρ is chosen according to the uniform distribution on $\mathcal{C}$.

$$M_{ij} = \begin{cases} p_0 & \text{if } i = j, \\ p_1 & \text{if } i \neq j \text{ and } i, j \text{ are in the same category}, \\ p_2 & \text{if } i \neq j \text{ and } i, j \text{ are in different categories}. \end{cases}$$

where p_0, p_1 and p_2 are all nonnegative, such that $p_0 + (q-1)p_1 + qp_2 = 1$.

Eigenvalues: $\lambda_1 = p_0 - p_1$, $\lambda_2 = p_0 + (q-1)p_1 - qp_2$, and $\lambda_3 = 1$.

Main Results

Improved flexibility comes along with increased complexity. The additional class of mutation complicates the discussion of the second largest eigenvalue in absolute value.

Theorem

When $q \geq 4$, for every d the Kesten-Stigum bound is not tight, i.e. the reconstruction is solvable for some λ even if $d\lambda^2 < 1$.

Theorem

When $q = 2$, for any $\kappa > 1$ if $\max\left\{\frac{|\lambda_1|}{|\lambda_2|}, \frac{|\lambda_2|}{|\lambda_1|}\right\} \geq \kappa$, there exists a $D = D(\kappa) > 0$ such that for $d \geq D$ the Kesten-Stigum bound is tight, in addition, there is non-reconstruction at the Kesten-Stigum bound, namely, $|\lambda| = d^{-1/2}$.

Outline

Notations

$f_n(i, \sigma(n))$: the posterior probability of state i at the root given the random configuration $\sigma(n)$ of the leaves.

$$f_n(j, \sigma^i(n)) \overset{\mathbb{D}}{\sim} \begin{cases} X^{(1)}(n) & \text{if } i = j, \\ X^{(2)}(n) & \text{if } i \neq j \text{ are in the same category}, \\ X^{(3)}(n) & \text{if } i \neq j \text{ are in different categories}. \end{cases}$$

$$x_n = \mathbf{E}\left(X^{(1)}(n) - \frac{1}{2q}\right), \quad z_n = \mathbf{E}\left(X^{(3)}(n) - \frac{1}{2q}\right),$$

$$u_n = \mathbf{E}\left(X^{(1)}(n) - \frac{1}{2q}\right)^2, \quad w_n = \mathbf{E}\left(X^{(3)}(n) - \frac{1}{2q}\right)^2.$$

The equivalent condition for non-reconstruction

Lemma

The non-reconstruction is equivalent to

$$\lim_{n\to\infty} x_n = 0.$$

Moment recursion

- Key Idea: analyze the recursive relation between $X^{(i)}(n)$ and $X^{(i)}(n+1)$ by the structure of the tree.

Moment recursion

- Key Idea: analyze the recursive relation between $X^{(i)}(n)$ and $X^{(i)}(n+1)$ by the structure of the tree.
- $Y_{ij}(n) = f_n\left(i, \sigma_j^1(n+1)\right),$

Moment recursion

- Key Idea: analyze the recursive relation between $X^{(i)}(n)$ and $X^{(i)}(n+1)$ by the structure of the tree.
- $Y_{ij}(n) = f_n\left(i, \sigma_j^1(n+1)\right)$,
- Markov random field property

$$X^{(1)}(n+1) = f_{n+1}(1, \sigma^1(n+1)) = \frac{Z_1}{\sum_{i=1}^{2q} Z_i}$$

and

$$X^{(3)}(n+1) = f_{n+1}(q+1, \sigma^1(n+1)) = \frac{Z_{q+1}}{\sum_{i=1}^{2q} Z_i},$$

- if $1 \leq i \leq q$

$$Z_i = \prod_{j=1}^{d} \left[1 + 2q\lambda_1 \left(Y_{ij} - \frac{1}{2q} \right) + 2(\lambda_1 - \lambda_2) \sum_{q+1 \leq \ell \leq 2q} \left(Y_{\ell j} - \frac{1}{2q} \right) \right]$$

- if $q + 1 \leq i \leq 2q$

$$Z_i = \prod_{j=1}^{d} \left[1 + 2q\lambda_1 \left(Y_{ij} - \frac{1}{2q} \right) + 2(\lambda_1 - \lambda_2) \sum_{1 \leq \ell \leq q} \left(Y_{\ell j} - \frac{1}{2q} \right) \right].$$

Nonlinear Dynamical System

Taking $\mathcal{X}_n = x_n + z_n$ and $\mathcal{Z}_n = -z_n$, we can obtain a nonlinear dynamical system.

$$\begin{cases} \mathcal{X}_{n+1} = d\lambda_1^2\mathcal{X}_n + \frac{d(d-1)}{2}\left(\frac{2q(q-3)}{q-1}\lambda_1^4\mathcal{X}_n^2 + 4q\lambda_1^2\lambda_2^2\mathcal{X}_n\mathcal{Z}_n\right) \\ \qquad\qquad +R_x + R_z + V_x \\ \\ \mathcal{Z}_{n+1} = d\lambda_2^2\mathcal{Z}_n + \frac{d(d-1)}{2}\left(\frac{q}{q-1}\lambda_1^4\mathcal{X}_n^2 - 4q\lambda_2^4\mathcal{Z}_n^2\right) - R_z + V_z \end{cases}$$

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

$$R_z = \mathbf{E}\left(\frac{Z_{q+1}}{\sum_{i=1}^{2q} Z_i} - \frac{1}{2q}\right)\frac{(\sum_{i=1}^{2q} Z_i - 2q)^2}{(2q)^2},$$

and

$$|V_x|, |V_z| \leq C_V x_n^2 \left(\left|\frac{u_n}{x_n} - \frac{1}{2q}\right| + \left|\frac{w_n}{x_n} - \frac{1}{2q}\right| + x_n\right)$$

with C_V a constant depending on q only.

Outline

x_n drops slowly

x_n does not drop from a very large value to a very small one.

Lemma

For any $a > 0$, there exists a constant $\gamma = \gamma(q, a) > 0$ such that

$$x_{n+1} \geq \gamma x_n,$$

for all n, if $\min\{|\lambda_1|, |\lambda_2|\} > a$,

Weak interactions between spins for large n

Fixed finite different vertices far away from the root can effect the root little.

Lemma

For any $\varepsilon > 0$ and positive integer k there exists $M = M(q, \varepsilon, k)$ such that for any collection of vertices $v_1, \ldots, v_k \in L(M)$,

$$\sup_{i, i_1, \ldots, i_k \in \mathcal{C}} \left| \mathbf{P}(\sigma_\rho = i \mid \sigma_{v_j} = i_j, 1 \leq j \leq k) - \frac{1}{2q} \right| \leq \varepsilon$$

Concentration Analysis

Theorem

Assume $\min\{|\lambda_1|, |\lambda_2|\} > \varrho$ *for some* $\varrho > 0$. *Given arbitrary* $\varepsilon, \alpha > 0$ *there exist constants* $C = C(q, \varepsilon, \alpha, \varrho) > 0$ *and* $N = N(q, \varepsilon, \alpha)$ *such that when* $n \geq N$,

$$\mathbf{P}\left(\left|\frac{Z_1}{\sum_{i=1}^{2q} Z_i} - \frac{1}{2q}\right| > \varepsilon\right) \leq Cx_n^{\alpha}$$

and

$$\mathbf{P}\left(\left|\frac{Z_{q+1}}{\sum_{i=1}^{2q} Z_i} - \frac{1}{2q}\right| > \varepsilon\right) \leq Cx_n^{\alpha}.$$

Theorem

Assume $|\lambda_2| > a$ *and* $|\lambda_1| = |\lambda_2|$ *or* $|\lambda_1|/|\lambda_2| \geq \kappa$ *for some* $\kappa > 1$*. For any* $\varepsilon > 0$*, there exist* $N = N(q, \kappa, \varepsilon)$ *and* $\delta = \delta(q, \kappa, a, \varepsilon) > 0$ *such that if* $n \geq N$ *and* $x_n \leq \delta$ *then*

$$\left|\frac{u_n}{x_n} - \frac{1}{2q}\right| < \varepsilon \quad \text{and} \quad \left|\frac{w_n}{x_n} - \frac{1}{2q}\right| < \varepsilon.$$

Proof of Reconstruction

$$\begin{cases} \mathcal{X}_{n+1} \approx d\lambda_1^2 \mathcal{X}_n + \frac{d(d-1)}{2}\left(\frac{2q(q-3)}{q-1}\lambda_1^4 \mathcal{X}_n^2 + 4q\lambda_1^2\lambda_2^2 \mathcal{X}_n \mathcal{Z}_n\right) \\ \mathcal{Z}_{n+1} \approx d\lambda_2^2 \mathcal{Z}_n + \frac{d(d-1)}{2}\left(\frac{q}{q-1}\lambda_1^4 \mathcal{X}_n^2 - 4q\lambda_2^4 \mathcal{Z}_n^2\right) \end{cases}$$

Finally, we investigate the stability of the system. When $q \geq 4$, even if $d\lambda_1^2 < 1$ for some λ_1, x_n will not converge to 0 and hence there is reconstruction beyond the Kesten-Stigum bound.

Outline

Large Degree Asymptotics

Define

(A) when $1 \leq i \leq q$

$$U_{ij} = \log \left[1 + 2q\lambda_1 \left(Y_{ij} - \frac{1}{2q} \right) + 2(\lambda_1 - \lambda_2) \sum_{q+1 \leq \ell \leq 2q} \left(Y_{\ell j} - \frac{1}{2q} \right) \right];$$

(B) when $q + 1 \leq i \leq 2q$

$$U_{ij} = \log \left[1 + 2q\lambda_1 \left(Y_{ij} - \frac{1}{2q} \right) + 2(\lambda_1 - \lambda_2) \sum_{1 \leq \ell \leq q} \left(Y_{\ell j} - \frac{1}{2q} \right) \right]$$

Gaussian Approximation

- Recall

$$x_{n+1} = \frac{\exp(\sum_{j=1}^{d} U_{ij})}{\sum_{i=1}^{2q} \exp(\sum_{j=1}^{d} U_{ij})} - \frac{1}{2q}$$

- Construct a multivariate Gaussian distribution

$$W_i = s\mu_i + \sqrt{s}U_i + t\nu_i + \sqrt{t}V_i$$

- Gaussian Approximation Function:

$$f(s,t) = \mathbf{E}\frac{e^{W_1}}{\sum_{i=1}^{2q} e^{W_i}} - \frac{1}{2q}$$

Non-reconstruction for $q = 2$

Lemma

For each $\varepsilon > 0$ *there exists an positive integer* $D = D(q, \varepsilon)$ *such that for all* n *when* $d > D$

$$\left|x_{n+1} - f(d\lambda_1^2 X_n, d\lambda_2^2 Z_n)\right| \leq \varepsilon.$$

REFERENCES I

Georgii, H. O.
Gibbs Measures and Phase Transition.
de Gruyter, Berlin, 1988

Berger, N., Kenyon, C., Mossel, E. and Peres, Y.
Glauber dynamics on trees and hyperbolic graphs.
Probab. Theory Related Fields, **131** 311–340, 2005

Bhamidi, S., Rajagopal, R. and Roch, S.
Network delay inference from additive metrics.
Random Structures Algorithms, **37** 176–203, 2010

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

REFERENCES II

Borgs, C., Chayes, J. T., Mossel, E. and Roch, S.
The Kesten-Stigum reconstruction bound is tight for roughly symmetric binary channels.
FOCS 518–530. IEEE Comput. Soc., Berkeley, CA, 2006

Evans, W., Kenyon, C., Peres, Y. and Schulman, L. J.
Broadcasting on trees and the Ising model.
Ann. Appl. Probab., **10** 410–433, 2000

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

REFERENCES III

Kesten, H. and Stigum, B. P.
Limit theorems for decomposable multi-dimesional Galton-Watson processes.
J. Math. Anal. Appl., **17** 309–338, 1966

Liu, W. and Ning, N.
Reconstruction for the Asymmetric Ising Model on Regular Trees.
Proceedings of IET-IEEE Electrical, Electronics, Engineering Trends, Communication, Optimization and Sciences (EEECOS), 2016.

Liu, W. and Ning, N.
The Tightness of the Kesten-Stigum Reconstruction Bound of Symmetric Model with Multiple Mutations,
Journal of Statistical Physics, **170** 617–641, 2018

REFERENCES IV

Martinelli, F., Sinclair, A. and Weitz, D.
Fast mixing for independent sets, colorings, and other models on trees.
Random Structures Algorithms, **31** 134–172, 2007

Mézard, M. and Montanari, A.
Reconstruction on trees and spin glass transition.
J. Stat. Phys., **124** 1317–1350, 2006

Mossel, E.
Reconstruction on trees: beating the second eigenvalue.
Ann. Appl. Probab., **11** 285–300, 2001

Mossel, E.
Phase transitions in phylogeny.
Trans. Amer. Math. Soc. **356** 2379–2404, 2004

Mossel, E.
Survey: information flow on trees.
Graphs, morphisms and statistical physics. DIMACS Ser. Discrete Math. Theoret. Comput. Sci., **63** 155–170. Amer. Math. Soc., Providence, RI, 2004

Sly, A.
Reconstruction for the Potts model.
Ann. Probab., **39** 1365–1406, 2011

Tetali, P., Vera, J. C., Vigoda, E. and Yang, L.
Phase transition for the mixing time of the Glauber dynamics for coloring regular trees.
Ann. Appl. Probab., **22** 2210–2239, 2012

Thank you!