A Study on the Non-Reconstruction Conjecture in Markov Trees

Yurong Liu

Abstract

Consider a d-ary tree T which simulates the process of broadcasting information from the root to other vertices, where each edge is a copy of an irreducible and aperiodic Markov chain M with reversible transition matrix $M \in \mathbb{R}^{n\times n}$ on state space , the goal is to reconstruct the value of the root given values of nodes at level n of the tree, where $n \to \infty$. This branching process is useful for modeling complex populations that exhibit dependencies between the states of individuals and their ancestors. It can be used to study a wide range of phenomena, including the spread of diseases in populations, the growth of organisms in ecosystems, and the diffusion of information and ideas. We are going to work on the non-reconstruction conjecture of this problem. The conjecture states that information on root cannot be reconstructed if $|\lambda_2(M)| < \frac{1}{d}$, where $\lambda_2(M)$ is the second largest eigenvalue of M. Our focus is on the scenario where M is symmetric.

1 Introduction

The study of information propagation has gained significant attention in recent years due to its wide-ranging applications in diverse domains such as epidemiology, ecology, and social network analysis. The ability to model the behavior of these systems, as well as the limitations of information recovery, can provide valuable insights into the underlying mechanisms driving their dynamics.

In this paper, we investigate a specific instance of information broadcasting in a d-ary tree, wherein the edges represent irreducible and aperiodic Markov chains with a symmetric transition matrix. The d-ary tree T serves as a natural model for representing the process of broadcasting information from a root node to the remaining vertices. Each edge in this tree is a copy of an irreducible and aperiodic Markov chain M with a reversible transition matrix $M \in \mathbb{R}^{n\times n}$ on the state space . Our objective is to reconstruct the value of the root node based on the values of nodes at level n of the tree, as $n \to \infty$. This branching process is particularly relevant in the context of modeling complex populations that exhibit dependencies between the states of individuals and their ancestors.

We dedicate our e orts to understanding the non-reconstruction conjecture as-optimization problems. It can also be linked to the reconstruction problem for the Potts model [5], a generalization of the Ising model used in statistical mechanics to describe the behavior of interacting particles in a lattice. In this setting, the d-ary tree with symmetric Markov chain edges can be viewed as a lattice structure, where each node represents a particle with one of the possible discrete states.

2.2 Existing reconstruction methods

Various reconstruction methods have been developed to address the problem of inferring the root state in a d-ary tree. One such method to use maximum likelihood estimation (MLE), which is consistent for inferring the tree topology [6]. In particular, we find the optimal assignment of states to the root node that maximizes the likelihood of the observed data. Another approach is the census method, which involves observing whether the census of the configuration at level n contains any significant information on the root variable. Reconstruction (and census) solvability when $d\,\lambda_2(M) > 1$ was initially demonstrated in [7], though it was expressed in the context of multi-type branching processes which we will later introduce in §3.3. The proofs of the non-reconstruction result when $d\,\lambda_2(M) \leq 1$ are harder as shown in [8], where it's also demonstrated that the asymptotic independence of the root in the census is determined by the spectral properties of M.

3 Preliminaries

3.1 Markov chains

In this section, we introduce the basic concepts and notations related to Markov chains, which will be employed throughout the paper to analyze the non-reconstruction conjecture in information broadcast over d-ary trees.

A Markov chain is a stochastic process that models the transition between states in a system, where the future state depends only on the current state and not on the past states. This property is known as the Markov property.

Definition 1. (Markov Chain) A Markov chain is a sequence of random variables $X_n, n \in \mathbb{N}$ taking values in a finite or countable state space and satisfying the Markov property: for any $n \in \mathbb{N}$ and any states $x_0, x_1, \ldots, x_{n+1} \in$ x 1

Definition 2. (Transition Matrix) Let $\mathcal{M}$ be a Markov chain with state space Ω. The transition matrix $M \in \mathbb{R}^{|\Omega| \times |\Omega|}$ of $\mathcal{M}$ is a matrix such that M_{ij} is the probability of transitioning from state i to state j:

$$M_{ij} = P(X_{n+1} = j | X_n = i), i, j \in \Omega,$$

where $\forall i, j \in \Omega, M_{ij} \geq 0$, and $\sum_{j=1}^{n} M_{ij} = 1$ for all $i \in \Omega$.

A stationary distribution is a probability distribution over the state space of a Markov chain that remains invariant under the transition probabilities.

Definition 3. (Stationary Distribution) Let $\mathcal{M}$ be a Markov chain with transition matrix M. A probability distribution π over the state space Ω is a stationary distribution of $\mathcal{M}$ if

$$\pi M = \pi.$$

Note that another way to express this is that π is an eigenvector with all its elements being nonnegative, and its associated eigenvalue is 1.

Example 1. Consider a Markov chain represented by a random walk on the nodes of an n-cycle. At each step, there is a $1/2$ probability of staying at the current node, a $1/4$ probability of moving left, and a $1/4$ probability of moving right. The uniform distribution, which assigns a probability of $1/n$ to each node, acts as a stationary distribution for this chain, because it remains constant after performing a single step in the chain.

For Markov chains, irreducibility and aperiodicity are essential properties that ensure the existence and uniqueness of a stationary distribution.

Definition 4. (Irreducibility) A Markov chain with transition matrix M is *irreducible* if there exists a sequence of transitions between any pair of states $i, j \in \Omega$ with positive probability

$$\forall i, j \in \Omega, \exists t \in \mathbb{N} \text{ s.t. } (M^t)_{ij} > 0.$$

Definition 5. (Aperiodicity) A Markov chain with transition matrix M is *aperiodic* if for all states $i \in \Omega$, the greatest common divisor of the set $\{t \in \mathbb{N} : (M^t)_{ii} > 0\}$ equals 1.

Theorem 1. If a Markov chain $\mathcal{M}$ is irreducible then it has a unique stationary distribution π.

A Markov chain is said to be *ergodic* if it is both irreducible and aperiodic. Hence we derive the definition of ergodicity as follows.

Theorem 2. (Convergence to stationary distribution) If a Markov chain $\mathcal{M}$ is ergodic, then there exists a unique stationary distribution π such that for any given (initial) distribution μ, $\lim_{t \to \infty} \mu M^t = \pi$.

Definition 6. (Reversibility) An ergodic Markov chain is *reversible* if the stationary distribution π satisfies the detailed balance equations: $\forall i, j \in \Omega, \pi_i M_{ij} = \pi_j M_{ji}$

3.2 Coupling

Coupling is a technique used in probability theory to study the convergence of Markov chains. It involves constructing two Markov chains on the same probability space that eventually couple or synchronize their states. We will employ this technique in subsequent proofs. In short, the term coupling in probability refers to creating a joint distribution from two separate distributions, and , with the resulting joint distribution having and as its marginals. This coupling can provide valuable insight into the di erence between the two distributions, measured by the total variation distance. Suppose ; are two distributions on , we want to de ne measures that enable us to compare and .

De nition 7. (Coupling) A coupling ! is a joint distribution on × such that

$$8y; \quad \underset{x\,2}{X} \; ! \,(x;y) = \quad (y);$$

$$8x; \quad \underset{y\,2}{X} \; ! \,(x;y) = \quad (x):$$

where ; are two distributions on .

Example 2. Consider a Markov chain on the state space = f 0; 1g with the following transition probability matrix M :

$$M = \begin{matrix} 0{:}7 & 0{:}3 \\ 0{:}6 & 0{:}4 \end{matrix}$$

We want to study the convergence of this Markov chain to its stationary distribution. To do this, we construct two copies of the Markov chain, say X and Y, with initial states x_0 and y_0, respectively, where $x_0 \neq y_0$. Now we de fne a coupling of these two chains such that:

- If $X_t = Y_t$: 1) If $X_t = Y_t = 0$, then $X_{t+1} = Y_{t+1}$ with probability 0.7 both

Note that this is only one possible coupling for the given Markov chain. Coupling works as long as the following conditions are satisfed:

- If X and Y are in the same state (i.e., $X_t = Y_t$), they stay synchronized (i.e., $X_{t+1} = Y_{t+1}$)
- If X and Y are in diferent states, they may synchronize with some probability

By constructing the coupled Markov chains X and Y, we can analyze the synchronization time (i.e., the time it takes for the chains to reach the same state) and use this information to study the convergence to the stationary distribution.

We also introduce a measure of the di erence between two probability distributions. It is de ned as the sum of the absolute di erences between the probabilities assigned to each event by the two distributions.

De nition 8. (Total Variation Distance) The **total variation distance** between probability distributions and is defned as

$$d_{TV} := \sup_{A\,2} j$$

Proof. For any event A and coupling (X, Y) for and ,

$$
\begin{aligned}
(A) - (A) &= P[X \in A] - P[Y \in A] \\
&= P[X \in A; X = Y] + P[X \in A; X \neq Y] - P[Y \in A; X = Y] - P[Y \in A; X \neq Y] \\
&= P[X \in A; X \neq Y] - P[Y \in A; X \neq Y] \\
&\leq P[X \neq Y]
\end{aligned}
$$

The intuition is that we want to nd a coupling (X, Y) s.t. $X \neq Y$ only if $(x) \neq (x)$ i.e. x is in the marginals of ! coincide with and . The second line involves three cases when we randomly select a point x in : 1) $X \in A; Y \in A$; 2) $X \in A; Y \notin A$; 3) $X \notin A; Y \in A$. In case 1), we set $X = Y$; in case 2) and 3), we set $X \neq Y$. Similarly, we can show that

$$(A) - (A) \leq P[X \neq Y];$$

and hence

$$d_{TV} = \sup_{A \in} |(A) - (A)| \leq P[X \neq Y];$$

□

3.3 Galton-Watson Branching Process

The Galton-Watson branching process (or GW-process for short) is a mathematical model that describes the evolution of a population over time. Formally, the GW-process can be de ned as a discrete-time branching process, where the number of o spring produced by each individual in the population is modeled as a random variable. This random variable is typically assumed to follow a certain probability distribution, such as the Poisson distribution or the geometric distribution, which determines the average number of o spring and the variance in the number of o spring. The size of the population at any given time is given by the sum of the number of o spring produced by each individual in the previous generation.

The GW-process is used to model a variety of real-world systems, including the spread of diseases, the growth of populations, and the evolution of species. By analyzing the behavior of the GW-process, it is possible to obtain information about the long-term behavior of the population, such as the probability of extinction or the average population size over time.

Example 3. Consider a branching process modeling population growth, where each individual can have 0, 1, or 2 o spring with probabilities 0.4, 0.4 and 0.2, respectively. Starting with a single individual (generation 0), the process unfolds in discrete generations. Each individual in generation n produces a random number of o spring (0, 1, or 2) according to the given probabilities, forming generation $n+1$. This Galton-Watson process models the evolution of the population over time, capturing growth or extinction dynamics.

3.3.1 Single-type Branching Process

The most common formulation of a branching process is Galton–Watson process.

Definition 9. A Galton-Watson process is a discrete-time Markov chain $\{M_n = 0, 1, 2, \ldots\}$ on , where M_n denote the number of individuals on n^{th} level, with transition function defined in terms of offspring distribution $\{p_k\}$, where $k = 0, 1, 2, \cdots$, $p_k \geq 0$, and $\sum p_k = 1$, by

$$P(i, j) = P\{M_{n+1}$$

3.3.2 Multi-type Branching Process

In many scenarios, the individuals in a branching process are not identical. Some examples of this include: 1) Population Genetics - where the inheritance of alleles can be modeled by a 3-type branching process that corresponds to the genotypes; 2) Physics - such as cosmic-ray cascades that involve both electrons and photons and can be modeled by a 2-type branching process. Multi-type branching process refers to a mathematical model that describes the evolution of a population in which individuals can give rise to offspring of multiple types, and the number and type of offspring is determined by a probability distribution that depends on the current state of the individual and its ancestry. In our case, we can form the multi-type branching process as [10].

Definition 10. A multi ($-$type) Galton-Watson process is a Markov chain $\{M_n^{\kappa} = 0; 1; 2; \ldots\}$ on $^{\kappa}$, where M_n is a -dimensional vector whose i^{th} entry gives the number of individuals of type i on the n^{th} level, with transition function

$$P(x; y) = P\{M_{n+1} = y | M_n = x\}; \quad x; y \in {}^{\kappa}:$$

Now let m_j

4 Problem Definition

We now turn to the reconstruction problem. When the distribution of the process on n^{th} level is independent of the root value as n goes to infinity, we say that the root is non-reconstructible . In this case, we have no way to reconstruct given this "same" distribution. Following this intuition, we can formally define non-reconstructibility as follows.

Definition 11. Given Markov chain M with transition matrix M and two trees generated from random roots that are independent, where distributions of level n are denoted as $_n$ and $_n$, then the root is non-reconstructible if

$$\lim_{n \to \infty} d_{TV}(_n; _n) = 0 \tag{6}$$

Following Lemma 1, suppose we create random variables $X; Y$ with probability distributions $_n$ and $_n$, then we have

$$\lim_{n \to \infty} P(X \neq Y) = 0$$

if the root is non-reconstructible.

5 Recap on $M^{2\times 2}$ transition matrix

Mossel [11] has showed that the information of the root can not be reconstructed for the d-ary tree and binary symmetric channel M where transition matrix

$$M = \begin{matrix} 1- \ _1 & _1 \\ 1- \ _2 & _2 \end{matrix} \tag{7}$$

when $|\ _2(M)| = |\ _2 - \ _1| \leq \frac{1}{d}$.

Theorem 5. Let M be in form (7). Take integer d s.t. $|d\ _2(M)| \leq 1$, then the root is non-reconstructible for the d-ary tree.

5.1 Proof I

We first introduce the random process called -percolation [11]. Denote the $d-$ary tree as $T = \{V; E\}$, where V represents the set of vertices (nodes) in T, and E represents the set of edges. Consider $: E \to \{0; 1\}$ which maps from the set of edges to $\{0; 1\}$. Given any $e \in E$, we define $P((e) = 1) =$.

Now we prove can prove Theorem 5 following Mossel [11].

Proof. Given transition matrix $M^{k\times k}$, we first show tha f 14.432 1.495 Td [( )]TJ/F19 6.9738 Tf[(Giv)20 97Gi parts: 1) copying the original distribution; 2) broadcast via matrix N

Consider $= {}_2(M) = j\ {}_1 - {}_2 j$.

If ${}_1 - {}_2 < 0$, then $= {}_2 - {}_1$. Let $N = I$, where I is the identity matrix $\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}$, and $v = \frac{(1-\delta_2;\delta_2)}{1-\lambda}$. Then

$$M = ({}_2 - {}_1)I + \begin{matrix} 1- & {}_2 & {}_1 \\ 1- & {}_2 & {}_1 \end{matrix} ;$$

so for each row vector M_{i} in M, we have

$$\begin{aligned} M_{i} &= I_{i} + (1 - \ ) \cdot \frac{(1 - {}_2;\ {}_1)}{1-} \\ &= N_{i} + (1 - \ )v: \end{aligned}$$

Then similarly, if ${}_1 - {}_2 > 0$, then $= {}_1 - {}_2$. Let $N = J$, where $J = \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}$, and $v = \frac{(1-\delta_1;\delta_2)}{1-\lambda}$. Then

$$M = ({}_1 - {}_2)J + \begin{matrix} 1- & {}_1 & {}_2 \\ 1- & {}_1 & {}_2 \end{matrix} ;$$

so for each row vector M_{i} in M, we also have

$$\begin{aligned} M_{i} &= J_{i} + (1 - \ ) \cdot \frac{(1 - {}_1;\ {}_2)}{1-} \\ &= N_{i} + (1 - \ )v: \end{aligned}$$

We now show that when $d \leq 1$, the root is non-reconstructible given transition matrix M. In fact, for any transition matrix that can be written in the form (8), the broadcast process is non-reconstructible.

We simulate the broadcast on $d-$ary tree $T = f V; Eg$ with root node 2 as a -pMrcolati 9.9666n20 9.9626 Tf 54.047 0 Td [(pro)-28(ces(.)-49(Noate)-019(that)-402wv)28(e)-019uset th

that the prbabiltity

Then according the de nition of $(v; v^0)$, we de ne the procedure as follows

$$v^0 = \begin{cases} N_v(\ (v)) & \text{if} \ \ ((v; v^0) = 1 \\ Y_v & \text{if} \ \ ((v; v^0) = 0 \end{cases}$$

Therefore, for any node $v \, 2 \, V$, we have probability to perform the transition by M, and probability $1 -$ by Y, and the two di erent processes are independent.

In this way, we obtain a coupling of the two distributions on n^{th} level of T. Let the set of vertices that has path to root node that contains only set of edges E^0 s.t. $(E^0) = 1$ be L, and let the set of vertices at n^{th} level be S_n. Let the probability distribution given root, say $_v$, at n^{th} level be $_n$. Then if $L \setminus S_n = ;$, we obtain same distribution on n^{th} level given any value of root . Then since

$$\max_{\zeta_v ; \zeta_y \, 2} P(\ _n \ 6= \ _n) \le P(L \setminus S_n = ;);$$

and since it has been proved in [12] that when $d \ \le 1$,

$$\lim_{n \, ! \, 1} P(L \setminus S_n = ;) = 0;$$

we have

$$\lim_{n \, ! \, 1} \max_{\zeta_v ; \zeta_y \, 2} P(\ _n \ 6= \ _n) = 0;$$

which implies that the root is non-reconstructible. □

5.2 Proof II

An alternative proof using coupling is proposed as follows.

Proof. Say $_n$ and $_n$ are distributions of n^{th} level of trees started with di erent root values. Let $X_n; Y_n$ be random variables with probability distributions $_n$ and $_n$. By Lemma 1, we have

$$d_{TV}(\ ; \ ) \le P(X \ 6= \ Y);$$

so $P(X \ 6= \ Y)$ is an upper bound of d_{TV}. Now since $M = \begin{matrix} 1 - & 1 & 1 \\ 1 - & 2 & 2 \end{matrix}$,

$_2(M) = j \ _1 - \ _2 j \le \frac{1}{}$ 9.9626 Tt Tf 15.1v.543 -21.917 Td [(d)]TJ/F22 6.9738 Tf7 1.495l(Y)]TJ/F22 6.9738 Tf

Thus given that $|{}_1 - {}_2| \le \frac{1}{d}$, we have

$$\lim_{n \to \infty} P(X_{n,i} \neq Y_{n,i}) = \lim_{n \to \infty} |{}_1 - {}_2|^n = 0;$$

which implies that

$$\lim_{n \to \infty} d_{TV}({}_n; {}_n) \le \lim_{n \to \infty} P(X_n \neq Y_n) = 0:$$

Therefore, we've showed that when n goes to in nity, $X_{n+1;i}$ and $Y_{n+1;i}$ always agree. $\square$

6 Extend to $M^{3\times 3}$ transition matrix

Now we extend Theorem 5 to 3×3 transition matrices, simulating the transitions as multi-type branching processes with 3 types. We start with the case when M is positive de nite (PSD).

6.1 Symmetric 3×3 transition matrix with 2 variables

In order to apply coupling, we rst consider the following case where transition matrix M is symmetric and reversible with 2 variables.

Corollary 1. Let

$$M = \begin{pmatrix} 1 - {}_1 - {}_2 & {}_1 & {}_2 \\ {}_1 & 1 - {}_1 - {}_2 & {}_2 \\ {}_2 & {}_2 & 1 - 2{}_2 \end{pmatrix} \tag{10}$$

be a positive de nite symmetric transition matrix. Then if we take integer d s.t. $|d\,{}_2(M)| \le 1$, the root is non-reconstructible for the d-ary tree.

Given M in form 6.3, we denote the three states as A, B, and C, corresponds to row 1, 2, and 3. Note that we have a choice for coupling the two broadcast processes. Now since we want to compare the broadcast distributions given two di erent root nodes, we de ne the coupled new states AA, BB, CC, AB, AC, and BC. Note that when it reaches state AA, BB, or CC, two coupled distribution "agrees" and hence extinct.

Hence we only look at types AB, AC, and BC, where two distributions disagree. Hence we let the coupling matrix be in the form

$$\begin{pmatrix} P(AB \to AB) & P(AB \to AC) & P(AB \to BC) \\ P(AC \to AB) & P(AC \to AC) & P(AC \to BC) \\ P(BC \to AB) & P(BC \to AC) & P(BC \to BC) \end{pmatrix}$$

We rst want to show that there exists a coupling s.t.

$${}_2(M) = \ \text{(coupling matrix)} \tag{11}$$

by doing a case analysis. Note that given ${}_1 + {}_2 + {}_3 = 1$, we use the sign of

$$
\begin{gathered}
{}_1 - {}_2 \\
{}_2 - 1{=}3 \\
2\,{}_1 + {}_2 - 1 \\
2\,{}_2 + {}_1 - 1
\end{gathered}
$$

to determine the sign of the entries of M. We consider ${}_1 - {}_2$ and ${}_2 - 1{=}3$ as major cases, and the other two in sub cases.

Case I: Let ${}_1 \leq {}_2$; ${}_2 \leq \frac{1}{3}$, then $2\,{}_1 + {}_2 - 1 \leq 2\,{}_2 + {}_1 - 1 \leq 0$) ${}_1 \leq 1 - {}_1 - {}_2$; ${}_2 \leq 1 - {}_1 - {}_2$.

	A B	A C	B C
A B	$1 - 2\,{}_1 - {}_2$	0	0
A C	$1 - {}_1 - 2\,{}_2$	0	0
B C	${}_2 - {}_1$	0	$1 - 3\,{}_2$

$j\ {}_2(M)j = \ = 1 - 2\,{}_1 - {}_2$

Case II: Let ${}_1 \leq {}_2$; ${}_2 > \frac{1}{3}$, then $2\,{}_1 + {}_2 - 1 > 0$ only if $2\,{}_2 + {}_1 - 1 > 0$.

i) ${}_2 \leq 1 - {}_1 - {}_2$, ${}_1 \leq 1 - {}_1 - {}_2$

	A B	A C	B C
A B	$1 - 2\,{}_1 - {}_2$	0	0
A C	0 1 − 1 1 11≤ 2		

$$_{2}(M) = \quad = 3\,_{2} - 1$$

Case III: Let $_{1} > \ _{2};\ _{2} \leq \frac{1}{3}$, then $2\,_{2} + \ _{1} - 1 > 0$ only if $2\,_{1} + \ _{2} - 1 > 0$.

i) $_{1} \leq 1 - \ _{1} - \ _{2},\ _{2} \leq 1 - \ _{1} -$

6.2 Extend to certain distributions

Claim 1. Given $d-$ary tree formed by broadcast process M and transition matrix M. Let M_n denote the vector of node counts for each type at level nodes, then we have

$$
\begin{aligned}
E[X] &= 1 + \sum_{k=1}^{d} \binom{d}{k} (1- \ )^{d-k} \ ^{k} \cdot (k \cdot E[X]) \\
&= 1 + \sum_{k=1}^{d} \frac{d}{k} \binom{d-1}{k-1} (1- \ )^{d-k} \ ^{k-1} \cdot k \cdot E[X] \\
&= 1 + \ E[X] d \sum_{k=1}^{d} \binom{\ -1}{k-1} (1- \ )^{d-k} \ ^{k-1} \\
&= 1 + d \ \cdot E[X]
\end{aligned}
$$

which implies that

$$E[X] = \frac{1}{1-d}$$

and thus

$$
\begin{aligned}
E[Y] = E[Xd - (X-1)] &= \frac{1}{1-d} \cdot d - \left(\frac{1}{1-d} - 1\right) \\
&= \frac{d-1+1-d}{1-d} \\
&= \frac{d-1}{1-d} + 1 \\
&= \frac{(1- \ )d}{1-d}
\end{aligned}
$$

□

Hence Theorem 6 follows.

Theorem 6. Given $d-$ary tree T_M formed by transition matrix M with second eigenvalue $\ _2$. Let T_C be a d-regular tree formed by coupling matrix $C = (1 - \ )M + \ I$. Let the expected number of children for T_M and T_C be $E[M]$ and $E[C]$ respectively, and let $E[M] < d; E[C] = d$. Then T_M is non-reconstructible if T_C is non-reconstructible.

Proof. Suppose T_C is non-reconstructible, then T_M is non-reconstructible since the it corresponds to the e ective parts of T_C. Since C is symmetric and the second eigenvalue of C is $(1- \ ) \ _2 + \ $ by construction, if $d|(1- \ ) \ _2 + \ | < 1$, then by Corollary 1, tree T_C formed by C is non-reconstructible. Now we want to show that if $d| \ _2| < 1$, then $d|(1- \ ) \ _2 + \ | < 1$.

By Lemma 2, we have

$$d = E[Y] = \frac{(1- \ )d}{1-d}:$$

Then

$$
\begin{aligned}
& d|\ _2| < 1 \\
\Rightarrow\ & E[M]|\ _2| < 1 \\
\Rightarrow\ & \frac{(1-\ )d}{1-\ d}|\ _2| < 1 \\
\Rightarrow\ & (1-\ )d|\ _2| < 1 - d|\ _2| \\
\Rightarrow\ & d|(1-\ )\ _2 + \ | < 1
\end{aligned}
$$

Therefore, if $d|\ _2| < 1$, $d|(1-\ )\ _2 + \ | < 1$, which implies that T_C is non-reconstructible, hence the effective part of T_C is non-reconstructible, and thus T_M is non-reconstructible.

□

Therefore, we are able to show the non-reconstructibility of the tree with broadcast matrix M when $E[\text{children}] < d$.

6.3 Generalized case for 3×3 matrix with certain distributions

Now since we've proved in § 6.1 that when the transition matrix is 3×3 and is PSD, coupling proves the conjecture that when $E[\text{number of children}] \cdot \ _2(M) < 1$, the root non-reconstructible, we want to extend it to trees with general offspring distributions. We try to prove it case by case after obtaining the coupling matrix following what we did in § 6.1. We start with the 3 3×3 transition matrix in following distribution. Given

$$M = \begin{pmatrix} 1-a-b & a & b \\ a & 1-a-c & c \\ b & c & 1-b-c \end{pmatrix}$$

whose eigenvalues are 1 and $1 - a - b - c \pm \sqrt{a^2 - ab + b^2 - ac - bc + c^2}$, we have 6 combinations of $a; b; c$ that forms the general cases, which are

$$
\begin{aligned}
a \geq b \geq c \\
a \geq c \geq b \\
b \geq a \geq c \\
b \geq c \geq a \\
c \geq a \geq b \\
c \geq b \geq a
\end{aligned}
$$

Now consider the expressions

$$
\begin{aligned}
2a + b; 2a + c \\
2b + a; 2b + c \\
2c + a; 2c + b
\end{aligned}
$$

Let $ab; ac; ba; bc; ca; cb$ be there abbreviations. Then WLOG, given any case, say $a \geq b \geq c$, we have 9 sub-cases given any general case. Given $a \geq b \geq c$, then if $ab \leq 1$, all the other expressions are all less or equal to 1. If $ab \geq 1$, either $ac \leq 1$ or $ba \leq 1$ leads to all the other expressions follows all less or equal to 1. Continue this way, we can have the cases listed below:

$$\begin{cases} ab \geq 1 & \begin{cases} ac \geq 1; ba \geq 1 & \begin{cases} bc \geq 1; ca \geq 1 & \begin{cases} cb \geq 1 \\ cb \leq 1 \end{cases} \\ bc \geq 1; ca \leq 1 & \\ bc \leq 1; ca \geq 1 & \\ bc \leq 1; ca \leq 1 & \end{cases} \\ ac \geq 1; ba \leq 1 & \\ ac \leq 1; ba \geq 1 & \\ ac \leq 1; ba \leq 1 & \end{cases} \\ ab \leq 1 & \end{cases}$$

Imagine it as a tree. Every leaf node means 1 case where all the expressions follows (in the order of $ab; ac; ba; bc; ca; cb$ for case $a \geq b \geq c$) have to be less or equal to 1.

Hence we have in total 54 cases.

Now similar to what we did for 2×2 matrices, for each case, we compare the second eigenvalue of the transition matrix and the spectral radius of the coupling matrix. Then we notice that when

$$2a + b;\ 2a + c;\ 2b + a > 1$$
$$2b + c;\ 2c + a;\ 2c + b < 1$$

the coupling matrix is as follows

	A B	A C	B C
A B	$2a + c - 1$	$b - c$	0
A C	$2b + a - 1$	0	$1 - 2b - c$
B C	0	$a - b$	$1 - a - 2c$

and we obtain the result

$$\lambda_2(M) \neq \rho(\text{coupling matrix}) \tag{12}$$

Hence there exists a case where coupling fails. For instance, let

$$M = \begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix}:$$

Then in the bad case where the coupling matrix is as follows

$$\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix};$$

we have

$${}_2(M) = \sqrt{\frac{1}{3}}$$

but

$$(\text{coupling matrix}) = \frac{2}{3}.$$

In our future work, we plan to investigate the non-symmetric case further and solve the bad case.

7 Acknowledgment

I would like to give my special thanks to Professor Daniel Stefankovic, a truly insightful and supportive advisor, for granting me the invaluable opportunity to participate in such an interesting theoretical computer science project. His guidance and mentorship throughout the process have been exceptional. Moreover, I would like to express my gratitude to the members of my review committee — Professor Daniel Stefankovic, Professor Sevak Mkrtchyan, and Professor Jonathan Pakianathan — for their time and valuable suggestions on my paper. Furthermore, I am deeply appreciative of the professors who have significantly contributed to my university studies, particularly Professor Steve Gonek and Professor Naomi Jochnowitz. Their guidance and support have been instrumental in my continued pursuit of mathematical studies.

References

[1] Elchanan Mossel. On the impossibility of reconstructing ancestral data and phylogenies. *Journal of computational biology*, 10(5):669–676, 2003.

[2] Constantinos Daskalakis, Elchanan Mossel, and Sébastien Roch. Optimal phylogenetic reconstruction. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 159–168, 2006.

[3] Antoine Gerschenfeld and Andrea Montanari. Reconstruction for models on random graphs. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pages 194–204. IEEE, 2007.

[4] Nayantara Bhatnagar, Allan Sly, and Prasad Tetali. Reconstruction threshold for the hardcore model. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques: 13th International*

Workshop, APPROX 2010, and 14th International Workshop, RANDOM 2010, Barcelona, Spain, September 1-3, 2010. Proceedings, pages 434{447. Springer, 2010.

[5] Fa-Yueh Wu. The potts model. Reviews of modern physics, 54(1):235, 1982.

[6] Mark Pagel. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. Systematic biology, 48(3):612{622, 1999.

[7]