
Large Degree Asymptotics and the Reconstruction Threshold of the Asymmetric Binary Channels

Wenjian Liu · Ning Ning

Abstract In this paper, we consider a broadcasting process in which information is propagated from a given root node on a noisy tree network, and answer the question that whether the symbols at the n th level of the tree contain non-vanishing information of the root as n goes to infinity. Although the reconstruction problem on the tree has been studied in numerous contexts including information theory, mathematical genetics and statistical physics, the existing literatures with rigorous reconstruction thresholds established are very limited. In the remarkable work of Borgs, Chayes, Mossel and Roch (*The Kesten-Stigum reconstruction bound is tight for roughly symmetric binary channels*. FOCS, IEEE Comput. Soc. (2006): 518–530. Berkeley, CA.), the exact threshold for the reconstruction problem for a binary asymmetric channel on the d -ary tree is established, provided that the asymmetry is sufficiently small, which is the first exact reconstruction threshold obtained in roughly a decade. In this paper, by means of refined analyses of moment recursion on a weighted version of the magnetization, and concentration investigations, we rigorously give a complete answer to the question of how small it needs to be to establish the tightness of the reconstruction threshold and further determine its asymptotics of large degrees.

Keywords Kesten-Stigum reconstruction bound · Markov random fields on trees · Distributional recursion · Nonlinear dynamical system

Mathematics Subject Classification (2000) 60K35 · 82B26 · 82B20

1 Introduction

1.1 Broadcasting Process and the Reconstruction Problem

We consider the following broadcasting process that can be considered as a communication tree network, as a model for propagation of a genetic property or as a tree-indexed Markov

Wenjian Liu

Dept. of Mathematics and Computer Science, Queensborough Community College, City University of New York

Research supported by CUNY Community College Research Grant #1541

E-mail: wjliu@qcc.cuny.edu

Ning Ning (Corresponding Author)

Dept. of Applied Mathematics, University of Washington, Seattle

E-mail: ningnin@uw.edu

chain. In this paper, we restrict our attention to the regular d -ary trees, which is an infinite rooted tree where every vertex has exactly d offspring, denoted by $\mathbb{T} = (\mathbb{V}, \mathbb{E}, \rho)$ with nodes \mathbb{V} edges \mathbb{E} and root $\rho \in \mathbb{V}$. A configuration on \mathbb{T} is an element of $\mathcal{C}^{\mathbb{T}}$ with \mathcal{C} being a finite characters set, that is, an assignment of a state in \mathcal{C} to each vertex. The state of the root ρ , denoted by σ_ρ , is chosen according to some initial distribution π on \mathcal{C} . This symbol is then propagated in the tree according to the probability transition matrix $\mathbf{M} = (M_{ij})_{i,j \in \mathcal{C}}$, which functions as the noisy communication channel on each edge. That is, for each vertex v having u as its parent, the spin at v is defined according to the probabilities

$$\mathbf{P}(\sigma_v = j \mid \sigma_u = i) = M_{ij}, \quad i, j \in \mathcal{C}.$$

The objective model taken into account is the asymmetric binary channel with the configuration set $\mathcal{C} = \{1, 2\}$, whose transition matrix is of the form

$$\mathbf{M} = \frac{1}{2} \begin{pmatrix} 1 + \theta & 1 - \theta \\ 1 - \theta & 1 + \theta \end{pmatrix} + \frac{\Delta}{2} \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix},$$

where $|\theta| + |\Delta| \leq 1$ and Δ is used to describe the deviation of \mathbf{M} from the symmetric channel. It is easy to see that \mathbf{M} has two eigenvalues, 1 and θ . Then we pick a state at the root according to the stationary distribution $\pi = (\pi_1, \pi_2)$ of \mathbf{M} , which is given by

$$\pi_1 = \frac{1}{2} - \frac{\Delta}{2(1-\theta)} \quad \text{and} \quad \pi_2 = \frac{1}{2} + \frac{\Delta}{2(1-\theta)},$$

and without loss of generality, it is convenient to assume that $\pi_1 \geq \pi_2$.

Recall that the classical Ising model, a mathematical model of ferromagnetism in statistical mechanics, consists of discrete variables that represent magnetic dipole moments of atomic spins that can be in one of two states (-1 or $+1$). Consider a set of lattice sites Λ , each with a set of adjacent sites (e.g. a graph) forming a lattice, and for each $k \in \Lambda$, there is a discrete variable $\sigma_k \in \{-1, +1\}$ representing the site's spin. The energy of a configuration σ is given by the Hamiltonian function

$$H(\sigma) = - \sum_{\langle i, j \rangle} J_{ij} \sigma_i \sigma_j - \mu \sum_j h_j \sigma_j,$$

where the notation $\langle i, j \rangle$ indicates that sites i and j are the nearest neighbors, J_{ij} denotes the interaction between two adjacent sites $i, j \in \Lambda$ and h_j models the external magnetic field interaction of site $j \in \Lambda$. In this literature, the current model corresponds to the general Ising model with external field on the tree.

The problem of reconstruction is to analyze whether there exists non-vanishing information on the letter transmitted from the root, given all the symbols received at the vertices of the n th generation, as n goes to infinity. We define the distance between probability measures in line with [Evans et al. \[2000\]](#). Let ν_+ and ν_- be two probability measures on the same space. Set $f = d\nu_+/d\nu$ and $g = d\nu_-/d\nu$ where $\nu := (\nu_+ + \nu_-)/2$. Inferring the root spin σ_ρ from the spin configurations on the finite vertex set is a basic problem of Bayesian hypothesis testing. The total variation distance, defined as $d_{TV}(\nu_+, \nu_-) := \frac{1}{2} \int |f - g| d\nu$, can be interpreted as the difference between the probabilities of correct and erroneous inferences. Denote $\sigma(n)$ as the spins at distance n from the root and $\sigma^i(n)$ as $\sigma(n)$ conditioned on $\sigma_\rho = i$. Then the reconstruction problem can be mathematical formulated as the following:

Definition 1 *The reconstruction problem for the infinite tree \mathbb{T} is **solvable** if for some $i, j \in \mathcal{C}$,*

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

*When the lim sup is 0, we will say that the model has **non-reconstruction** on \mathbb{T} .*

1.2 Background and Applications

The reconstruction problem arises naturally in statistical physics, where the reconstruction threshold is identified as the threshold for extremality of the infinite-volume Gibbs measure with free boundary conditions (see [Georgii \[2011\]](#)). In [Berger et al. \[2005\]](#), [Martinelli et al. \[2007\]](#), [Tetali et al. \[2012\]](#), the reconstruction bound is found to have a crucial determination effect on the efficiency of the Glauber dynamics on trees and random graphs. The reconstruction threshold is also believed to play an important role in a variety of other contexts, including phylogenetic reconstruction in evolutionary biology ([Mossel \[2004a\]](#), [Daskalakis et al. \[2006\]](#), [Roch \[2006\]](#)), communication theory in the study of noisy computation ([Evans et al. \[2000\]](#)), clustering problem in the setting of the stochastic block model ([Mossel et al. \[2012, 2013\]](#), [Neeman and Netrapalli \[2014\]](#)), and network tomography ([Bhamidi et al. \[2010\]](#)). For detailed explanation on the reconstruction problem in mixing, phylogeny and replicas, we refer to Section 1.3 in [Bernussou and Abatut \[1977\]](#). For other applications of reconstruction, we refer to Section 1.4 in [Sly \[2011\]](#) and Section 1.3 in [Liu et al. \[2018\]](#), as well as the references therein.

In this paper, we focus on analyzing the tightness of the reconstruction bound on the tree for asymmetric binary channels, which corresponds to the asymmetric Ising model on the tree in statistical physics term. Well known, the reconstruction problem is closely related to λ , the second largest eigenvalue in absolute value of the transition probability matrix, which is θ in the current model under investigation. [Kesten and Stigum \[1966, 1967\]](#) showed that the reconstruction problem is solvable if $d\lambda^2 > 1$, which is known as the Kesten-Stigum bound. However in the case of larger noise, i.e. $d\lambda^2 < 1$, one may wonder whether reconstruction problem is still solvable, that is collecting and analyzing the whole set of symbols received at the n th generation to retrieve information transmitted from the root.

First consider the symmetric channel. It was shown in [Bleher et al. \[1995\]](#) that the reconstruction problem is solvable if and only if $d\lambda^2 > 1$ in the binary model. For all other models, it was also known and easy to prove that $d\lambda^2 > 1$ implies solvability, while proving non-reconstructibility turned out to be harder. Although coupling arguments easily yield non-reconstruction, these arguments are typically not rigorous. A natural approach to establish non-reconstructibility is to analyze recursions in terms of random variables, each of whose values is the expectation of the chain at a vertex, given the state at the leaves of the subtree below it, and the corresponding probabilities. Although the reconstruction problem on the tree has been studied in numerous contexts, the existing literatures with rigorous reconstruction thresholds established are very limited. [Sly \[2011\]](#) proved the first exact reconstruction threshold in a nonbinary model by establishing the Kesten–Stigum bound for the 3-state Potts model on regular trees of large degree, and further established that the Kesten–Stigum bound is not tight for the q -state Potts model when $q \geq 5$, which confirms much of the picture conjectured earlier by [Mézard and Montanari \[2006\]](#). [Liu et al. \[2018\]](#) considered a $2q$ -state symmetric model, with two categories of q states in each category, and 3 transition probabilities (the probability to remain in the same state, the

probability to change states but remain in the same category, and the probability to change categories) and showed that the Kesten-Stigum reconstruction bound is not tight when $q \geq 4$.

Next let us turn to the existing results regarding the asymmetric channel. Mossel [2001, 2004b] showed that the Kesten-Stigum bound is not the bound for reconstruction in the binary asymmetric model with sufficiently large asymmetry or in the symmetric Potts model with sufficiently many characters, which shed the light on exploring the tightness of the Kesten-Stigum bound. Furthermore, Proposition 12 in Mossel [2001] implies that for any asymmetric channel, given d and π , the reconstructibility is monotone in $|\theta|$, say, there exist the thresholds $\theta^- < 0 < \theta^+$ such that, there is non-reconstruction when $\theta \in (\theta^-, \theta^+)$, while it is reconstructible when $\theta < \theta^-$ or $\theta > \theta^+$. Therefore, the Kesten-Stigum bound mentioned above implies immediately

$$\theta^+ \leq d^{-1/2} \quad \text{and} \quad \theta^- \geq -d^{-1/2},$$

but exact thresholds for non-solvability had not been known. The breakthrough result in Borgs et al. [2006] established the exact threshold for the reconstruction problem with the binary asymmetric channel on the d -ary tree, provided that the asymmetry is sufficiently small, which is the first exact reconstruction threshold obtained in roughly a decade. However, this beautiful result only rigorously proved the existence of Δ to satisfy the reconstruction criterion, does not answer the question that how small the asymmetry needs to be, therefore rigorously estimating the range of asymmetry to keep Kesten-Stigum bound tight is a natural question, which will be answered in the next section.

1.3 Main Results and Proof Sketch

In this section, we will present a critical condition of the stationary initial distribution π to keep the tightness of the Kesten-Stigum bound, by means of refined recursive equations of vector-valued distributions and concentration analyses. Furthermore, when the Kesten-Stigum bound is not tight, we provide a new reconstruction threshold $C_\pi \in (0, 1)$ for sufficiently large d . Since $d\theta^2 > 1$ always implies reconstruction, it suffices to consider $d\theta^2 \leq 1$ in the following context.

Theorem 1.1 *For every d and π such that $\pi_1\pi_2 < \frac{1}{6}$, the Kesten-Stigum bound is not tight. In other words, the reconstruction problem is solvable for some θ , even if $d\theta^2 < 1$.*

The proof to Theorem 1.1 above is given in Section 5. The proofs to Theorem 1.2 and Theorem 1.3 below are given in Section 6.3 and Section 6.4 respectively.

Theorem 1.2 *For every π such that $\pi_1\pi_2 < \frac{1}{6}$, there exists an asymptotic result of the reconstruction threshold, that is, when d goes to infinity,*

$$\lim_{d \rightarrow \infty} d(\theta^\pm)^2 = C_\pi,$$

where C_π is a constant taking values in $(0, 1)$ and depends only on π .

Theorem 1.3 *For every π such that $\pi_1\pi_2 > \frac{1}{6}$, there exists a $D = D(\pi) > 0$, such that for $d > D$ the Kesten-Stigum bound is sharp, that is*

$$\theta^+ = d^{-1/2} \quad \text{and} \quad \theta^- = -d^{-1/2}.$$

Furthermore, there is non-reconstruction at the Kesten-Stigum bound, when $\theta = \theta^+$ or θ^- .

The idea to establish Theorem 1.1, Theorem 1.2 and Theorem 1.3 is the following. One standard way to classify reconstruction and non-reconstruction is to analyze the quantity x_n : the probability of giving a correct guess of the root given the spins $\sigma(n)$ at distance n from the root, minus the probability of guessing the root according to stationary initial distribution. Non-reconstruction means that the mutual information between the root and the spins at distance n goes to zero as n tends to infinity. In Lemma 3, we rigorously show that x_n is always positive and the non-reconstruction is equivalent to

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

To analyze whether the reconstruction holds, inspired by Chayes et al. [1986], Borgs et al. [2006] and Sly [2011], we establish the distributional recursion and moment recursion, and then the recursive relation between the n th and the $(n+1)$ th generation's structure of the tree leads to a corresponding nonlinear dynamical system. In the mean time, we show that the interactions between spins become very weak, if they are sufficiently far away from each other. Therefore, under this weak interacting situation, i.e. x_n being sufficiently small, the concentration analysis is successfully developed and an approximation to the dynamical system is nicely established:

$$x_{n+1} \approx d\theta^2 x_n + \frac{1 - 6\pi_1\pi_2}{\pi_1\pi_2^2} \frac{d(d-1)}{2} \theta^4 x_n^2.$$

The sign of coefficient of the quadratic term which is determined by $1 - 6\pi_1\pi_2$, plays a crucial role in the asymptotic behavior of x_n . When $1 - 6\pi_1\pi_2 > 0$, equivalently $\Delta^2 > (1 - \theta)^2/3$, if $d\theta^2$ is sufficiently close to 1, then x_n does not converge to 0 and then there is reconstruction beyond the Kesten-Stigum bound. Then our focus is to find this new reconstruction threshold, which is executed in the following three steps: Step one, we rigorously show that, when degree d is large, the interactions between spins become very weak; Step two, using the Central Limit Theorem, we approximate the corresponding collection of small independent samples, to show that the reconstruction function can be asymptotically given by a new Gaussian approximation function $g(s)$, that is, $x_{n+1} \approx g(d\theta^2 x_n)$; Step three, we explore the first several major terms of the Maclaurin series of $g(s)$, and rigorously establish the reconstruction threshold by discussing the fixed point of $g(s)$. On the other hand, when $1 - 6\pi_1\pi_2 < 0$, the analysis of large degree asymptotics yields $g(s) < s$, which implies $\lim_{n \rightarrow \infty} x_n = 0$, that is, there is non-reconstruction.

2 Preliminaries

2.1 Notations

Let u_1, \dots, u_d be the children of ρ and \mathbb{T}_v be the subtree of descendants of $v \in \mathbb{T}$. Denote the n th level of the tree as $L(n) = \{v \in \mathbb{V} : d(\rho, v) = n\}$, where $d(\cdot, \cdot)$ is the graph distance on \mathbb{T} . With the notations above, let $\sigma(n)$ and $\sigma_j(n)$ be the spins on L_n and $L(n) \cap \mathbb{T}_{u_j}$ respectively. For a configuration A on $L(n)$, define the posterior function f_n by

$$f_n(i, A) = \mathbf{P}(\sigma_\rho = i \mid \sigma(n) = A).$$

By the recursive nature of the tree, for a configuration A on $L(n+1) \cap \mathbb{T}_{u_j}$, an equivalent form is given by

$$f_n(i, A) = \mathbf{P}(\sigma_{u_j} = i \mid \sigma_j(n+1) = A).$$

Next, with $i = 1, 2$, define

$$X_i = X_i(n) = f_n(i, \sigma(n)), \quad X^+ = X^+(n) = f_n(1, \sigma^1(n)), \quad X^- = X^-(n) = f_n(2, \sigma^2(n)),$$

and for $1 \leq j \leq d$,

$$Y_j = Y_j(n) = f_n(1, \sigma_j^1(n+1)),$$

where it is clear that the random variables $\{Y_j\}_{1 \leq j \leq d}$ are independent and identical in distribution. It is apparent that

$$X_1(n) + X_2(n) = 1 \tag{2.1}$$

and

$$\mathbf{E}(X_1) = \pi_1, \quad \mathbf{E}(X_2) = \pi_2. \tag{2.2}$$

We introduce the objective quantities in this paper:

$$x_n = \mathbf{E}(X^+(n) - \pi_1) \quad \text{and} \quad z_n = \mathbf{E}(X^+(n) - \pi_1)^2.$$

2.2 Preparations

Before proceeding to the analysis, it is convenient to firstly derive some very useful identities concerning x_n .

Lemma 1 *For any $n \in \mathbb{N} \cup \{0\}$, we have*

$$x_n = \frac{1}{\pi_1} \mathbf{E}(X_1 - \pi_1)^2 = \mathbf{E}(X^+(n) - \pi_1)^2 + \frac{\pi_2}{\pi_1} \mathbf{E}(X^-(n) - \pi_2)^2 \geq z_n \geq 0.$$

Proof. By Bayes' rule, we have

$$\begin{aligned} \mathbf{E}X^+ &= \mathbf{E}f_n(1, \sigma^1(n)) \\ &= \sum_A f_n(1, A) \mathbf{P}(\sigma(n) = A \mid \sigma_\rho = 1) \\ &= \frac{1}{\pi_1} \sum_A \mathbf{P}(\sigma_\rho = 1 \mid \sigma(n) = A) \mathbf{P}(\sigma(n) = A) f_n(1, A) \\ &= \frac{1}{\pi_1} \sum_A f_n(1, A)^2 \mathbf{P}(\sigma(n) = A) \\ &= \frac{1}{\pi_1} \mathbf{E}(X_1^2) \end{aligned}$$

and similarly,

$$\mathbf{E}X^- = \mathbf{E}f_n(2, \sigma^2(n)) = \frac{1}{\pi_2} \mathbf{E}(X_2^2).$$

Then it follows from equation (2.2) that

$$x_n = \frac{1}{\pi_1} (\mathbf{E}(X_1^2) - \pi_1^2) = \frac{1}{\pi_1} \mathbf{E}(X_1 - \pi_1)^2. \tag{2.3}$$

Next by equation (2.1), one has

$$x_n = \frac{1}{\pi_1} \mathbf{E}(X_2 - \pi_2)^2 = \frac{\pi_2}{\pi_1} (\mathbf{E}X^-(n) - \pi_2).$$

Therefore, the quantitative relation between x_n and z_n holds:

$$\begin{aligned} x_n &= \frac{1}{\pi_1} \mathbf{E}(X_1 - \pi_1)^2 \\ &= \frac{1}{\pi_1} [\mathbf{P}(\sigma_\rho = 1) \mathbf{E}((X_1 - \pi_1)^2 | \sigma_\rho = 1) + \mathbf{P}(\sigma_\rho = 2) \mathbf{E}((X_2 - \pi_2)^2 | \sigma_\rho = 2)] \\ &= \frac{1}{\pi_1} [\pi_1 \mathbf{E}(X^+(n) - \pi_1)^2 + \pi_2 \mathbf{E}(X^-(n) - \pi_2)^2] \\ &= \mathbf{E}(X^+(n) - \pi_1)^2 + \frac{\pi_2}{\pi_1} \mathbf{E}(X^-(n) - \pi_2)^2 \\ &\geq z_n \end{aligned}$$

□

With the preceding results, we calculate the means and variances of Y_j .

Lemma 2 For each $1 \leq j \leq d$, we have

$$\mathbf{E}(Y_j - \pi_1) = \theta x_n \quad \text{and} \quad \mathbf{E}(Y_j - \pi_1)^2 = \pi_1 x_n + \theta(z_n - \pi_1 x_n).$$

Proof. If $\sigma_{u_j}^1 = 1$, Y_j is distributed according to $X^+(n)$, while if $\sigma_{u_j}^1 = 2$, Y_j is distributed according to $1 - X^-(n)$. Therefore we have

$$\mathbf{E}(Y_j - \pi_1) = \mathbf{P}(\sigma_{u_j}^1 = 1) \mathbf{E}(X^+(n) - \pi_1) + \mathbf{P}(\sigma_{u_j}^1 = 2) \mathbf{E}(1 - X^-(n) - \pi_1) = M_{11} x_n - M_{12} \frac{\pi_1}{\pi_2} x_n = \theta x_n$$

and similarly we have

$$\begin{aligned} \mathbf{E}(Y_j - \pi_1)^2 &= \mathbf{P}(\sigma_{u_j}^1 = 1) \mathbf{E}(X^+(n) - \pi_1)^2 + \mathbf{P}(\sigma_{u_j}^1 = 2) \mathbf{E}(1 - X^-(n) - \pi_1)^2 \\ &= M_{11} \mathbf{E}(X^+(n) - \pi_1)^2 + M_{12} \mathbf{E}(X^-(n) - \pi_2)^2 \\ &= M_{11} z_n + M_{12} \frac{\pi_1}{\pi_2} (x_n - z_n) \\ &= \pi_1 x_n + \theta(z_n - \pi_1 x_n), \end{aligned}$$

as desired. □

2.3 An Equivalent Condition for Non-reconstruction

If the reconstruction problem is solvable, then $\sigma(n)$ contains significant information on the root variable, which may be formulated in several equivalent ways (see Mossel [2001], Proposition 14). As a result, in order to analyze the reconstruction, it suffices to investigate the asymptotic behavior of x_n as n goes to infinity.

Lemma 3 *The non-reconstruction is equivalent to*

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

Proof. The maximum-likelihood algorithm, which is the optimal reconstruction algorithm of σ_ρ given $\sigma(n)$, is successful with probability

$$\Delta_n = \mathbf{E} \max\{X_1(n), X_2(n)\}.$$

Therefore, the inequality of $x_n + \pi_1 \leq \Delta_n$ holds, which is an analogous result to that of [Mézard and Montanari \[2006\]](#), by noting that the algorithm that chooses σ_ρ randomly according to probabilities X_i is correct with probability $x_n + \pi_1$. On the other hand, recalling the assumption that $\pi_1 \geq \pi_2$, by the Cauchy-Schwartz inequality together with the identities equation (2.1) and equation (2.3), one can conclude

$$\begin{aligned} \Delta_n &= \pi_1 + \mathbf{E} \max\{X_1(n) - \pi_1, X_2(n) - \pi_1\} \\ &\leq \pi_1 + \mathbf{E} \max\{X_1(n) - \pi_1, X_2(n) - \pi_2\} \\ &= \pi_1 + \mathbf{E}|X_1(n) - \pi_1| \\ &\leq \pi_1 + \left(\mathbf{E}(X_1(n) - \pi_1)^2\right)^{1/2} \\ &\leq \pi_1 + \pi_1^{1/2} x_n^{1/2}. \end{aligned} \tag{2.4}$$

Hence, one has

$$x_n \leq \Delta_n - \pi_1 \leq \pi_1^{1/2} x_n^{1/2},$$

implying that x_n converging to 0 is equivalent to Δ_n converging to π_1 , which is equivalent to non-reconstruction (see [Mossel \[2001\]](#)). \square

3 Moment Recursion

3.1 Distributional Recursion

In the last section, it is known that the asymptotic behavior of x_n as n goes to infinity plays a crucial role, however it is still too difficult and not necessary to get the explicit expression for x_n . In fact, we only need to investigate the recursive formula of x_n , from which it is possible to illustrate the trend of x_n as n goes to infinity. Thus the key idea is to analyze the recursive relation between $X^+(n)$ and $X^+(n+1)$ by the structure of the tree. Suppose that A is a configuration on $L(n+1)$ and let A_j be its restriction on $\mathbb{T}_{u_j} \cap L(n+1)$. Then from the Markov random field property, we have

$$\begin{aligned} f_{n+1}(1, A) &= \frac{N_1}{N_1 + N_2} \\ &= \frac{\pi_1 \prod_{j=1}^d \left[\frac{M_{11}}{\pi_1} f_n(1, A_j) + \frac{M_{12}}{\pi_2} f_n(2, A_j) \right]}{\pi_1 \prod_{j=1}^d \left[\frac{M_{11}}{\pi_1} f_n(1, A_j) + \frac{M_{12}}{\pi_2} f_n(2, A_j) \right] + \pi_2 \prod_{j=1}^d \left[\frac{M_{21}}{\pi_1} f_n(1, A_j) + \frac{M_{22}}{\pi_2} f_n(2, A_j) \right]} \tag{3.1} \\ &= \frac{\pi_1 \prod_{j=1}^d \left[1 + \frac{\theta}{\pi_1} (f_n(1, A_j) - \pi_1) \right]}{\pi_1 \prod_{j=1}^d \left[1 + \frac{\theta}{\pi_1} (f_n(1, A_j) - \pi_1) \right] + \pi_2 \prod_{j=1}^d \left[1 - \frac{\theta}{\pi_2} (f_n(1, A_j) - \pi_1) \right]}, \end{aligned}$$

where

$$\begin{aligned} N_1 &= \pi_1 \prod_{j=1}^d [M_{11} \mathbf{P}(\sigma_j(n+1) = A_j \mid \sigma_{u_j} = 1) + M_{12} \mathbf{P}(\sigma_j(n+1) = A_j \mid \sigma_{u_j} = 2)] \\ &= \pi_1 \prod_{j=1}^d \left[\frac{M_{11}}{\pi_1} f_n(1, A_j) + \frac{M_{12}}{\pi_2} f_n(2, A_j) \right] \mathbf{P}(\sigma_j(n+1) = A_j) \end{aligned}$$

and

$$\begin{aligned} N_2 &= \pi_2 \prod_{j=1}^d \left[M_{21} \mathbf{P}(\sigma_j(n+1) = A_j \mid \sigma_{u_j} = 1) + M_{22} \mathbf{P}(\sigma_j(n+1) = A_j \mid \sigma_{u_j} = 2) \right] \\ &= \pi_2 \prod_{j=1}^d \left[\frac{M_{21}}{\pi_1} f_n(1, A_j) + \frac{M_{22}}{\pi_2} f_n(2, A_j) \right] \mathbf{P}(\sigma_j(n+1) = A_j). \end{aligned}$$

Next conditioning the root being 1 and setting $A = \sigma^1(n+1)$, we have

$$X^+(n+1) = \frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2},$$

where

$$Z_1 = \prod_{j=1}^d \left[1 + \frac{\theta}{\pi_1} (f_n(1, A_j) - \pi_1) \right] = \prod_{j=1}^d \left[1 + \frac{\theta}{\pi_1} (Y_j(n) - \pi_1) \right]$$

and

$$Z_2 = \prod_{j=1}^d \left[1 - \frac{\theta}{\pi_2} (f_n(1, A_j) - \pi_1) \right] = \prod_{j=1}^d \left[1 - \frac{\theta}{\pi_2} (Y_j(n) - \pi_1) \right].$$

3.2 Main Expansion of x_{n+1}

With the help of those preliminary results, we are about to figure out the recursive relation regarding x_{n+1} , specifically, its main expansion, which would play a crucial rule in further discussions. Firstly we take care of the approximating means and variances of Z_i and the Taylor series approximations.

Lemma 4 *For each positive integer k , there exists a $C = C(\pi, k)$ which only depends on π and k , such that for each $0 \leq \ell, m \leq k$,*

$$\mathbf{E} Z_1^\ell Z_2^m \leq C, \quad \left| \mathbf{E} Z_1^\ell Z_2^m - 1 - du \right| \leq C x_n^2, \quad \left| \mathbf{E} Z_1^\ell Z_2^m - 1 - du - \frac{d(d-1)}{2} u^2 \right| \leq C x_n^3,$$

where

$$u = \mathbf{E} \left[1 + \frac{\theta}{\pi_1} (Y_1(n) - \pi_1) \right]^\ell \left[1 - \frac{\theta}{\pi_2} (Y_1(n) - \pi_1) \right]^m - 1.$$

Proof. Since $\{Y_j\}_{1 \leq j \leq d}$ are independent and identical in distribution, we have

$$\begin{aligned} \mathbf{E}Z_1^\ell Z_2^m &= \prod_{j=1}^d \mathbf{E} \left[1 + \frac{\theta}{\pi_1}(Y_j(n) - \pi_1) \right]^\ell \left[1 - \frac{\theta}{\pi_2}(Y_j(n) - \pi_1) \right]^m \\ &= \left(\mathbf{E} \left[1 + \frac{\theta}{\pi_1}(Y_1(n) - \pi_1) \right]^\ell \left[1 - \frac{\theta}{\pi_2}(Y_1(n) - \pi_1) \right]^m \right)^d = (1+u)^d. \end{aligned}$$

It follows from $0 \leq Y_1 \leq 1$ that $|Y_1(n) - \pi_1| \leq 1$, and then when $i \geq 2$, we have $|Y_1(n) - \pi_1|^i \leq (Y_1(n) - \pi_1)^2$. Therefore Lemma 2 implies that

$$\begin{aligned} |u| &= \left| \mathbf{E} \left[1 + \frac{\theta}{\pi_1}(Y_1(n) - \pi_1) \right]^\ell \left[1 - \frac{\theta}{\pi_2}(Y_1(n) - \pi_1) \right]^m - 1 \right| \\ &= \left| \sum_{i=1}^{\ell+m} c_i \mathbf{E}[\theta(Y_1 - \pi_1)]^i \right| \\ &\leq |c_1| \theta^2 x_n + \sum_{i=2}^{\ell+m} |c_i| \theta^2 [\pi_1 x_n + \theta(z_n - \pi_1 x_n)] \\ &\leq c \theta^2 x_n, \end{aligned}$$

where $\{c_i\}_{i=1}^{\ell+m}$ and c depend on π and k only. Consequently, we have $d|u| \leq c x_n$ by means of $d\theta^2 \leq 1$. Using the binomial expansion and the Remainder Theorem, we have

$$\left| (1+u)^d - \sum_{i=0}^h \binom{d}{i} u^i \right| \leq \sum_{i=h+1}^{\infty} \frac{d^i}{i!} |u|^i \leq e^c c^{h+1} x_n^{h+1}.$$

Taking $h = 0, 1, 2$ respectively and $C = \max_{h \in \{0,1,2\}} \{e^c c^{h+1}\}$ complete the proof. \square

Next we aim to figure out the recursive relation of x_{n+1} by virtue of the following identity

$$\frac{a}{s+r} = \frac{a}{s} - \frac{ar}{s^2} + \frac{r^2}{s^2} \frac{a}{s+r}. \quad (3.2)$$

Specifically, plugging $a = \pi_1 Z_1$, $r = \pi_1 Z_1 + \pi_2 Z_2 - 1$ and $s = 1$ in equation (3.2) yields

$$\begin{aligned} x_{n+1} &= \mathbf{E}X^+(n+1) - \pi_1 \\ &= \mathbf{E}(\pi_1 Z_1) - \mathbf{E}[\pi_1 Z_1(\pi_1 Z_1 + \pi_2 Z_2 - 1)] + \mathbf{E} \left[(\pi_1 Z_1 + \pi_2 Z_2 - 1)^2 \frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} \right] - \pi_1. \end{aligned} \quad (3.3)$$

In the following, we analyze terms in equation (3.3), using the notation O_π to emphasize that the constant associated with the O -term depends on π only

$$\begin{aligned} \mathbf{E}Z_1 &= \mathbf{E} \prod_{j=1}^d \left[1 + \frac{\theta}{\pi_1}(Y_j - \pi_1) \right] \\ &= 1 + \frac{d\theta}{\pi_1} \mathbf{E}(Y_1 - \pi_1) + \frac{d(d-1)}{2} \left[\frac{\theta}{\pi_1} \mathbf{E}(Y_1 - \pi_1) \right]^2 + O_\pi(x_n^3) \\ &= 1 + \frac{d\theta^2}{\pi_1} x_n + \frac{d(d-1)}{2} \frac{\theta^4}{\pi_1^2} x_n^2 + O_\pi(x_n^3), \end{aligned}$$

$$\begin{aligned}
\mathbf{E}Z_1^2 &= \mathbf{E} \prod_{j=1}^d \left[1 + \frac{\theta}{\pi_1} (Y_j - \pi_1) \right]^2 \\
&= 1 + d \left\{ \mathbf{E} \left[1 + \frac{\theta}{\pi_1} (Y_1 - \pi_1) \right]^2 - 1 \right\} + \frac{d(d-1)}{2} \left\{ \mathbf{E} \left[1 + \frac{\theta}{\pi_1} (Y_1 - \pi_1) \right]^2 - 1 \right\}^2 + O_\pi(x_n^3) \\
&= 1 + d \left[\frac{\theta^2}{\pi_1} (3 - \theta) x_n + \frac{\theta^3}{\pi_1^2} z_n \right] + \frac{d(d-1)}{2} \left[\frac{\theta^2}{\pi_1} (3 - \theta) x_n + \frac{\theta^3}{\pi_1^2} z_n \right]^2 + O_\pi(x_n^3),
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{E}Z_1 Z_2 &= \mathbf{E} \prod_{j=1}^d \left[1 + \frac{\theta}{\pi_1} (Y_j - \pi_1) \right] \left[1 - \frac{\theta}{\pi_2} (Y_j - \pi_1) \right] \\
&= 1 + d \left\{ \mathbf{E} \left[1 + \frac{\theta}{\pi_1} (Y_1 - \pi_1) \right] \left[1 - \frac{\theta}{\pi_2} (Y_1 - \pi_1) \right] - 1 \right\} \\
&\quad + \frac{d(d-1)}{2} \left\{ \mathbf{E} \left[1 + \frac{\theta}{\pi_1} (Y_1 - \pi_1) \right] \left[1 - \frac{\theta}{\pi_2} (Y_1 - \pi_1) \right] - 1 \right\}^2 + O_\pi(x_n^3) \\
&= 1 + d \left[\theta^2 \left(\frac{1}{\pi_1} + \frac{\theta - 2}{\pi_2} \right) x_n - \frac{\theta^3}{\pi_1 \pi_2} z_n \right] + \frac{d(d-1)}{2} \left[\theta^2 \left(\frac{1}{\pi_1} + \frac{\theta - 2}{\pi_2} \right) x_n - \frac{\theta^3}{\pi_1 \pi_2} z_n \right]^2 + O_\pi(x_n^3).
\end{aligned}$$

Then the preceding results yield

$$\begin{aligned}
&\mathbf{E} \pi_1 Z_1 (\pi_1 Z_1 + \pi_2 Z_2 - 1) \\
&= \pi_1^2 \mathbf{E} Z_1^2 + \pi_1 \pi_2 \mathbf{E} Z_1 Z_2 - \pi_1 \mathbf{E} Z_1 \\
&= \pi_1^2 \frac{d(d-1)}{2} \left[\frac{\theta^2}{\pi_1} (3 - \theta) x_n + \frac{\theta^3}{\pi_1^2} z_n \right]^2 + \pi_1 \pi_2 \frac{d(d-1)}{2} \left[\theta^2 \left(\frac{1}{\pi_1} + \frac{\theta - 2}{\pi_2} \right) x_n - \frac{\theta^3}{\pi_1 \pi_2} z_n \right]^2 \\
&\quad - \frac{d(d-1)}{2} \frac{\theta^4}{\pi_1} x_n^2 + O_\pi(x_n^3).
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\mathbf{E}Z_2 &= \mathbf{E} \prod_{j=1}^d \left[1 - \frac{\theta}{\pi_2} (Y_j - \pi_1) \right] \\
&= 1 - \frac{d\theta}{\pi_2} \mathbf{E}(Y_1 - \pi_1) + \frac{d(d-1)}{2} \left[\frac{\theta}{\pi_2} \mathbf{E}(Y_1 - \pi_1) \right]^2 + O_\pi(x_n^3) \\
&= 1 - \frac{d\theta^2}{\pi_2} x_n + \frac{d(d-1)}{2} \frac{\theta^4}{\pi_2^2} x_n^2 + O_\pi(x_n^3),
\end{aligned}$$

$$\begin{aligned}
\mathbf{E}Z_2^2 &= \mathbf{E} \prod_{j=1}^d \left[1 - \frac{\theta}{\pi_2} (Y_j - \pi_1) \right]^2 \\
&= 1 + d \left\{ \mathbf{E} \left[1 - \frac{\theta}{\pi_2} (Y_1 - \pi_1) \right]^2 - 1 \right\} + \frac{d(d-1)}{2} \left\{ \mathbf{E} \left[1 - \frac{\theta}{\pi_2} (Y_1 - \pi_1) \right]^2 - 1 \right\}^2 + O_\pi(x_n^3) \\
&= 1 + d \left\{ \frac{\theta^2}{\pi_2} \left[\frac{\pi_1}{\pi_2} (1 - \theta) - 2 \right] x_n + \frac{\theta^3}{\pi_2^2} z_n \right\} + \frac{d(d-1)}{2} \left\{ \frac{\theta^2}{\pi_2} \left[\frac{\pi_1}{\pi_2} (1 - \theta) - 2 \right] x_n + \frac{\theta^3}{\pi_2^2} z_n \right\}^2 + O_\pi(x_n^3),
\end{aligned}$$

and then

$$\begin{aligned}
& \mathbf{E}(\pi_1 Z_1 + \pi_2 Z_2 - 1)^2 \\
&= \pi_1^2 \mathbf{E}(Z_1^2) + \pi_2^2 \mathbf{E}(Z_2^2) + 2\pi_1 \pi_2 \mathbf{E}(Z_1 Z_2) - 2\pi_1 \mathbf{E}(Z_1) - 2\pi_2 \mathbf{E}(Z_2) + 1 \\
&= \pi_1^2 \frac{d(d-1)}{2} \left[\frac{\theta^2}{\pi_1} (3-\theta)x_n + \frac{\theta^3}{\pi_1^2} z_n \right]^2 + \pi_2^2 \frac{d(d-1)}{2} \left\{ \frac{\theta^2}{\pi_2} \left[\frac{\pi_1}{\pi_2} (1-\theta) - 2 \right] x_n + \frac{\theta^3}{\pi_2^2} z_n \right\}^2 \\
&\quad + \pi_1 \pi_2 d(d-1) \left[\theta^2 \left(\frac{1}{\pi_1} + \frac{\theta-2}{\pi_2} \right) x_n - \frac{\theta^3}{\pi_1 \pi_2} z_n \right]^2 - d(d-1) \frac{\theta^4}{\pi_1} x_n^2 - d(d-1) \frac{\theta^4}{\pi_2} x_n^2 + O_\pi(x_n^3).
\end{aligned} \tag{3.4}$$

As a consequence, we have

$$\begin{aligned}
x_{n+1} &= \mathbf{E}(\pi_1 Z_1) - \mathbf{E}\pi_1 Z_1 (\pi_1 Z_1 + \pi_2 Z_2 - 1) + \pi_1 \mathbf{E}(\pi_1 Z_1 + \pi_2 Z_2 - 1)^2 - \pi_1 + S \\
&= d\theta^2 x_n + \frac{d(d-1)}{2} \theta^4 x_n^2 \left\{ \left(\frac{2}{\pi_1} - \frac{2}{\pi_2} \right) - \pi_2 \left[3 + \frac{\theta}{\pi_1} \left(\frac{z_n}{x_n} - \pi_1 \right) \right]^2 + \pi_1 \pi_2 (\pi_1 - \pi_2) \right. \\
&\quad \times \left. \left[\left(\frac{1}{\pi_1} - \frac{2}{\pi_2} \right) - \frac{\theta}{\pi_1 \pi_2} \left(\frac{z_n}{x_n} - \pi_1 \right) \right]^2 + \pi_1 \left[\frac{\pi_1}{\pi_2} - 2 + \frac{\theta}{\pi_2} \left(\frac{z_n}{x_n} - \pi_1 \right) \right]^2 \right\} + S + O_\pi(x_n^3) \\
&= d\theta^2 x_n + \frac{1-6\pi_1\pi_2}{\pi_1\pi_2^2} \frac{d(d-1)}{2} \theta^4 x_n^2 + R + S,
\end{aligned} \tag{3.5}$$

where

$$|R| \leq C_R x_n^2 \left(\frac{d(d-1)}{2} |\theta|^5 \left| \frac{z_n}{x_n} - \pi_1 \right| + x_n \right) \tag{3.6}$$

with C_R a constant depending only on π , and

$$S = \mathbf{E}(\pi_1 Z_1 + \pi_2 Z_2 - 1)^2 \left(\frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1 \right) \tag{3.7}$$

which will be handled in the following concentration investigation.

4 Concentration Analysis

Noting that $Z_1, Z_2 \geq 0$, we have $0 \leq \frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} \leq 1$. It is concluded from equations (3.3) and (3.4) that

$$|x_{n+1} - d\theta^2 x_n| \leq C x_n^2 \leq \varepsilon x_n, \tag{4.1}$$

where $C = C(\pi)$ depends only on π . In equation (4.1), the first inequality follows from Lemma 1 which states that $0 \leq z_n \leq x_n$, and the last inequality holds if $x_n < \delta$ for $\delta = \delta(\pi, \varepsilon)$ small enough. The following lemma ensures that x_n does not drop too fast.

Lemma 5 *For any $\rho > 0$, there exists a constant $\gamma = \gamma(\pi, \rho) > 0$, such that for all n when $|\theta| > \rho$,*

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

Proof. For a configuration A on $\mathbb{T}_{u_1} \cap L(n+1)$, we have

$$\begin{aligned}
f_{n+1}^*(1, A) &:= \mathbf{P}(\sigma_\rho = 1 \mid \sigma_1(n+1) = A) \\
&= \pi_1 \frac{\mathbf{P}(\sigma_1(n+1) = A \mid \sigma_\rho = 1)}{\mathbf{P}(\sigma_1(n+1) = A)} \\
&= \pi_1 \frac{M_{11} \mathbf{P}(\sigma_1(n+1) = A \mid \sigma_{u_1} = 1) + M_{12} \mathbf{P}(\sigma_1(n+1) = A \mid \sigma_{u_1} = 2)}{\mathbf{P}(\sigma_1(n+1) = A)} \\
&= \pi_1 \left[\frac{M_{11}}{\pi_1} f_n(1, A) + \frac{M_{12}}{\pi_2} (1 - f_n(1, A)) \right] \\
&= \pi_1 \left[1 + \frac{\theta}{\pi_1} (f_n(1, A) - \pi_1) \right],
\end{aligned}$$

and then

$$\mathbf{E} f_{n+1}^*(1, \sigma_1^1(n+1)) = \pi_1 + \theta^2 x_n.$$

Therefore, it follows from equation (2.4) that

$$\pi_1 + \theta^2 x_n \leq \Delta_{n+1} \leq \pi_1 + \pi_1^{1/2} x_{n+1}^{1/2},$$

namely,

$$x_{n+1} \geq \frac{1}{\pi_1} \theta^4 x_n^2 \geq \rho^4 x_n^2. \quad (4.2)$$

Next choosing $\varepsilon = \rho^2$, equation (4.1) indicates that there exists a $\delta = \delta(\pi, \varepsilon) > 0$, such that if $x_n < \delta$ then

$$x_{n+1} \geq (d\theta^2 - \varepsilon)x_n \geq (d-1)\rho^2 x_n \geq \rho^2 x_n.$$

On the other hand, if $x_n \geq \delta$, equation (4.2) becomes $x_{n+1} \geq \rho^4 \delta x_n$. Finally taking $\gamma = \min\{\rho^2, \rho^4 \delta\}$ completes the proof. \square

Actually, it can be seen from equation (3.5) that the estimates of R and S would play a key role in the recursive expression of x_{n+1} , hence we will verify that $\frac{\pi_1 Z_1}{\pi_1 + \pi_2 Z_2}$ and $\frac{z_n}{x_n}$ are both sufficiently around π_1 , analogous to the concentration analysis result in Sly [2011]. In the following lemma, we firstly establish a technical uniqueness result where the set of vertices which can be conditioned is limited to a set of k vertices.

Lemma 6 *For any $\varepsilon > 0$ and positive integer k , there exists $M = M(\pi, \varepsilon, k)$, such that for any collection of vertices $v_1, \dots, v_k \in L(M)$,*

$$\sup_{i_1, \dots, i_k \in \mathcal{C}} \left| \mathbf{P}(\sigma_\rho = 1 \mid \sigma_{v_j} = i_j, 1 \leq j \leq k) - \pi_1 \right| \leq \varepsilon$$

Proof. Denote the entries of the transition matrix at distance s as

$$U_s = M_{1,1}^s \quad \text{and} \quad V_s = M_{2,2}^s,$$

and it is natural that $M_{1,2}^s = 1 - U_s$ and $M_{2,1}^s = 1 - V_s$. As a result, it follows that

$$\begin{cases} U_s = M_{11} U_{s-1} + M_{12} (1 - V_{s-1}) \\ V_s = M_{21} (1 - U_{s-1}) + M_{22} V_{s-1}, \end{cases}$$

which yields a second order recursive formula

$$U_s - (1 + \theta)U_{s-1} + \theta U_{s-2} = 0$$

with the initial conditions $U_0 = 1$ and $U_1 = M_{11} = \pi_1 + \pi_2\theta$. Then the general solutions are given by

$$U_s = \pi_1 + \pi_2\theta^s \quad \text{and} \quad V_s = \pi_2 + \pi_1\theta^s.$$

Consequently, under the condition of $d\theta^2 \leq 1$, we have

$$\pi_1 - \pi_2 d^{-s/2} \leq M_{1,1}^s \leq \pi_1 + \pi_2 d^{-s/2},$$

$$\pi_2 - \pi_1 d^{-s/2} \leq M_{2,2}^s \leq \pi_2 + \pi_1 d^{-s/2},$$

$$\pi_2 - \pi_2 d^{-s/2} \leq M_{1,2}^s \leq \pi_2 + \pi_2 d^{-s/2},$$

$$\pi_1 - \pi_1 d^{-s/2} \leq M_{2,1}^s \leq \pi_1 + \pi_1 d^{-s/2}.$$

For fixed π , d and k , define

$$B(s) := \max \left\{ \frac{\pi_1 + \pi_2 d^{-s/2}}{\pi_1 - \pi_2 d^{-s/2}}, \frac{\pi_2 + \pi_1 d^{-s/2}}{\pi_2 - \pi_1 d^{-s/2}}, \frac{1 + d^{-s/2}}{1 - d^{-s/2}} \right\}$$

and let $N = N(\pi, \varepsilon, k)$ be a sufficiently large integer such that

$$B^k(N) \leq 1 + \varepsilon,$$

where the last inequality holds by the fact that $d^{-s/2} \leq 2^{-s/2} \rightarrow 0$ as $s \rightarrow \infty$ which implies $B(s) \rightarrow 1$ uniformly for all d .

Now fix an integer M such that $M > kN$ and choose any $v_1, \dots, v_k \in L(M)$. For $0 \leq \ell \leq M$, define $n(\ell)$ as the number of vertices of distance ℓ from the root with a decedent in the set $\{v_1, \dots, v_k\}$, that is

$$n(\ell) = \#\{v \in L(\ell) : |\mathbb{T}_v \cap \{v_1, \dots, v_k\}| > 0\}.$$

Then according to the definition, it is trivial to see that $n(\ell)$ is an increasing integer valued function with respect to ℓ from $n_0 = 1$ to $n_M = k$, which, by the pigeonhole principle, implies that there must exist some ℓ such that $n(\ell) = n(\ell + N)$. Next, denote $\{w_1, \dots, w_{n(\ell)}\}$ and $\{\bar{w}_1, \dots, \bar{w}_{n(\ell)}\}$ as vertices in sets $\{v \in L(\ell + N) : |\mathbb{T}_v \cap \{v_1, \dots, v_k\}| > 0\}$ and $\{v \in L(\ell) : |\mathbb{T}_v \cap \{v_1, \dots, v_k\}| > 0\}$ respectively, such that w_j is the descendent of \bar{w}_j , and then

$$\mathbf{P}(\sigma_{w_j} = i_2 \mid \sigma_{\bar{w}_j} = i_1) = M_{i_1, i_2}^N.$$

By Bayes' Rule and the Markov random field property, for any $i_1, \dots, i_{n(\ell)} \in \mathcal{C}$, we have

$$\begin{aligned}
& \frac{\mathbf{P}(\sigma_\rho = 1 \mid \sigma_{w_j} = i_j, 1 \leq j \leq n(\ell))}{\mathbf{P}(\sigma_\rho = 2 \mid \sigma_{w_j} = i_j, 1 \leq j \leq n(\ell))} \\
&= \frac{\pi_1 \mathbf{P}(\sigma_{w_j} = i_j, 1 \leq j \leq n(\ell) \mid \sigma_\rho = 1)}{\pi_2 \mathbf{P}(\sigma_{w_j} = i_j, 1 \leq j \leq n(\ell) \mid \sigma_\rho = 2)} \\
&= \frac{\pi_1 \sum_{h_1, \dots, h_{n(\ell)} \in \mathcal{C}} \mathbf{P}(\forall j \sigma_{w_j} = i_j \mid \forall j \sigma_{\bar{w}_j} = h_j) \mathbf{P}(\forall j \sigma_{\bar{w}_j} = h_j \mid \sigma_\rho = 1)}{\pi_2 \sum_{h_1, \dots, h_{n(\ell)} \in \mathcal{C}} \mathbf{P}(\forall j \sigma_{w_j} = i_j \mid \forall j \sigma_{\bar{w}_j} = h_j) \mathbf{P}(\forall j \sigma_{\bar{w}_j} = h_j \mid \sigma_\rho = 2)} \\
&= \frac{\pi_1 \sum_{h_1, \dots, h_{n(\ell)} \in \mathcal{C}} \mathbf{P}(\forall j \sigma_{\bar{w}_j} = h_j \mid \sigma_\rho = 1) \prod_{j=1}^{n(\ell)} M_{h_j i_j}^N}{\pi_2 \sum_{h_1, \dots, h_{n(\ell)} \in \mathcal{C}} \mathbf{P}(\forall j \sigma_{\bar{w}_j} = h_j \mid \sigma_\rho = 2) \prod_{j=1}^{n(\ell)} M_{h_j i_j}^N} \\
&\leq \frac{\pi_1}{\pi_2} B^{n(\ell)}(N) \frac{\sum_{h_1, \dots, h_{n(\ell)} \in \mathcal{C}} \mathbf{P}(\forall j \sigma_{\bar{w}_j} = h_j \mid \sigma_\rho = 1)}{\sum_{h_1, \dots, h_{n(\ell)} \in \mathcal{C}} \mathbf{P}(\forall j \sigma_{\bar{w}_j} = h_j \mid \sigma_\rho = 2)} \\
&\leq \frac{\pi_1}{\pi_2} B^k(N) \\
&\leq \frac{\pi_1}{\pi_2} (1 + \varepsilon),
\end{aligned}$$

which implies that

$$\pi_1 - \varepsilon \leq \mathbf{P}(\sigma_\rho = 1 \mid \sigma_{w_j} = i_j, 1 \leq j \leq n(\ell)) \leq \pi_1 + \varepsilon.$$

Hence, for the reason that σ_ρ is conditionally independent of the collection $\sigma_{v_1}, \dots, \sigma_{v_k}$ given $\sigma_{w_1}, \dots, \sigma_{w_{n(\ell)}}$, one has

$$\sup_{i_1, \dots, i_k \in \mathcal{C}} \left| \mathbf{P}(\sigma_\rho = 1 \mid \sigma_{v_j} = i_j, 1 \leq j \leq k) - \pi_1 \right| \leq \sup_{i_1, \dots, i_{n(\ell)} \in \mathcal{C}} \left| \mathbf{P}(\sigma_\rho = 1 \mid \sigma_{w_j} = i_j, 1 \leq j \leq n(\ell)) - \pi_1 \right| \leq \varepsilon.$$

□

Lemma 7 *Assume $|\theta| > \rho$ for some $\rho > 0$. Given arbitrary $\varepsilon, \alpha > 0$, there exist constants $C = C(\pi, \varepsilon, \alpha, \rho) > 0$ and $N = N(\pi, \varepsilon, \alpha)$, such that whenever $n \geq N$,*

$$\mathbf{P} \left(\left| \frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1 \right| > \varepsilon \right) \leq C x_n^\alpha.$$

Proof. Fix k an integer with $k > \alpha$. Choose M to hold with bound $\varepsilon/2$ in Lemma 6. Let $v_1, \dots, v_{|L(M)|}$ denote the vertices in $L(M)$ and define

$$W(v) = f_{n+1-M}(1, \sigma_v^1(n+1)),$$

where $\sigma_v^1(n+1)$ denotes the spins of vertices in $\mathbb{T}_v \cap L(n+1)$. Then $W(v)$ would be distributed as

$$W(v) \sim \begin{cases} X^+(n+1-M) & \text{if } \sigma_v^1 = 1, \\ 1 - X^-(n+1-M) & \text{if } \sigma_v^1 = 2. \end{cases} \quad (4.3)$$

The recursion formula in equation (3.1) together with the fact that $1 - W(v) = f_{n+1-M}(2, \sigma_v^1(n+1))$, yield a function

$$H(W_1, \dots, W_{|L(M)|}) = f_{n+1}(1, \sigma^1(n+1)) = \frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2},$$

where $W_i = W(v_i)$ for $1 \leq i \leq |L(M)|$. There is no difficulty in finding that when all the entries W_i are identically π_1 one has

$$H(W_1, \dots, W_{|L(M)|}) = \pi_1,$$

and H is a continuous function of the vector $(W_i)_{1 \leq i \leq |L(M)|}$. Therefore, by Lemma 6, if there are at most k vertices in $L(M)$ such that $W(v) \neq \pi_1$, then

$$|H(W_1, \dots, W_{|L(M)|}) - \pi_1| < \frac{\varepsilon}{2},$$

and there exists some $\delta = \delta(\pi, \varepsilon) > 0$ such that if

$$\#\{v \in L(M) : |W(v) - \pi_1| > \delta\} \leq k$$

then

$$|H(W_1, \dots, W_{|L(M)|}) - \pi_1| < \varepsilon.$$

Next, by the Chebyshev's inequality together with equation (4.3), the following result holds:

$$\begin{aligned} \mathbf{P}(|W(v) - \pi_1| > \delta) &\leq \delta^{-2} [\mathbf{E}(X^+(n+1-M) - \pi_1)^2 + \mathbf{E}(X^-(n+1-M) - \pi_2)^2] \\ &\leq \frac{\delta^{-2}}{\min\{1, \frac{\pi_2}{\pi_1}\}} x_{n+1-M}. \end{aligned}$$

Random variables $|W(v) - \pi_1|$ for distinct v are conditionally independent given $\sigma(M)$, so there exist suitable constants $C(\pi, \varepsilon, \alpha, \rho)$ and $N(\pi, \varepsilon, \alpha)$, such that when $n \geq N$, one has

$$\begin{aligned} \mathbf{P}\left(\left|\frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1\right| > \varepsilon\right) &\leq \mathbf{P}(\#\{v \in L(M) : |W(v) - \pi_1| > \delta\} > k) \\ &= \sum_A \mathbf{P}(\#\{v \in L(M) : |W(v) - \pi_1| > \delta\} > k \mid \sigma(M) = A) \mathbf{P}(\sigma(M) = A) \\ &\leq \sum_A \mathbf{P}\left[\mathbf{B}\left(|L(M)|, \frac{\delta^{-2}}{\min\{1, \frac{\pi_2}{\pi_1}\}} x_{n+1-M}\right) > k\right] \mathbf{P}(\sigma(M) = A) \\ &\leq C'(\pi, \varepsilon, \alpha, \rho) x_{n+1-M}^\alpha \\ &\leq C x_n^\alpha, \end{aligned}$$

where $\mathbf{B}(\cdot, \cdot)$ denotes the binomial distribution and the last inequality holds due to Lemma 5. \square

Now, we are able to bound S and R in equation (3.7) using the preceding concentration results.

Proposition 1 *Assume $|\theta| > \rho$ for some $\rho > 0$. For any $\varepsilon > 0$, there exist $N = N(\pi, \varepsilon)$ and $\delta = \delta(\pi, \varepsilon, \rho) > 0$, such that if $n \geq N$ and $x_n \leq \delta$ then $|S| \leq \varepsilon x_n^2$.*

Proof. For any $\eta > 0$, using the Cauchy-Schwartz inequality and by Lemma 7, one has

$$\begin{aligned}
|S| &= \left| \mathbf{E}(\pi_1 Z_1 + \pi_2 Z_2 - 1)^2 \left(\frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1 \right) \right| \\
&\leq \mathbf{E} \left((\pi_1 Z_1 + \pi_2 Z_2 - 1)^2 \left| \frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1 \right| ; \left| \frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1 \right| \leq \eta \right) \\
&\quad + \mathbf{E} \left((\pi_1 Z_1 + \pi_2 Z_2 - 1)^2 \left| \frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1 \right| ; \left| \frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1 \right| > \eta \right) \\
&\leq \eta \mathbf{E}(\pi_1 Z_1 + \pi_2 Z_2 - 1)^2 + \mathbf{E}(\pi_1 Z_1 + \pi_2 Z_2 - 1)^2 \mathbf{I} \left(\left| \frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1 \right| > \eta \right) \\
&\leq \eta \mathbf{E}(\pi_1 Z_1 + \pi_2 Z_2 - 1)^2 + \mathbf{P} \left(\left| \frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1 \right| > \eta \right)^{1/2} [\mathbf{E}(\pi_1 Z_1 + \pi_2 Z_2 - 1)^4]^{1/2}.
\end{aligned}$$

Also, it follows from equation (3.4) and Lemma 4 respectively that

$$\mathbf{E}(\pi_1 Z_1 + \pi_2 Z_2 - 1)^2 \leq C_1(\pi) x_n^2$$

and

$$(\mathbf{E}(\pi_1 Z_1 + \pi_2 Z_2 - 1)^4)^{1/2} \leq C_2(\pi).$$

Taking $\alpha = 6$ in Lemma 7, there exist $C_3 = C_3(\pi, \eta, \rho)$ and $N = N(\pi, \eta)$, such that if $n \geq N$ then

$$\mathbf{P} \left(\left| \frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1 \right| > \eta \right) \leq C_3^2 x_n^6.$$

Finally taking $\eta = \varepsilon / (2C_1)$ and $\delta = \varepsilon / (2C_2 C_3)$, we have that if $n \geq N$ and $x_n \leq \delta$ then

$$|S| \leq \eta C_1 x_n^2 + C_2 C_3 x_n^3 \leq \varepsilon x_n^2.$$

□

Proposition 2 Assume $|\theta| > \rho$ for some $\rho > 0$. For any $\varepsilon > 0$, there exist $N = N(\pi, \varepsilon)$ and $\delta = \delta(\pi, \varepsilon, \rho)$, such that if $n \geq N$ and $x_n \leq \delta$ then

$$\left| \frac{z_n}{x_n} - \pi_1 \right| \leq \varepsilon.$$

Proof. Plugging $a = (Z_1 - Z_2)^2$, $r = (\pi_1 Z_1 + \pi_2 Z_2)^2 - 1$ and $s = 1$ in the identity equation (3.2), we have

$$\begin{aligned}
z_{n+1} &= \mathbf{E} \left(\frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1 \right)^2 \\
&= \pi_1^2 \pi_2^2 \mathbf{E} \frac{(Z_1 - Z_2)^2}{1 + (\pi_1 Z_1 + \pi_2 Z_2)^2 - 1} \\
&= \pi_1^2 \pi_2^2 \left\{ \mathbf{E}(Z_1 - Z_2)^2 - \mathbf{E}(Z_1 - Z_2)^2 [(\pi_1 Z_1 + \pi_2 Z_2)^2 - 1] + \mathbf{E}[(\pi_1 Z_1 + \pi_2 Z_2)^2 - 1]^2 \frac{(Z_1 - Z_2)^2}{(\pi_1 Z_1 + \pi_2 Z_2)^2} \right\}.
\end{aligned}$$

Next we will estimate these expectation terms one by one with the O_π -constants depend only on π :

$$\begin{aligned}
\mathbf{E}(Z_1 - Z_2)^2 &= \mathbf{E}(Z_1^2 + Z_2^2 - 2Z_1Z_2) \\
&= 1 + d \left[\frac{\theta^2}{\pi_1}(3-\theta)x_n + \frac{\theta^3}{\pi_1^2}z_n \right] + 1 + d \left\{ \frac{\theta^2}{\pi_2} \left[\frac{\pi_1}{\pi_2}(1-\theta) - 2 \right] x_n + \frac{\theta^3}{\pi_2^2}z_n \right\} \\
&\quad - 2 \left\{ 1 + d \left[\theta^2 \left(\frac{1}{\pi_1} + \frac{\theta-2}{\pi_2} \right) x_n - \frac{\theta^3}{\pi_1\pi_2}z_n \right] \right\} + O_\pi(x_n^2) \\
&= d\theta^2 \left(\frac{1-\theta}{\pi_1\pi_2^2}x_n + \frac{\theta}{\pi_1^2\pi_2^2}z_n \right) + O_\pi(x_n^2), \\
\mathbf{E}(Z_1 - Z_2)^2 [(\pi_1Z_1 + \pi_2Z_2)^2 - 1] &= O_\pi(x_n^2), \\
\pi_1^2\pi_2^2\mathbf{E}[(\pi_1Z_1 + \pi_2Z_2)^2 - 1]^2 \frac{(Z_1 - Z_2)^2}{(\pi_1Z_1 + \pi_2Z_2)^2} &\leq \mathbf{E}[(\pi_1Z_1 + \pi_2Z_2)^2 - 1]^2 = O_\pi(x_n^2),
\end{aligned}$$

where we used the fact that $\pi_1^2\pi_2^2(Z_1 - Z_2)^2/(\pi_1Z_1 + \pi_2Z_2)^2 \leq 1$ in the last inequality. Therefore, the recursion formula of z_{n+1} can be written as

$$z_{n+1} = d\theta^2[\pi_1(1-\theta)x_n + \theta z_n] + O_\pi(x_n^2).$$

In the rest of the proof, we let $\{C_i\}_{i=1,2,3,4}$ be constants depend only on π . It follows from equation (4.1) that

$$\left| \frac{d\theta^2 x_n}{x_{n+1}} - 1 \right| \leq C_1 \frac{x_n^2}{x_{n+1}},$$

and in view of $d\theta^2 \geq 1/2$, there exists $\delta_1 = \delta_1(\pi) > 0$, such that if $x_n \leq \delta_1$ then

$$\frac{x_n}{x_{n+1}} = \frac{x_n}{d\theta^2 x_n + O_\pi(x_n^2)} \leq \frac{x_n}{9d\theta^2 x_n/10} = \frac{10}{9} \frac{1}{d\theta^2} \leq \frac{20}{9}. \quad (4.4)$$

Consequently,

$$\begin{aligned}
\left| \frac{z_{n+1}}{x_{n+1}} - \left[\pi_1(1-\theta) + \theta \frac{z_n}{x_n} \right] \right| &= \left| \frac{z_{n+1}}{x_{n+1}} - \frac{d\theta^2 x_n}{x_{n+1}} \left[\pi_1(1-\theta) + \theta \frac{z_n}{x_n} \right] \right| + \left| \left(\frac{d\theta^2 x_n}{x_{n+1}} - 1 \right) \left[\pi_1(1-\theta) + \theta \frac{z_n}{x_n} \right] \right| \\
&\leq C_2 \frac{x_n^2}{x_{n+1}} + C_3 \frac{x_n^2}{x_{n+1}} \\
&\leq C_4 x_n.
\end{aligned} \quad (4.5)$$

For any $k \in \mathbb{N}$, by equation (4.1), there exists a $\delta_2 = \delta_2(\pi, k)$, such that if $x_n \leq \delta_2$ then $x_{n+\ell} \leq 2\delta_2 \leq \delta_1$ for any $1 \leq \ell \leq k$. Now iterating k times the inequality (4.5) yields

$$\begin{aligned}
&\left| \frac{z_{n+k}}{x_{n+k}} - \left[\pi_1(1-\theta^k) + \theta^k \frac{z_n}{x_n} \right] \right| \\
&\leq \sum_{\ell=1}^k \left| \pi_1(1-\theta^{k-\ell}) + \theta^{k-\ell} \frac{z_{n+\ell}}{x_{n+\ell}} - \pi_1(1-\theta^{k-\ell+1}) - \theta^{k-\ell+1} \frac{z_{n+\ell-1}}{x_{n+\ell-1}} \right| \\
&\leq \sum_{\ell=1}^k |\theta|^{k-\ell} \left| \frac{z_{n+\ell}}{x_{n+\ell}} - \left[\pi_1(1-\theta) + \theta \frac{z_{n+\ell-1}}{x_{n+\ell-1}} \right] \right| \\
&\leq C_4 \sum_{\ell=1}^k |\theta|^{k-\ell} x_{n+\ell-1}.
\end{aligned}$$

Therefore, noting that $|\theta| \leq d^{-1/2} \leq 2^{-1/2}$, and taking $k = k(\varepsilon)$ large enough and $\delta_3 = \delta_3(\pi, \varepsilon, k) = \delta_3(\pi, \varepsilon)$ sufficiently small, we obtain that if $x_n \leq \delta_3$ then

$$\left| \frac{z_{n+k}}{x_{n+k}} - \pi_1 \right| \leq |\theta|^k + 2\delta_3 C_4 \sum_{\ell=1}^k |\theta|^{k-\ell} = |\theta|^k + 2\delta_3 C_4 \frac{1 - |\theta|^k}{1 - |\theta|} \leq \varepsilon, \quad (4.6)$$

where the first inequality relies on the fact that $|z_n/x_n - \pi_1| < 1$. At last, choosing $N = N(\pi, \varepsilon) > k$ and $\delta = \delta(\pi, \varepsilon, \rho) = \gamma^k \delta_3$, and noting that by Lemma 5 if $n \geq N$ and $x_n \leq \delta$ then $x_{n-k} \leq \gamma^{-k} x_n \leq \delta_3$, the previous result in equation (4.6) completes the proof. \square

5 Proof of Theorem 1.1

To accomplish the proof, it suffices to show that when $d\theta^2$ is close enough to 1, x_n does not converge to 0. For convenience, we suppose that $d\theta^2 \geq 1/2$. For any fixed d and π , there is $|\theta| \geq (2d)^{-1/2}$, and we take $\rho = (2d)^{-1/2}$ in Lemma 5 to generate $\gamma = \gamma(\pi, d) > 0$. When $1 - 6\pi_1\pi_2 > 0$, by Proposition 1 and Proposition 2, there exist $N = N(\pi)$ and $\delta = \delta(\pi, d) > 0$, such that if $n \geq N$ and $x_n \leq \delta$, then the remainders in equation (3.5) could be evaluated respectively as

$$|R| \leq \frac{1}{6} \frac{1 - 6\pi_1\pi_2}{\pi_1\pi_2^2} \frac{d(d-1)}{2} \theta^4 x_n^2 \quad (5.1)$$

and

$$|S| \leq \frac{1}{6} \frac{1 - 6\pi_1\pi_2}{\pi_1\pi_2^2} \frac{d(d-1)}{2} \theta^4 x_n^2. \quad (5.2)$$

As a consequence,

$$x_{n+1} \geq d\theta^2 x_n + \frac{1}{2} \frac{(1 - 6\pi_1\pi_2)}{\pi_1\pi_2^2} \frac{d(d-1)}{2} \theta^4 x_n^2. \quad (5.3)$$

Furthermore, in light of $x_0 = 1 - \pi_1 = \pi_2$ and Lemma 5, for all n we have

$$x_n \geq \pi_2 \gamma^n. \quad (5.4)$$

Define $\varepsilon = \varepsilon(\pi, d) = \min\{\pi_2 \gamma^N, \delta \gamma\} > 0$. Then equation (5.4) implies that $x_n \geq \varepsilon$ when $n \leq N$. Next, by choosing suitable $|\theta| < d^{-1/2}$, we achieve

$$d\theta^2 + \frac{1}{2} \frac{(1 - 6\pi_1\pi_2)}{\pi_1\pi_2^2} \frac{d(d-1)}{2} \theta^4 \varepsilon \geq 1, \quad (5.5)$$

for the reason that ε is independent of θ . Now, suppose $x_n \geq \varepsilon$ for some $n \geq N$. If $x_n \geq \gamma^{-1}\varepsilon$, then Lemma 5 gives $x_{n+1} \geq \gamma x_n \geq \varepsilon$. If $\varepsilon \leq x_n \leq \gamma^{-1}\varepsilon \leq \delta$, then by equation (5.3) and equation (5.5), we have

$$x_{n+1} \geq d\theta^2 x_n + \frac{1}{2} \frac{(1 - 6\pi_1\pi_2)}{\pi_1\pi_2^2} \frac{d(d-1)}{2} \theta^4 x_n^2 \geq x_n \left[d\theta^2 + \frac{1}{2} \frac{(1 - 6\pi_1\pi_2)}{\pi_1\pi_2^2} \frac{d(d-1)}{2} \theta^4 \varepsilon \right] \geq x_n \geq \varepsilon.$$

Hence it can be shown by induction that $x_n \geq \varepsilon$ for all n , namely, the Kesten-Stigum bound is not tight.

6 Large Degree Asymptotics

6.1 Gaussian Approximation

For $1 \leq j \leq d$, define

$$U_j = \log \left[1 + \frac{\theta}{\pi_1} (Y_j - \pi_1) \right] \quad \text{and} \quad V_j = \log \left[1 - \frac{\theta}{\pi_2} (Y_j - \pi_1) \right].$$

Lemma 8 *There exist positive constants $C = C(\pi)$ and $D = D(\pi)$, such that when $d > D$,*

$$\begin{aligned} \left| d\mathbf{E}U_j - \frac{d\theta^2}{2\pi_1} x_n \right| &\leq Cd^{-1/2}, & \left| d\mathbf{E}V_j + \frac{1+\pi_2}{2\pi_2^2} d\theta^2 x_n \right| &\leq Cd^{-1/2}, \\ \left| d\mathbf{Var}(U_j) - \frac{d\theta^2}{\pi_1} x_n \right| &\leq Cd^{-1/2}, & \left| d\mathbf{Var}(V_j) - \frac{\pi_1}{\pi_2^2} d\theta^2 x_n \right| &\leq Cd^{-1/2}, \\ \left| d\mathbf{Cov}(U_j, V_j) + \frac{d\theta^2}{\pi_2} x_n \right| &\leq Cd^{-1/2}. \end{aligned}$$

Proof. Starting with the Taylor series expansion of $\log(1+w)$, there exists a constant $W > 0$, such that when $|w| < W$,

$$\left| \log(1+w) - w + \frac{w^2}{2} \right| \leq |w|^3. \quad (6.1)$$

Taking $D = D(\pi)$ sufficiently large, when $d > D$, we have that $|\theta| \leq d^{-\frac{1}{2}}$ is small enough to guarantee equation (6.1) for $w = \theta(Y_j - \pi_1)/\pi_1$ and then

$$\begin{aligned} \left| \mathbf{E}U_j - \frac{\theta^2}{2\pi_1} x_n \right| &\leq \mathbf{E} \left| U_j - w + \frac{w^2}{2} \right| + \left| \mathbf{E}w - \frac{w^2}{2} - \frac{\theta^2}{2\pi_1} x_n \right| \\ &\leq \mathbf{E} \frac{\theta^3}{\pi_1^3} |Y_j - \pi_1|^3 + \frac{\theta^3}{2\pi_1^2} |z_n - \pi_1 x_n| \\ &\leq \frac{\theta^3}{\pi_1^3} + \frac{\theta^3}{2\pi_1^2} \\ &\leq C(\pi) d^{-3/2} \end{aligned}$$

for some constant $C = C(\pi)$, where the third inequality follows from $0 \leq z_n \leq x_n \leq 1$. The rest estimates follow similarly. \square

Next, define a 2-dimensional vector $\mu = (\mu_1, \mu_2)$ with $\mu_1 = \frac{1}{2\pi_1}$, $\mu_2 = -\frac{1+\pi_2}{2\pi_2^2}$, and a 2×2 -covariance matrix

$$\Sigma = \begin{pmatrix} \frac{1}{\pi_1} & -\frac{1}{\pi_2} \\ -\frac{1}{\pi_2} & \frac{\pi_1}{\pi_2^2} \end{pmatrix},$$

which is a positive semi-definite symmetric 2×2 -matrix. Let (G_1, G_2) possess the Gaussian distribution $\mathbf{N}(\mu, \Sigma)$, then the following lemma can be established by the Central Limit Theorem, the Gaussian approximation and the Portmanteau Theorem.

Lemma 9 *Let $\psi : \mathbb{R}^2 \mapsto \mathbb{R}$ be a differentiable bounded function. For any $\varepsilon > 0$, there exists $D = D(\pi, \psi, \varepsilon) > 0$, such that if $d > D$ then*

$$\left| \mathbf{E} \psi \left(\sum_{j=1}^d U_j, \sum_{j=1}^d V_j \right) - \mathbf{E} \psi(G_1, G_2) \right| \leq \varepsilon.$$

Next, define

$$\psi(w_1, w_2) = \frac{\pi_1 e^{w_1}}{\pi_1 e^{w_1} + \pi_2 e^{w_2}}$$

and then

$$\begin{aligned} x_{n+1} &= \mathbf{E} \frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1 \\ &= \mathbf{E} \frac{\pi_1 \exp\left(\sum_{j=1}^d U_j\right)}{\pi_1 \exp\left(\sum_{j=1}^d U_j\right) + \pi_2 \exp\left(\sum_{j=1}^d V_j\right)} - \pi_1 \\ &= \mathbf{E} \psi \left(\sum_{j=1}^d U_j, \sum_{j=1}^d V_j \right) - \pi_1. \end{aligned}$$

If (W_1, W_2) has the Gaussian distribution $\mathbf{N}(0, \Sigma)$, then $(s\mu_1 + \sqrt{s}W_1, s\mu_2 + \sqrt{s}W_2)$ is distributed according to $\mathbf{N}(s\mu, s\Sigma)$. At last, define

$$g(s) = \mathbf{E} \psi(s\mu_1 + \sqrt{s}W_1, s\mu_2 + \sqrt{s}W_2) - \pi_1 = \mathbf{E} \frac{\pi_1 \exp(s\mu_1 + \sqrt{s}W_1)}{\pi_1 \exp(s\mu_1 + \sqrt{s}W_1) + \pi_2 \exp(s\mu_2 + \sqrt{s}W_2)} - \pi_1.$$

Therefore, Lemma 9 implies the following approximation to x_{n+1} .

Lemma 10 *For arbitrary $\varepsilon > 0$, there exists a $D = D(\pi, \varepsilon) > 0$, such that whenever $d > D$,*

$$|x_{n+1} - g(d\theta^2 x_n)| \leq \varepsilon.$$

6.2 Asymptotic Estimation of the Reconstruction Threshold

In order to estimate x_{n+1} , it suffices to investigate the properties of $g(s)$ on the interval $[0, \pi_2]$, considering that $0 \leq x_n \leq \pi_2$ and $d\theta^2 \leq 1$.

Lemma 11 *The function $g(s)$ is continuously differentiable and increasing on the interval $(0, \pi_2]$.*

Proof. When $s > 0$, it is concluded that

$$\begin{aligned} & \mathbf{E} \left| \frac{\partial}{\partial s} \frac{\pi_1 \exp(s\mu_1 + \sqrt{s}W_1)}{\pi_1 \exp(s\mu_1 + \sqrt{s}W_1) + \pi_2 \exp(s\mu_2 + \sqrt{s}W_2)} \right| \\ &= \mathbf{E} \left| \frac{\partial}{\partial s} \left(1 + \frac{\pi_2}{\pi_1} \exp(s(\mu_2 - \mu_1) + \sqrt{s}(W_2 - W_1)) \right)^{-1} \right| \\ &= \mathbf{E} \left| \frac{\frac{\pi_2}{\pi_1} \exp(s(\mu_2 - \mu_1) + \sqrt{s}(W_2 - W_1))}{\left(1 + \frac{\pi_2}{\pi_1} \exp(s(\mu_2 - \mu_1) + \sqrt{s}(W_2 - W_1)) \right)^2} \left(\mu_2 - \mu_1 + \frac{W_2 - W_1}{2\sqrt{s}} \right) \right| \\ &\leq \frac{1}{4} \mathbf{E} \left| \mu_2 - \mu_1 + \frac{W_2 - W_1}{2\sqrt{s}} \right| \\ &< \infty, \end{aligned}$$

by the fact that $\left| \frac{\pi_2}{\pi_1} e^t / \left(1 + \frac{\pi_2}{\pi_1} e^t \right)^2 \right| \leq 1/4$ holds for any $t \in \mathbb{R}$. Then we establish the differentiability with respect to s .

Now, let (W'_1, W'_2) be an independent copy of (W_1, W_2) . Thus if $0 \leq s' \leq s$, it is feasible to construct equivalent distributions such as

$$\sqrt{s}(W_1, W_2) \sim \sqrt{s'}(W_1, W_2) + \sqrt{s-s'}(W'_1, W'_2).$$

In view of $(W_1, W_2) \sim \mathbf{N}(0, \Sigma)$, it follows that $\mathbf{E}(W_2 - W_1) = 0$ and

$$\mathbf{Var}(W_2 - W_1) = \mathbf{E}W_2^2 + \mathbf{E}W_1^2 - 2\mathbf{E}W_1W_2 = \frac{1}{\pi_1} + \frac{\pi_1}{\pi_2^2} - 2\left(-\frac{1}{\pi_2}\right) = \frac{1}{\pi_1\pi_2^2},$$

which implies that $W_2 - W_1$ and $W'_2 - W'_1$ are both distributed as $\mathbf{N}(0, a)$, with $a = 1/\pi_1\pi_2^2$.

Next, it is well known that if W has the distribution $\mathbf{N}(\mu, \sigma^2)$, the expectation of the exponential random variable could be estimated as

$$\mathbf{E}e^W = e^{\mu + \frac{\sigma^2}{2}}, \quad (6.2)$$

based on which, we are able to estimate the conditional expectation given W_1 and W_2 :

$$\mathbf{E} \left[\exp(\sqrt{s'}(W_2 - W_1) + \sqrt{s-s'}(W'_2 - W'_1)) \mid \{W_1, W_2\} \right] = \exp \left[\sqrt{s'}(W_2 - W_1) + \frac{a}{2}(s-s') \right].$$

Then applying Jensen's inequality, and considering that the function $(1+x)^{-1}$ is convex and

$$\mu_2 - \mu_1 = -(1 + \pi_2)/(2\pi_2^2) - 1/(2\pi_1) = -1/(2\pi_1\pi_2^2) = -a/2,$$

we have

$$\begin{aligned} g(s) &= \mathbf{E} \left(1 + \frac{\pi_2}{\pi_1} \exp(s(\mu_2 - \mu_1) + \sqrt{s}(W_2 - W_1)) \right)^{-1} - \pi_1 \\ &= \mathbf{E} \left(1 + \frac{\pi_2}{\pi_1} \exp \left(\frac{-as}{2} + \sqrt{s'}(W_2 - W_1) + \sqrt{s-s'}(W'_2 - W'_1) \right) \right)^{-1} - \pi_1 \\ &\geq \mathbf{E} \left(1 + \frac{\pi_2}{\pi_1} \exp \left(-\frac{as}{2} \right) \mathbf{E} \left[\exp(\sqrt{s'}(W_2 - W_1) + \sqrt{s-s'}(W'_2 - W'_1)) \mid \{W_1, W_2\} \right] \right)^{-1} - \pi_1 \\ &= \mathbf{E} \left(1 + \frac{\pi_2}{\pi_1} \exp \left(-\frac{as'}{2} \right) \mathbf{E} \exp(\sqrt{s'}(W_2 - W_1)) \right)^{-1} - \pi_1 \\ &= \mathbf{E} \left(1 + \frac{\pi_2}{\pi_1} \exp(s'(\mu_2 - \mu_1) + \sqrt{s'}(W_2 - W_1)) \right)^{-1} - \pi_1 \\ &= g(s'), \end{aligned}$$

as desired. □

It is necessary to discuss the Taylor expansion of $g(s)$ in the small neighborhoods of $s = 0$.

Lemma 12 *For small $s > 0$, we have*

$$g(s) = s + \frac{1 - 6\pi_1\pi_2}{2\pi_1\pi_2^2} s^2 + \frac{1 - 24\pi_1\pi_2 + 90\pi_1^2\pi_2^2}{6\pi_1^2\pi_2^4} s^3 + O_\pi(s^4).$$

Proof. Define $W = s(\mu_2 - \mu_1) + \sqrt{s}(W_2 - W_1)$. By the results in Lemma 11, it is apparent that $W \sim \mathbf{N}(-as/2, as)$. Therefore by equation (6.2) the following moments can be calculated:

$$\begin{aligned}\mathbf{E}(e^W - 1) &= e^{-\frac{as}{2} + \frac{as}{2}} - 1 = e^0 - 1 = 0, \\ \mathbf{E}(e^W - 1)^2 &= e^{as} - 1 = as + \frac{a^2 s^2}{2} + \frac{a^3 s^3}{6} + O(s^4), \\ \mathbf{E}(e^W - 1)^3 &= e^{3as} - 3e^{as} + 2 = 3a^2 s^2 + 4a^3 s^3 + O(s^4), \\ \mathbf{E}(e^W - 1)^4 &= e^{6as} - 4e^{3as} + 6e^{as} - 3 = 3a^2 s^2 + 19a^3 s^3 + O(s^4), \\ \mathbf{E}(e^W - 1)^5 &= e^{10as} - 5e^{6as} + 10e^{3as} - 10e^{as} + 4 = 30a^3 s^3 + O(s^4), \\ \mathbf{E}(e^W - 1)^6 &= e^{15as} - 6e^{10as} + 15e^{6as} - 20e^{3as} + 15e^{as} - 5 = 15a^3 s^3 + O(s^4), \\ \mathbf{E}(e^W - 1)^7 &= e^{21as} - 7e^{15as} + 21e^{10as} - 35e^{6as} + 35e^{3as} - 21e^{as} + 6 = O(s^4).\end{aligned}$$

Next starting from the identity

$$\frac{1}{1 + \pi_2(e^W - 1)} = \sum_{n=0}^6 (-1)^n \pi_2^n (e^W - 1)^n - \pi_2^7 (e^W - 1)^7 \frac{1}{1 + \pi_2(e^W - 1)},$$

we obtain the power series of $g(s)$ as

$$\begin{aligned}\frac{g(s) + \pi_1}{\pi_1} &= \frac{1}{\pi_1} \mathbf{E} \frac{\pi_1 \exp(s\mu_1 + \sqrt{s}W_1)}{\pi_1 \exp(s\mu_1 + \sqrt{s}W_1) + \pi_2 \exp(s\mu_2 + \sqrt{s}W_2)} \\ &= \mathbf{E} \frac{1}{1 + \pi_2(e^W - 1)} \\ &= \mathbf{E} \left[\sum_{n=0}^6 (-1)^n \pi_2^n (e^W - 1)^n - \pi_2^7 (e^W - 1)^7 \frac{1}{1 + \pi_2(e^W - 1)} \right] \\ &= 1 + \frac{1}{\pi_1} \left(s + \frac{1 - 6\pi_1\pi_2}{2\pi_1\pi_2^2} s^2 + \frac{1 - 24\pi_1\pi_2 + 90\pi_1^2\pi_2^2}{6\pi_1^2\pi_2^4} s^3 + O(s^4) \right),\end{aligned}$$

that is,

$$g(s) = s + \frac{1 - 6\pi_1\pi_2}{2\pi_1\pi_2^2} s^2 + \frac{1 - 24\pi_1\pi_2 + 90\pi_1^2\pi_2^2}{6\pi_1^2\pi_2^4} s^3 + O(s^4).$$

□

6.3 Proof of Theorem 1.2

In this section, we precisely rephrase Theorem 1.2 and give its rigorous proof.

Theorem 6.1 *When $\pi_1\pi_2 < \frac{1}{6}$, define*

$$\omega^* = \inf\{\omega > 0 : \exists s \in (0, \pi_2), g(\omega s) = s\}.$$

Then $0 < \omega^ < 1$, and for any $\delta > 0$ there exists a $D = D(\pi, \delta)$, such that if $d > D$ then the model has reconstruction when $d\theta^2 \geq \omega^* + \delta$, but does not have reconstruction when $d\theta^2 \leq \omega^* - \delta$. In other words,*

$$\lim_{d \rightarrow \infty} d(\theta^\pm)^2 = \omega^*.$$

Proof. It follows from Lemma 12 that when $1 - 6\pi_1\pi_2 > 0$, there exists $0 < \bar{s} < \pi_2$ such that $g(\bar{s}) > \bar{s}$. Moreover, noting that $g(0 \cdot \bar{s}) = g(0) = 0 < \bar{s}$, the Intermediate Value Theorem implies the existence of $0 < \bar{\omega} < 1$ such that $g(\bar{\omega}\bar{s}) = \bar{s}$. Consequently, ω^* does exist and $0 \leq \omega^* \leq \bar{\omega} < 1$. Furthermore, for any $\omega^* < \omega < 1$, it follows from Lemma 11 that the set $\{0 < s < \pi_2 : g(\omega s) \geq s\}$ is a non-empty compact set bounded away from 0. Then it is further established by the continuity of $g(s)$ that the set

$$\{0 < s < \pi_2 : g(\omega^* s) = s\} = \bigcap_{\omega^* < \omega < 1} \{0 < s < \pi_2 : g(\omega s) \geq s\}$$

is non-empty and compact. Hence it implies immediately that $0 < \omega^* < 1$.

Next, taking $s^* \in \{0 < s < \pi_2 : g(\omega^* s) = s\}$ and considering $d\theta^2 = \omega^* + \delta$, one has

$$g\left[(\omega^* + \delta)\left(s^* \frac{\omega^*}{\omega^* + \delta}\right)\right] = g(s^* \omega^*) = s^* > s^* \frac{\omega^*}{\omega^* + \delta}.$$

Define $\varepsilon = \varepsilon(\pi, \delta) = s^* - s^* \omega^* / (\omega^* + \delta) > 0$. By Lemma 10 there exists a $D = D(\pi, \varepsilon) = D(\pi, \delta)$, such that if $d > D$ and $x_n \geq s^* \omega^* / (\omega^* + \delta)$ then

$$x_{n+1} \geq g(d\theta^2 x_n) - \varepsilon \geq g\left[(\omega^* + \delta)\left(s^* \frac{\omega^*}{\omega^* + \delta}\right)\right] - \left(s^* - s^* \frac{\omega^*}{\omega^* + \delta}\right) = s^* \frac{\omega^*}{\omega^* + \delta},$$

where the second inequality follows from Lemma 11. Consequently, it is shown by induction, and noting the initial value $x_0 = \pi_2 > s^*$, that $x_n \geq s^* \omega^* / (\omega^* + \delta)$ for all n , which further establishes reconstruction. At last, Proposition 12 in Mossel [2001] implies that the reconstruction is solvable for any $d\theta^2 \geq \omega^* + \delta$.

On the other hand, when $d\theta^2 = \omega^* - \delta$, we have $g(d\theta^2 s) \leq d\theta^2 s / \omega^*$. Taking $\eta = (1 - \omega^*) / 2 > 0$ in equation (4.1), there exists a constant $\zeta = \zeta(\pi)$, such that if $x_n < \zeta$ then

$$x_{n+1} \leq d\theta^2 x_n + \eta x_n \leq \frac{1}{2}(1 + \omega^*) x_n,$$

where the fact that $(1 + \omega^*) / 2 < 1$ implies that $\lim_{n \rightarrow \infty} x_n = 0$ and then there is non-reconstruction. So, it suffices to find some m , such that $x_m < \zeta$, which could be accomplished by choosing

$$\varepsilon = \left(1 - \frac{d\theta^2}{\omega^*}\right) \frac{\zeta}{2}$$

in Lemma 10. Then, there exists a sufficiently large $D = D(\pi, \varepsilon) = D(\pi, \delta)$, such that if $d > D$ and $x_n \geq \zeta$ then

$$x_{n+1} \leq g(d\theta^2 x_n) + \varepsilon \leq \frac{d\theta^2}{\omega^*} x_n + \varepsilon \leq \frac{1}{2} \left(1 + \frac{d\theta^2}{\omega^*}\right) x_n.$$

Then the fact that $(1 + d\theta^2 / \omega^*) / 2 < 1$ guarantees the existence of m satisfying $x_m < \zeta$, as desired. Finally using Proposition 12 in Mossel [2001] again, one can conclude non-reconstruction for any $d\theta^2 \leq \omega^* - \delta$. \square

6.4 Proof of Theorem 1.3

When $1 - 6\pi_1\pi_2 < 0$, the proof of Theorem 1.3 would resemble Theorem 1.1 in establishing a similar recursive inequality as equation (5.3), under the condition that $x_n \leq \delta$ and $n \geq N$ for suitable $\delta = \delta(\pi, d)$ and $N = N(\pi)$. However, there still exists a crucial discrepancy between these two proofs, that is, Theorem 1.3 relies heavily on large d . Before we proceed, let us firstly give the following lemma:

Lemma 13 *For any $0 < \varepsilon < 1$ and $\alpha > 1$, there exist $C = C(\pi, \varepsilon, \alpha)$ and $D = D(\pi, \varepsilon, \alpha)$ such that if $d > D$ then*

$$P\left(\left|\frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1\right| > \varepsilon\right) \leq Cx_n^\alpha. \quad (6.3)$$

Furthermore, there exist $D = D(\pi, \varepsilon)$ and $\delta = \delta(\pi, \varepsilon)$ such that if $d > D$ and $x_n \leq \delta$ then

$$|S| \leq \varepsilon x_n^2. \quad (6.4)$$

Proof. For any $1 \leq j \leq d$, define

$$w_j = \theta \frac{Y_j - \pi_1}{\pi_1} - \left(\theta \frac{Y_j - \pi_1}{\pi_1}\right)^2 \quad \text{and} \quad M = -\frac{\log(1 - \varepsilon)}{4\alpha}.$$

From equation (6.1) and $|\theta| \leq d^{-\frac{1}{2}}$, one can find a suitable $D = D(\pi, M) > 0$ such that when $d \geq D$ we have $\theta < \pi_1$, $U_j \geq w_j$ and $|w_j - \mathbf{E}w_j| \leq 2M$. It is concluded from Lemma 2 that $|d\mathbf{E}w_j| \leq C_1 x_n$ and

$$\sum_{j=1}^d \mathbf{E}(w_j - \mathbf{E}w_j)^2 \leq 2d\mathbf{E}\left(\theta \frac{Y_1 - \pi_1}{\pi_1}\right)^2 + 2d\mathbf{E}\left(\theta \frac{Y_1 - \pi_1}{\pi_1}\right)^4 \leq 4d\mathbf{E}\left(\theta \frac{Y_1 - \pi_1}{\pi_1}\right)^2 \leq C_2 x_n, \quad (6.5)$$

where C_1 and C_2 denote the constants depending only on π . In the following context, it is convenient to presume

$$x_n \leq -\frac{\log(1 - \varepsilon)}{2C_1}, \quad (6.6)$$

for the reason that equation (6.3) would be trivial otherwise. Therefore, it follows from equation (6.6) and the Bennet's inequality that

$$\begin{aligned}
\mathbf{P}(Z_1 \leq 1 - \varepsilon) &= \mathbf{P}\left(\sum_{j=1}^d U_j \leq \log(1 - \varepsilon)\right) \\
&\leq \mathbf{P}\left(\sum_{j=1}^d w_j \leq \log(1 - \varepsilon)\right) \\
&\leq \mathbf{P}\left(-\sum_{j=1}^d (w_j - \mathbf{E}w_j) \geq -\log(1 - \varepsilon) + d\mathbf{E}w_1\right) \\
&\leq \mathbf{P}\left(\sum_{j=1}^d [-(w_j - \mathbf{E}w_j)] \geq -\frac{\log(1 - \varepsilon)}{2}\right) \\
&\leq \exp\left[\left(-\frac{C_2 x_n}{4M^2}\right)\left(-\frac{M\log(1 - \varepsilon)}{C_2 x_n}\right)\left(\log\frac{-M\log(1 - \varepsilon)}{C_2 x_n} - 1\right)\right] \\
&\leq \exp\left[\left(\frac{\log(1 - \varepsilon)}{4M}\right)\left(\log\frac{-M\log(1 - \varepsilon)}{C_2} - 1\right)\right] x_n^{-\frac{\log(1 - \varepsilon)}{4M}} \\
&\leq C_3 x_n^\alpha,
\end{aligned}$$

where C_3 depends only on π, ε, α . Similarly one can show that $\mathbf{P}(Z_1 \geq 1 + \varepsilon) < C_4 x_n^\alpha$ and then

$$\mathbf{P}(|Z_1 - 1| > \varepsilon) \leq (C_3 + C_4) x_n^\alpha$$

for some $C_4 = C_4(\pi, \varepsilon, \alpha)$. Similarly one can also show that $\mathbf{P}(|Z_2 - 1| > \varepsilon) \leq C_5 x_n^\alpha$. On the other hand, there exists $\eta = \eta(\pi, \varepsilon) > 0$ such that if $|Z_i - 1| \leq \eta$ for $i = 1, 2$ then

$$\left|\frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1\right| \leq \varepsilon.$$

Finally, we have

$$\begin{aligned}
P\left(\left|\frac{\pi_1 Z_1}{\pi_1 Z_1 + \pi_2 Z_2} - \pi_1\right| > \varepsilon\right) &\leq \mathbf{P}(\{|Z_1 - 1| > \eta\} \cup \{|Z_2 - 1| > \eta\}) \\
&\leq \mathbf{P}(|Z_1 - 1| > \eta) + \mathbf{P}(|Z_2 - 1| > \eta) \\
&\leq C x_n^\alpha,
\end{aligned}$$

where $C = C(\pi, \varepsilon, \alpha)$. Then we can achieve equation (6.4) by modifying the proof of Proposition 1. \square

Lemma 14 *When $1 - 6\pi_1\pi_2 < 0$, for any $0 < s \leq \pi_2$ we have*

$$g(s) < s.$$

Proof of Theorem 1.3. Similar to the proof of Theorem 1.1, we will analyze R and S in equation (3.5) respectively, under the condition that $1 - 6\pi_1\pi_2 < 0$. Taking $D_1 = C_R^2(6\pi_1\pi_2^2)^2/(6\pi_1\pi_2 - 1)^2$, if

$d > D_1$ which implies $|\theta| \leq d^{-1/2} \leq D_1^{-1/2}$, by equation (3.6) and the inequality that $|z_n/x_n - \pi_1| \leq 1$, we obtain

$$C_R \frac{d(d-1)}{2} |\theta|^5 \left| \frac{z_n}{x_n} - \pi_1 \right| x_n^2 \leq \frac{1}{6} \frac{6\pi_1\pi_2 - 1}{\pi_1\pi_2^2} \frac{d(d-1)}{2} \theta^4 x_n^2. \quad (6.7)$$

Moreover, according to Lemma 13, there exist $D_2 = D_2(\pi) > D_1$ and $\delta = \delta(\pi) > 0$ independent of d , such that if $d > D_2$ and $x_n < \delta$ then an analogue of equation (5.2) holds as

$$|S| \leq \frac{1}{4} \frac{6\pi_1\pi_2 - 1}{\pi_1\pi_2^2} \frac{d(d-1)}{2} \theta^4 x_n^2, \quad (6.8)$$

and then by equation (6.7) we have

$$|R| \leq \frac{1}{4} \frac{6\pi_1\pi_2 - 1}{\pi_1\pi_2^2} \frac{d(d-1)}{2} \theta^4 x_n^2. \quad (6.9)$$

Next we claim that there is a positive integer m such that $x_m < \delta$. Define $\varepsilon = \varepsilon(\pi, \delta) = \varepsilon(\pi) = \frac{1}{2} \min_{s \geq \delta} (s - g(s))$. Since the function $s - g(s)$ is continuous and positive on $[\delta, \pi_2]$ by Lemma 14, we have $\varepsilon > 0$. Then by Lemma 10, there exists a $D = D(\pi, \varepsilon) = D(\pi) > D_2 > 0$, such that when $d > D$,

$$|x_{n+1} - g(d\theta^2 x_n)| < \varepsilon,$$

thus if $x_n \geq \delta$ then

$$x_{n+1} < g(d\theta^2 x_n) + \varepsilon \leq g(x_n) + \varepsilon \leq x_n - 2\varepsilon + \varepsilon \leq x_n - \varepsilon,$$

where the second inequality follows from Lemma 11 which claims that $g(s)$ is increasing on $[0, \pi_2]$. Therefore, there must exist an $m \in \mathbb{Z}^+$, such that $x_m < \delta$, as desired.

When $d > D$, it can be shown by induction, equation (6.8) and equation (6.9) that when $n \geq m$,

$$x_{n+1} \leq d\theta^2 x_n - \frac{1}{2} \frac{(6\pi_1\pi_2 - 1)}{\pi_1\pi_2^2} \frac{d(d-1)}{2} \theta^4 x_n^2 \leq d\theta^2 x_n \leq x_n < \delta. \quad (6.10)$$

Therefore, the limit L defined as $L = \lim_{n \rightarrow \infty} x_n \geq 0$ does exist, since the sequence $\{x_n\}_{n \geq m}$ is bounded and decreasing. Thus, passing to the limit on both sides of equation (6.10) gives

$$L \leq d\theta^2 L - \frac{1}{2} \frac{(6\pi_1\pi_2 - 1)}{\pi_1\pi_2^2} \frac{d(d-1)}{2} \theta^4 L^2,$$

which implies $L = 0$, hence non-reconstruction. □

Acknowledgement

We give special thanks to the journal editor and two anonymous reviewers who provided us with many constructive and helpful comments. We also give special thanks to Sebastien Roch for his inspiring discussions and reading of some proofs in the first version of this paper. We truly appreciate the warm encouragements on finishing this complete version and helpful discussions on this topic as well as future extensions, from colleagues at the 2017 & 2018 Columbia-Princeton Probability Day, 2017 Northeast Probability Seminar, 2018 Frontier Probability Days, 2017 & 2018 Finger Lakes Probability Seminar, and 2017 & 2018 Seminar on Stochastic Processes.

References

- Noam Berger, Claire Kenyon, Elchanan Mossel, and Yuval Peres. Glauber dynamics on trees and hyperbolic graphs. *Probability Theory and Related Fields*, 131(3):311–340, 2005.
- Jacques Bernussou and Jean-Louis Abatut. *Point mapping stability*. Pergamon, 1977.
- Shankar Bhamidi, Ram Rajagopal, and Sébastien Roch. Network delay inference from additive metrics. *Random Structures & Algorithms*, 37(2):176–203, 2010.
- Pavel M Bleher, Jean Ruiz, and Valentin A Zagrebnov. On the purity of the limiting Gibbs state for the Ising model on the bethe lattice. *Journal of Statistical Physics*, 79(1-2):473–482, 1995.
- Christian Borgs, Jennifer Chayes, Elchanan Mossel, and Sébastien Roch. The Kesten-Stigum reconstruction bound is tight for roughly symmetric binary channels. In *Foundations of Computer Science, 2006. FOCS'06. 47th Annual IEEE Symposium on*, pages 518–530. IEEE, 2006.
- JT Chayes, L Chayes, James P Sethna, and DJ Thouless. A mean field spin glass with short-range interactions. *Communications in Mathematical Physics*, 106(1):41–89, 1986.
- 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. ACM, 2006.
- William Evans, Claire Kenyon, Yuval Peres, and Leonard J Schulman. Broadcasting on trees and the Ising model. *Annals of Applied Probability*, pages 410–433, 2000.
- Hans-Otto Georgii. *Gibbs measures and phase transitions*, volume 9. Walter de Gruyter, 2011.
- Harry Kesten and Bernt P Stigum. Additional limit theorems for indecomposable multidimensional Galton-Watson processes. *The Annals of Mathematical Statistics*, 37(6):1463–1481, 1966.
- Harry Kesten and Bernt P Stigum. Limit theorems for decomposable multi-dimensional Galton-Watson processes. *Journal of Mathematical Analysis and Applications*, 17(2):309–338, 1967.
- Wenjian Liu, Sreenivasa Rao Jammalamadaka, and Ning Ning. The tightness of the Kesten-Stigum reconstruction bound of symmetric model with multiple mutations. *Journal of Statistical Physics*, 170(3):617–641, 2018.
- Fabio Martinelli, Alistair Sinclair, and Dror Weitz. Fast mixing for independent sets, colorings, and other models on trees. *Random Structures & Algorithms*, 31(2):134–172, 2007.
- Marc Mézard and Andrea Montanari. Reconstruction on trees and spin glass transition. *Journal of statistical physics*, 124(6):1317–1350, 2006.
- Elchanan Mossel. Reconstruction on trees: beating the second eigenvalue. *Annals of Applied Probability*, pages 285–300, 2001.

-
- Elchanan Mossel. Phase transitions in phylogeny. *Transactions of the American Mathematical Society*, 356(6):2379–2404, 2004a.
- Elchanan Mossel. Survey-information flow on trees. *DIMACS series in discrete mathematics and theoretical computer science*, 63:155–170, 2004b.
- Elchanan Mossel, Joe Neeman, and Allan Sly. Stochastic block models and reconstruction. *arXiv preprint arXiv:1202.1499*, 2012.
- Elchanan Mossel, Joe Neeman, and Allan Sly. A proof of the block model threshold conjecture. *Combinatorica*, pages 1–44, 2013.
- Joe Neeman and Praneeth Netrapalli. Non-reconstructability in the stochastic block model. *arXiv preprint arXiv:1404.6304*, 2014.
- Sebastien Roch. A short proof that phylogenetic tree reconstruction by maximum likelihood is hard. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 3(1):92–94, 2006.
- Allan Sly. Reconstruction for the Potts model. *The Annals of Probability*, 39:1365–1406, 2011.
- Prasad Tetali, Juan C Vera, Eric Vigoda, and Linji Yang. Phase transition for the mixing time of the Glauber dynamics for coloring regular trees. *Ann. Appl. Probab.*, 22:2210–2239., 2012.