J Stat Phys (2018) 170:617–641
https://doi.org/10.1007/s10955-017-1937-1

# The Tightness of the Kesten–Stigum Reconstruction Bound of Symmetric Model with Multiple Mutations

**Wenjian Liu**[1] · **Sreenivasa Rao Jammalamadaka**[2] · **Ning Ning**[2]

Received: 26 January 2017 / Accepted: 30 November 2017 / Published online: 7 December 2017
© Springer Science+Business Media, LLC, part of Springer Nature 2017

**Abstract** It is well known that reconstruction problems, as the interdisciplinary subject, have been studied in numerous contexts including statistical physics, information theory and computational biology, to name a few. We consider 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. We construct a nonlinear second-order dynamical system based on this model and show that the Kesten–Stigum reconstruction bound is not tight when $q \geq 4$.

**Keywords** Kesten–Stigum reconstruction bound · Markov random fields on trees · Distributional recursion · Dynamical system

**Mathematics Subject Classification** 60K35 · 82B26 · 82B20

## 1 Introduction

### 1.1 Preliminaries

We start with the following broadcasting process that stands as a discrete, irreducible, aperiodic, and reversible Markov chain. Let $\mathbb{T} = (\mathbb{V}, \mathbb{E}, \rho)$ be a tree with nodes $\mathbb{V}$, edges $\mathbb{E}$ and

✉ Wenjian Liu
wjliu@qcc.cuny.edu

Sreenivasa Rao Jammalamadaka
rao@pstat.ucsb.edu

Ning Ning
ning@pstat.ucsb.edu

1 Department of Mathematics and Computer Science, Queensborough Community College, The City University of New York, New York City, NY, USA

2 Department of Statistics and Applied Probability, University of California, Santa Barbara, CA, USA

root $\rho \in \mathbb{V}$. Each edge of the tree acts as a channel on a finite characters set $\mathcal{C}$, whose elements are configurations on $\mathbb{T}$, denoted by $\sigma$. We set a probability transition matrix $\mathbf{M} = (M_{ij})$ as the noisy communication channel on each edge. The state of the root $\rho$, denoted by $\sigma_\rho$, is chosen according to an initial distribution $\pi$ on $\mathcal{C}$, and then propagated in the tree as follows: for each vertex $v$ having $u$ as its parent , the spin at $v$ is defined according to the probabilities

$$\mathbf{P}(\sigma_v = j \mid \sigma_u = i) = M_{ij}$$

with $i, j \in \mathcal{C}$. Roughly speaking, reconstruction is to answer the question that considering all the symbols received at the vertices of the $n$th generation, does this configuration contain non-vanishing information transmitted by the root, as $n$ goes to $\infty$?

In this paper, we will restrict our attention to $d$-ary trees, i.e. the infinite rooted tree where every vertex has exactly $d$ offspring (every vertex has degree $d+1$ except the root which has degree $d$). Let $\sigma(n)$ denote the spins at distance $n$ from the root and let $\sigma^i(n)$ denote $\sigma(n)$ conditioned on $\sigma_\rho = i$. Consider a characters set $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$, consisting of two categories $\mathcal{C}_1 = \{1, \ldots, q\}$ and $\mathcal{C}_2 = \{q+1, \ldots, 2q\}$ with $q \geq 2$, and the state of the root $\rho$ is chosen according to the uniform distribution on $\mathcal{C}$. Moreover, a $2q \times 2q$ probability transition matrix $\mathbf{M} = (M_{ij})_{2q\times 2q}$ is defined as follows:

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

where $p_0$, $p_1$ and $p_2$ are all nonnegative, such that $p_0 + (q-1)p_1 + qp_2 = 1$. It can be verified that the eigenvalues of $\mathbf{M}$ are $\lambda_1 = p_0 - p_1$, $\lambda_2 = p_0 + (q-1)p_1 - qp_2$, and $\lambda_3 = p_0 + (q-1)p_1 + qp_2 = 1$. Therefore, there are two candidates $\lambda_1$ and $\lambda_2$ for the second largest eigenvalue in absolute value, say, $\lambda$, which plays a crucial role in the reconstruction problem.

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

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

where $d_{TV}$ is the total variation distance. When the lim sup is 0, we say the model has *non-reconstruction* on $\mathbb{T}$.

### 1.2 Background

Beyond the basic interest in determining the reconstruction threshold of a Markov random field in probability, this problem is relevant to statistical physics, biology (Daskalakis et al. [12], Mossel [33]), and information theory (Bhamidi et. al. [5], Evans et al. [15]), where one is interested in computing the information capacity of the tree network. Most closely related to the origins of this work, for spin systems in statistical physics, the threshold for reconstruction is equivalent to the threshold for extremality of the infinite-volume Gibbs measure induced by free-boundary conditions, see Georgh [17]. The reconstruction threshold also has an important effect in the efficiency of the Glauber dynamics on trees and random graphs. It is well known that when the model is reconstructible, the mixing time for the Glauber dynamics on trees is $n^{1+\Omega(1)}$, while it is slower than at higher temperature when the mixing time is $O(n \log n)$. The corresponding bound is tight for the Ising model, namely, the mixing time is $O(n \log n)$ when $d\lambda^2 < 1$. In Martinelli et al. [30], this result is extended to the log Sobolev constant and it is also shown that for measures on trees, a super-linear decay of point-to-set correlations implies an $\Omega(1)$ spectral gap for the Glauber dynamics with free boundary conditions. A similar transition takes place in the colouring model as shown in Tetali et al. [46]. Sharp bounds of this type are not known for the hardcore model, although it is conjectured that the Glauber dynamics should again be $O(n \log n)$ in the non-reconstruction regime.

For any channel $\mathbf{M}$, it is well known that the reconstruction problem is connected closely to $\lambda$, the second largest eigenvalue in absolute value of $\mathbf{M}$. An important general bound was obtained by Kesten and Stigum [21,22]: the reconstruction problem is solvable if $d|\lambda|^2 > 1$ ($\lambda$ may be a complex number), which is known as the Kesten–Stigum bound. On the other hand, for larger noise ($d|\lambda|^2 < 1$) one may wonder whether reconstruction is possible, by exploiting the whole set of symbols received at the $n$th generation, through a clever use of the correlations between the symbols received on the leaves. The answer depends on the channel.

For the binary symmetric channel, it was shown in Bleher et al. [7] that the reconstruction problem is solvable if and only if $d\lambda^2 > 1$. For all other channels, however, it would be a little challenging to prove the non-reconstructibility. Mossel [32,34] showed that the Kesten–Stigum bound is not the bound for reconstruction in the binary asymmetric model with sufficiently large asymmetry or in the Potts model with sufficiently many characters, which sheds the light on exploring the tightness of the Kesten–Stigum bound. The first exact reconstruction threshold in roughly a decade, was obtained by Borgs et al. [8], in which the authors displayed a delicate analysis of the moment recursion on a weighted version of the magnetization, and thus achieved a breakthrough result.

A particularly important example is provided by $q$-state symmetric channels, i.e. Potts models in the terminology of statistical mechanics, with the transition matrix

$$\mathbf{M} = \begin{pmatrix} p_0 & p_1 & \cdots & p_1 \\ p_1 & p_0 & \cdots & p_1 \\ \vdots & \vdots & \ddots & \vdots \\ p_1 & p_1 & \cdots & p_0 \end{pmatrix}$$

and $\lambda = p_0 - p_1$. This model was completely investigated by Sly [43] by means of the recursive structure of the tree, and more importantly, Sly showed that non-reconstruction is equivalent to $\lim_{n\to\infty} x_n = 0$, where $x_n = \mathbf{EP}(\sigma_\rho = 1 \mid \sigma^1(n)) - \frac{1}{q}$. Thus the key idea was to analyze the recursion relationship between $x_n$ and $x_{n+1}$. This work then went on to engage the refined recursive equations of vector-valued distributions and concentration analyses, to confirm much of the picture conjectured earlier by Mézard and Montanari [31].

Inspired by the popular K80 model proposed by Kimura [23], which distinguishes between transitions and transversions, we analyze the case that transition matrix has two mutation classes with $q$ states in each class. Improved flexibility comes along with increased complexity, which is mainly due to the fact that the additional class of mutation complicates the discussion of the second largest eigenvalue in absolute value. However, by introducing additional auxiliary quantities $y_n$ and $z_n$ besides $x_n$ defined in Sect. 2.1, we succeed in investigating the tightness of the Kesten–Stigum bound.

### 1.3 Applications

The reconstruction problem arises naturally in many fields including statistical physics, where the Ising model and the Potts model are popular and have been studied extensively from different angles, see [2,10,13,14,16,19,20,27–29,36,38,40,41,44,45,47,48]. In this article, we focus on the reconstruction threshold on trees, which plays an important role in the dynamic phase transitions in certain glassy systems subject to random constraints. For random colorings on the Erdös–Rényi random graph with average connectivity $d$, Achlioptas and Coja-Oghlan [1] proved that there is a phase transition, from the situation that most of the mass is contained in one giant component, to the case that the space of solutions breaks into exponentially many smaller clusters. This phase transition has been proved corresponding to known bounds on the reconstruction threshold for proper colorings on trees, see e.g. Mossel and Peres [35], Semerjian [39] and Sly [42].

In computational biology, the broadcast model is the main model for the evolution of base pairs of DNA. Phylogenetic reconstruction is a major task of systematic biology, which is to construct the ancestry tree of a collection of species, given the information of present species. The corresponding reconstruction threshold answers the question whether the ancestral DNA information can be reconstructed from a known phylogenetic tree. This threshold is also crucial to determine the number of samples required, in the sense that, only enumerations of each type of spin at the leaves are collected, regardless of their positions on the leaves. Interested readers on Phylogenetic tree reconstruction are referred to Roch [37] and Daskalakis et al. [12].

The popular K80 model [23], has some obvious advantages over other models in Phylogeny reconstruction, which is favored by both Akaike Information Criterion and Bayesian Information Criterion (see Sect. 2.2.2 in Cadotte and Davies [9]). The K80 model distinguishes between transitions ($A \leftrightarrow G$, i.e. from purine to purine, or $C \leftrightarrow T$, i.e. from pyrimidine to pyrimidine) and transversions (from purine to pyrimidine or vice versa). Inspired by this and related literatures, we analyze the case that the transition matrix has two mutation classes and $q$ states in each class. We believe that the $q$-state symmetric Potts model as a generalization of 2-state symmetric Ising model, cannot fully represent the spirit of the classical 2-state symmetric Ising model in terms of dichotomy, and this is one of the areas this work can contribute to.

A tree is a connected undirected graph with no simple circuits. In other words, an undirected graph is a tree if and only if there is a unique simple path between any two of its vertices. The theory that the reconstruction threshold on trees corresponds to the reconstruction threshold on locally treelike graphs, is verified in Gerschenfeld and Montanari [18]. The strong and increasing interest in the study of the properties of social networks, is a result of the rapid and global emergence of online social networks and their meteoric adoption by millions of Internet users. When it comes to Socio–psychological mechanisms of generation and dissemination of network, our model's advantage in providing more flexibility to mimic psychological behaviors is obvious. For example, our model and the construction threshold can be used to effectively identify community effect in social networks and customer loyalty in marketing research, especially for different firms or organizations who want to promote their products or philosophies. In this sense, many possible extensions can be made on research on graph structures with psychological factors involved, such as the work by Liu, Ying and Shakkottai [24] on analyzing the formation and propagation of opinions across networks by an Ising model based approach, the work by Bisconti et al. [6] on reconstruction of a real world social network using the Potts model and Loopy Belief Propagation, etc.

### 1.4 Main Results and Proof Sketch

Because non-reconstruction happens at most $d|\lambda|^2 = 1$, without loss of generality, it would be convenient to presume $1/2 \le d|\lambda|^2 \le 1$ in the following context.

**Main Theorem** *Assume* $0 < |\lambda_2| \le |\lambda_1|$. *When* $q \ge 4$, *for every* $d$ *the Kesten–Stigum bound is not tight, i.e. the reconstruction is solvable for some* $\lambda_1$ *even if* $d\lambda_1^2 < 1$.

The ideas and techniques used to prove the Main Theorem can be seen as the following. One standard to classify reconstruction and nonreconstruction 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 randomly which is $\frac{1}{2q}$ in this case. Nonreconstruction means that the mutual information between the root and the spins at distance $n$ goes to 0 as $n$ tends to infinity. It can be established that the nonreconstruction is equivalent to

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

Our analysis is similar to Borgs et al. [8], Chayes et al. [11] in the context of spin–glasses, and Sly [43]. However, the two classes of mutation complicates the discussion of $\lambda$, the second largest eigenvalue in absolute value of the transition matrix, which makes the problem much more challenging. In this case, it is necessary to consider the corresponding quantities similar to $x_n$, viz. wrong guess but right group $y_n$, and wrong guess and even wrong group $z_n$. In Sect. 2.2, we investigate the properties and relations between $x_n$, $y_n$ and $z_n$. By these preliminary results, we focus on the analysis of $x_n$ and $z_n$ in the sequel.

In order to research the reconstruction, according to the Markov random field property, we establish the distributional recursion and moment recursion, by analyzing the recursive relation between the $n$th and the $(n+1)$th generations' structure of the tree. Furthermore, we display that the interactions between spins become very weak, if they are sufficiently far away from each other. Therefore, we can obtain a nonlinear dynamical system. If $x_n$ is small, we are able to develop the concentration analysis and achieve the approximation to the dynamical system:

$$\begin{cases} x_{n+1} \approx d\lambda_1^2 x_n + (d\lambda_1^2 - d\lambda_2^2)z_n + \frac{d(d-1)}{2}\left(\frac{q(2q-5)}{q-1}\lambda_1^4(x_n+z_n)^2 - 4q\lambda_1^2\lambda_2^2(x_n+z_n)z_n \right. \\ \qquad\qquad \left. -4q\lambda_2^4 z_n^2\right) \\ z_{n+1} \approx d\lambda_2^2 z_n - \frac{d(d-1)}{2}\left(\frac{q}{q-1}\lambda_1^4(x_n+z_n)^2 - 4q\lambda_2^4 z_n^2\right). \end{cases}$$

Finally, we investigate the stability of the system and then establish the threshold of $q$ relevant to the reconstruction. When $q \ge 4$, even if $d\lambda_1^2 < 1$ for some $\lambda_1$, $x_n$ will not converge to 0 and hence there is reconstruction beyond the Kesten–Stigum bound. More detailed definitions and interpretations can be seen in the next section.

## 2 Second Order Recursion Relation

### 2.1 Notations

Let $u_1, \ldots, u_d$ be the children of $\rho$ and $\mathbb{T}_v$ be the subtree of descendants of $v \in \mathbb{T}$. Furthermore, if we set $d(\cdot,\cdot)$ as the graph-metric distance on $\mathbb{T}$, denote the $n$th level of the tree by $L_n = \{v \in \mathbb{V} : d(\rho, v) = n\}$ and then let $\sigma_j(n)$ be the spins on $L_n \cap \mathbb{T}_{u_j}$. For a configuration $A$ on $L_n$, define the posterior function

$$f_n(i, A) = \mathbf{P}(\sigma_\rho = i \mid \sigma(n) = A). \tag{2.1}$$

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

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

Now for any $1 \le i \le 2q$, define a collection of random variables

$$X_i(n) = f_n(i, \sigma(n))$$

to describe the posterior probability of state $i$ at the root given the random configuration $\sigma(n)$ of the leaves, and analogously,

$$X^{(1)}(n) = f_n(1, \sigma^1(n)), \quad X^{(2)}(n) = f_n(2, \sigma^1(n)), \quad X^{(3)}(n) = f_n(q+1, \sigma^1(n)).$$

By symmetry, the collections $\{f_n(i, \sigma^1(n)) : 2 \le i \le q\}$ and $\{f_n(i, \sigma^1(n)) : q+1 \le i \le 2q\}$ are exchangeable respectively; in addition, $f_n(j, \sigma^i(n))$ is distributed as

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

Finally, denote the first and second central moments of $X^{(1)}(n)$, $X^{(2)}(n)$ and $X^{(3)}(n)$, which would be the principal quantities in our analysis, as

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

and

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

### 2.2 Preliminaries

For any $i = 1, \cdots, 2q$ and nonnegative $n \in \mathbb{Z}$, it is concluded from the symmetric property of the tree that

$$\mathbf{E}X_i(n) = \frac{1}{2q}$$

is always true.

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

$$x_n = \mathbf{E}\sum_{i=1}^{2q}\left(X_i(n) - \frac{1}{2q}\right)^2 \ge 0, \quad z_n \le 0, \quad \text{and} \quad x_n + z_n \ge 0.$$

*Proof* First, by Bayes' rule, we have

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

and

$$0 \le \mathbf{E}\sum_{i=1}^{2q}\left(X_i(n) - \frac{1}{2q}\right)^2 = \sum_{i=1}^{2q}\mathbf{E}X_i^2(n) - \frac{2}{2q}\sum_{i=1}^{2q}\mathbf{E}X_i(n) + \frac{1}{2q} = x_n. \qquad (2.2)$$

Next, we consider the covariance matrix of random variables $\left\{X_i(n) - \frac{1}{2q}\right\}_1^{2q}$ and express covariances in terms of $x_n$, $y_n$ and $z_n$. Similarly, we obtain

$$y_n + \frac{1}{2q} = 2q \sum_A \mathbf{P}(\sigma(n) = A) f_n(1, A) f_n(2, A) = 2q\mathbf{E}X_1(n)X_2(n),$$

so for any $i_1 < i_2$ in the same category, it is concluded from the symmetric property of the tree that

$$\mathbf{E}\left(X_{i_1}(n) - \frac{1}{2q}\right)\left(X_{i_2}(n) - \frac{1}{2q}\right) = \mathbf{E}\left(X_1 - \frac{1}{2q}\right)\left(X_2 - \frac{1}{2q}\right) = \frac{y_n}{2q}.$$

Similarly, if $i_1$ and $i_2$ are from different categories, we have

$$\mathbf{E}\left(X_{i_1}(n) - \frac{1}{2q}\right)\left(X_{i_2}(n) - \frac{1}{2q}\right) = \frac{z_n}{2q}.$$

Therefore, the covariance matrix is given by

$$\Sigma_X(n) = \begin{pmatrix} \frac{x_n}{2q} & \frac{y_n}{2q} & \cdots & \frac{y_n}{2q} & \frac{z_n}{2q} & \frac{z_n}{2q} & \cdots & \frac{z_n}{2q} \\ \frac{y_n}{2q} & \frac{x_n}{2q} & \cdots & \frac{y_n}{2q} & \frac{z_n}{2q} & \frac{z_n}{2q} & \cdots & \frac{z_n}{2q} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{y_n}{2q} & \frac{y_n}{2q} & \cdots & \frac{x_n}{2q} & \frac{z_n}{2q} & \frac{z_n}{2q} & \cdots & \frac{z_n}{2q} \\ \frac{z_n}{2q} & \frac{z_n}{2q} & \cdots & \frac{z_n}{2q} & \frac{x_n}{2q} & \frac{y_n}{2q} & \cdots & \frac{y_n}{2q} \\ \frac{z_n}{2q} & \frac{z_n}{2q} & \cdots & \frac{z_n}{2q} & \frac{y_n}{2q} & \frac{x_n}{2q} & \cdots & \frac{y_n}{2q} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{z_n}{2q} & \frac{z_n}{2q} & \cdots & \frac{z_n}{2q} & \frac{y_n}{2q} & \frac{y_n}{2q} & \cdots & \frac{x_n}{2q} \end{pmatrix}_{2q \times 2q}$$

whose eigenvalues are 0, $\frac{x_n+(q-1)y_n-qz_n}{2q}$ and $\frac{x_n-y_n}{2q}$. It is well known that the covariance matrix of a multivariate probability distribution is always positive semi-definite, which implies that all eigenvalues are nonnegative, say, $x_n + (q-1)y_n - qz_n \geq 0$ and $x_n - y_n \geq 0$. It suffices to complete the proof, by these results and the fact of $x_n + (q-1)y_n + qz_n = 0$.

**Lemma 2** *For any $n \in \mathbb{N} \cup \{0\}$, the following hold:*

(i) $x_n = u_n + (q-1)v_n + qw_n$;
(ii) $\mathbf{E}\left(X^{(1)}(n) - \frac{1}{2q}\right)\left(X^{(2)}(n) - \frac{1}{2q}\right) = v_n + \frac{y_n - x_n}{2q}$;
(iii) $\mathbf{E}\left(X^{(1)}(n) - \frac{1}{2q}\right)\left(X^{(3)}(n) - \frac{1}{2q}\right) = w_n + \frac{z_n - x_n}{2q}$;
(iv) $\mathbf{E}\left(X^{(2)}(n) - \frac{1}{2q}\right)\left(X^{(3)}(n) - \frac{1}{2q}\right) = -\frac{w_n}{q-1} - \frac{z_n}{2q(q-1)} - \frac{y_n}{2q}$;
(v) $\mathbf{E}\left(f_n(q+1, \sigma^1(n)) - \frac{1}{2q}\right)\left(f_n(2q, \sigma^1(n)) - \frac{1}{2q}\right) = -\frac{w_n}{q-1} - \frac{z_n}{2(q-1)}$;
(vi) $\mathbf{E}\left(f_n(2, \sigma^1(n)) - \frac{1}{2q}\right)\left(f_n(q, \sigma^1(n)) - \frac{1}{2q}\right) = -\frac{2v_n}{q-2} - \frac{z_n}{2(q-1)} + \frac{qw_n}{(q-1)(q-2)}$.

*Proof* By the total probability formula and using Lemma 1, we can prove (i) as follows:

$$\begin{aligned}
x_n &= \mathbf{E}\sum_{i=1}^{2q}\left(X_i(n)-\frac{1}{2q}\right)^2\\
&= \sum_{j=1}^{2q}\mathbf{E}\left(\sum_{i=1}^{2q}\left(X_i(n)-\frac{1}{2q}\right)^2 \mid \sigma_\rho = j\right)\mathbf{P}(\sigma_\rho = j)\\
&= \sum_{j=1}^{2q}\frac{1}{2q}\left[\mathbf{E}\left(X^{(1)}(n)-\frac{1}{2q}\right)^2 + (q-1)\mathbf{E}\left(X^{(2)}(n)-\frac{1}{2q}\right)^2\right.\\
&\quad \left.+\, q\mathbf{E}\left(X^{(3)}(n)-\frac{1}{2q}\right)^2\right]\\
&= \mathbf{E}\left(X^{(1)}(n)-\frac{1}{2q}\right)^2 + (q-1)\mathbf{E}\left(X^{(2)}(n)-\frac{1}{2q}\right)^2\\
&\quad +q\mathbf{E}\left(X^{(3)}(n)-\frac{1}{2q}\right)^2\\
&= u_n + (q-1)v_n + qw_n.
\end{aligned}$$

Applying the same technique, we obtain

$$\begin{aligned}
\mathbf{E}X^{(1)}(n)X^{(2)}(n) &= \sum_A \mathbf{P}(\sigma_\rho = 1 \mid \sigma(n) = A)\mathbf{P}(\sigma_\rho = 2 \mid \sigma(n) = A)\mathbf{P}\big(\sigma(n)\\
&= A \mid \sigma_\rho = 1\big)\\
&= \sum_A [\mathbf{P}(\sigma_\rho = 1 \mid \sigma(n) = A)]^2\mathbf{P}(\sigma(n) = A \mid \sigma_\rho = 2)\\
&= \mathbf{E}\left(X^{(2)}(n)\right)^2
\end{aligned}$$

and hence (ii) follows:

$$\mathbf{E}\left(X^{(1)}(n)-\frac{1}{2q}\right)\left(X^{(2)}(n)-\frac{1}{2q}\right) = \mathbf{E}\left(X^{(2)}-\frac{1}{2q}\right)^2 + \frac{y_n-x_n}{2q} = v_n + \frac{y_n-x_n}{2q}.$$

Similarly, (iii) turns out to be true due to

$$\mathbf{E}X^{(1)}(n)X^{(3)}(n) = \mathbf{E}\left(X^{(3)}(n)\right)^2.$$

The statement (iv), (v) and (vi) can be handled in the same way, using the symmetry,

$$\begin{aligned}
\mathbf{E}X^{(2)}(n)X^{(3)}(n) &= \sum_A \mathbf{P}(\sigma_\rho = 2 \mid \sigma(n) = A)\mathbf{P}(\sigma_\rho = q+1 \mid \sigma(n) = A)\mathbf{P}\,\big(\sigma(n)\\
&= A \mid \sigma_\rho = 1\big)\\
&= \sum_A \mathbf{P}(\sigma_\rho = 1 \mid \sigma(n) = A)\mathbf{P}(\sigma_\rho = 2 \mid \sigma(n) = A)\mathbf{P}\,\big(\sigma(n)\\
&= A \mid \sigma_\rho = q+1\big)\\
&= \sum_A \mathbf{P}(\sigma_\rho = q+1 \mid \sigma(n) = A)\mathbf{P}(\sigma_\rho = 2q \mid \sigma(n) = A)\mathbf{P}\,\big(\sigma(n)
\end{aligned}$$

$$= A \mid \sigma_\rho = 1\big)$$
$$= \mathbf{E} f_n(q+1, \sigma^1(n)) f_n(2q, \sigma^1(n)). \tag{2.3}$$

To obtain $\mathbf{E}X^{(2)}(n)X^{(3)}(n)$, recall that

$$\begin{aligned} z_n + \frac{1}{2q} &= \mathbf{E}X^{(3)}(n) \\ &= \mathbf{E} f_n(q+1, \sigma^1(n)) \sum_{i=1}^{2q} f_n(i, \sigma^1(n)) \\ &= \mathbf{E}X^{(1)}(n)X^{(3)}(n) + (q-1)\mathbf{E}X^{(2)}(n)X^{(3)}(n) \\ &\quad + \mathbf{E}(X^{(3)})^2 + (q-1)\mathbf{E} f_n(q+1, \sigma^1(n)) f_n(2q, \sigma^1(n)) \\ &= \mathbf{E}X^{(1)}(n)X^{(3)}(n) + 2(q-1)\mathbf{E}X^{(2)}(n)X^{(3)}(n) \\ &\quad + \mathbf{E}(X^{(3)})^2, \end{aligned}$$

which implies that

$$\mathbf{E}\left(X^{(2)}(n) - \frac{1}{2q}\right)\left(X^{(3)}(n) - \frac{1}{2q}\right) = -\frac{w_n}{q-1} - \frac{z_n}{2q(q-1)} - \frac{y_n}{2q}. \tag{2.4}$$

Thus, (2.3) together with (2.4) gives

$$\mathbf{E}\left(f_n(q+1, \sigma^1(n)) - \frac{1}{2q}\right)\left(f_n(2q, \sigma^1(n)) - \frac{1}{2q}\right) = -\frac{w_n}{q-1} - \frac{z_n}{2(q-1)}.$$

As the preceding discussion, consider

$$\begin{aligned} y_n + \frac{1}{2q} &= \mathbf{E} f_n(2, \sigma^1(n)) \sum_{i=1}^{2q} f(i, \sigma^1(n)) \\ &= 2\mathbf{E}(X^{(2)}(n))^2 + (q-2)\mathbf{E} f_n(2, \sigma^1(n)) f_n(q, \sigma^1(n)) \\ &\quad + q\mathbf{E}X^{(2)}(n)X^{(3)}(n), \end{aligned}$$

and thus

$$\begin{aligned} &\mathbf{E}\left(f_n(2, \sigma^1(n)) - \frac{1}{2q}\right)\left(f_n(q, \sigma^1(n)) - \frac{1}{2q}\right) \\ &= \frac{1}{q-2}\left(y_n + \frac{1}{2q} - 2\mathbf{E}(X^{(2)}(n))^2 - q\mathbf{E}X^{(2)}(n)X^{(3)}(n)\right) \\ &\quad - \frac{2}{2q}\left(y_n + \frac{1}{2q}\right) + \frac{1}{4q^2} \\ &= -\frac{2v_n}{q-2} - \frac{z_n}{2(q-1)} + \frac{q w_n}{(q-1)(q-2)}. \end{aligned}$$

### 2.3 Means and Covariances of $Y_{ij}$

Define

$$Y_{ij}(n) = f_n\left(i, \sigma^1_j(n+1)\right),$$

and it is apparent that the random vectors $(Y_{ij})_{i=1}^{2q}$ are independent, for $j = 1, \ldots, d$, by the symmetries of the model. The central moments of $Y_{ij}$ would play a key role in further analysis, and therefore it is necessary to figure them out in the first place. For each $1 \le j \le d$, we rely on the total probability formula to conclude:

(i) when $i = 1$,

$$\begin{aligned}\mathbf{E}\left(Y_{1j}(n) - \frac{1}{2q}\right) &= p_0\mathbf{E}\left(X^{(1)}(n) - \frac{1}{2q}\right)\\ &\quad + (q-1)p_1\mathbf{E}\left(X^{(2)}(n) - \frac{1}{2q}\right)\\ &\quad + qp_2\mathbf{E}\left(X^{(3)}(n) - \frac{1}{2q}\right)\\ &= \lambda_1 x_n + (\lambda_1 - \lambda_2)z_n;\end{aligned}$$

(ii) for $2 \le i \le q$,

$$\begin{aligned}\mathbf{E}\left(Y_{ij}(n) - \frac{1}{2q}\right) &= p_1\mathbf{E}\left(X^{(1)}(n) - \frac{1}{2q}\right)\\ &\quad + [p_0 + (q-2)p_1]\mathbf{E}\left(X^{(2)}(n) - \frac{1}{2q}\right)\\ &\quad + qp_2\mathbf{E}\left(X^{(3)}(n) - \frac{1}{2q}\right)\\ &= -\frac{\lambda_1}{q-1}x_n - \frac{\lambda_1 + (q-1)\lambda_2}{q-1}z_n;\end{aligned}$$

(iii) for $q+1 \le i \le 2q$, by means of the identity $\sum_{i=1}^{2q} Y_{ij}(n) \equiv 1$, it follows immediately that

$$\mathbf{E}\left(Y_{ij}(n) - \frac{1}{2q}\right) = -\frac{1}{q}\sum_{i=1}^{q}\mathbf{E}\left(Y_{ij}(n) - \frac{1}{2q}\right) = \lambda_2 z_n;$$

(iv) resembling the discussion of (i), (ii) and (iii), it is further concluded that when $i = 1$,

$$\mathbf{E}\left(Y_{1j}(n) - \frac{1}{2q}\right)^2 = \frac{1 + \lambda_2 - 2\lambda_1}{2q}x_n + \lambda_1 u_n + (\lambda_1 - \lambda_2)w_n;$$

(v) for $2 \le i \le q$,

$$\mathbf{E}\left(Y_{ij}(n) - \frac{1}{2q}\right)^2 = \left(\frac{1}{2q} + \frac{\lambda_2}{2q} + \frac{\lambda_1}{q(q-1)}\right)x_n - \frac{\lambda_1}{q-1}u_n - \frac{\lambda_1 + (q-1)\lambda_2}{q-1}w_n;$$

(vi) for $q+1 \le i \le 2q$,

$$\mathbf{E}\left(Y_{ij}(n) - \frac{1}{2q}\right)^2 = \frac{1 - \lambda_2}{2q}x_n + \lambda_2 w_n;$$

(vii) for $2 \le i \le q$,

$$\begin{aligned}\mathbf{E}\left(Y_{1j}(n) - \frac{1}{2q}\right)\left(Y_{ij}(n) - \frac{1}{2q}\right) &= \frac{(q+2)\lambda_1 - \lambda_2 - 1}{2q(q-1)}x_n - \frac{z_n}{2(q-1)}\\ &\quad - \frac{\lambda_1}{q-1}u_n - \frac{(q+1)\lambda_1 - \lambda_2}{q-1}w_n;\end{aligned}$$

(viii) for $q+1 \le i \le 2q$,

$$\mathbf{E}\left(Y_{1j}(n)-\frac{1}{2q}\right)\left(Y_{ij}(n)-\frac{1}{2q}\right)=-\frac{\lambda_1}{2q}x_n+\frac{z_n}{2q}+\lambda_1 w_n;$$

(ix) for $1 < i_1 < i_2 \le q$,

$$\begin{aligned}&\mathbf{E}\left(Y_{i_1 j}(n)-\frac{1}{2q}\right)\left(Y_{i_2 j}(n)-\frac{1}{2q}\right)\\&=\left[-\frac{2(q+2)\lambda_1+(q-2)\lambda_2}{2q(q-1)(q-2)}\right.\\&\quad\left.-\frac{1}{2q(q-1)}\right]x_n-\frac{z_n}{2(q-1)}\\&\quad+\frac{2\lambda_1}{(q-1)(q-2)}u_n+\frac{2(q+1)\lambda_1+(q-2)\lambda_2}{(q-1)(q-2)}w_n;\end{aligned}$$

(x) for $1 < i_1 \le q < i_2 \le 2q$,

$$\mathbf{E}\left(Y_{i_1 j}(n)-\frac{1}{2q}\right)\left(Y_{i_2 j}(n)-\frac{1}{2q}\right)=\frac{\lambda_1}{2q(q-1)}x_n+\frac{z_n}{2q}-\frac{\lambda_1}{q-1}w_n;$$

(xi) for $q+1 \le i_1 < i_2 \le 2q$,

$$\mathbf{E}\left(Y_{i_1 j}(n)-\frac{1}{2q}\right)\left(Y_{i_2 j}(n)-\frac{1}{2q}\right)=\frac{\lambda_2-1}{2q(q-1)}x_n-\frac{z_n}{2(q-1)}-\frac{\lambda_2}{q-1}w_n.$$

### 2.4 Distributional Recursion

The key method of this paper is to analyze the relation between $X^{(1)}(n)$, $X^{(3)}(n)$ and $X^{(1)}(n+1)$, $X^{(3)}(n+1)$ using the recursive structure of the tree. Take $A=\sigma^1(n+1)$ and then the following relations follow from the Markov random field property:

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

and

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

where

(A) for $1 \le i \le q$,

$$\begin{aligned}Z_i=Z_i(n)=\prod_{j=1}^{d}\Bigg[&1+2q(p_0-p_2)\left(Y_{ij}-\frac{1}{2q}\right)\\&+2q(p_1-p_2)\sum_{1\le \ell\ne i\le q}\left(Y_{\ell j}-\frac{1}{2q}\right)\Bigg]\\=\prod_{j=1}^{d}\Bigg[&1+2q(p_0-p_1)\left(Y_{ij}-\frac{1}{2q}\right)\end{aligned}$$

$$
\begin{aligned}
&\quad - 2q(p_1 - p_2) \sum_{q+1 \le \ell \le 2q} \left( Y_{\ell j} - \frac{1}{2q} \right) \Bigg] \\
&= \prod_{j=1}^{d} \Bigg[ 1 + 2q\lambda_1 \left( Y_{ij} - \frac{1}{2q} \right) \\
&\quad + 2(\lambda_1 - \lambda_2) \sum_{q+1 \le \ell \le 2q} \left( Y_{\ell j} - \frac{1}{2q} \right) \Bigg],
\end{aligned}
$$

(B) for $q + 1 \le i \le 2q$,

$$
\begin{aligned}
Z_i = Z_i(n) &= \prod_{j=1}^{d} \Bigg[ 1 + 2q(p_0 - p_2) \left( Y_{ij} - \frac{1}{2q} \right) \\
&\quad + 2q(p_1 - p_2) \sum_{q+1 \le \ell \ne i \le 2q} \left( Y_{\ell j} - \frac{1}{2q} \right) \Bigg] \\
&= \prod_{j=1}^{d} \Bigg[ 1 + 2q(p_0 - p_1) \left( Y_{ij} - \frac{1}{2q} \right) \\
&\quad - 2q(p_1 - p_2) \sum_{1 \le \ell \le q} \left( Y_{\ell j} - \frac{1}{2q} \right) \Bigg] \\
&= \prod_{j=1}^{d} \Bigg[ 1 + 2q\lambda_1 \left( Y_{ij} - \frac{1}{2q} \right) \\
&\quad + 2(\lambda_1 - \lambda_2) \sum_{1 \le \ell \le q} \left( Y_{\ell j} - \frac{1}{2q} \right) \Bigg].
\end{aligned}
$$

To continue the proof, it is necessary to firstly reveal some identities concerning $Z_i(n)$.

**Lemma 3** *For any nonnegative $n \in \mathbb{Z}$ and $1 \le i \le 2q$, we have*

$$\mathbf{E} Z_1(n) Z_i(n) = \mathbf{E} Z_i(n)^2,$$

*and given any $2 \le i_1 \le q < q + 1 \le i_2 \le 2q$, we have*

$$\mathbf{E} Z_{i_1}(n) Z_{i_2}(n) = \mathbf{E} Z_{q+1}(n) Z_{2q}(n).$$

*Proof* When $i = 1$, the result is trivial. If $2 \le i \le 2q$, for any configurations $A = (A_1, \ldots, A_d)$ on the $(n+1)$th level, where $A_j$ denote the spins on $L_{n+1} \cap \mathbb{T}_{u_j}$ ,we have

$$Z_i(A) = 2q \frac{\mathbf{P}(\sigma(n+1) = A)}{\prod_{j=1}^{d} \mathbf{P}(\sigma_j(n+1) = A_j)} \mathbf{P}(\sigma_\rho = i \mid \sigma(n+1) = A).$$

By the symmetry of the tree, we have

$$\begin{aligned}
\mathbf{E}Z_1 Z_i &= (2q)^2 \sum_A \left( \frac{\mathbf{P}(\sigma(n+1)=A)}{\prod_{j=1}^d \mathbf{P}(\sigma_j(n+1)=A_j)} \right)^2 \mathbf{P}(\sigma_\rho = 1 \mid \sigma(n+1)=A) \\
&\quad \times \mathbf{P}(\sigma_\rho = i \mid \sigma(n+1)=A)\mathbf{P}(\sigma(n+1)=A \mid \sigma_\rho = 1) \\
&= (2q)^2 \sum_A \left( \frac{\mathbf{P}(\sigma(n+1)=A)}{\prod_{j=1}^d \mathbf{P}(\sigma_j(n+1)=A_j)} \right)^2 \mathbf{P}^2 \left(\sigma_\rho = 1 \mid \sigma(n+1)\right. \\
&= \left. A\right) \times \mathbf{P}(\sigma(n+1)=A \mid \sigma_\rho = i) \\
&= (2q)^2 \sum_A \left( \frac{\mathbf{P}(\sigma(n+1)=A)}{\prod_{j=1}^d \mathbf{P}(\sigma_j(n+1)=A_j)} \right)^2 \mathbf{P}^2 \left(\sigma_\rho = i \mid \sigma(n+1)\right. \\
&= \left. A\right) \mathbf{P}(\sigma(n+1)=A \mid \sigma_\rho = 1) \\
&= \mathbf{E}Z_i^2.
\end{aligned}$$

Similarly, given arbitrary $2 \le i_1 \le q < i_2 \le 2q$,

$$\begin{aligned}
\mathbf{E}Z_{i_1} Z_{i_2} &= (2q)^2 \sum_A \left( \frac{\mathbf{P}(\sigma(n+1)=A)}{\prod_{j=1}^d \mathbf{P}(\sigma_j(n+1)=A_j)} \right)^2 \mathbf{P}(\sigma_\rho = 1 \mid \sigma(n+1)=A) \\
&\quad \times \mathbf{P}(\sigma_\rho = i_1 \mid \sigma(n+1)=A)\mathbf{P}(\sigma(n+1)=A \mid \sigma_\rho = i_2) \\
&= (2q)^2 \sum_A \left( \frac{\mathbf{P}(\sigma(n+1)=A)}{\prod_{j=1}^d \mathbf{P}(\sigma_j(n+1)=A_j)} \right)^2 \mathbf{P}(\sigma_\rho = q+1 \mid \sigma(n+1)=A) \\
&\quad \times \mathbf{P}(\sigma_\rho = 2q \mid \sigma(n+1)=A)\mathbf{P}(\sigma(n+1)=A \mid \sigma_\rho = 1) \\
&= \mathbf{E}Z_{q+1} Z_{2q}.
\end{aligned}$$

Next approximate the means and variances of monomials of $Z_i$ by expanding them using Taylor series. The following relations hold, where the symbol $O_q$ emphasizes that the corresponding constant depends only on $q$.

(i) When $i = 1$,

$$\begin{aligned}
\mathbf{E}Z_1 &= 1 + d\left[2q\lambda_1^2 x_n + 2q(\lambda_1^2 - \lambda_2^2)z_n\right] \\
&\quad + \frac{d(d-1)}{2}\left[2q\lambda_1^2 x_n + 2q(\lambda_1^2 - \lambda_2^2)z_n\right]^2 + O_q(x_n^3);
\end{aligned}$$

(ii) For $2 \le i \le q$,

$$\begin{aligned}
\mathbf{E}Z_i &= 1 + d\left[-\frac{2q\lambda_1^2}{q-1}x_n - \left(\frac{2q\lambda_1^2}{q-1} + 2q\lambda_2^2\right) z_n\right] \\
&\quad + \frac{d(d-1)}{2}\left[-\frac{2q\lambda_1^2}{q-1}x_n - \left(\frac{2q\lambda_1^2}{q-1} + 2q\lambda_2^2\right) z_n\right]^2 + O_q(x_n^3);
\end{aligned}$$

(iii) For $q+1 \le i \le 2q$,

$$\mathbf{E}Z_i = 1 + d(2q\lambda_2^2 z_n) + \frac{d(d-1)}{2}(2q\lambda_2^2 z_n)^2 + O_q(x_n^3).$$

Consider covariances of $(Z_i)$. By Lemma 3, it is known that $\mathbf{E}Z_1 Z_i = \mathbf{E}Z_i^2$, therefore we obtain the following results:

(a) when $i = 1$,

$$\mathbf{E}Z_1^2 = 1 + d\Pi_1 + \frac{d(d-1)}{2}\Pi_1^2 + O_q(x_n^3),$$

where

$$\begin{aligned}\Pi_1 &= \mathbf{E}\left[1 + 2q\lambda_1\left(Y_{1j} - \frac{1}{2q}\right) + 2(\lambda_1 - \lambda_2)\sum_{q+1\le\ell\le 2q}\left(Y_{\ell j} - \frac{1}{2q}\right)\right]^2 - 1\\ &= 6q\lambda_1^2 x_n + 6q(\lambda_1^2 - \lambda_2^2)z_n + 4q^2\lambda_1^3\left(u_n - \frac{x_n}{2q}\right)\\ &\quad + 12q^2\lambda_1^2(\lambda_1 - \lambda_2)\left(w_n - \frac{x_n}{2q}\right);\end{aligned}$$

(b) for $2 \le i \le q$,

$$\mathbf{E}Z_i^2 = 1 + d\Pi_2 + \frac{d(d-1)}{2}\Pi_2^2 + O_q(x_n^3),$$

where

$$\begin{aligned}\Pi_2 &= \mathbf{E}\left[1 + 2q\lambda_1\left(Y_{ij} - \frac{1}{2q}\right) + 2(\lambda_1 - \lambda_2)\sum_{q+1\le\ell\le 2q}\left(Y_{\ell j} - \frac{1}{2q}\right)\right]^2 - 1\\ &= \frac{2q(q-3)}{q-1}\lambda_1^2 x_n + \left(\frac{2q(q-3)}{q-1}\lambda_1^2 - 6q\lambda_2^2\right)z_n - \frac{4q^2}{q-1}\lambda_1^3\left(u_n - \frac{x_n}{2q}\right)\\ &\quad - 4q^2\frac{3\lambda_1 + (q-3)\lambda_2}{q-1}\lambda_1^2\left(w_n - \frac{x_n}{2q}\right);\end{aligned}$$

(c) for $q + 1 \le i \le 2q$,

$$\mathbf{E}Z_i^2 = 1 + d\Pi_3 + \frac{d(d-1)}{2}\Pi_3^2 + O_q(x_n^3),$$

where

$$\begin{aligned}\Pi_3 &= \mathbf{E}\left[1 + 2q\lambda_1\left(Y_{ij} - \frac{1}{2q}\right) + 2(\lambda_2 - \lambda_1)\sum_{q+1\le\ell\le 2q}\left(Y_{\ell j} - \frac{1}{2q}\right)\right]^2 - 1\\ &= 2q\lambda_1^2 x_n + 2q(\lambda_1^2 + \lambda_2^2)z_n + 4q^2\lambda_1^2\lambda_2\left(w_n - \frac{x_n}{2q}\right);\end{aligned}$$

(d) for $2 \le i_1 < i_2 \le q$,

$$\mathbf{E}Z_{i_1}Z_{i_2} = \mathbf{E}Z_2 Z_q = 1 + d\Pi_4 + \frac{d(d-1)}{2}\Pi_4^2 + O_q(x_n^3),$$

where

$$\begin{aligned}\Pi_4 &= \mathbf{E}\left[1+2q\lambda_1\left(Y_{2j}-\frac{1}{2q}\right)+2(\lambda_1-\lambda_2)\sum_{q+1\le\ell\le 2q}\left(Y_{\ell j}-\frac{1}{2q}\right)\right]\\ &\quad\times\left[1+2q\lambda_1\left(Y_{qj}-\frac{1}{2q}\right)+2(\lambda_1-\lambda_2)\sum_{q+1\le\ell\le 2q}\left(Y_{\ell j}-\frac{1}{2q}\right)\right]-1\\ &= -\frac{6q\lambda_1^2}{q-1}x_n-\left(\frac{6q\lambda_1^2}{q-1}+6q\lambda_2^2\right)z_n+\frac{8q^2\lambda_1^3}{(q-1)(q-2)}\left(u_n-\frac{x_n}{2q}\right)\\ &\quad+4q^2\frac{6\lambda_1+(3q-6)\lambda_2}{(q-1)(q-2)}\lambda_1^2\left(w_n-\frac{x_n}{2q}\right);\end{aligned}$$

(e) for $q+1\le i\le 2q$,

$$\mathbf{E}Z_2Z_i=\mathbf{E}Z_2Z_{q+1}=1+d\Pi_5+\frac{d(d-1)}{2}\Pi_5^2+O_q(x_n^3),$$

where

$$\begin{aligned}\Pi_5 &= \mathbf{E}\left[1+2q\lambda_1\left(Y_{2j}-\frac{1}{2q}\right)+2(\lambda_1-\lambda_2)\sum_{q+1\le\ell\le 2q}\left(Y_{\ell j}-\frac{1}{2q}\right)\right]\\ &\quad\times\left[1+2q\lambda_1\left(Y_{(q+1)j}-\frac{1}{2q}\right)+2(\lambda_1-\lambda_2)\sum_{1\le\ell\le q}\left(Y_{\ell j}-\frac{1}{2q}\right)\right]-1\\ &= -\frac{2q\lambda_1^2}{q-1}x_n+\left(-\frac{2q\lambda_1^2}{q-1}+2q\lambda_2^2\right)z_n-\frac{4q^2}{q-1}\lambda_1^2\lambda_2\left(w_n-\frac{x_n}{2q}\right).\end{aligned}$$

## 2.5 Main Expansion of $x_{n+1}$ and $z_{n+1}$

In this section, we aim to figure out the second order recursive relation between $x_{n+1}$ and $z_{n+1}$, by virtue of the following identity

$$\frac{a}{s+r}=\frac{a}{s}-\frac{ar}{s^2}+\frac{r^2}{s^2}\frac{a}{s+r}. \tag{2.5}$$

Specifically, taking $a=Z_1$, $s=2q$ and $r=\sum_{i=1}^{2q}Z_i-2q$, (2.5) yields

$$\begin{aligned}x_{n+1}+\frac{1}{2q} &= \mathbf{E}\frac{Z_1}{\sum_{i=1}^{2q}Z_i}=\mathbf{E}\frac{Z_1}{2q}-\mathbf{E}\frac{Z_1(\sum_{i=1}^{2q}Z_i-2q)}{(2q)^2}\\ &\quad+\mathbf{E}\frac{Z_1}{\sum_{i=1}^{2q}Z_i}\frac{(\sum_{i=1}^{2q}Z_i-2q)^2}{(2q)^2};\end{aligned} \tag{2.6}$$

$$\begin{aligned}z_{n+1}+\frac{1}{2q} &= \mathbf{E}\frac{Z_{q+1}}{\sum_{i=1}^{2q}Z_i}=\mathbf{E}\frac{Z_{q+1}}{2q}-\mathbf{E}\frac{Z_{q+1}(\sum_{i=1}^{2q}Z_i-2q)}{(2q)^2}\\ &\quad+\mathbf{E}\frac{Z_{q+1}}{\sum_{i=1}^{2q}Z_i}\frac{(\sum_{i=1}^{2q}Z_i-2q)^2}{(2q)^2}.\end{aligned} \tag{2.7}$$

Finally, plugging the results of Sect. 2.4 into (2.6) and (2.7) and taking substitutions of $\mathcal{X}_n = x_n + z_n$ and $\mathcal{Z}_n = -z_n$, we obtain a two dimensional recursive formula of the linear diagonal canonical form:

$$\begin{cases} \mathcal{X}_{n+1} = d\lambda_1^2\mathcal{X}_n + \frac{d(d-1)}{2}\left(\frac{2q(q-3)}{q-1}\lambda_1^4\mathcal{X}_n^2 + 4q\lambda_1^2\lambda_2^2\mathcal{X}_n\mathcal{Z}_n\right) + R_x + R_z + V_x \\ \mathcal{Z}_{n+1} = d\lambda_2^2\mathcal{Z}_n + \frac{d(d-1)}{2}\left(\frac{q}{q-1}\lambda_1^4\mathcal{X}_n^2 - 4q\lambda_2^4\mathcal{Z}_n^2\right) - R_z + V_z \end{cases} \tag{2.8}$$

where

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

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

and

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

with $C_V$ a constant depending on $q$ only.

## 3 Proof of the Main Theorem

If the reconstruction problem is solvable, then $\sigma(n)$ contains significant information on the root variable. This may be expressed in several equivalent ways (Mossel [32], Proposition 14).

**Lemma 4** *The nonreconstruction is equivalent to*

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

In order to study the stability of dynamical system (2.8), we expect $R_x$, $R_z$ and $V_x$, $V_z$ to be just small perturbations, for example, of the order $o(x_n^2)$. It is known that fixed finite different vertices far away from the root can affect the root little, based on which, it is possible to explore further the concentration analysis. We can verify that $\frac{Z_1}{\sum_{i=1}^{2q} Z_i}$ and $\frac{Z_{q+1}}{\sum_{i=1}^{2q} Z_i}$ are both sufficiently around $\frac{1}{2q}$, and thus are able to bound $R_x$ and $R_z$ in (2.8).

**Lemma 5** *Assume* $\min\{|\lambda_1|, |\lambda_2|\} > \varrho$*, for some* $\varrho > 0$*. For any* $\varepsilon > 0$*, there exist* $N = N(q, \varepsilon)$ *and* $\delta = \delta(q, \varepsilon, \varrho) > 0$*, such that if* $n \ge N$ *and* $x_n \le \delta$*, then*

$$|R_x| \le \varepsilon x_n^2 \quad \text{and} \quad |R_z| \le \varepsilon x_n^2.$$

*Proof* For any $\eta > 0$ and $1 \le i \le 2q$, applying Cauchy-Schwartz inequality,

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

$$
\begin{aligned}
&\leq \eta \mathbf{E}\left(\frac{(\sum_{i=1}^{2q} Z_i - 2q)^2}{(2q)^2}; \left|\frac{Z_1}{\sum_{i=1}^{2q} Z_i} - \frac{1}{2q}\right| \leq \eta\right) \\
&\quad + \mathbf{E}\left(\frac{(\sum_{i=1}^{2q} Z_i - 2q)^2}{(2q)^2} \mathbf{I}\left(\left|\frac{Z_1}{\sum_{i=1}^{2q} Z_i} - \frac{1}{2q}\right| > \eta\right)\right) \\
&\leq \eta \mathbf{E}\left(\frac{(\sum_{i=1}^{2q} Z_i - 2q)^2}{(2q)^2}\right) \\
&\quad + \left(\mathbf{E}\frac{(\sum_{i=1}^{2q} Z_i - 2q)^4}{(2q)^4}\right)^{1/2} \left(\mathbf{P}\left(\left|\frac{Z_1}{\sum_{i=1}^{2q} Z_i} - \frac{1}{2q}\right| > \eta\right)\right)^{1/2}.
\end{aligned}
$$

We can derive from the calculation for distributional recursion that

$$
\mathbf{E}\left(\frac{(\sum_{i=1}^{2q} Z_i - 2q)^2}{(2q)^2}\right) \leq C_1(q) x_n^2 \quad \text{and} \quad \mathbf{E}\frac{(\sum_{i=1}^{2q} Z_i - 2q)^4}{(2q)^4} \leq C_2(q).
$$

Similar to Lemma 2.11 of Sly [43] and Lemma 4.3 of Liu and Ning [25], there exist $C_3 = C_3(q, \eta, \varrho)$ and $N = N(q, \eta)$, such that when $n \geq N$,

$$
\mathbf{P}\left(\left|\frac{Z_1}{\sum_{i=1}^{2q} Z_i} - \frac{1}{2q}\right| > \eta\right) \leq C_3 x_n^6.
$$

Thus there exists $C_4 = C_4(q, \eta, \varrho)$, such that

$$
|R_x| = \left|\mathbf{E}\left(\frac{Z_1}{\sum_{i=1}^{2q} Z_i} - \frac{1}{2q}\right) \frac{(\sum_{i=1}^{2q} Z_i - 2q)^2}{(2q)^2}\right| \leq \eta C_1 x_n^2.
$$

Finally, it suffices to take $C_1\eta = \varepsilon/2$, and then if $x_n \leq \delta$, we have $R_x \leq \varepsilon x_n^2$. Similar analysis gives $R_z \leq \varepsilon x_n^2$.

Prior to establishing the concentration results involving $V_x$ and $V_z$, we need to firstly show that the value of $x_n$ does not drop too fast to be non-reconstruction.

**Lemma 6** *For any $\varrho > 0$, there exists a constant $\gamma = \gamma(q, \varrho) > 0$, such that*

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

*for all $n$, if $\min\{|\lambda_1|, |\lambda_2|\} > \varrho$.*

*Proof* Similarly to (2.1), for a configuration $A$ on $\mathbb{T}_{u_1} \cap L(n+1)$, define the posterior function

$$
\begin{aligned}
g_{n+1}(1, A) &= \mathbf{P}(\sigma_\rho = 1 \mid \sigma_1(n+1) = A) \\
&= \frac{1}{2q} + p_0\left(f_n(1, A) - \frac{1}{2q}\right) \\
&\quad + p_1 \sum_{i=2}^{q}\left(f_n(i, A) - \frac{1}{2q}\right) \\
&\quad + p_2 \sum_{i=q+1}^{2q}\left(f_n(i, A) - \frac{1}{2q}\right)
\end{aligned}
$$

$$= \frac{1}{2q} + \lambda_1 \left( f_n(1, A) - \frac{1}{2q} \right) + \frac{\lambda_1 - \lambda_2}{q} \sum_{i=q+1}^{2q} \left( f_n(i, A) - \frac{1}{2q} \right)$$

and then

$$\begin{aligned} \mathbf{E} g_{n+1}(1, \sigma_1^1(n+1)) &= \frac{1}{2q} + \lambda_1 \mathbf{E} \left( Y_{11}(n) - \frac{1}{2q} \right) \\ &\quad + \frac{\lambda_1 - \lambda_2}{q} q \mathbf{E} \left( Y_{(q+1)1}(n) - \frac{1}{2q} \right) \\ &= \frac{1}{2q} + \lambda_1^2 x_n + (\lambda_1^2 - \lambda_2^2) z_n. \end{aligned}$$

The estimator that chooses a state with probability $f_{n+1}(i, \sigma_1(n+1))$ correctly reconstructs the root with probability $\frac{1}{2q} + \lambda_1^2 x_n + (\lambda_1^2 - \lambda_2^2) z_n$. Apparently this probability must be less than the maximum-likelihood estimator (Mézard and Montanari [31]). Therefore, the following inequalities hold:

$$\begin{aligned} \frac{1}{2q} + \lambda_1^2 x_n + (\lambda_1^2 - \lambda_2^2) z_n &\le \mathbf{E} \max_{1 \le i \le 2q} X_i(n+1) \\ &\le \frac{1}{2q} + \left( \mathbf{E} \max_i \left( X_i(n+1) - \frac{1}{2q} \right)^2 \right)^{1/2} \\ &\le \frac{1}{2q} + \left( \mathbf{E} \sum_{i=1}^{2q} \left( X_i(n+1) - \frac{1}{2q} \right)^2 \right)^{1/2} \\ &= \frac{1}{2q} + x_{n+1}^{1/2}. \end{aligned}$$

On one hand, if $\lambda_1^2 \ge \lambda_2^2$, then it is concluded from $x_n + z_n \ge 0$ in Lemma 2.2 that

$$\lambda_2^2 x_n \le \lambda_2^2 x_n + \left(\lambda_1^2 - \lambda_2^2\right)(x_n + z_n) = \lambda_1^2 x_n + \left(\lambda_1^2 - \lambda_2^2\right) z_n \le x_{n+1}^{1/2}.$$

On the other hand, if $\lambda_1^2 \le \lambda_2^2$ then $\lambda_1^2 x_n \le x_{n+1}^{1/2}$, since $z_n \le 0$. To sum up, we always have

$$\min\{\lambda_1^2, \lambda_2^2\} x_n \le x_{n+1}^{1/2}.$$

Next choose $\varepsilon = \varrho^2$. It can be concluded from (2.8), Lemma 5, as well as the inequalities

$$\left| \frac{u_n}{x_n} - \frac{1}{2q} \right| \le 1, \quad \left| \frac{w_n}{x_n} - \frac{1}{2q} \right| \le 1, \tag{3.1}$$

that there exists a $\delta = \delta(q, \varepsilon) > 0$, such that, when $x_n < \delta$, one has

$$x_{n+1} \ge (d \min\{\lambda_1^2, \lambda_2^2\} - \varepsilon) x_n \ge (d-1) \varrho^2 x_n \ge \varrho^2 x_n.$$

On the contrary, if $x_n \ge \delta$, one has $x_{n+1} \ge (\min\{\lambda_1^2, \lambda_2^2\} x_n)^2 \ge \varrho^4 \delta x_n$. Finally, taking $\gamma = \min\{\varrho^2, \varrho^4 \delta\}$ completes the proof.

The following lemma improves the result of Lemma 1 by establishing the strict positivity of the sum of $x_n$ and $z_n$.

**Lemma 7** *Assume* $\lambda_1 \neq 0$. *For any nonnegative* $n \in \mathbb{Z}$, *we always have*

$$x_n + z_n > 0.$$

*Proof* In Lemma 2.2 we have proved that $x_n + z_n \geq 0$, so here it suffices to exclude the equality. Now let us apply reductio ad absurdum and assume $x_n + z_n = 0$ for some $n \in \mathbb{N}$. It follows that for $i \neq j$ in the same configuration set, one has

$$\mathbf{E}(X_i(n) - X_j(n))^2 = 2EX_i^2(n) - 2EX_i(n)X_j(n) = \frac{x_n + z_n}{q-1} = 0.$$

Therefore, $X_1(n) = X_2(n) = \cdots = X_q(n)$ and $X_{q+1}(n) = X_{q+2}(n) = \cdots = X_{2q}(n)$ a.s., that is, for any configuration combination $A$ on the $n$th level, we always have

$$\mathbf{P}(\sigma_\rho = 1 \mid \sigma(n) = A) = \mathbf{P}(\sigma_\rho = 2 \mid \sigma(n) = A).$$

Denote the leftmost vertex on the $n$th level by $v_n(1)$, and it follows

$$\mathbf{P}(\sigma_\rho = 1 \mid \sigma_{v_n(1)} = 1) = \mathbf{P}(\sigma_\rho = 2 \mid \sigma_{v_n(1)} = 1).$$

Define the transition matrices at distance $s$ by

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

and then it is convenient to figure out the iterative formulae for them

$$\begin{cases} U_s = p_0 U_{s-1} + (q-1)p_1 V_{s-1} + qp_2 W_{s-1} \\ V_s = p_1 U_{s-1} + [p_0 + (q-2)p_1]V_{s-1} + qp_2 W_{s-1} \\ W_s = p_2 U_{s-1} + (q-1)p_2 V_{s-1} + [p_0 + (q-1)p_1]W_{s-1}. \end{cases}$$

To evaluate this three order recursive system, starting with the difference of the first two equation

$$U_s - V_s = \lambda_1(U_{s-1} - V_{s-1}),$$

and then in light of the initial conditions $U_0 = 1$ and $V_0 = W_0 = 0$, it follows that

$$U_s - V_s = \lambda_1^s. \tag{3.2}$$

Finally, from the reversible property of the channel, we can conclude that

$$\lambda_1^n = U_n - V_n = \mathbf{P}(\sigma_\rho = 1 \mid \sigma_{v_n(1)} = 1) - \mathbf{P}(\sigma_\rho = 2 \mid \sigma_{v_n(1)} = 1) = 0,$$

i.e., $\lambda_1 = 0$, a contradiction to the assumption of $\lambda_1 \neq 0$.

The following result provides the crucial concentration estimates of $u_n - \frac{x_n}{2q}$ and $w_n - \frac{x_n}{2q}$, when $x_n$ is small.

**Lemma 8** *Assume* $|\lambda_2| > \varrho$, *and* $|\lambda_1| = |\lambda_2|$ *or* $|\lambda_1|/|\lambda_2| \geq \kappa$ *for some* $\kappa > 1$. *For any* $\varepsilon > 0$, *there exist* $N = N(q, \kappa, \varepsilon)$ *and* $\delta = \delta(q, \kappa, \varrho, \varepsilon) > 0$, *such that if* $n \geq N$ *and* $x_n \leq \delta$, *one has*

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

*Proof* Applying the identity (2.5), we have

$$
\begin{aligned}
u_{n+1} &= \mathbf{E}\frac{\left(Z_1-\frac{1}{2q}\sum_{i=1}^{2q}Z_i\right)^2}{\left(\sum_{i=1}^{2q}Z_i\right)^2}\\
&= \frac{1}{4q^2}\mathbf{E}\left(Z_1-\frac{1}{2q}\sum_{i=1}^{2q}Z_i\right)^2\\
&\quad-\frac{1}{16q^4}\mathbf{E}\left(Z_1-\frac{1}{2q}\sum_{i=1}^{2q}Z_i\right)^2\left(\left(\sum_{i=1}^{2q}Z_i\right)^2-4q^2\right)\\
&\quad+\frac{1}{16q^4}\mathbf{E}\frac{\left(Z_1-\frac{1}{2q}\sum_{i=1}^{2q}Z_i\right)^2}{\left(\sum_{i=1}^{2q}Z_i\right)^2}\left(\left(\sum_{i=1}^{2q}Z_i\right)^2-4q^2\right).
\end{aligned}
\tag{3.3}
$$

The first expectation of (3.3) will contribute the major terms of the expansion:

$$
\begin{aligned}
&\mathbf{E}\left(Z_1-\frac{1}{2q}\sum_{i=1}^{2q}Z_i\right)^2\\
&= \mathbf{E}(Z_1-1)^2-\frac{2}{2q}\mathbf{E}(Z_1-1)\left(\sum_{i=1}^{2q}Z_i-2q\right)\\
&\quad+\frac{1}{4q^2}E\left(\sum_{i=1}^{2q}Z_i-2q\right)^2\\
&= 2dq\lambda_1^2x_n+2dq(\lambda_1^2-\lambda_2^2)z_n+4dq^2\lambda_1^3\left(u_n-\frac{x_n}{2q}\right)\\
&\quad+12dq^2\lambda_1^2(\lambda_1-\lambda_2)\left(w_n-\frac{x_n}{2q}\right)+O_q(x_n^2).
\end{aligned}
$$

Similarly, we can bound both the second and third terms of (3.3) by $O_q(x_n^2)$:

$$
\mathbf{E}\left(Z_1-\frac{1}{2q}\sum_{i=1}^{2q}Z_i\right)^2\left(\left(\sum_{i=1}^{2q}Z_i\right)^2-4q^2\right)=O_q(x_n^2),
$$

and

$$
\mathbf{E}\left(\left(\sum_{i=1}^{2q}Z_i\right)^2-4q^2\right)^2=O_q(x_n^2).
$$

Note that all the $O_q$ terms in the context only depend on $q$. Substituting these bounds into (3.3) gives

$$
u_{n+1}=\frac{x_{n+1}}{2q}+d\lambda_1^3\left(u_n-\frac{x_n}{2q}\right)+3d\lambda_1^2(\lambda_1-\lambda_2)\left(w_n-\frac{x_n}{2q}\right)+O_q(x_n^2), \tag{3.4}
$$

by means of $x_{n+1} = d\lambda_1^2 x_n + d(\lambda_1^2 - \lambda_2^2)z_n + O_q(x_n^2)$. Moreover, the similar expansion of $w_{n+1}$ would be

$$w_{n+1} = \frac{1}{4q^2}\mathbf{E}(Z_{q+1}-1)^2 + O_q(x_n^2) = \frac{x_{n+1}}{2q} + d\lambda_1^2\lambda_2\left(w_n - \frac{x_n}{2q}\right) + O_q(x_n^2),$$

and thus

$$\frac{w_{n+1}}{x_{n+1}} - \frac{1}{2q} = d\lambda_1^2\lambda_2 \frac{x_n}{x_{n+1}}\left(\frac{w_n}{x_n} - \frac{1}{2q}\right) + O_q\left(\frac{x_n^2}{x_{n+1}}\right). \tag{3.5}$$

Next display the discussion in the $\mathcal{X}O\mathcal{Z}$ plane. First consider the case of $|\lambda_1|/|\lambda_2| \geq \kappa$ for $\kappa > 1$. In a small neighborhood of $(0, 0)$, since $d\lambda_2^2 < \kappa^2 d|\lambda_2^2| \leq d\lambda_1^2 < 1$ and $\mathcal{X}_n > 0$, the discrete trajectories approach to the origin point "tangentially" to the $\mathcal{X}$-axis, if $x_n$ is small enough for some $n$ (see Bernussou [4] for reference). Besides, the conclusion of Lemma 7 excludes the trajectory along $\mathcal{Z}$-axis. Then for some $M > 1$, there exist absolute constants $N_1 = N_1(q, \kappa, M)$ and $\delta_1 = \delta_1(q, \kappa, M)$, such that if $n \geq N_1$ and $x_n \leq \delta_1$, we have

$$\mathcal{X}_n \geq M\mathcal{Z}_n \quad \text{and} \quad \frac{1}{M(M+1)}d\lambda_1^2 x_n + O_q(x_n^2) > 0,$$

where the remainder term $O_q(x_n^2)$ comes from the expansion of $x_{n+1}$. Consequently, it follows $x_n + z_n = \mathcal{X}_n \geq \frac{M}{M+1}x_n$, which yields, in connection with $z_n \leq 0$ in Lemma 2.2,

$$\begin{aligned}\frac{x_n}{x_{n+1}} &= \frac{x_n}{d\lambda_1^2 x_n + d(\lambda_1^2 - \lambda_2^2)z_n + O_q(x_n^2)} \leq \frac{x_n}{\frac{M}{M+1}d\lambda_1^2 x_n + O_q(x_n^2)} \leq \frac{x_n}{\left(1 - \frac{1}{M}\right)d\lambda_1^2 x_n} \\ &= \frac{M}{M-1}\frac{1}{d\lambda_1^2}.\end{aligned} \tag{3.6}$$

The second case taken into account is $|\lambda_1| = |\lambda_2|$. In view of $1/2 \leq d\lambda^2 = d\lambda_1^2 \leq 1$, there also exist absolute constants $N_2 = N_2(q, M)$ and $\delta_2 = \delta_2(q, M)$, such that if $n \geq N_2$ and $x_n \leq \delta_2$, one has

$$\frac{x_n}{x_{n+1}} = \frac{x_n}{d\lambda_1^2 x_n + O_q(x_n^2)} \leq \frac{x_n}{\left(1 - \frac{1}{M}\right)d\lambda_1^2 x_n} = \frac{M}{M-1}\frac{1}{d\lambda_1^2}.$$

For fixed $k$, it is known from (2.8) that

$$|x_{n+1} - (d\lambda_1^2\mathcal{X}_n + d\lambda_2^2\mathcal{Z}_n)| \leq C(q)x_n^2,$$

and then there exists $\delta_3 = \delta_3(q, \kappa, M, k) < \min\{\delta_1, \delta_2\}$, such that if $x_n < \delta_3$ then one has $x_{n+\ell} < 2\delta_3$, for any $1 \leq \ell \leq k$. Therefore, for any positive integer $k$, (3.5) yields

$$\begin{aligned}\frac{w_{n+k}}{x_{n+k}} - \frac{1}{2q} &= d\lambda_1^2\lambda_2 \frac{x_{n+k-1}}{x_{n+k}}\left(\frac{w_{n+k-1}}{x_{n+k-1}} - \frac{1}{2q}\right) + O_q\left(x_{n+k-1}\frac{x_{n+k-1}}{x_{n+k}}\right) \\ &= (d\lambda_1^2\lambda_2)^k\left(\prod_{\ell=1}^{k}\frac{x_{n+\ell-1}}{x_{n+\ell}}\right)\left(\frac{w_n}{x_n} - \frac{1}{2q}\right) + R,\end{aligned}$$

where

$$(d\lambda_1^2\lambda_2)^k\left(\prod_{\ell=1}^{k}\frac{x_{n+\ell-1}}{x_{n+\ell}}\right) \leq (d\lambda_1^2|\lambda_2|)^k\left(\frac{M}{M-1}\frac{1}{d\lambda_1^2}\right)^k = \left(\frac{M}{M-1}|\lambda_2|\right)^k$$

and

$$|R| \leq 2C\delta_3 \left( \sum_{i=1}^{k} \left( \frac{M}{M-1} \frac{1}{d\lambda_1^2} \right)^i (d\lambda_1^2 |\lambda_2|)^{i-1} \right) \leq \delta_3 \frac{1 - \left( \frac{M}{M-1} |\lambda_2| \right)^k}{1 - \left( \frac{M}{M-1} |\lambda_2| \right)} \frac{M}{M-1} \frac{1}{d\lambda_1^2},$$

with $C$ denoting the $O_q$ constant in (3.5). From the identity (i) in Lemma 2, we obtain $0 \leq \frac{w_n}{x_n} \leq \frac{1}{q}$, which implies

$$\left| \frac{w_n}{x_n} - \frac{1}{2q} \right| \leq \frac{1}{2q}.$$

Noting that $|\lambda_2| \leq |\lambda_1| \leq d^{-1/2} \leq 1/\sqrt{2}$, it is possible to achieve $\frac{M}{M-1}|\lambda_2| < 1$ by choosing $M = 4$. Therefore, it is feasible to take $k = k(\varepsilon)$ sufficiently large and $\delta_4 = \delta_4(q, \kappa, k, \varepsilon) = \delta_4(q, \kappa, \varepsilon) < \delta_3$ sufficiently small to guarantee

$$\left| \frac{w_{n+k}}{x_{n+k}} - \frac{1}{2q} \right| < \varepsilon.$$

Finally, in view of $|\lambda_2| > \varrho$, there exists $\gamma = \gamma(q, \varrho)$ by Lemma 6 satisfying $x_{n-k} \leq \gamma^{-k} x_n$. Thus choose $N = N(q, \kappa, \varepsilon, k) = N(q, \kappa, \varepsilon) > \max\{N_1 + k, N_2 + k\}$ and $\delta = \gamma^k \delta_4$, such that, if $x_n \leq \delta$ and $n \geq N$, one has

$$\left| \frac{w_n}{x_n} - \frac{1}{2q} \right| < \varepsilon. \tag{3.7}$$

Finally, the second part of the lemma follows by plugging (3.7) into (3.4) and proceeding similarly as above.

*Proof of the Main Theorem* First, consider $\varrho < |\lambda_2| < |\lambda_1|$, for any fixed $\varrho > 0$. By Lemma 4, it suffices to show that when $d\lambda_1^2$ is close enough to 1, $\mathcal{X}_n$ does not converge to 0. Therefore, it implies that $x_n$ does not converge to 0 either, considering $0 \leq \mathcal{X}_n = x_n + z_n \leq x_n$. It is convenient to make $|\lambda_2| > \varrho$ fixed and just $\lambda_1$ varying, and then without loss of generality, assume $d\lambda_1^2 > \frac{1+d\lambda_2^2}{2}$. Consequently choose $\kappa = \kappa(d, \lambda_2) = \left( \frac{1+d\lambda_2^2}{2d\lambda_2^2} \right)^{1/2} > 1$ and thus $|\lambda_1|/|\lambda_2| \geq \kappa$.

As in Lemma 8, display our proof in the $\mathcal{X}O\mathcal{Z}$ plane. Under the condition of $q \geq 4$ and (2.8), it is apparent that

$$\begin{aligned} \mathcal{X}_{n+1} &= d\lambda_1^2 \mathcal{X}_n + \frac{d(d-1)}{2} \left( \frac{2q(q-3)}{q-1} \lambda_1^4 \mathcal{X}_n^2 + 4q\lambda_1^2\lambda_2^2 \mathcal{X}_n \mathcal{Z}_n \right) \\ &\quad + R_x + R_z + V_x \\ &\geq d\lambda_1^2 \mathcal{X}_n + \frac{d(d-1)}{2} \frac{2q(q-3)}{q-1} \lambda_1^4 \mathcal{X}_n^2 - |R_x| - |R_z| \\ &\quad - C_V x_n^2 \left( \left| \frac{u_n}{x_n} - \frac{1}{2q} \right| + \left| \frac{w_n}{x_n} - \frac{1}{2q} \right| + x_n \right), \end{aligned}$$

where the last inequality comes from $|\lambda_1| \leq d^{-1/2} < 1$. Then by Lemma 8 and Lemma 5, there exist $N = N(q, \kappa, \varrho)$ and $\delta = \delta(q, d, \kappa, \varrho) > 0$, such that if $n \geq N$ and $x_n \leq \delta$, then in the small neighborhood of the origin point (0, 0), we have $\mathcal{X}_n \geq \mathcal{Z}_n$ and $\mathcal{X}_n \geq \frac{x_n}{2}$. Meanwhile, the following estimates hold:

$$x_n \leq \frac{1}{48C_V}\frac{d(d-1)}{2}\frac{2q(q-3)}{q-1}\lambda_1^4;$$
$$\left|\frac{u_n}{x_n}-\frac{1}{2q}\right|, \left|\frac{w_n}{x_n}-\frac{1}{2q}\right| \leq \frac{1}{48C_V}\frac{d(d-1)}{2}\frac{2q(q-3)}{q-1}\lambda_1^4;$$
$$|R_x|, |R_z| \leq \frac{1}{32}\frac{d(d-1)}{2}\frac{2q(q-3)}{q-1}\lambda_1^4 x_n^2 \leq \frac{1}{8}\frac{d(d-1)}{2}\frac{2q(q-3)}{q-1}\lambda_1^4 \mathcal{X}_n^2.$$

Therefore, the quadratic term of $\mathcal{X}_n^2$ is "large" enough to control the remainder terms:

$$\begin{aligned}\mathcal{X}_{n+1} &\geq d\lambda_1^2\mathcal{X}_n + \frac{1}{2}\frac{d(d-1)}{2}\frac{2q(q-3)}{q-1}\lambda_1^4\mathcal{X}_n^2\\ &= \mathcal{X}_n\left[d\lambda_1^2 + \frac{1}{2}\frac{d(d-1)}{2}\frac{2q(q-3)}{q-1}\lambda_1^4\mathcal{X}_n\right].\end{aligned} \tag{3.8}$$

Take $\varepsilon = \min\{\frac{1}{4}\gamma^N, \gamma\delta\} > 0$, where $\gamma = \gamma(q, \varrho) > 0$ is the constant in Lemma 6. Because $q \geq 4$ and $\varepsilon$ is independent of $\lambda_1$, we can choose $|\lambda_1| < d^{-1/2}$ to make

$$d\lambda_1^2 + \frac{1}{2}\frac{d(d-1)}{2}\frac{2q(q-3)}{q-1}\lambda_1^4\varepsilon > 1. \tag{3.9}$$

Since $x_0 = 1 - \frac{1}{2q} > \frac{1}{2}$, it is concluded that $x_n > \frac{1}{2}\gamma^n \geq 2\varepsilon$ when $n \leq N$, in addition, $\mathcal{X}_N \geq \frac{\mathcal{X}_N+\mathcal{Z}_N}{2} = \frac{x_N}{2} \geq \varepsilon$. Now suppose $\mathcal{X}_n \geq \varepsilon$ for some $n \geq N$. Then display our discussion of $\mathcal{X}_n$ as follows:

(a) If $\mathcal{X}_n \geq 2\gamma^{-1}\varepsilon$, then

$$\mathcal{X}_{n+1} \geq \frac{x_{n+1}}{2} \geq \frac{\gamma x_n}{2} \geq \frac{\gamma\mathcal{X}_n}{2} \geq \varepsilon;$$

(b) If $\varepsilon \leq \mathcal{X}_n \leq 2\gamma^{-1}\varepsilon$, then $x_n \leq \frac{\mathcal{X}_n}{2} \leq \gamma^{-1}\varepsilon \leq \delta$, and thus it follows from (3.8) and (3.9) that

$$x_{n+1} \geq \mathcal{X}_{n+1} \geq \mathcal{X}_n\left[d\lambda_1^2 + \frac{1}{2}\frac{d(d-1)}{2}\frac{2q(q-3)}{q-1}\lambda_1^4\mathcal{X}_n\right] \geq \mathcal{X}_n \geq \varepsilon.$$

Finally, we have $x_n \geq \mathcal{X}_n \geq \varepsilon$ for all $n$, by induction. Consequently, it is established that the Kesten–Stigum bound is not tight.

The second case to be considered is $|\lambda_1| = |\lambda_2|$, under which there are two equal multipliers in this nonlinear second order point mapping and the origin point must be a star node. Although the principal axis is undetermined, just by the comparison of the quadratic terms and for $q \geq 4$, it is concluded that

$$\begin{aligned}&\frac{d(d-1)}{2}\left(\frac{2q(q-3)}{q-1}\lambda_1^4\mathcal{X}_n^2 + 4q\lambda_1^4\mathcal{X}_n\mathcal{Z}_n\right)\\ &\quad - \frac{d(d-1)}{2}\left(\frac{q}{q-1}\lambda_1^4\mathcal{X}_n^2 - 4q\lambda_1^4\mathcal{Z}_n^2\right)\\ &= \frac{d(d-1)}{2}\left(\frac{2q^2-7q}{q-1}\lambda_1^4\mathcal{X}_n^2 + 4q\lambda_1^4\mathcal{X}_n\mathcal{Z}_n + 4q\lambda_1^4\mathcal{Z}_n^2\right)\\ &\geq \frac{d(d-1)}{2}\lambda_1^4 x_n^2,\end{aligned}$$

and thus the decay rate of $\mathcal{X}_n$ is much slower than $\mathcal{Z}_n$ if $x_n$ is sufficiently small. Therefore, in light of the preceding discussion, there still exist $N = N(q)$ and $\delta = \delta(q)$, such that if $n \geq N$

and $x_n \leq \delta$, we have $\mathcal{X}_n \geq \mathcal{Z}_n$ and thus $x_n = \mathcal{X}_n + \mathcal{Z}_n \leq 2\mathcal{X}_n$. Then the rest discussion would be similar as the first part. □

**Acknowledgements** We would like to thank two anonymous reviewers who provided us with many constructive and helpful comments.

## References

1. Achlioptas, D., Coja-Oghlan, A.: Algorithmic barriers from phase transitions. In IEEE 49th Annual IEEE Symposium on Foundations of Computer Science, FOCS'08, pp. 793-802. IEEE (2008)
2. Baxter, R.J.: The Riemann surface of the chiral Potts model free energy function. J. Stat. Phys. **112**(1–2), 1–26 (2003)
3. Berger, N., Kenyon, C., Mossel, E., Peres, Y.: Glauber dynamics on trees and hyperbolic graphs. Probab. Theory Related Fields **131**, 311–340 (2005)
4. Bernussou, J.: Point Mapping Stability. Pergamon Press, New York (1977)
5. Bhamidi, S., Rajagopal, R., Roch, S.: Network delay inference from additive metrics. Random Struct. Algorithms **37**, 176–203 (2010)
6. Bisconti, C., Corallo, A., Fortunato, L., Gentile, A. A., Massafra, A., Pell, P.: Reconstruction of a real world social network using the Potts model and Loopy belief propagation. Front. Psychol. (2015). https://doi.org/10.3389/fpsyg.2015.01698
7. Bleher, P.M., Ruiz, J., Zagrebnov, V.A.: On the purity of the limiting Gibbs state for the Ising model on the Bethe lattice. J. Statist. Phys. **79**, 473–482 (1995)
8. Borgs, C., Chayes, J. T., Mossel, E., Roch, S.: The Kesten-Stigum reconstruction bound is tight for roughly symmetric binary channels. In: IEEE Comput. Soc. FOCS, Berkeley, CA, pp. 518–530 (2006)
9. Cadotte, M.W., Davies, T.J.: Phylogenies in Ecology: A Guide to Concepts and Methods. Princeton University Press, Princeton (2016)
10. Cassandro, M., Ferrari, P.A., Merola, I., Presutti, E.: Geometry of contours and Peierls estimates in d= 1 Ising models with long range interactions. J. Math. Phy. **46**(5), 053305 (2005)
11. Chayes, J.T., Chayes, L., Sethna, J.P., Thouless, D.J.: A mean field spin glass with short-range interactions. Commun. Math. Phys. **106**(1), 41–89 (1986)
12. Daskalakis, C., Mossel, E., Roch, S.: Optimal phylogenetic reconstruction. In: STOC'06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing, pp. 159-168. ACM, New York (2006)
13. Derrida, B., Bray, A.J., Godreche, C.: Non-trivial exponents in the zero temperature dynamics of the 1D Ising and Potts models. J. Phys. A **27**(11), L357 (1994)
14. Dhar, D.: The relaxation to equilibrium in one-dimensional Potts models. J. Indian Inst. Sci. **75**(3), 297 (2013)
15. Evans, W., Kenyon, C., Peres, Y., Schulman, L.J.: Broadcasting on trees and the Ising model. Ann. Appl. Probab. **10**, 410–433 (2000)
16. Ferrari, P.A., Fernndez, R., Garcia, N.L.: Perfect simulation for interacting point processes, loss networks and Ising models. Stoch. Process. Appl. **102**(1), 63–88 (2002)
17. Georgii, H.O.: Gibbs Measures and Phase Transition. de Gruyter, Berlin (1988)
18. Gerschenfeld, A., and Montanari, A.: Reconstruction for models on random graphs. In: 48th Annual IEEE Symposium on Foundations of Computer Science, FOCS'07. IEEE (2007)
19. Giuliani, A., Mastropietro, V.: Universal finite size corrections and the central charge in non-solvable Ising models. Commun. Math. Phys. **324**(1), 179–214 (2013)
20. Giuliani, A., Seiringer, R.: Periodic striped ground states in Ising models with competing interactions. Commun. Math. Phys. **347**(3), 983–1007 (2016)
21. Kesten, H., Stigum, B.P.: Additional limit theorems for indecomposable multidimensional Galton-Watson processes. Ann. Math. Statist. **37**, 1463–1481 (1966)
22. Kesten, H., Stigum, B.P.: Limit theorems for decomposable multi-dimesional Galton-Watson processes. J. Math. Anal. Appl. **17**, 309–338 (1966)
23. Kimura, M.: A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mole. Evol. **16**(2), 111–120 (1980)
24. Liu, S., Ying, L., Shakkottai, S.: Influence maximization in social networks: an ising-model-based approach. In 2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton), pp. 570–576. IEEE (2010)

25. Liu, W., and Ning, N.: Reconstruction for the asymmetric Ising model on regular trees. In: Proceedings of 3rd International Conference on Electrical, Electronics, Engineering Trends, Communication, Optimization and Sciences (EEECOS) (IET-IEEE), pp. 458–465 (2016)
26. Ma, J., Ratan, A., Raney, B.J., Suh, B.B., Miller, W., Haussler, D.: The infinite sites model of genome evolution. Proc. Natl. Acad. Sci. USA **105**, 14254–14261 (2008)
27. Maes, C., Redig, F., Van Moffaert, A.: The restriction of the Ising model to a layer. J. Stat. Phys. **96**(1), 69–107 (1999)
28. Maes, C., Velde, K.V.: The fuzzy Potts model. J. Phys. A. **28**(15), 4261 (1995)
29. Majumdar, S.N., Dhar, D.: Equivalence between the Abelian sandpile model and the q 0 limit of the Potts model. Physica A **185**(1–4), 129–145 (1992)
30. Martinelli, F., Sinclair, A., Weitz, D.: Fast mixing for independent sets, colorings, and other models on trees. Random Struct. Algorithms **31**, 134–172 (2007)
31. Mézard, M., Montanari, A.: Reconstruction on trees and spin glass transition. J. Stat. Phys. **124**, 1317–1350 (2006)
32. Mossel, E.: Reconstruction on trees: beating the second eigenvalue. Ann. Appl. Probab. **11**, 285–300 (2001)
33. Mossel, E.: Phase transitions in phylogeny. Trans. Am. Math. Soc. **356**, 2379–2404 (2004)
34. Mossel, E.: Survey: information flow on trees. In: Graphs, Morphisms and Statistical Physics. DIMACS Ser. Discrete Math. Theoret. Comput. Sci. vol. 63, pp. 155–170. American. Math. Soc., Providence, RI (2004)
35. Mossel, E., Yuval, P.: Information flow on trees. Ann. Appl. Probab. **13**(3), 817–844 (2003)
36. Olejarz, J., Krapivsky, P.L., Redner, S.: Zero-temperature coarsening in the 2d Potts model. J. Stat. Mech. **2013**(06), P06018 (2013)
37. Roch, S.: A short proof that phylogenetic tree reconstruction by maximum likelihood is hard. IEEE/ACM Trans. Comput. Biol. Bioinform. **3**(1), 92 (2006)
38. Saul, L., Kardar, M.: Exact integer algorithm for the two-dimensionalJ Ising spin glass. Phys. Rev. E **48**(5), R3221 (1993)
39. Semerjian, G.: On the freezing of variables in random constraint satisfaction problems. J. Stat. Phys. **130**(2), 251–293 (2008)
40. Sire, C., Majumdar, S.N.: Coarsening in the q-state Potts model and the Ising model with globally conserved magnetization. Phys. Rev. E. **52**(1), 244 (1995)
41. Sire, C., Majumdar, S.N.: Correlations and coarsening in the q-state Potts model. Phys. Rev. Lett. **74**(21), 4321 (1995)
42. Sly, A.: Reconstruction of random colourings. Commun. Math. Phys. **288**(3), 943–961 (2009)
43. Sly, A.: Reconstruction for the Potts model. Ann. Probab. **39**, 1365–1406 (2011)
44. Spirin, V., Krapivsky, P.L., Redner, S.: Fate of zero-temperature Ising ferromagnets. Phys. Rev. E **63**(3), 036118 (2001)
45. Spohn, H., Dmcke, R.: Quantum tunneling with dissipation and the Ising model over. J. Stat. Phys. **41**(3–4), 389–423 (1985)
46. Tetali, P., Vera, J.C., Vigoda, E., Yang, L.: Phase transition for the mixing time of the Glauber dynamics for coloring regular trees. Ann. Appl. Probab. **22**, 2210–2239 (2012)
47. Tracy, C.A.: Universality class of a Fibonacci Ising model. J. Stat. Phys. **51**(3), 481–490 (1988)
48. Weeks, J.D., Gilmer, G.H., Leamy, H.J.: Structural transition in the Ising-model interface. Phys. Rev. Lett. **31**(8), 549 (1973)