Filomat 31:2 (2017), 273–283
DOI 10.2298/FIL1702273H

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

The Asymptotic Behavior for Markov Chains in a Finite i.i.d Random Environment Indexed by Cayley Trees

Huilin Huang[a]

[a]*College of Mathematics and Information Science, Wenzhou University, Zhejiang, 325035, PR China*

Abstract. We firstly define a Markov chain indexed by a homogeneous tree in a finite i.i.d random environment. Then, we prove the strong law of large numbers and Shannon-McMillan theorem for finite Markov chains indexed by a homogeneous tree in the finite i.i.d random environment.

1. Introduction

A tree T is a graph which is connected and contains no loops. Given any two vertices $\alpha \neq \beta \in T$, let $\overline{\alpha\beta}$ be the unique path connecting α and β. Define the graph distance $d(\alpha, \beta)$ to be the number of edges contained in the path $\overline{\alpha\beta}$.

Let T be an infinite tree with root 0. The set of all vertices with distance n from the root is called the n-th generation of T, which is denoted by L_n. We denote by $T^{(n)}$ the union of the first n generations of T. For each vertex t, there is a unique path from 0 to t, and $|t|$ for the number of edges on this path. We denote the first predecessor of t by ^1t. The degree of a vertex is defined to be the number of neighbors of it. If every vertex of the tree has degree $d+1$, we say it Cayley tree, which is denoted by $T_{C,d}$. Thus the root vertex has $d+1$ neighbors in the first generation and every other vertex has d neighbors in the next generation. For any two vertices s and t of tree T, write $s \leq t$ if s is on the unique path from the root 0 to t. We denote by $s \wedge t$ the vertex farthest from 0 satisfying $s \wedge t \leq s$ and $s \wedge t \leq t$. $X^A = \{X_t, t \in A\}$ and denote by $|A|$ the number of vertices of A.

In the following, we always let T denote the Cayley tree $T_{C,d}$. At first we extend the definition of tree-indexed Markov chains, which is put forward by Benjamini and Peres([1]), to the case of Markov chain indexed by Cayley tree in a finite i.i.d random environment.

Definition 1.1(*T*-indexed homogeneous Markov chains(see[1])) *Let T be an infinite Cayley tree, $\{X_t, t \in T\}$ be a stochastic process defined on probability space $(\Omega, \mathcal{F}, \boldsymbol{P})$ and with a finite state space $\mathcal{X}$. Let*

$$p = \{p(i), i \in \mathcal{X}\} \tag{1.1}$$

2010 *Mathematics Subject Classification*. 60C05, 05C80.
Keywords. Markov chain, Cayley tree, random environment, strong law of large number, Shannon-McMillan theorem.
Received: 15 September 2014; Accepted: 03 April 2015
Communicated by Svetlana Janković
Research supported by Natural Science Foundation of Zhejiang Province (Grant No.LY17A010013), National Natural Science Foundation of China(Grant No.11201344), Wenzhou Bureau of science and technology(Grant No.R20160005)
Email address: huilin_huang@sjtu.org (Huilin Huang)

be a distribution on $\mathcal{X}$, and

$$A = (a(i,j)), \quad i, j \in \mathcal{X} \tag{1.2}$$

be a transition probability matrix on $\mathcal{X}^2$. For every vertex $t \in T \setminus \{0\}$, suppose that

$$\begin{aligned} &\boldsymbol{P}(X_t = j | X_{1_t} = i \ \text{ and } \ X_s = x_s \ \text{ for } \ t \wedge s \leq {}^1t) \\ &= \boldsymbol{P}(X_t = j | X_{1_t} = i) = a(i,j)) \ \ \forall i, j \in \mathcal{X}, \end{aligned} \tag{1.3}$$

and

$$\boldsymbol{P}(X_0 = i) = p(i) \ \ \forall i \in \mathcal{X}.$$

Thus we call $\{X_t, t \in T\}$ to be an $\mathcal{X}$-valued homogeneous Markov chain indexed by infinite Cayley tree with the initial distribution (1.1) *and transition probability matrix A whose elements are determined by* (1.3)*.*

The subject of tree-indexed processes in deterministic environments has been deeply studied and made abundant achievements. Benjamini and Peres ([1]) have given the notion of the tree-indexed Markov chains and studied the recurrence and ray-recurrence for them. Berger and Ye ([2]) have studied the existence of entropy rate for some stationary random fields on a homogeneous tree. Ye and Berger (see [3],[4]) by using Pemantle's result([5]) and a combinatorial approach, have studied the Shannon-McMillan theorem with convergence in probability for a PPG-invariant and ergodic random field on a homogeneous tree. Yang and Liu ([6]) and Yang([7]) have studied a strong law of large numbers for Markov chains fields on a homogeneous tree (a particular case of tree-indexed Markov chains and PPG-invariant random fields). Yang and Ye([8]) have established the Shannon-McMillan theorem with convergence almost surely for nonhomogeneous Markov chains on a homogeneous tree. Huang and Yang (see [9]) has studied the Shannon-McMillan theorem in the sense of almost surely for finite homogeneous Markov chains indexed by a uniformly bounded infinite tree.

All above results are concentrated on the model of Markov chain indexed by trees in deterministic environments. In this article, we want to know the story of the Markov chain indexed by trees in a finite state i.i.d. random environment and mainly establish the strong law of large numbers and Shannon-McMillan theorem with convergence almost surely for this model. Our idea is inspired by the popular work of Cogburn([10]). The result of this paper is achieved by refining the method of [6-9], where a sequence of strong law of numbers and Shannon-McMillan theorem were proved for Markov chains indexed by trees with finite states in deterministic enviroment. In fact, our present outcomes can imply the similar results in the work [7].

The rest of this article is organized as follows. In the following of this section we formulate a model of Markov chain indexed by trees in a finite state i.i.d. random environment. In section 2 we construct a non-negative martingale in Lemma 2.1 and prove a strong limit theorem based on Doob martingale convergence theorem in Lemma 2.2. In section 3, we study the strong law of large numbers for Markov chain indexed by a homogeneous tree in the finite i.i.d. random environment by using Lemma 2.2. At last, we establish the Shannon-McMillan theorem for finite Markov chains indexed by a homogeneous tree with such finite i.i.d random environment in section 4.

Definition 1.2 *Let T be an infinite Cayley tree, both $\mathcal{X}$ and Θ two finite state spaces, and $\{\xi_t, t \in T\}$ be a Θ-valued random field indexed by T. For every vertex $t \in T \setminus \{0\}$, if*

$$\begin{aligned} &\boldsymbol{P}(X_t = j | X_{1_t} = i, \ \text{ and } \ X_s = x_s, \ \text{ for } \ t \wedge s \leq {}^1t; \xi_l, l \in T) \\ &= P_{\xi_{1_t}}(i,j), \qquad\quad a.s. \end{aligned} \tag{1.4}$$

and

$$\boldsymbol{P}(X_0 = i | \xi_l, l \in T) = P(i|\xi_0) \ \ \forall i \in \mathcal{X}, a.s. \tag{1.5}$$

For each $i, j \in \mathcal{X}$, $P_x = (P_x(i,j))_{i,j \in \mathcal{X}}, x \in \Theta$ is a family of stochastic matrices, then $\{X_t, t \in T\}$ is called to be a Markov chain indexed by tree T in a random environment $\{\xi_t, t \in T\}$. The ξ_t's are called the environmental process or control *process indexed by tree T. Moreover, if $\{\xi_t, t \in T\}$ is a T-indexed i.i.d random processes, then we call $\{X_t, t \in T\}$ to be a Markov chain indexed by tree T in an i.i.d random environment.*

Beginning in Section 2 we will assume $\xi = \{\xi_t, t \in T\}$ is a T-indexed i.i.d random fields taking values in $\Theta^T = \{0, 1, \cdots, r-1\}^T$ with the distribution as follows:

$$P(\xi_t = \theta) = \Lambda(\theta). \tag{1.6}$$

We also denote the probability of going from i to j in one step in the θth environment by $P_\theta(i, j)$. In this case it is easy to derive that $\{(\xi_t, X_t), t \in T\}$ is a Markov chain indexed by T with initial distribution $q = (q(\theta, i))$ and one-step transition function on $\Theta \times \mathcal{X}$ determined by

$$P(\alpha, i; \beta, j) = \Lambda(\beta) P_\alpha(i, j), \tag{1.7}$$

where $q(\theta, i) = \mathbf{P}(\xi_0 = \theta, X_0 = i)$. Then we call $\{(\xi_t, X_t), t \in T\}$ to be the bichain indexed by tree T. We also denote the distribution vector of ξ_t by

$$\Lambda = (\Lambda(0), \Lambda(1), \cdots, \Lambda(r-1)). \tag{1.8}$$

For simplicity we always suppose that the component of Λ is positive, that is, for any $\theta \in \Theta$, $\Lambda(\theta) > 0$.
Remark: If

$$P(\xi_t = \theta) = \Lambda(\theta) = 1,$$

then our model of Markov chain indexed by homogeneous tree in a finite state i.i.d. environment is reduced to the model of Markov chain indexed by homogeneous trees(see [7]).

2. Some Useful Statements

Let $\{X_t, t \in T\}$ be a Markov chain indexed by an infinite Cayley tree T in a finite state i.i.d. random environment $\{\xi_t, t \in T\}$, which is defined as definition 1.2. Let $g_t(\alpha, i, \beta, j)$ be functions defined on $(\Theta \times \mathcal{X})^2$. Let λ be a real number,$L_0 = \{0\}$. For every finite $n \in \mathbf{N}$, $\mathcal{F}_n = \sigma(\xi^{T^{(n)}}, X^{T^{(n)}})$, now we define a stochastic sequence as follows:

$$\varphi_n(\lambda, \omega) = \frac{e^{\lambda \sum_{t \in T^{(n)} \setminus \{0\}} g_t(\xi_{1_t}, X_{1_t}, \xi_t, X_t)}}{\prod_{t \in T^{(n)} \setminus \{0\}} E[e^{\lambda g_t(\xi_{1_t}, X_{1_t}, \xi_t, X_t)} | \xi_{1_t}, X_{1_t}]}. \tag{2.1}$$

At first we come to prove the following fact.
Lemma 2.1 *$\{\varphi_n(\lambda, \omega), \mathcal{F}_n, n \geq 1\}$ is a nonnegative martingale.*
Proof of Lemma 2.1**:** Because of the fact that $\{(\xi_t, X_t), t \in T\}$ is a bichain indexed by tree T, it is easy to see

$$\mathbf{P}(\xi^{T^{(n)}} = \alpha^{T^{(n)}}, X^{T^{(n)}} = x^{T^{(n)}}) = q(\alpha_0, x_0) \prod_{t \in T^{(n)} \setminus \{0\}} P(\alpha_{1_t}, x_{1_t}; \alpha_t, x_t), \tag{2.2}$$

then, obviously we have

$$\begin{aligned} & \mathbf{P}(\xi^{L_n} = \alpha^{L_n}, X^{L_n} = x^{L_n} | \xi^{T^{(n-1)}} = \alpha^{T^{(n-1)}}, X^{T^{(n-1)}} = x^{T^{(n-1)}}) \\ = & \frac{\mathbf{P}(\xi^{T^{(n)}} = \alpha^{T^{(n)}}, X^{T^{(n)}} = x^{T^{(n)}})}{\mathbf{P}(\xi^{T^{(n-1)}} = \alpha^{T^{(n-1)}}, X^{T^{(n-1)}} = x^{T^{(n-1)}})} \\ = & \prod_{t \in L_n} \mathbf{P}(\xi_t = \alpha_t, X_t = x_t | \xi_{1_t} = \alpha_{1_t}, X_{1_t} = x_{1_t}). \end{aligned} \tag{2.3}$$

Furthermore, we have

$$\begin{aligned}
& E[e^{\lambda\sum_{t\in L_n} g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)}|\mathcal{F}_{n-1}] \\
= & \sum_{\alpha^{L_n},x^{L_n}} e^{\lambda\sum_{t\in L_n} g_t(\xi_{1_t},X_{1_t},\alpha_t,x_t)}\mathbf{P}(\xi^{L_n}=\alpha^{L_n},X^{L_n}=x^{L_n}|\xi^{T^{(n-1)}},X^{T^{(n-1)}}) \\
= & \sum_{\alpha^{L_n},x^{L_n}}\prod_{t\in L_n} e^{\lambda g_t(\xi_{1_t},X_{1_t},\alpha_t,x_t)}\mathbf{P}(\xi_t=\alpha_t,X_t=x_t|\xi_{1_t},X_{1_t}) \\
= & \prod_{t\in L_n}\sum_{(\alpha_t,x_t)\in\Theta\times\mathcal{X}} e^{\lambda g_t(\xi_{1_t},X_{1_t},\alpha_t,x_t)}\mathbf{P}(\xi_t=\alpha_t,X_t=x_t|\xi_{1_t},X_{1_t}) \\
= & \prod_{t\in L_n} E[e^{\lambda g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)}|\xi_{1_t},X_{1_t}]\quad a.s..
\end{aligned}\tag{2.4}$$

On the other hand, we also have

$$\varphi_n(\lambda,\omega)=\varphi_{n-1}(\lambda,\omega)\frac{e^{\lambda\sum_{t\in L_n} g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)}}{\prod_{t\in L_n} E[e^{\lambda g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)}|\xi_{1_t},X_{1_t}]}.\tag{2.5}$$

Combining (2.4) and (2.5), we get

$$E[\varphi_n(\lambda,\omega)|\mathcal{F}_{n-1}]=\varphi_{n-1}(\lambda,\omega)\quad a.s..$$

Thus we complete the proof of Lemma 2.1.

Lemma 2.2 *Let $\{X_t,t\in T\}$ be a Markov chain indexed by an infinite Cayley tree T in a finite state i.i.d. random environment $\{\xi_t,t\in T\}$. $\{g_t(\alpha,i,\beta,j),t\in T\}$ are functions defined as above, denote*

$$R_n(\omega)=\sum_{t\in T^{(n)}\setminus\{0\}} E[g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)|\xi_{1_t},X_{1_t}],\tag{2.6}$$

Let $b>0$, denote

$$M(\omega)=\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{0\}} E[g_t^2(\xi_{1_t},X_{1_t},\xi_t,X_t)e^{b|g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)|}|\xi_{1_t},X_{1_t}],\tag{2.7}$$

suppose that

$$D(b)=\{\omega:M(\omega)<\infty\}\tag{2.8}$$

and

$$H_n(\omega)=\sum_{t\in T^{(n)}\setminus\{0\}} g_t(\xi_{1_t},X_{1_t},\xi_t,X_t).\tag{2.9}$$

Then we get

$$\lim_{n\to\infty}\frac{H_n(\omega)-R_n(\omega)}{|T^{(n)}|}=0\qquad a.s. on\ D(b).\tag{2.10}$$

Proof: By **Lemma 2.1**, we have known that $\{\varphi_n(\lambda,\omega),\mathcal{F}_n,n\geq 1\}$ is a nonnegative martingale. According to Doob martingale convergence theorem , we have

$$\lim_n\varphi_n(\lambda,\omega)=\varphi(\lambda,\omega)<\infty\qquad a.s.,$$

which implies that

$$\limsup_{n\to\infty} \frac{\ln \varphi_n(\lambda,\omega)}{|T^{(n)}|} \le 0 \quad a.s. . \tag{2.11}$$

Let $H_n(\omega)$ be defined as (2.9), combining(2.1) and (2.11), we arrive at

$$\limsup_{n\to\infty} \frac{1}{|T^{(n)}|}\{\lambda H_n(\omega) - \sum_{t\in T^{(n)}\setminus\{0\}} \ln[E[e^{\lambda g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)}|\xi_{1_t},X_{1_t}]]\} \le 0 \;\; a.s. \tag{2.12}$$

Let $\lambda > 0$. Dividing two sides of above inequality by λ, we get

$$\limsup_{n\to\infty} \frac{1}{|T^{(n)}|}\{H_n(\omega) - \sum_{t\in T^{(n)}\setminus\{0\}} \frac{\ln[E[e^{\lambda g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)}|\xi_{1_t},X_{1_t}]]}{\lambda}\} \le 0 \;\; a.s. \tag{2.13}$$

For case $0 < \lambda \le b$, combining with (2.13), the inequalities $\ln x \le x - 1 (x > 0)$ and $0 \le e^x - 1 - x \le 2^{-1}x^2 e^{|x|}$, it follows that

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{|T^{(n)}|}[H_n(\omega) - \sum_{t\in T^{(n)}\setminus\{0\}} E[g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)|\xi_{1_t},X_{1_t}]] \\
\le\;& \limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{0\}} \{\frac{\ln[E[e^{\lambda g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)}|\xi_{1_t},X_{1_t}]]}{\lambda} - E[g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)|\xi_{1_t},X_{1_t}]\} \\
\le\;& \limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{0\}} \{\frac{E[e^{\lambda g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)}|\xi_{1_t},X_{1_t}] - 1}{\lambda} - E[g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)|\xi_{1_t},X_{1_t}]\} \\
\le\;& \frac{\lambda}{2} \limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{0\}} E[g_t^2(\xi_{1_t},X_{1_t},\xi_t,X_t)e^{\lambda|g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)|}|\xi_{1_t},X_{1_t}] \\
\le\;& \frac{\lambda}{2} \limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{0\}} E[g_t^2(\xi_{1_t},X_{1_t},\xi_t,X_t)e^{b|g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)|}|\xi_{1_t},X_{1_t}] \\
\le\;& \frac{\lambda}{2} M(\omega) \quad a.s. on \;\; D(b).
\end{aligned} \tag{2.14}$$

Letting $\lambda \to 0^+$ in (2.14), combining with (2.6) we have

$$\limsup_{n\to\infty} \frac{H_n(\omega) - R_n(\omega)}{|T^{(n)}|} \le 0 \quad a.s. on \;\; D(b) \tag{2.15}$$

Let $-b \le \lambda < 0$. Similarly to the analysis of the case $0 < \lambda \le b$. It follows from (2.12) that

$$\liminf_{n\to\infty} \frac{H_n(\omega) - R_n(\omega)}{|T^{(n)}|} \ge \frac{\lambda}{2} M(\omega) \;\; a.s. on \;\; D(b).$$

Letting $\lambda \to 0^-$, we can arrive at

$$\liminf_{n\to\infty} \frac{H_n(\omega) - R_n(\omega)}{|T^{(n)}|} \ge 0 \quad a.s. on \;\; D(b). \tag{2.16}$$

Combining (2.15) and (2.16), it is easy to see that the conclusion (2.10) is true.

3. Strong Law of Large Numbers

For every finite $n \in \mathbf{N}$, let $\{X_t, t \in T\}$ be a Markov chain indexed by an infinite Cayley tree T in the finite state i.i.d. random environment $\{\xi_t, t \in T\}$. Now we define several stochastic sequences as follows:

$$S_n(\alpha, i) = \sum_{t \in T^{(n)}} \delta_\alpha(\xi_t)\delta_i(X_t) \qquad \forall(\alpha, i) \in \Theta \times \mathcal{X}; \tag{3.1}$$

$$S_n(i) = \sum_{t \in T^{(n)}} \delta_i(X_t) \qquad \forall i \in \mathcal{X}; \tag{3.2}$$

$$S_n(i, j) = \sum_{t \in T^{(n)} \setminus \{0\}} \delta_i(X_{1_t})\delta_j(X_t) \qquad \forall(i, j) \in \mathcal{X}^2; \tag{3.3}$$

$$S_n(\alpha, i, j) = \sum_{t \in T^{(n)} \setminus \{0\}} \delta_\alpha(\xi_{1_t})\delta_i(X_{1_t})\delta_j(X_t) \qquad \forall(\alpha, i, j) \in \Theta \times \mathcal{X}^2; \tag{3.4}$$

$$S_n(\alpha, i, \theta) = \sum_{t \in T^{(n)} \setminus \{0\}} \delta_\alpha(\xi_{1_t})\delta_i(X_{1_t})\delta_\theta(\xi_t) \qquad \forall(i, \alpha, \theta) \in \Theta \times \mathcal{X} \times \Theta; \tag{3.5}$$

$$S_n(\alpha, \theta, j) = \sum_{t \in T^{(n)} \setminus \{0\}} \delta_\alpha(\xi_{1_t})\delta_\theta(\xi_t)\delta_j(X_t) \qquad \forall(\alpha, \theta, j) \in \Theta^2 \times \mathcal{X}; \tag{3.6}$$

$$S_n(i, \theta, j) = \sum_{t \in T^{(n)} \setminus \{0\}} \delta_i(X_{1_t})\delta_\theta(\xi_t)\delta_j(X_t) \qquad \forall(i, \theta, j) \in \mathcal{X} \times \Theta \times \mathcal{X}; \tag{3.7}$$

$$S_n(\alpha, i, \theta, j) = \sum_{t \in T^{(n)} \setminus \{0\}} \delta_\alpha(\xi_{1_t})\delta_i(X_{1_t})\delta_\theta(\xi_t)\delta_j(X_t) \qquad \forall(\alpha, i, \theta, j) \in \Theta \times \mathcal{X} \times \Theta \times \mathcal{X}, \tag{3.8}$$

here and thereafter $\delta_\cdot(\cdot)$ denotes the Kronecker function. What we are interested in are the strong limit laws of those random sequences which are defined as above.

In the rest of this article, we always suppose that the following **ergodic condition** holds: the stochastic matrix $P = (P(\alpha, i; \beta, j))$ is ergodic, and $\pi = (\pi(\alpha, i))_{(\alpha,i) \in \Theta \times \mathcal{X}}$ is the stationary distribution which is determined by P. If the stochastic matrix $P = P(\alpha, i; \beta, j)$ is ergodic, and $\pi = (\pi(\alpha, i))_{(\alpha,i) \in \Theta \times \mathcal{X}}$ is the stationary distribution which is determined by P, then we have

$$\pi P = \pi; \Sigma_{\alpha \in \Theta, i \in \mathcal{X}} \pi(\alpha, i) = 1.$$

Thus, combining with (1.7) and above relations, it follows that

$$\Sigma_{\alpha \in \Theta, i \in \mathcal{X}} \pi(\alpha, i) P(\alpha, i; \beta, j) = \pi(\beta, j).$$

That is

$$\Sigma_{\alpha \in \Theta, i \in \mathcal{X}} \pi(\alpha, i) P_\alpha(i, j) = \pi(\beta, j)/\Lambda(\beta). \tag{3.9}$$

Since the left hand side is independent of the factor β, then we can denote

$$\mu(j) =: \pi(\beta, j)/\Lambda(\beta).$$

where $\mu(j)$ is also independent of β. In a word, for any $\beta \in \Theta$ and $j \in \mathcal{X}$, we have

$$\pi(\beta, j) = \mu(j)\Lambda(\beta). \tag{3.10}$$

It is easy to derive that $\mu = (u(i), i \in \mathcal{X})$ is a probability distribution. Let's re-substitute (3.10) into (3.9), we get to

$$\Sigma_{i\in\mathcal{X}}\mu(i)\Sigma_{\alpha\in\Theta}\Lambda(\alpha)P_\alpha(i, j) = \mu(j). \tag{3.11}$$

Thus we might as well denote

$$a(i, j) =: \Sigma_{\alpha\in\Theta}\Lambda(\alpha)P_\alpha(i, j). \tag{3.12}$$

Then, (3.11) is equivalent to the following relation

$$\Sigma_{i\in\mathcal{X}}\mu(i)a(i, j) = \mu(j). \tag{3.13}$$

Since μ is a probability distribution vector, then it is the stationary distribution of stochastic matrix $A = (a(i, j))$.

Theorem 3.1 *Let $\{X_t, t \in T\}$ be a Markov chain indexed by an infinite Cayley tree T in an i.i.d environment $\{\xi_t, t \in T\}$ which is defined as definition 1.2. Suppose that the stochastic matrix $P = (P(\alpha, i; \beta, j))$ satisfies the* ***ergodic condition****, $S_n(\alpha, i)$ and $S_n(\alpha, i, \theta, j)$ are defined as* (3.1) *and* (3.8)*, then we have*

$$\lim_{n\to\infty} \frac{S_n(\alpha, i)}{|T^{(n)}|} = \Lambda(\alpha)\mu(i) \qquad \forall(\alpha, i) \in \Theta\times\mathcal{X}; \ \ a.s., \tag{3.14}$$

$$\lim_{n\to\infty} \frac{S_n(\alpha, i, \theta, j)}{|T^{(n)}|} = \Lambda(\alpha)\Lambda(\theta)\mu(i)P_\alpha(i, j) \qquad \forall(\alpha, i, \theta, j) \in \Theta\times\mathcal{X}\times\Theta\times\mathcal{X}, \quad a.s.. \tag{3.15}$$

Proof of Theorem 3.1: At first, we come to prove (3.14). In **Lemma 2.2**, we can take

$$g_t(\xi_{1_t}, X_{1_t}, \xi_t, X_t) = \delta_\alpha(\xi_t)\delta_i(X_t),$$

obviously in (2.8) we have $D(b) = \Omega$, then we get

$$\begin{aligned} H_n(\omega) &= \Sigma_{t\in T^{(n)}\setminus\{0\}} g_t(\xi_{1_t}, X_{1_t}, \xi_t, X_t) \\ &= \Sigma_{t\in T^{(n)}\setminus\{0\}} \delta_\alpha(\xi_t)\delta_i(X_t) \\ &= S_n(\alpha, i) - \delta_\alpha(\xi_0)\delta_i(X_0) \end{aligned} \tag{3.16}$$

$$\begin{aligned} R_n(\omega) &= \sum_{t\in T^{(n)}\setminus\{0\}} E[g_t(\xi_{1_t}, X_{1_t}, \xi_t, X_t)|\xi_{1_t}, X_{1_t}] \\ &= \sum_{t\in T^{(n)}\setminus\{0\}} \sum_{\theta\in\Theta}\sum_{j\in\mathcal{X}} \delta_\alpha(\theta)\delta_i(j)P(\xi_{1_t}, X_{1_t}; \theta, j) \\ &= \sum_{t\in T^{(n)}\setminus\{0\}} P(\xi_{1_t}, X_{1_t}; \alpha, i) \\ &= \sum_{\theta\in\Theta}\sum_{j\in\mathcal{X}}\sum_{t\in T^{(n)}\setminus\{0\}} \delta_\theta(\xi_{1_t})\delta_j(X_{1_t})P(\theta, j; \alpha, i) \\ &= d\sum_{\theta\in\Theta}\sum_{j\in\mathcal{X}} S_{n-1}(\theta, j)P(\theta, j; \alpha, i) + P(\xi_0, X_0; \alpha, i) \end{aligned} \tag{3.17}$$

Noting that $\lim_{n\to\infty}\frac{|T^{(n)}|}{|T^{(n-1)}|}=d$, combining (2.10), (3.16) and (3.17), we easily arrive at

$$\begin{aligned}&\lim_{n\to\infty}\frac{H_n(\omega)-R_n(\omega)}{|T^{(n)}|}\\=\ &\lim_{n\to\infty}\{\frac{S_n(\alpha,i)}{|T^{(n)}|}-\frac{\sum_{\theta\in\Theta}\sum_{j\in\mathcal{X}}S_{n-1}(\theta,j)P(\theta,j;\alpha,i)}{|T^{(n-1)}|}\}=0\ \ a.s..\end{aligned}\tag{3.18}$$

Now we come to prove the equation (3.14) followed by the procedure of induction to (3.18). Multiplying both sides of equality (3.18) by $P(\alpha,i;\beta,l)$ and adding them together, by using (3.18) again, we get

$$\begin{aligned}&\lim_{n\to\infty}[\sum_{\alpha\in\Theta}\sum_{i\in\mathcal{X}}\frac{S_n(\alpha,i)}{|T^{(n)}|}P(\alpha,i;\beta,l)-\frac{S_{n+1}(\beta,l)}{|T^{(n+1)}|}]\\&+\lim_{n\to\infty}[\frac{S_{n+1}(\beta,l)}{|T^{(n+1)}|}-\sum_{\alpha\in\Theta}\sum_{i\in\mathcal{X}}\sum_{\theta\in\Theta}\sum_{j\in\mathcal{X}}\frac{S_{n-1}(\theta,j)P(\theta,j;\alpha,i)P(\alpha,i;\beta,l)}{|T^{(n-1)}|}]\\=\ &\lim_{n\to\infty}\{\frac{S_{n+1}(\beta,l)}{|T^{(n+1)}|}-\sum_{\theta\in\Theta}\sum_{j\in\mathcal{X}}\frac{S_{n-1}(\theta,j)P^{(2)}(\theta,j;\beta,l)}{|T^{(n-1)}|}\}=0\ \ a.s.,\end{aligned}\tag{3.19}$$

where $P^{(N)}(\theta,j;\beta,l)$ is the $N-$step transition probability determined by the transition matrix P. By induction we have

$$\lim_{n\to\infty}\{\frac{S_{n+N}(\beta,l)}{|T^{(n+N)}|}-\sum_{\theta\in\Theta}\sum_{j\in\mathcal{X}}\frac{S_{n-1}(\theta,j)P^{(N+1)}(\theta,j;\beta,l)}{|T^{(n-1)}|}\}=0.\ \ a.s.\tag{3.20}$$

Since

$$\lim_{N\to\infty}P^{(N+1)}(\theta,j;\beta,l)=\pi(\beta,l),\qquad\qquad \sum_{\theta\in\Theta}\sum_{j\in\mathcal{X}}S_{n-1}(\theta,j)=|T^{(n-1)}|,\tag{3.21}$$

then, (3.14) is followed from (3.20), (3.21) and (3.10).

Now we come to prove the equality (3.15). By using **Lemma 2.2** again, taking

$$g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)=\delta_\alpha(\xi_{1_t})\delta_i(X_{1_t})\delta_\theta(\xi_t)\delta_j(X_t),$$

then we get

$$\begin{aligned}H_n(\omega)\ &=\ \Sigma_{t\in T^{(n)}\setminus\{0\}}g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)\\&=\ \Sigma_{t\in T^{(n)}\setminus\{0\}}\delta_\alpha(\xi_{1_t})\delta_i(X_{1_t})\delta_\theta(\xi_t)\delta_j(X_t)\\&=\ S_n(\alpha,i,\theta,j),\end{aligned}\tag{3.22}$$

$$\begin{aligned}R_n(\omega)\ &=\ \sum_{t\in T^{(n)}\setminus\{0\}}E[g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)|\xi_{1_t},X_{1_t}]\\&=\ \sum_{t\in T^{(n)}\setminus\{0\}}\sum_{\beta\in\Theta}\sum_{l\in\mathcal{X}}\delta_\alpha(\xi_{1_t})\delta_i(X_{1_t})\delta_\theta(\beta)\delta_j(l)P(\xi_{1_t},X_{1_t};\beta,l)\\&=\ \sum_{t\in T^{(n)}\setminus\{0\}}\delta_\alpha(\xi_{1_t})\delta_i(X_{1_t})P(\alpha,i;\theta,j)\\&=\ dS_{n-1}(\alpha,i)P(\alpha,i;\theta,j)+\delta_\alpha(\xi_0)\delta_i(X_0)P(\alpha,i;\theta,j)\end{aligned}\tag{3.23}$$

Combining the equalities (3.14), (3.22) and (3.23) and by using $\lim_{n\to\infty}\frac{|T^{(n)}|}{|T^{(n-1)}|}=d$ and (2.10) again, we conclude that (3.15) is true. The proof of this theorem is completed.

Corollary 3.2 *Under the same conditions of **Theorem 3.1**, we have*

$$\lim_{n\to\infty}\frac{S_n(i)}{|T^{(n)}|}=\mu(i)\qquad \forall i\in\mathcal{X}; \tag{3.24}$$

$$\lim_{n\to\infty}\frac{S_n(i,j)}{|T^{(n)}|}=\mu(i)a(i,j)\qquad \forall (i,j)\in\mathcal{X}^2; \tag{3.25}$$

$$\lim_{n\to\infty}\frac{S_n(\alpha,i,j)}{|T^{(n)}|}=\Lambda(\alpha)\mu(i)P_\alpha(i,j)\qquad \forall (\alpha,i,j)\in\Theta\times\mathcal{X}^2; \tag{3.26}$$

$$\lim_{n\to\infty}\frac{S_n(\alpha,i,\theta)}{|T^{(n)}|}=\Lambda(\alpha)\mu(i)\Lambda(\theta)\qquad \forall (\alpha,i,\theta)\in\Theta\times\mathcal{X}\times\Theta; \tag{3.27}$$

$$\lim_{n\to\infty}\frac{S_n(\alpha,\theta,j)}{|T^{(n)}|}=\Lambda(\alpha)\Lambda(\theta)\sum_{i\in\mathcal{X}}\mu(i)P_\alpha(i,j)\qquad \forall (\alpha,\theta,j)\in\Theta^2\times\mathcal{X}; \tag{3.28}$$

$$\lim_{n\to\infty}\frac{S_n(i,\theta,j)}{|T^{(n)}|}=\Lambda(\theta)\mu(i)a(i,j)\qquad \forall (i,\theta,j)\in\mathcal{X}\times\Theta\times\mathcal{X}. \tag{3.29}$$

All above equations holds under the sense of convergence almost surely .

Proof of Corollary 3.2: Since $S_n(i)=\sum_{\alpha\in\Theta}S_n(\alpha,i)$, then (3.24) can be derived from (3.14) directly. Similarly, $S_n(i,j)=\sum_{\alpha\in\Theta,\theta\in\Theta}S_n(\alpha,i,\theta,j)$, we can easily arrive at (3.25) by using (3.15) and (3.12). At last, (3.26) – (3.29) can also be derived from (3.15) without any difficulty.

4. Shannon-McMillan Theorem

Let T be an infinite Cayley tree and $\{X_t, t\in T\}$ a Markov chain indexed by tree T in a random environment $\{\xi_t, t\in T\}$. Since $\{\xi_t, X_t, t\in T\}$ is a bichain indexed by tree T. Now we define

$$\begin{aligned} f_n(\omega) &= -\frac{1}{|T^{(n)}|}\ln \mathbf{P}(\xi^{T^{(n)}}, X^{T^{(n)}}) \\ &= -\frac{1}{|T^{(n)}|}[\ln q(\xi_0,X_0)+\sum_{t\in T^{(n)}\setminus\{0\}}\ln \mathbf{P}(\xi_t,X_t|\xi_{1_t},X_{1_t})] \end{aligned} \tag{4.1}$$

The convergence of $f_n(\omega)$ to a constant in a sense (L_1 convergence, convergence in probability, a.s. convergence) is called the **Shannon-McMillan theorem** in information theory.

Theorem 4.1 *Let $\{X_t, t\in T\}$ be a Markov chain indexed by an infinite Cayley tree T in an i.i.d environment $\{\xi_t, t\in T\}$ which is defined as definition 1.2. Suppose that the stochastic matrix $P=(P(\alpha,i;\beta,j))$ satisfies the **ergodic condition** and $q(\alpha,i)>0$ for all $(\alpha,i)\in\Theta\times\mathcal{X}$, then we have*

$$\lim_{n\to\infty}f_n(\omega)=-\sum_{\alpha\in\Theta}\sum_{i\in\mathcal{X}}\sum_{\theta\in\Theta}\sum_{j\in\mathcal{X}}\Lambda(\alpha)\Lambda(\theta)\mu(i)P_\alpha(i,j)\ln P(\alpha,i;\theta,j)\qquad a.s.. \tag{4.2}$$

Proof of Theorem 4.1: At first, we assert that

$$\lim_{n\to\infty}\{f_n(\omega)+\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{0\}}\sum_{\alpha\in\Theta}\sum_{i\in\mathcal{X}}\sum_{\theta\in\Theta}\sum_{j\in\mathcal{X}}\delta_\alpha(\xi_{1_t})\delta_i(X_{1_t})P(\alpha,i;\theta,j)\ln P(\alpha,i;\theta,j)\}=0 \qquad a.s.. \tag{4.3}$$

Taking $b=\frac{1}{2}, g_t(\xi_{1_t},X_{1_t},\xi_t,X_t)=-\ln P(\xi_t,X_t|\xi_{1_t},X_{1_t})$ in **Lemma 2.2**, by using the elementary inequality $(\ln x)^2x^{\frac{1}{2}}\le 16e^{-2}$ for every $0<x\le 1$, then we have

$$\begin{aligned}
M(\omega) &= \limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{0\}}\sum_{\alpha\in\Theta,i\in\mathcal{X}}(\ln P(\xi_{1_t},X_{1_t};\alpha,i))^2e^{\frac{1}{2}|\ln P(\xi_{1_t},X_{1_t};\alpha,i)|}P(\xi_{1_t},X_{1_t};\alpha,i)\\
&= \limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{0\}}\sum_{\alpha\in\Theta,i\in\mathcal{X}}(\ln P(\xi_{1_t},X_{1_t};\alpha,i))^2(P(\xi_{1_t},X_{1_t};\alpha,i))^{\frac{1}{2}}\\
&\le \limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{0\}}\sum_{\alpha\in\Theta,i\in\mathcal{X}}16e^{-2}
\end{aligned}$$

Noting that both Θ and $\mathcal{X}$ are finite sets, thus it follows that

$$M(\omega)<\infty. \tag{4.4}$$

This implies that $D(\frac{1}{2})=\Omega$ in (2.8). We also note that

$$\begin{aligned}
&\frac{H_n(\omega)-R_n(\omega)}{|T^{(n)}|}\\
=\;&\frac{\ln q(\xi_0,X_0)}{|T^{(n)}|}+f_n(\omega)\\
&+\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{0\}}\sum_{\alpha\in\Theta}\sum_{i\in\mathcal{X}}\sum_{\theta\in\Theta}\sum_{j\in\mathcal{X}}\delta_\alpha(\xi_{1_t})\delta_i(X_{1_t})P(\alpha,i;\theta,j)\ln P(\alpha,i;\theta,j)
\end{aligned} \tag{4.5}$$

Combining (4.4),(4.5) and (2.10), we conclude that (4.3) is true.

Then we have

$$\begin{aligned}
&|f_n(\omega)+\sum_{\alpha\in\Theta}\sum_{i\in\mathcal{X}}\sum_{\theta\in\Theta}\sum_{j\in\mathcal{X}}\Lambda(\alpha)\mu(i)P(\alpha,i;\theta,j)\ln P(\alpha,i;\theta,j)|\\
\le\;&|f_n(\omega)+\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}\setminus\{0\}}\sum_{\alpha\in\Theta}\sum_{i\in\mathcal{X}}\sum_{\theta\in\Theta}\sum_{j\in\mathcal{X}}\delta_\alpha(\xi_{1_t})\delta_i(X_{1_t})P(\alpha,i;\theta,j)\ln P(\alpha,i;\theta,j)|\\
&+\sum_{\alpha\in\Theta}\sum_{i\in\mathcal{X}}|\frac{dS_{n-1}(\alpha,i)+\delta_\alpha(\xi_0)\delta_i(X_0)}{|T^{(n)}|}-\Lambda(\alpha)\mu(i)|\cdot|\sum_{\theta\in\Theta}\sum_{j\in\mathcal{X}}P(\alpha,i;\theta,j)\ln P(\alpha,i;\theta,j)|\\
=:\;&I_1+I_2
\end{aligned} \tag{4.6}$$

On the one hand, I_1 disappears as n tends to infinity by (4.3) with convergence almost surely. On the other hand, by using (3.14) we have

$$\lim_{n\to\infty}\frac{dS_{n-1}(\alpha,i)+\delta_\alpha(\xi_0)\delta_i(X_0)}{|T^{(n)}|}=\lim_{n\to\infty}\frac{S_{n-1}(\alpha,i)}{|T^{(n-1)}|}=\Lambda(\alpha)\mu(i).\ \ a.s.$$

Both Θ and $\mathcal{X}$ are finite sets, it follows that I_2 approximates to zero as n tends to infinity with convergence almost surely. Thus we complete the proof of theorem 4.1.

References

[1] Benjamini, I. and Peres, Y.. Markov chains indexed by trees. *Ann Probab*, **22: 219-243** (1994)
[2] Berger, T. and Ye , Z.. Entropic aspects of random fields on trees. *IEEE Trans Inform Theory*, **36: 1006-1018**.(1990)
[3] Ye ,Z., and Berger, T.. Ergodic , regulary and asymptotic equipartition property of random fields on trees. *J. Combin Inform System Sci*, **21: 157-184**. (1996)
[4] Ye ,Z. and Berger, T., Information Measures for Discrete Random Fields. *Science, Beijing*,(1998)
[5] Pemantle, R., Antomorphism invariant measure on trees. *Ann Prob*, **20: 1549-1566** (1992.)
[6] Yang, W.G. and Liu, W.. Strong law of large numbers for Markov chains fields on a Bethe tree.*Statist Prob Lett*, **49:245-250** (2000)
[7] Yang,W.G.. Some limit properties for Markov chains indexed by a homogeneous tree. *Statist Prob Lett*, **65: 241-250**(2003)
[8] Yang,W.G. and Ye Z.X.. The asymptotic equipartition property for nonhomogeneous Markov chains indexed by homogeneous trees. *IEEE Trans Inform Theory*, **53(9):3275-3280**(2007)
[9] Huang H.L. and Yang W.G.. Strong law of large numbers for Markov chains indexed by an infinite tree with uniformly bounded degree.*Science in China Series A: Mathematics*,**51:195-202** (2008)
[10] Cogburn, R.. Markov chains in random environments: the case of Markovian environments. *Ann. Probability* ,**8: 908-916** (1980)