*Journal of Mathematical Research & Exposition*
Nov., 2010, Vol. 30, No. 6, pp. 976–984
DOI:10.3770/j.issn:1000-341X.2010.06.004
Http://jmre.dlut.edu.cn

# Strong Law of Large Numbers and Asymptotic Equipartition Probability for Nonsymmetric Markov Chain Indexed by Cayley Tree

**Yan DONG**[1,*] **Wei Guo YANG**[2]

1. *Department of Mathematics, Shanghai Jiaotong University, Shanghai* 200240*, P. R. China;*
2. *Faculty of Science, Jiangsu University, Jiangsu* 212013*, P. R. China*

**Abstract** In this paper, we study the strong law of large numbers for the frequencies of occurrence of states and ordered couples of states for nonsymmetric Markov chain (NSMC) indexed by Cayley tree with any finite states. The asymptotic equipartition properties with almost everywhere (a.e.) convergence for NSMC indexed by Cayley tree are obtained. This article generalizes a recent result.

**Keywords** strong law of large numbers; nonsymmetric Markov chain; Cayley tree; asymptotic equipartition property.

**Document code** A

**MR(2000) Subject Classification** 60F10; 60F15

**Chinese Library Classification** O211.6

## 1. Introduction

A tree is a graph $G=\{T,E\}$ which is connected and contains no circuits. Thus G is a tree if and only if given any two vertices $\sigma\neq t\in T$, there exists a unique path $\sigma=z_1,z_2,\ldots,z_m=t$ from $\sigma$ to $t$ with distinct $z_1,z_2,\ldots,z_m$. The distance between $\sigma$ and $t$ is defined to be $m-1$, the number of edges in the path connecting $\sigma$ and $t$. Select a vertex as the root (denoted by $o$). For any two vertices $\sigma$ and $t$ of tree $T$, we write $\sigma\leq t$ if $\sigma$ is on the unique path from the root $o$ to $t$. We denote by $\sigma\wedge t$ the vertex farthest from $o$ satisfying $\sigma\wedge t\leq t$ and $\sigma\wedge t\leq\sigma$. For any vertex $t$ of tree $T$, we denote by $|t|$ the distance between $o$ and $t$. The set of all vertices with distance $n$ from the root $o$ is called the $n$-th level of $T$. For any vertex $t$ of tree $T$, we denote the first predecessor of t by $1_t$, and refer $t$ to be the son of $1_t$. In this article, we only investigate the tree $T$, on which each vertex has two sons on the next level. In this case, there are two branches on the next level for each vertex. In order to distinguish them, we call them the left branch and the right branch, respectively. For any vertex $t$ of tree $T$, we call the son of t on the left branch the left vertex, and the son of $t$ on the right branch the right vertex.

Received October 27, 2008; Accepted January 18, 2010
Supported by the National Natural Science Foundation of China (Grant No. 10571076).
* Corresponding author
E-mail address: jddy2008@126.com (Y. DONG)

Denote by $T^l$ the subgraph of $T$ containing all the left vertices of tree $T$, $T^r$ the subgraph of $T$ containing all the right vertices of tree $T$. We also denote by $T^{(n)}$ the subtree comprised of level 0 (the root $o$) through level $n$, by $T_n^l$ the set of all the left vertices on $T^{(n)}$, and by $T_n^r$ the set of all the right vertices on $T^{(n)}$. In addition, we denote by $L_n$ the set of all vertices on level $n$.

Let $S$ be the subgraph of $T$, $X^S = \{X_t, t \in S\}$, and denote by $|S|$ the number of vertices of $S$, $x^S$ the realization of $X^S$.

**Definition 1** *Let $G = \{0, 1, \ldots, b-1\}$ be a finite state space, $X = \{X_t, t \in T\}$ be a collection of $G$-valued random variables defined on probability space $(\Omega, \mathcal{F}, P)$. Let*

$$p = (p(x), x \in G) \tag{1}$$

*be a distribution on $G$,*

$$P_l = (P_l(y|x)), \quad x, y \in G \tag{2}$$

$$P_r = (P_r(y|x)), \quad x, y \in G \tag{3}$$

*be two stochastic matrices on $G^2$. If for any vertex $t \in T$, $\sigma_i \in T$ satisfies $\sigma_i \wedge t \leq 1_t$, $1 \leq i \leq n$, we have*

$$\begin{aligned} &P(X_t = y, t \in T^l | X_{1_t} = x, X_{\sigma_i} = x_i,\ 1 \leq i \leq n) \\ &\quad = P(X_t = y, t \in T^l | X_{1_t} = x) = P_l(y|x), \quad \forall x, y, x_1, \ldots, x_n \in G, \end{aligned} \tag{4}$$

$$\begin{aligned} &P(X_t = y, t \in T^r | X_{1_t} = x, X_{\sigma_i} = x_i,\ 1 \leq i \leq n) \\ &\quad = P(X_t = y, t \in T^r | X_{1_t} = x) = P_r(y|x), \quad \forall x, y, x_1, \ldots, x_n \in G, \end{aligned} \tag{5}$$

*and*

$$P(X_o = x) = p(x), \quad \forall x \in G \tag{6}$$

*then $X = \{X_t, t \in T\}$ will be called $G$-valued nonsymmetric Markov chain indexed by Cayley tree $T$ with the initial distribution (1) and the transition matrices (2) and (3), or called $T$-indexed nonsymmetric Markov chain with state space $G$.*

The above definition is the extensions of the definition of $T$-indexed Markov chain [2] and the definition of nonsymmetric Markov chain on Cayley tree [3].

The subject of tree-indexed processes is rather young. Benjamini and Peres [2] gave the notion of the tree-indexed Markov chains and studied the recurrence and ray-recurrence for them. Berger and Ye [3] studied the existence of entropy rate for some stationary random fields on a homogeneous tree. Pemantle [5] proved a mixing property and a weak law of large numbers for a PPG-invariant and ergodic random field on a homogeneous tree. Ye and Berger [8, 9], by using Pemantle's result and a combinatorial approach, studied the asymptotic equipartition property (AEP) in the sense of convergence in probability for a PPG-invariant and ergodic random field on a homogeneous tree (a particular case of tree-indexed Markov chain and PPG-invariant random fields). Yang and Liu [6] studied a strong law of large numbers for the frequency of occurrence of states for Markov chains field on a homogeneous tree (a particular case of tree-indexed Markov chains and PPG-invariant random fields). Yang [7] studied some strong limit theorems for countable homogeneous Markov chains indexed by a homogeneous tree and the strong law of large numbers and the asymptotic equipartition property (AEP) for finite homogeneous Markov chains indexed by a homogeneous tree. Recently, Huang and Yang [4] studied strong law of large numbers and asymptotic equipartition property (AEP) for Markov chains indexed by an infinite tree with uniformly bounded degree. Bao and Ye [1] studied strong law of large numbers and asymptotic equipartition property (AEP) for nonsymmetric Markov chain fields on Cayley tree.

Bao and Ye [1] studied strong law of large numbers and asymptotic equipartition property (AEP) for nonsymmetric Markov chain fields on Cayley tree with two states, and in that paper, $P_l$ and $P_r$ have a unique common invariant probability measure. In this paper, we generalize the nonsymmetric Markov chain to be any finite States, and only require the stochastic matrix $P = \frac{1}{2}(P_l + P_r)$ to be ergodic. Rather different from [1], the approach used in this paper is the improvement of [7], that is, by constructing a martingale, we first study the local convergence theorem for finite nonsymmetric Markov chain indexed by Cayley tree, then obtain some limit theorems for the frequencies of occurrence of states and ordered couples of states for nonsymmetric Markov chain indexed by that tree, and finally, we obtain the strong law of large numbers and asymptotic equipartition property (AEP) with a.e. convergence for finite nonsymmetric Markov chain indexed by Cayley tree. In fact, our present outcomes can generalize the result of [1].

## 2. Strong limit theorem

**Lemma 1** *Let $X = \{X_t, t \in T\}$ be a $T$-indexed nonsymmetric Markov chain with state space $G$ as defined in Definition 1, and $\{g_t(x,y), t \in T\}$ be functions defined on $G^2$. Let $L_0 = \{o\}$, $\mathcal{F}_n = \sigma(X^{T^{(n)}})$,*

$$t_n(\lambda,\omega) = \frac{e^{\lambda \sum_{t\in T^{(n)}\setminus\{o\}} g_t(X_{1_t},X_t)}}{\prod_{t\in T^{(n)}\setminus\{o\}} E[e^{\lambda g_t(X_{1_t},X_t)}|X_{1_t}]}, \tag{7}$$

*where $\lambda$ is a real number. Then $\{t_n(\lambda,\omega), \mathcal{F}_n, n \geq 1\}$ is a nonnegative martingale.*

**Proof** Let

$$\delta^l(t) = \begin{cases} 1, & t \in T^l, \\ 0, & t \in T^r, \end{cases} \qquad \delta^r(t) = \begin{cases} 1, & t \in T^r, \\ 0, & t \in T^l. \end{cases} \tag{8}$$

Obviously, we have

$$P(x^{T^{(n)}}) = P(X^{T^{(n)}} = x^{T^{(n)}}) = P[X_o = x_o] \prod_{t\in T^{(n)}\setminus\{o\}} [P_l(x_t|x_{1_t})]^{\delta^l(t)}[P_r(x_t|x_{1_t})]^{\delta^r(t)}. \tag{9}$$

Hence

$$P(X^{L_n} = x^{L_n}|X^{T^{(n-1)}} = x^{T^{(n-1)}}) = \frac{P(x^{T^{(n)}})}{P(x^{T^{(n-1)}})} = \prod_{t\in L_n} [P_l(x_t|x_{1_t})]^{\delta^l(t)}[P_r(x_t|x_{1_t})]^{\delta^r(t)}. \tag{10}$$

Then

$$E[e^{\lambda \sum_{t\in L_n} g_t(X_{1_t},X_t)}|\mathcal{F}_{n-1}] = \sum_{x^{L_n}} e^{\lambda \sum_{t\in L_n} g_t(X_{1_t},x_t)} P(X^{L_n} = x^{L_n}|X^{T^{(n-1)}})$$

$$
\begin{aligned}
&= \sum_{x^{L_n}} e^{\lambda \sum_{t\in L_n} g_t(X_{1_t},x_t)} \prod_{t\in L_n} [P_l(x_t|X_{1_t})]^{\delta^l(t)} [P_r(x_t|X_{1_t})]^{\delta^r(t)} \\
&= \prod_{t\in L_n} \sum_{x_t} e^{\lambda g_t(X_{1_t},x_t)} [P_l(x_t|X_{1_t})]^{\delta^l(t)} [P_r(x_t|X_{1_t})]^{\delta^r(t)} \\
&= \prod_{t\in L_n} E[e^{\lambda g_t(X_{1_t},X_t)}|X_{1_t}] \quad \text{a.e. .} \qquad (11)
\end{aligned}
$$

On the other hand, we also have

$$
t_n(\lambda,\omega) = t_{n-1}(\lambda,\omega) \frac{e^{\lambda \sum_{t\in L_n} g_t(X_{1_t},X_t)}}{\prod_{t\in L_n} E[e^{\lambda g_t(X_{1_t},X_t)}|X_{1_t}]}. \qquad (12)
$$

Combining (11) and (12), we arrive at

$$
E[t_n(\lambda,\omega)|\mathcal{F}_{n-1}] = t_{n-1}(\lambda,\omega) \quad \text{a.e. .}
$$

Thus we conclude the proof of the lemma. □

**Theorem 1** *Let $X = \{X_t, t\in T\}$ and $\{g_t(x,y), t\in T\}$ be defined as Lemma 1, and*

$$
G_n(\omega) = \sum_{t\in T^{(n)}\setminus\{0\}} E[g_t(X_{1_t},X_t)|X_{1_t}], \qquad (13)
$$

*$\{a_n, n\ge 1\}$ be a sequence of nonnegative random variables. Let $\alpha > 0$. Set*

$$
B = \{\lim_{n\to\infty} a_n = \infty\}, \qquad (14)
$$

*and*

$$
D(\alpha) = \Big\{ \lim_{n\to\infty} \sup \frac{1}{a_n} \sum_{t\in T^{(n)}\setminus\{0\}} E[g_t^2(X_{1_t},X_t) e^{\alpha|g_t(X_{1_t},X_t)|}|X_{1_t}] = M(\omega) < \infty \Big\} \bigcap B, \qquad (15)
$$

$$
H_n(\omega) = \sum_{t\in T^{(n)}\setminus\{0\}} g_t(X_{1_t},X_t). \qquad (16)
$$

*Then*

$$
\lim_{n\to\infty} \frac{H_n(\omega) - G_n(\omega)}{a_n} = 0 \quad \text{a.e.} \ \ \omega \in D(\alpha). \qquad (17)
$$

**Proof** The proof of the theorem is similar to that in [4], so we omit it. □

**Corollary 1** *Let $X = \{X_t, t\in T\}$ be a $T$-indexed nonsymmetric Markov chain with state space $G$ as defined in Definition 1. If $\{g_t(x,y), t\in T\}$ are the uniformly bounded functions defined on $G^2$, then*

$$
\lim_{n\to\infty} \frac{H_n(\omega) - G_n(\omega)}{|T^{(n)}|} = 0 \quad \text{a.e. .} \qquad (18)
$$

**Proof** Let $a_n = |T^{(n)}|$ in Theorem 1. We notice that $\{g_t(x,y), t\in T\}$ are uniformly bounded, and then $D(\alpha) = \Omega$ for $\forall \alpha > 0$. This corollary follows from Theorem 1 directly. □

**Definition 2** *Let*

$$
S_k(T^{(n)}) = |\{t\in T^{(n)}; X_t = k\}|, \ S_k(T_n^l) = |\{t\in T_n^l; X_t = k\}|, \ S_k(T_n^r) = |\{t\in T_n^r; X_t = k\}|.
$$

*Obviously, we have*

$$S_k(T^{(n)}) = \sum_{t\in T^{(n)}} I_k(X_t), \tag{19}$$

$$S_k(T_n^l) = \sum_{t\in T^{(n)}\setminus\{0\}} I_k(X_t)\delta^l(t), \tag{20}$$

$$S_k(T_n^r) = \sum_{t\in T^{(n)}\setminus\{0\}} I_k(X_t)\delta^r(t), \tag{21}$$

*where*

$$I_k(i) = \begin{cases} 1, & i = k, \\ 0, & i \neq k. \end{cases} \tag{22}$$

**Theorem 2** *Let $X = \{X_t, t \in T\}$ be a T-indexed nonsymmetric Markov chain with state space $G$ as defined in Definition 1. Let $S_k(T_n^l)$ and $S_k(T_n^r)$ be defined as (20) and (21), respectively. Then we have*

$$\lim_{n\to\infty}\left\{\frac{S_k(T_n^l)}{|T^{(n)}|} - \sum_{l=0}^{b-1}\frac{S_l(T^{(n-1)})}{2|T^{(n-1)}|}P_l(k|l)\right\} = 0, \quad \text{a.e.}, \tag{23}$$

$$\lim_{n\to\infty}\left\{\frac{S_k(T_n^r)}{|T^{(n)}|} - \sum_{l=0}^{b-1}\frac{S_r(T^{(n-1)})}{2|T^{(n-1)}|}P_r(k|l)\right\} = 0, \quad \text{a.e.}\ . \tag{24}$$

**Proof** Let $g_t(x,y) = I_k(y)\delta^l(t)$ in Theorem 1. Then we have

$$H_n(\omega) = \sum_{t\in T^{(n)}\setminus\{0\}} I_k(X_t)\delta^l(t) = S_k(T_n^l), \tag{25}$$

$$\begin{aligned} G_n(\omega) &= \sum_{t\in T^{(n)}\setminus\{0\}}\sum_{x_t\in G} I_k(x_t)\delta^l(t)P(x_t|X_{1_t}) \\ &= \sum_{t\in T^{(n)}\setminus\{0\}} \delta^l(t)P_l(k|X_{1_t}) = \sum_{t\in T^{(n)}\setminus\{0\}}\sum_{l=0}^{b-1}\delta^l(t)I_l(X_{1_t})P_l(k|l) \\ &= \sum_{l=0}^{b-1}\sum_{t\in T^{(n)}\setminus\{0\}}\delta^l(t)I_l(X_{1_t})P_l(k|l) = \sum_{l=0}^{b-1}S_l(T^{(n-1)})P_l(k|l). \end{aligned} \tag{26}$$

Let $a_n = |T^{(n)}|$. Obviously $\{g_t(x,y), t \in T\}$ are uniformly bounded functions defined on $G^2$, then from Corollary 1, we have

$$\lim_{n\to\infty}\frac{1}{|T^{(n)}|}\Big[S_k(T_n^l) - \sum_{l=0}^{b-1}S_l(T^{(n-1)})P_l(k|l)\Big] = 0 \ \ \text{a.e.}\ .$$

That is

$$\lim_{n\to\infty}\left\{\frac{S_k(T_n^l)}{|T^{(n)}|} - \sum_{l=0}^{b-1}\frac{S_l(T^{(n-1)})}{2|T^{(n-1)}|}P_l(k|l)\right\} = 0 \ \ \text{a.e.}\ . \tag{27}$$

Similarly, let $g_t(x,y) = I_k(y)\delta^r(t)$. We have

$$\lim_{n\to\infty}\left\{\frac{S_k(T_n^r)}{|T^{(n)}|} - \sum_{l=0}^{b-1}\frac{S_l(T^{(n-1)})}{2|T^{(n-1)}|}P_r(k|l)\right\} = 0 \ \ \text{a.e.}\ . \tag{28}$$

Thus we have completed the proof of this theorem. □

## 3. Strong law of large numbers and Shannon-McMillan theorem

**Theorem 3** *Let $X = \{X_t, t \in T\}$ be a $T$-indexed nonsymmetric Markov chain with state space $G$ as defined in Definition 1. Let $S_k(T^{(n)})$ be defined as (19). Suppose that stochastic matrix $P = \frac{1}{2}(P_l + P_r)$ is ergodic. Then we have*

$$\lim_{n\to\infty} \frac{S_k(T^{(n)})}{|T^{(n)}|} = \pi(k) \quad \text{a.e.}, \tag{29}$$

*where $(\pi(0), \ldots, \pi(b-1))$ is the invariant probability measure determined by $P$.*

**Proof** Obviously

$$S_k(T_n^l) + S_k(T_n^r) = S_k(T^{(n)}) - I_k(X_0). \tag{30}$$

By (23),(24) and (30), we have

$$\lim_{n\to\infty} \Big\{ \frac{S_k(T^{(n)})}{|T^{(n)}|} - \sum_{l=0}^{b-1} \frac{S_l(T^{(n-1)})}{|T^{(n-1)}|} P(k|l) \Big\} = 0 \quad \text{a.e. .} \tag{31}$$

Multiplying the $k$-th equality of (31) by $P(m|k)$, then adding them together and using (31) once again, we have

$$\lim_{n\to\infty} \Big\{ \sum_{l=0}^{b-1} \frac{S_k(T^{(n)}) P(m|k)}{|T^{(n)}|} - \frac{S_m(T^{(n+1)})}{|T^{(n+1)}|} + \frac{S_m(T^{(n+1)})}{|T^{(n+1)}|} - \sum_{l=0}^{b-1} \frac{S_l(T^{(n-1)})}{|T^{(n-1)}|} P^{(2)}(m|l) \Big\}$$
$$= \lim_{n\to\infty} \Big\{ \frac{S_m(T^{(n+1)})}{|T^{(n+1)}|} - \sum_{l=0}^{b-1} \frac{S_l(T^{(n-1)})}{|T^{(n-1)}|} P^{(2)}(m|l) \Big\} = 0 \quad \text{a.e.},$$

where $P^{(N)}(m|l)$ is the $N$-step transition probability determined by the stochastic matrix $P$. By induction we have

$$\lim_{n\to\infty} \Big\{ \frac{S_m(T^{(n+N)})}{|T^{(n+N)}|} - \sum_{l=0}^{b-1} \frac{S_l(T^{(n-1)})}{|T^{(n-1)}|} P^{(N+1)}(m|l) \Big\} = 0 \quad \text{a.e. .} \tag{32}$$

Noticing that:

$$\frac{1}{|T^{(n-1)}|} \sum_{l=0}^{b-1} S_l(T^{(n-1)}) = 1, \tag{33}$$

and

$$\lim_{n\to\infty} P^{(N)}(k|l) = \pi(k), \quad k \in G, \tag{34}$$

then (29) follows from (32), (33) and (34). □

**Corollary 2** *Let $X = \{X_t, t \in T\}$ be a $T$-indexed nonsymmetric Markov chain with state space $G$ as defined in Definition 1. Let $S_k(T_n^l)$ and $S_k(T_n^r)$ be defined as (20) and (21), respectively. Suppose that stochastic matrix $P = \frac{1}{2}(P_l + P_r)$ is ergodic. Then we have*

$$\lim_{n\to\infty} \frac{S_k(T_n^l)}{|T^{(n)}|} = \frac{1}{2} \sum_{l=0}^{b-1} \pi(l) P_l(k|l), \quad \text{a.e.}, \tag{35}$$

$$\lim_{n\to\infty}\frac{S_k(T_n^r)}{|T^{(n)}|}=\frac{1}{2}\sum_{l=0}^{b-1}\pi(l)P_r(k|l), \quad \text{a.e. .} \tag{36}$$

**Proof** This corollary follows from Theorems 2 and 3 directly. □

**Corollary 3** ([1]) *Let $G=\{0,1\}$, $X=\{X_t, t\in T\}$ be the $T$-indexed nonsymmetric Markov chain with state space $G$. Let $S_k(T_n^l)$ and $S_k(T_n^r)$ be defined as (20) and (21), respectively. Suppose that $P_l$ and $P_r$ are two strictly positive $2\times 2$ stochastic matrices defined on $G^2$ with a unique common invariant probability measure $(\pi(0),\pi(1))$. Then we have*

$$\lim_{n\to\infty}\frac{S_k(T_n^l)}{|T^{(n)}|}=\frac{1}{2}\sum_{l=0}^{1}\pi(l)P_l(k|l), \quad \text{a.e.,} \tag{37}$$

$$\lim_{n\to\infty}\frac{S_k(T_n^r)}{|T^{(n)}|}=\frac{1}{2}\sum_{l=0}^{1}\pi(l)P_r(k|l), \quad \text{a.e. .} \tag{38}$$

**Proof** Obviously, $\pi$ is the unique invariant probability measure for $P$. Let $b=2$ in Corollary 2. This corollary can be obtained from Corollary 2 directly. □

**Definition 3** *Let*

$$S_{k,m}(T^{(n)}\setminus\{0\})=|\{t\in T^{(n)}:(X_{1_t},X_t)=(k,m)\}|,$$

$$S^l_{k,m}(T^{(n)}\setminus\{0\})=|\{t\in T_n^l:(X_{1_t},X_t)=(k,m)\}|,$$

$$S^r_{k,m}(T^{(n)}\setminus\{0\})=|\{t\in T_n^r:(X_{1_t},X_t)=(k,m)\}|.$$

*Obviously, we have*

$$S_{k,m}(T^{(n)}\setminus\{0\})=\sum_{t\in T^{(n)}\setminus\{0\}} I_m(X_t)I_k(X_{1_t}), \tag{39}$$

$$S^l_{k,m}(T^{(n)}\setminus\{0\})=\sum_{t\in T^{(n)}\setminus\{0\}} I_m(X_t)I_k(X_{1_t})\delta^l(t), \tag{40}$$

$$S^r_{k,m}(T^{(n)}\setminus\{0\})=\sum_{t\in T^{(n)}\setminus\{0\}} I_m(X_t)I_k(X_{1_t})\delta^r(t), \tag{41}$$

$$S_{k,m}(T^{(n)}\setminus\{0\})=S^l_{k,m}(T^{(n)}\setminus\{0\})+S^r_{k,m}(T^{(n)}\setminus\{0\}). \tag{42}$$

**Theorem 4** *Let $X=\{X_t, t\in T\}$ be a $T$-indexed nonsymmetric Markov chain with state space $G$ as defined in Definition 1. Let $S^l_{k,m}(T^{(n)}\setminus\{0\})$, $S^r_{k,m}(T^{(n)}\setminus\{0\})$ be defined as (40) and (41), respectively. Then we have*

$$\lim_{n\to\infty}\frac{S^l_{k,m}(T^{(n)}\setminus\{0\})}{|T^{(n)}|}=\frac{1}{2}\pi(k)P_l(m|k), \quad \text{a.e.,} \tag{43}$$

$$\lim_{n\to\infty}\frac{S^r_{k,m}(T^{(n)}\setminus\{0\})}{|T^{(n)}|}=\frac{1}{2}\pi(k)P_r(m|k), \quad \text{a.e. .} \tag{44}$$

**Proof** Let $g_t(x,y) = I_k(x)I_m(y)\delta^l(t), a_n = |T^{(n)}|$. Then we have

$$H_n(\omega) = \sum_{t\in T^{(n)}\setminus\{0\}} I_m(X_t)I_k(X_{1_t})\delta^l(t) = S^l_{k,m}(T^{(n)}\setminus\{0\}), \tag{45}$$

$$\begin{aligned} G_n(\omega) &= \sum_{t\in T^{(n)}\setminus\{0\}} \sum_{x_t} I_k(X_{1_t})I_m(x_t)\delta^l(t)P(x_t|X_{1_t}) \\ &= \sum_{t\in T^{(n)}\setminus\{0\}} I_k(X_{1_t})\delta^l(t)P_l(m|k) \\ &= \sum_{t\in T^{(n-1)}} I_k(X_t)P_l(m|k) = S_k(T^{(n-1)})P_l(m|k). \end{aligned} \tag{46}$$

Obviously, $\{g_t(x,y), t\in T\}$ are uniformly bounded, then from Corollary 1, we have

$$\lim_{n\to\infty}\Big\{\frac{S^l_{k,m}(T^{(n)}\setminus\{0\})}{|T^{(n)}|} - \frac{S_k(T^{(n-1)})}{2|T^{(n-1)}|}P_l(m|k)\Big\} = 0 \ \text{ a.e. .} \tag{47}$$

Then (43) follows from (29) and (47).

Similarly, let $g_t(x,y) = I_k(x)I_m(y)\delta^r(t), a_n = |T^{(n)}|$, (44) can be obtained in the same way.

Thus we have completed the proof of this theorem. □

**Corollary 4** *Let $X = \{X_t, t\in T\}$ be a $T$-indexed nonsymmetric Markov chain with state space $G$ as defined in Definition 1, $S_{k,m}(T^{(n)}\setminus\{0\})$ be defined as (39). Then we have*

$$\lim_{n\to\infty}\frac{S_{k,m}(T^{(n)}\setminus\{0\})}{|T^{(n)}|} = \frac{1}{2}\pi(k)(P_l(m|k)+P_r(m|k)) \ \text{ a.e. .} \tag{48}$$

**Proof** Combing (42)–(44), we obtain (48) directly. □

Let $T$ be a tree, $(X_t)_{t\in T}$ be a stochastic process indexed by tree $T$ with state space $G$. Denote $P(x^{T^{(n)}}) = P(X^{T^{(n)}} = x^{T^{(n)}})$. Let

$$f_n(\omega) = -\frac{1}{|T^{(n)}|}\ln P(X^{T^{(n)}}). \tag{49}$$

Then $f_n(\omega)$ is called the entropy density of $X^{T^{(n)}}$. If $(X_t)_{t\in T}$ is a $T$-indexed nonsymmetric Markov chain with state space $G$, then we have

$$P(x^{T^{(n)}}) = P(X^{T^{(n)}} = x^{T^{(n)}}) = P(X_0 = x_0)\prod_{t\in T^l_n} P_l(x_t|x_{1_t})\prod_{t\in T^r_n} P_r(x_t|x_{1_t}). \tag{50}$$

Obviously, by (49) and (50), we have

$$f_n(\omega) = -\frac{1}{|T^{(n)}|}\Big\{\ln P(X_0) + \sum_{t\in T^l_n}\ln P_l(X_t|X_{1_t}) + \sum_{t\in T^r_n}\ln P_r(X_t|X_{1_t})\Big\}. \tag{51}$$

By (51), we have

$$\begin{aligned} \lim_{n\to\infty} f_n(\omega) = &-\lim_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^l_n}\sum_{k=0}^{b-1}\sum_{m=0}^{b-1} I_k(X_{1_t})I_m(X_t)\ln P_l(m|k) - \\ &\lim_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^r_n}\sum_{k=0}^{b-1}\sum_{m=0}^{b-1} I_k(X_{1_t})I_m(X_t)\ln P_r(m|k) \end{aligned}$$

$$= -\sum_{k=0}^{b-1}\sum_{m=0}^{b-1}\lim_{n\to\infty}\frac{S^l_{k,m}(T^{(n)}\setminus\{0\})}{|T^{(n)}|}\ln P_l(m|k) - \sum_{k=0}^{b-1}\sum_{m=0}^{b-1}\lim_{n\to\infty}\frac{S^r_{k,m}(T^{(n)}\setminus\{0\})}{|T^{(n)}|}\ln P_r(m|k). \tag{52}$$

The convergence of $f_n(\omega)$ to a constant in a sense ($L_1$ convergence, convergence in Probability, a.e. convergence) is called the Shannon-McMillan theorem or the entropy theorem or the AEP in information theory.

**Theorem 5** *Let $X = \{X_t, t \in T\}$ be a $T$-indexed nonsymmetric Markov chain with state space $G$ as defined in Definition 1, $f_n(\omega)$ be defined as (52). Then*

$$\lim_{n\to\infty} f_n(\omega) = -\frac{1}{2}\sum_{k=0}^{b-1}\sum_{m=0}^{b-1}\pi(k)P_l(m|k)\ln P_l(m|k) - \frac{1}{2}\sum_{k=0}^{b-1}\sum_{m=0}^{b-1}\pi(k)P_r(m|k)\ln P_r(m|k) \quad \text{a.e. .} \tag{53}$$

**Proof** This theorem can be obtained from (52) and Theorem 4 directly. □

**Corollary 5** ([7]) *Let $\{X_t, t \in T\}$ be $G$-valued Markov chain indexed by Cayley tree $T$ with the initial distribution (1) and the transition matrix $P = (P(y|x), x, y \in G)$. Suppose that $P$ is an ergodic stochastic matrix. Let $f_n(\omega)$ be the entropy density of $X^{T^{(n)}}$. Then we have*

$$\lim_{n\to\infty} f_n(\omega) = -\sum_{k=0}^{b-1}\sum_{m=0}^{b-1}\pi(k)P(m|k)\ln P(m|k) \quad \text{a.e.}, \tag{54}$$

*where $(\pi(0), \ldots, \pi(b-1))$ is the invariant probability measure determined by $P$.*

**Proof** Let $P_l = P_r$ in Theorem 5. (54) can be obtained from (53) directly. □

## References


[1] BAO Zhenhua, YE Zhongxing. *Strong law of large numbers and asymptotic equipartition property for nonsymmetric Markov chain fields on Cayley trees* [J]. Acta Math. Sci. Ser. B Engl. Ed., 2007, **27**(4): 829–837.

[2] BENJAMINI I, PERES Y. *Markov chains indexed by trees* [J]. Ann. Probab., 1994, **22**(1): 219–243.

[3] BERGER T, YE Zhongxing. *Entropic aspects of random fields on trees* [J]. IEEE Trans. Inform. Theory, 1990, **36**(5): 1006–1018.

[4] HUANG Huilin, YANG Weiguo. *Strong law of large numbers for Markov chains indexed by an infinite tree with uniformly bounded degree* [J]. Sci. China Ser. A, 2008, **51**(2): 195–202.

[5] PEMANTLE R. *Automorphism invariant measures on trees* [J]. Ann. Probab., 1992, **20**(3): 1549–1566.

[6] YANG Weiguo, LIU Wen. *Strong law of large numbers for Markov chains field on a Bethe tree* [J]. Statist. Probab. Lett., 2000, **49**(3): 245–250.

[7] YANG Weiguo. *Some limit properties for Markov chains indexed by a homogeneous tree* [J]. Statist. Probab. Lett., 2003, **65**(3): 241–250.

[8] YE Zhongxing, BERGER T. *Ergodicity, regularity and asymptotic equipartition property of random fields on trees* [J]. J. Combin. Inform. System Sci., 1996, **21**(2): 157–184.

[9] YE Zhongxing, BERGER T. *Information Measures for Discrete Random Fields* [M]. Beijing: Science Press, 1998.