International Scholarly Research Network
ISRN Applied Mathematics
Volume 2012, Article ID 536530, 9 pages
doi:10.5402/2012/536530

Research Article

Strong Law of Large Numbers of the Offspring Empirical Measure for Markov Chains Indexed by Homogeneous Tree

Huilin Huang

College of Mathematics and Information Science, Wenzhou University, Zhejiang 325035, China

Correspondence should be addressed to Huilin Huang, huilin_huang@sjtu.org

Received 24 December 2011; Accepted 26 January 2012

Academic Editors: K. Karamanos and E. Yee

We study the limit law of the offspring empirical measure and for Markov chains indexed by homogeneous tree with almost everywhere convergence. Then we prove a Shannon-McMillan theorem with the convergence almost everywhere.

1. Introduction

A tree is a graph $G = \{T, E\}$ which is connected and contains no circuits, where T and E denote the vertex set and the edge set, respectively. 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 G be an infinite tree with root 0. The set of all vertices with distance n from the root is called the nth 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 second predecessor of t by ${}^2t$ and denote by ${}^nt$ the nth predecessor of t. The degree of a vertex is defined to be the number of neighbors of it. If the degree sequence of a tree is uniformly bounded, we call the tree a uniformly bounded tree. Let d be a positive integer. If every vertex of the tree has d neighbors in the next generation, we say it Cayley tree, which is denoted by $T_{C,d}$. Thus on Cayley tree, every vertex has degree $d+1$ except that the root index which has degree d. 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.

Definition 1.1 (see [1]). Let G be an infinite Cayley tree $T_{C,d}$, S a finite state space, and $\{X_t, t \in T\}$ be a collection of S-valued random variables defined on probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Let

$$p = \{p(x), x \in S\} \tag{1.1}$$

be a distribution on S and

$$P = (P(y \mid x)), \quad x, y \in S, \tag{1.2}$$

be a stochastic matrix on S^2. If for any vertex t,

$$\begin{aligned} &\mathbf{P}\Big(X_t = y \mid X_{1_t} = x \text{ and } X_s \text{ for } t \wedge s \leq^1 t\Big) \\ &\quad = \mathbf{P}(X_t = y \mid X_{1_t} = x) = P(y \mid x) \quad \forall x, y \in S, \\ &\mathbf{P}(X_0 = x) = p(x) \quad \forall x \in S, \end{aligned} \tag{1.3}$$

$\{X_t, t \in T\}$ will be called S-valued *Markov chains indexed by an infinite tree* G with the initial distribution (1.1) and transition matrix (1.2) or called tree-indexed Markov chains with state-space S. Furthermore, if transition matrix P is ergodic, then we call $\{X_t, t \in T\}$ an *ergodic Markov chains* indexed by the infinite tree T.

The above definition is the extension of the definitions of Markov chain fields on trees (see [1, page 456] and [2]). In this paper, we always suppose that the tree-indexed Markov chain is *ergodic*.

The subject of tree-indexed processes is rather young. Benjamini and Peres [3] have given the notion of the tree-indexed Markov chains and studied the recurrence and ray-recurrence for them. Berger and Ye [4] have studied the existence of entropy rate for some stationary random fields on a homogeneous tree. Ye and Berger (see [5, 6]), by using Pemantle's result [7] 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 [8] have studied a strong law of large numbers for the frequency of occurrence of states for Markov chains fields on a homogeneous tree (a particular case of tree-indexed Markov chains and PPG-invariant random fields). Takacs (see [9]) have studied the strong law of large numbers for the univariate functions of finite Markov chains indexed by an infinite tree with uniformly bounded degree. Subsequently, Huang and Yang (see [10]) has studied the Shannon-McMillan theorem of finite homogeneous Markov chains indexed by a uniformly bounded infinite tree. Dembo et al., (see [11]) has showed the large deviation principle holds for the empirical offspring measure of Markov chains on random trees and demonstrated the explicit rate function, which is defined in terms of specific relative entropy (see [12]) and Cramér's rate function.

In this paper, we study the strong law of large numbers for the offspring empirical measure and the Shannon-McMillan theorem with a.e. convergence for Markov chain fields on tree $T_{C,d}$ by using a method similar to that of [10].

2. Statements of the Results

For every vertex $t \in T$, the random vector of offspring states is defined as

$$\mathbf{C}_t = (X_1(t), X_2(t), \ldots, X_d(t)) \in S^d. \tag{2.1}$$

Let $\mathbf{c} = (c_1, c_2, \ldots, c_d)$ be a d-dimensional vector on S^d.

Now we also let the distribution (1.1) serve as the initial distribution. Define the offspring transition kernel Q from S to S^d. We define the law $\mathbf{P}$ of a tree-indexed process X by the following rules.

(i) The state of the root random variable X_0 is determined by distribution (1.1).

(ii) For every vertex $t \in T$ with state x, the offspring states are given independently of everything else, by the offspring law $Q(\cdot \mid x)$ on S^d, where

$$Q(\mathbf{c} \mid x) := Q(\mathbf{C}_t = (c_1, c_2, \ldots, c_d) \mid X_t = x) = \prod_{i=1}^{d} P(c_i \mid x). \tag{2.2}$$

Here the last equation holds because of the property of conditional independence.

For every finite $n \in \mathbf{N}$, let $\{X_t, t \in T\}$ be S-valued Markov chains indexed by an infinite tree T. Now we define the offspring empirical measure

$$L_n(x, \mathbf{c}) = \frac{\sum_{t \in T^{(n)}} \mathbf{I}\{(X_t, \mathbf{C}_t) = (x, \mathbf{c})\}}{\left|T^{(n)}\right|} \quad \forall (x, \mathbf{c}) \in S \times S^d. \tag{2.3}$$

For any state $x \in S$, $S_n(x)$ is the empirical measure, which is defined as follows:

$$S_n(x) = \frac{\sum_{t \in T^{(n)}} \mathbf{I}\{X_t = x\}}{\left|T^{(n)}\right|} \quad \forall x \in S, \tag{2.4}$$

where $\mathbf{I}\{\cdot\}$ denotes the indicator function as usual and $\mathbf{c} = (c_1, c_2, \ldots, c_d)$.

In the rest of this paper, we consider the limit law of the random sequence of $\{L_n(x, \mathbf{c}), n \geq 1\}$, which is defined as above.

Theorem 2.1. *Let G be a Cayley tree $T_{C,d}$, S a finite state space, and $\{X_t, t \in T\}$ be tree-indexed Markov chain with initial distribution* (1.1) *and ergodic transition matrix P. Let $L_n(x, \mathbf{c})$ be defined as* (2.3). *Thus one has*

$$\lim_{n \to \infty} L_n(x, \mathbf{c}) = \pi(x) Q(\mathbf{c} \mid x) \quad \text{a.e.}, \tag{2.5}$$

where π is the stationary distribution of the ergodic matrix P, that is, $\pi = \pi P$, and $\sum_{x \in S} \pi(x) = 1$.

Corollary 2.2. *Under the condition of Theorem 2.1, suppose that $f(x,\mathbf{c})$ is any function defined on $S\times S^d$. Denote*

$$H_n(\omega)=\sum_{t\in T^{(n)}} f(X_t,\mathbf{C}_t). \tag{2.6}$$

Then

$$\lim_{n\to\infty}\frac{H_n(\omega)}{\left|T^{(n)}\right|}=\sum_{(x,\mathbf{c})\in S\times S^d}\pi(x)Q(\mathbf{c}\mid x)f(x,\mathbf{c}).\quad \text{a.e.} \tag{2.7}$$

Proof. Noting that

$$\begin{aligned} H_n(\omega)&=\sum_{t\in T^{(n)}} f(X_t,\mathbf{C}_t)\\ &=\sum_{(x,\mathbf{c})\in S\times S^d}\sum_{t\in T^{(n)}}\mathbf{I}\{(X_t,\mathbf{C}_t)=(x,\mathbf{c})\}f(x,\mathbf{c}), \end{aligned} \tag{2.8}$$

thus by using Theorem 2.1 we get

$$\begin{aligned} \lim_{n\to\infty}\frac{H_n(\omega)}{\left|T^{(n)}\right|}&=\sum_{(x,\mathbf{c})\in S\times S^d} f(x,\mathbf{c})\lim_{n\to\infty}L_n(x,\mathbf{c})\\ &=\sum_{(x,\mathbf{c})\in S\times S^d}\pi(x)Q(\mathbf{c}\mid x)f(x,\mathbf{c}).\quad \text{a.e.} \end{aligned} \tag{2.9}$$

Let $G=\{T,E\}$ be a tree graph, $(X_t)_{t\in T}$ be a stochastic process indexed by tree G with state space S. Denote $Y_t=(X_t,\mathbf{C}_t)$ to be the offspring processes derived by $(X_t)_{t\in T}$. It is easy to see that

$$\begin{aligned} \mathbf{P}\left(y^{T^{(n)}}\right)&=\mathbf{P}\left(Y^{T^{(n)}}=y^{T^{(n)}}\right)\\ &=p(x_0)\prod_{t\in T^{(n+1)}\setminus\{0\}}P(x_t\mid x_{1_t})\\ &=p(x_0)\prod_{t\in T^{(n)}}Q(\mathbf{c}_t\mid x_t), \end{aligned} \tag{2.10}$$

where $\mathbf{c}_t\in S^d$. Let

$$f_n(\omega)=-\frac{1}{\left|T^{(n)}\right|}\ln\mathbf{P}\left(Y^{T^{(n)}}\right). \tag{2.11}$$

$f_n(\omega)$ will be called the entropy density of $Y^{T^{(n)}}$. If $(X_t)_{t\in T}$ is a tree-indexed Markov chain with state space S defined by Definition 1.1, we have by (2.10)

$$f_n(\omega) = -\frac{1}{|T^{(n)}|}\left[\ln p(X_0) + \sum_{t\in T^{(n)}} \ln Q(\mathbf{C}_t \mid X_t)\right]. \tag{2.12}$$

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. Here from Corollary 2.2, if we let

$$f(x,\mathbf{c}) = -\ln Q(\mathbf{c} \mid x), \tag{2.13}$$

we can easily obtain the Shannon-McMillan theorem with a.e. convergence for Markov chain fields on tree $T_{C,d}$. □

Corollary 2.3. *Under the condition of Corollary 2.2, let $f_n(\omega)$ be defined as* (2.12). *Then*

$$\lim_{n\to\infty} f_n(\omega) = -\sum_{(x,\mathbf{c})\in S\times S^d} \pi(x)Q(\mathbf{c} \mid x)\ln Q(\mathbf{c} \mid x). \quad \text{a.e.} \tag{2.14}$$

3. Proof of Theorem 2.1

Let $T_{C,d}$ be a Cayley tree, S a finite state space, and $\{X_t, t \in T\}$ tree-indexed Markov chain with any initial distribution (1.1) and ergodic transition matrix P. Let $g_t(X_t, \mathbf{C}_t)$ be functions defined on $S \times S^d$. Letting λ be a real number, $L_0 = \{0\}$, $\mathcal{F}_n = \sigma(X^{T^{(n)}})$, now we can define a nonnegative martingale as follows:

$$t_n(\lambda,\omega) = \frac{e^{\lambda \sum_{t\in T^{(n-1)}} g_t(X_t,\mathbf{C}_t)}}{\prod_{t\in T^{(n-1)}} E\left[e^{\lambda g_t(X_t,\mathbf{C}_t)} \mid X_t\right]}. \tag{3.1}$$

At first we come to prove the above fact.

Theorem 3.1. $\{t_n(\lambda,\omega), \mathcal{F}_n, n \geq 1\}$ *is a nonnegative martingale.*

Proof of Theorem 3.1. Note that, by Markov property and the property of conditional independence, we have

$$\begin{aligned} E\left[e^{\lambda \sum_{t\in L_n} g_t(X_t,\mathbf{C}_t)} \mid \mathcal{F}_n\right] &= \sum_{x^{L_{n+1}}} e^{\lambda \sum_{t\in L_n} g_t(X_t,\mathbf{c}_t)} \mathbf{P}\left(X^{L_{n+1}} = x^{L_{n+1}} \mid X^{T^{(n)}}\right) \\ &= \sum_{x^{L_{n+1}}} \prod_{t\in L_n} e^{\lambda g_t(X_t,\mathbf{c}_t)} Q(\mathbf{c}_t \mid X_t) \\ &= \prod_{t\in L_n} \sum_{\mathbf{c}_t\in S^d} e^{\lambda g_t(X_t,\mathbf{c}_t)} Q(\mathbf{c}_t \mid X_t) \\ &= \prod_{t\in L_n} E\left[e^{\lambda g_t(X_t,\mathbf{C}_t)} \mid X_t\right] \quad \text{a.e.} \end{aligned} \tag{3.2}$$

On the other hand, we also have

$$t_{n+1}(\lambda,\omega) = t_n(\lambda,\omega) \frac{e^{\lambda \sum_{t\in L_n} g_t(X_t,\mathbf{C}_t)}}{\prod_{t\in L_n} E\left[e^{\lambda g_t(X_t,\mathbf{C}_t)} \mid X_t\right]}. \tag{3.3}$$

Combining (3.2) and (3.3), we get

$$E[t_{n+1}(\lambda,\omega) \mid \mathcal{F}_n] = t_n(\lambda,\omega) \quad \text{a.e.} \tag{3.4}$$

Thus we complete the proof of this theorem. □

Theorem 3.2. *Let $(X_t)_{t\in T}$ and $\{g_t(x,\mathbf{c}), t \in T\}$ be defined as above, and denote*

$$G_n(\omega) = \sum_{t\in T^{(n)}} E\left[g_t(X_t,\mathbf{C}_t) \mid X_t\right]. \tag{3.5}$$

Let $\alpha > 0$, denote

$$D(\alpha) = \left\{ \limsup_{n\to\infty} \frac{1}{\left|T^{(n)}\right|} \sum_{t\in T^{(n)}} E\left[g_t^2(X_t,\mathbf{C}_t) e^{\alpha|g_t(X_t,\mathbf{C}_t)|} \mid X_t\right] = M(\omega) < \infty \right\}, \tag{3.6}$$

$$H_n(\omega) = \sum_{t\in T^{(n)}} g_t(X_t,\mathbf{C}_t). \tag{3.7}$$

Then

$$\lim_{n\to\infty} \frac{H_n(\omega) - G_n(\omega)}{\left|T^{(n)}\right|} = 0 \quad \text{a.e. on } D(\alpha). \tag{3.8}$$

Proof. By Theorem 3.1, we have known that $\{t_n(\lambda,\omega), \mathcal{F}_n, n \geq 1\}$ is a nonnegative martingale. According to Doob martingale convergence theorem, we have

$$\lim_n t_n(\lambda,\omega) = t(\lambda,\omega) < \infty \quad \text{a.e.} \tag{3.9}$$

so that

$$\limsup_{n\to\infty}\frac{\ln t_{n+1}(\lambda,\omega)}{|T^{(n)}|}\le 0 \quad \text{a.e.} \tag{3.10}$$

Combining (3.1), (3.7), and (3.10), we arrive at

$$\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\left\{\lambda H_n(\omega)-\sum_{t\in T^{(n)}}\ln\left[E\left[e^{\lambda g_t(X_t,\mathbf{C}_t)}\mid X_t\right]\right]\right\}\le 0 \quad \text{a.e.} \tag{3.11}$$

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

$$\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\left\{H_n(\omega)-\sum_{t\in T^{(n)}}\frac{\ln\left[E\left[e^{\lambda g_t(X_t,\mathbf{C}_t)}\mid X_t\right]\right]}{\lambda}\right\}\le 0 \quad \text{a.e.} \tag{3.12}$$

By (3.12) and inequalities $\ln x\le x-1\ (x>0),\ 0\le e^x-1-x\le 2^{-1}x^2e^{|x|}$, as $0<\lambda\le\alpha$, it follows that

$$\begin{aligned}
&\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\left[H_n(\omega)-\sum_{t\in T^{(n)}}E\left[g_t(X_t,\mathbf{C}_t)\mid X_t\right]\right]\\
&\quad\le\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}}\left\{\frac{\ln\left[E\left[e^{\lambda g_t(X_t,\mathbf{C}_t)}\mid X_t\right]\right]}{\lambda}-E\left[g_t(X_t,\mathbf{C}_t)\mid X_t\right]\right\}\\
&\quad\le\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}}\left\{\frac{E\left[e^{\lambda g_t(X_t,\mathbf{C}_t)}\mid X_t\right]-1}{\lambda}-E\left[g_t(X_t,\mathbf{C}_t)\mid X_t\right]\right\}\\
&\quad\le\frac{\lambda}{2}\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}}E\left[g_t^2(X_t,\mathbf{C}_t)e^{\lambda|g_t(X_t,\mathbf{C}_t)|}\mid X_t\right]\\
&\quad\le\frac{\lambda}{2}\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{t\in T^{(n)}}E\left[g_t^2(X_t,\mathbf{C}_t)e^{\alpha|g_t(X_t,\mathbf{C}_t)|}\mid X_t\right]\\
&\quad\le\frac{\lambda}{2}M(\omega)\quad \text{a.e. } \omega\in D(\alpha).
\end{aligned} \tag{3.13}$$

Letting $\lambda\to 0^+$ in (3.13), by (3.5) we have

$$\limsup_{n\to\infty}\frac{H_n(\omega)-G_n(\omega)}{|T^{(n)}|}\le 0 \quad \text{a.e.}\ \ \omega\in D(\alpha). \tag{3.14}$$

Let $-\alpha\le\lambda<0$. Similarly to the analysis of the case $0<\lambda\le\alpha$, it follows from (3.12) that

$$\liminf_{n\to\infty}\frac{H_n(\omega)-G_n(\omega)}{|T^{(n)}|}\ge\frac{\lambda}{2}M(\omega) \quad \text{a.e.}\ \ \omega\in D(\alpha). \tag{3.15}$$

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

$$\liminf_{n\to\infty} \frac{H_n(\omega) - G_n(\omega)}{\left|T^{(n)}\right|} \geq 0 \quad \text{a.e.} \quad \omega \in D(\alpha). \tag{3.16}$$

Combining (3.14) and (3.16), we obtain (3.8) directly. □

Corollary 3.3. *Under the conditions of Theorem 3.2, one has*

$$\lim_{n\to\infty} [L_n(x, \mathbf{c}) - S_n(x)Q(\mathbf{c} \mid x)] = 0 \quad \text{a.e.,} \tag{3.17}$$

where π is the stationary distribution of the ergodic matrix P, that is, $\pi = \pi P$, and $\Sigma_{x\in S}\pi(x) = 1$.

Proof. For any $t \in T$, let

$$\begin{aligned} g_t(X_t, \mathbf{C}_t) &= \mathbf{I}\{(X_t, \mathbf{C}_t) = (x, \mathbf{c})\} \\ &= \mathbf{I}\{X_t = x\} \cdot \mathbf{I}\{\mathbf{C}_t = \mathbf{c}\}. \end{aligned} \tag{3.18}$$

Then we have

$$\begin{aligned} G_n(\omega) &= \sum_{t\in T^{(n)}} E\left[g_t(X_t, \mathbf{C}_t) \mid X_t\right] \\ &= \sum_{t\in T^{(n)}} \sum_{\mathbf{c}_t\in S^d} \mathbf{I}\{X_t = x\} \cdot \mathbf{I}\{\mathbf{c}_t = \mathbf{c}\} Q(\mathbf{c}_t \mid X_t) \\ &= \sum_{t\in T^{(n)}} \mathbf{I}\{X_t = x\} Q(\mathbf{c} \mid x) \\ &= \left|T^{(n)}\right| \cdot S_n(x) Q(\mathbf{c} \mid x) \end{aligned} \tag{3.19}$$

$$\begin{aligned} H_n(\omega) &= \sum_{t\in T^{(n)}} g_t(X_t, \mathbf{C}_t) \\ &= \sum_{t\in T^{(n)}} \mathbf{I}\{(X_t, \mathbf{C}_t) = (x, \mathbf{c})\} \\ &= \left|T^{(n)}\right| \cdot L_n(x, \mathbf{c}). \end{aligned} \tag{3.20}$$

Combing (3.19) and (3.20), we can derive our conclusion by Theorem 3.2.

In our proof, we will use Lemma 3.4. □

Lemma 3.4 (see [10]). *Let $T_{C,d}$ be a Cayley tree, S a finite state space, and $\{X_t, t \in T\}$ tree-indexed Markov chain with any initial distribution* (1.1) *and ergodic transition matrix P. Let $S_n(x)$ be defined as* (2.4). *Thus one has*

$$\lim_{n \to \infty} S_n(x) = \pi(x) \quad \text{a.e.} \tag{3.21}$$

Proof of Theorem 2.1. Combining Corollary 3.3 and Lemma 3.4, we arrive at our conclusion directly. □

Acknowledgment

This work was supported by National Natural Science Foundation of China (Grant no. 11071104).

References

[1] J. G. Kemeny, J. L. Snell, and A. W. Knapp, *Denumerable Markov Chains*, Springer, New York, NY, USA, Second edition, 1976.
[2] F. Spitzer, "Markov random fields on an infinite tree," *Annals of Probability*, vol. 3, no. 3, pp. 387–398, 1975.
[3] I. Benjamini and Y. Peres, "Markov chains indexed by trees," *Annals of Probability*, vol. 22, no. 1, pp. 219–243, 1994.
[4] T. Berger and Z. X. Ye, "Entropic aspects of random fields on trees," *IEEE Transactions on Information Theory*, vol. 36, no. 5, pp. 1006–1018, 1990.
[5] Z. Ye and T. Berger, "Ergodicity, regularity and asymptotic equipartition property of random fields on trees," *Journal of Combinatorics, Information & System Sciences*, vol. 21, no. 2, pp. 157–184, 1996.
[6] Z. Ye and T. Berger, *Information Measures for Discrete Random Fields*, Science Press, Beijing, China, 1998.
[7] R. Pemantle, "Automorphism invariant measures on trees," *Annals of Probability*, vol. 20, no. 3, pp. 1549–1566, 1992.
[8] W. Yang and W. Liu, "Strong law of large numbers for Markov chains field on a Bethe tree," *Statistics & Probability Letters*, vol. 49, no. 3, pp. 245–250, 2000.
[9] C. Takacs, "Strong law of large numbers for branching Markov chains," *Markov Processes and Related Fields*, vol. 8, no. 1, pp. 107–116, 2001.
[10] H. Huang and W. Yang, "Strong law of large numbers for Markov chains indexed by an infinite tree with uniformly bounded degree," *Science in China A*, vol. 51, no. 2, pp. 195–202, 2008.
[11] A. Dembo, P. Mörters, and S. Sheffield, "Large deviations of Markov chains indexed by random trees," *Annales de l'Institut Henri Poincaré B*, vol. 41, no. 6, pp. 971–996, 2005.
[12] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, vol. 38 of *Applications of Mathematics*, Springer, New York, NY, USA, 2nd edition, 1998.