J. Appl. Math. & Informatics Vol. **30**(2012), No. 1 - 2, pp. 83 - 92
Website: http://www.kcam.biz

SOME GENERALIZED SHANNON-MCMILLAN THEOREMS FOR NONHOMOGENEOUS MARKOV CHAINS ON SECOND-ORDER GAMBLING SYSTEMS INDEXED BY AN INFINITE TREE WITH UNIFORMLY BOUNDED DEGREE†

KANGKANG WANG* AND ZURUN XU

Abstract. In this paper, a generalized Shannon-McMillan theorem for the nonhomogeneous Markov chains indexed by an infinite tree which has a uniformly bounded degree is discussed by constructing a nonnegative martingale and analytical methods. As corollaries, some Shannon-Mcmillan theorems for the nonhomogeneous Markov chains indexed by a homogeneous tree and the nonhomogeneous Markov chain are obtained. Some results which have been obtained are extended.

AMS Mathematics Subject Classification : 60F15, 65F15.
Key words and phrases : Shannon-McMillan theorem, the infinite tree, Markov chains, the entropy density, uniformly bounded degree.

1. Introduction

A tree is a graph $G = \{T, E\}$ which is connected and contains no circuits. 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 $\overline{\alpha\beta}$.

In this paper, we mainly consider an infinite tree which has uniformly bounded degree, that is, the numbers of neighbors of any vertices in this tree are uniformly bounded. When the context permits, this type of trees are all denoted simply by T. For a better explanation of the tree T, we take Cayley tree $T_{C,N}$ for example. It's a special case of the tree T, the root o of Cayley tree has N neighbors and all the other vertices of it have $N+1$ neighbors each (see Fig.1).

Let T be an infinite tree with a root o, the set of all vertices with the distance n from the root is called the n-th generation of T, which is denoted by L_n.

Received December 9, 2010. Revised July 20, 2011. Accepted August 12, 2011. *Corresponding author. †This work was supported by the research grant of Higher Schools' Natural Science Basic Research of Jiangsu Province of China (09KJD110002).

In other words, L_n represents the set of all vertices on the level n. We denote by $T^{(n)}$ the union of the first n generations of T. Denote by t the t-th vertex from the root to the upper part, from the left side to the right side on the tree. For each vertex t, there is a unique path from o to t, and $|t|$ for the number of the edges on this path. We denote the first predecessor of t by 1_t, the second predecessor of t by 2_t, and denote by n_t the n-th predecessor of t. For any two vertices s and t of the tree T, write $s \leq t$ if s is on the unique path from the root o to t. We denote $s \wedge t$ the vertex nearest from o satisfying $s \wedge t \leq s$ and $s \wedge t \leq t$. $X^A = \{X_t, t \in A\}$ and $|A|$ denote by the number of the vertices of A.

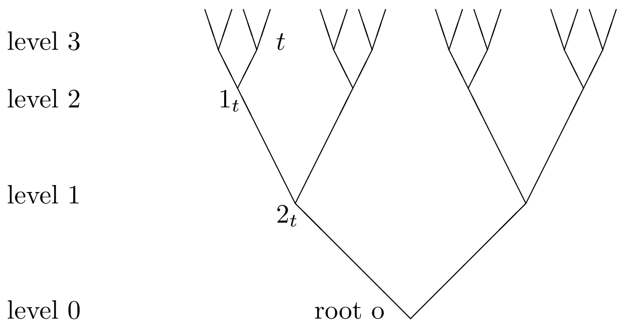

Fig.1 An infinite tree $T_{C,2}$

Definition 1 ([10])**.** Let T be an infinite tree, $S = \{s_0, s_1, s_2, \cdots, s_{N-1}\}$ a finite state space, $\{X_t, t \in T\}$ be a collection of S-valued random variables defined on the probability space $\{\Omega, F, P\}$. Let

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

be a distribution on S, and

$$P_t = (P_t(y|x)), \ x, y \in S, \ t \in T. \tag{2}$$

be a series of strictly positive stochastic matrices on S^2. If for any vertex t,

$$\begin{aligned} & P(X_t = y|X_{1_t} = x, \ and \ X_s \ for \ t \wedge s \leq 1_t) \qquad (3) \\ & = P(X_t = y|X_{1_t} = x) = P_t(y|x) \ \forall x, y \in S, \end{aligned}$$

and

$$P(X_0 = x) = p(x), \ \forall x \in S. \tag{4}$$

$\{X_t, t \in T\}$ will be called $S-$valued Markov chains indexed by an infinite tree with the initial distribution (1) and transition matrices (2).

The above definition is an extension of the definitions of Markov chain fields on trees (see[10]).

Two special finite tree-indexed Markov chains are introduced in Kemeny et al.(1976[10]), Spitzer (1975[11]), and there the finite transition matrix is assumed to be positive and reversible to its stationary distribution, and this tree-indexed Markov chains ensure that the cylinder probabilities are independent of the direction we travel along a path. In this paper, we have no such assumption.

It is easy to see that when $\{X_t, t \in T\}$ is a T-indexed Markov chain,

$$P(x^{T^{(n)}}) = P(X^{T^{(n)}} = x^{T^{(n)}}) = P(X_0 = x_0) \prod_{t \in T^{(n)} \setminus \{o\}} P_t(x_t | x_{1_t}). \quad (5)$$

Let T be a tree, $\{X_t, t \in T\}$ be a stochastic process indexed by the tree T with the state space S. $|T^{(n)}|$ represents the number of all the vertices from level 0 to level n. Denote

$$P(x^{T^{(n)}}) = P(X^{T^{(n)}} = x^{T^{(n)}}).$$

Let

$$f_n(\omega) = -\frac{1}{|T^{(n)}|} \log P(X^{T^{(n)}}). \quad (6)$$

$f_n(\omega)$ will be called the entropy density of $X^{T^{(n)}}$, where log is the natural logarithm. If $\{X_t, t \in T\}$ is a T-indexed Markov chain with the state space S defined by Definition 1, we have by (5)

$$f_n(\omega) = -\frac{1}{|T^{(n)}|} [\log P(X_0) + \sum_{t \in T^{(n)} \setminus \{o\}} \log P_t(X_t | X_{1_t})]. \quad (7)$$

Definition 2. Let $\{f_n(x_0, \cdots, x_n), n \geq 0\}$ and $\{g_n(x_0, \cdots, x_n), n \geq 0\}$ be two sequences of real-valued functions defined on $S^{n+1} (n = 1, 2, \cdots)$, which will be called the generalized selection functions if $\{f_n, n \geq 0\}$ take values in an interval of $[s, a] (s > 0, a > 1)$, $\{g_n, n \geq 0\}$ take values in an interval of $[0, b]$. We let

$$Y_0 = y \ (y \textit{ is an arbitrary real number}), \ Y_t = f_{|t|}(X_{1_t}, X_{2_t}, X_{3_t}, \cdots, X_0), \quad |t| \geq 1, \quad (8)$$

$$b_0 = h \ (h \textit{ is an arbitrary real number}), \ b_t = g_{|t|}(X_{1_t}, X_{2_t}, X_{3_t}, \cdots, X_0), \quad |t| \geq 1, \quad (9)$$

where $|t|$ stands for the number of the edges on the path from the ro ot o to t. Then $\{Y_t^{b_t}, \ t \in T^{(n)}\}$ is called the generalized second-order gambling system or the generalized second-order random selection system indexed by an infinite tree with uniformly bounded degree. The traditional random selection system $\{Y_n, \ n \geq 0\}$[17] takes values in the set of $\{0, 1\}$.

In order to explain the real meaning of the notion of the random selection, we consider the traditional gambling model. Let $\{X_n, n \geq 0\}$ be a second-order nonhomogeneous Markov chain, and $\{g_n(x, y), n \geq 1\}$ be a real-valued function sequence defined on S^2. Interpret X_n as the result of the nth trial, the type of which may change at each step. Let $\mu_n = Y_n g_n(X_{n-1}, X_n)$ denote the gain of the bettor at the nth trial, where Y_n represents the bet size, $g_n(X_{n-1}, X_n)$ is determined by the gambling rules, and $\{Y_n, n \geq 0\}$ is called a gambling system or a random selection system. The bettor's strategy is to determine $\{Y_n, n \geq 0\}$ by the results of the last trial. Let the entrance fee that the bettor pays at the

nth trial be b_n. Also suppose that b_n depends on X_{n-1} as $n \geq 1$, and b_1 is a constant. Thus $\sum_{k=2}^{n} Y_k g_k(X_{k-1}, X_k)$ represents the total gain in the first n trials, $\sum_{k=1}^{n} b_k$ the accumulated entrance fees, and $\sum_{k=1}^{n} [Y_k g_k(X_{k-1}, X_k) - b_k]$ the accumulated net gain. Motivated by the classical definition of "fairness" of game of chance (see Kolmogorov[16]), we introduce the following definition:

Definition 3. The game is said to be fair, if for almost all $\omega \in \{\omega : \sum_{k=2}^{\infty} Y_k = \infty\}$, the accumulated net gain in the first n trial is to be of smaller order of magnitude than the accumulated stake $\sum_{k=2}^{n} Y_k$ as n tends to infinity, that is

$$\lim_{n\to\infty} \frac{1}{\sum_{k=1}^{n} Y_k} \sum_{k=1}^{n} [Y_k g_k(X_{k-1}, X_k) - b_k] = 0 \ a.s. \ on \ \{\omega : \sum_{k=1}^{\infty} Y_k = \infty\}.$$

Definition 4. Let $\{Y_t^{b_t}, t \in T^{(n)}\}$ be a generalized gambling system defined by (8), we call

$$S_n(\omega) = -\frac{1}{\sum_{t\in T^{(n)}} Y_t^{b_t}} [Y_0^{b_0} \log P(X_0) + \sum_{t\in T^{(n)}\setminus\{o\}} Y_t^{b_t} \log P_t(X_t|X_{1_t})] \quad (10)$$

the generalized entropy density of the T-indexed Markov chain $\{X_t, t \in T\}$ on the generalized gambling system. Obviously, the generalized entropy density $S_n(\omega)$ is just the general entropy density $f_n(\omega)$ if $Y_t \equiv 1,\ t \in T^{(n)}$.

The convergence of $f_n(\omega)$ in a sense (L_1 convergence, convergence in probability, or almost sure convergence) is called the Shannon-McMillan theorem or the asymptotic equipartition property(AEP) in the information theory. Shannon-McMillan theorems on the Markov chain have been studied extensively (see[1],[2]). In the recent years, with the development of information theory scholars get to study the Shannon-McMillan theorems for random field on the tree graph(see[3]). The tree models have recently drawn increasing interest from specialists in physics, probability and information theory. Berger and Ye (see[4]) have studied the existence of entropy rate for G-invariant random fields. Recently, Ye and Berger (see[5]) have also studied the ergodic property and Shannon-McMillan theorem for PPG-invariant random fields on trees. But their results only relate to convergence in probability. Liu and Yang (see[6],[7]) have recently studied a.s. convergence of Shannon-McMillan theorem for Markov chains indexed by a homogeneous tree and the generalized Cayley tree. Yang and Ye (see[8]) have studied the asymptotic equipartition property for nonhomogeneous Markov chains indexed by the homogeneous tree. Wang (see[13]) have also studied the asymptotic equipartition property for mth-order nonhomogeneous Markov chains.

The conception of random selection derives from gambling. We consider a sequence of Bernoulli trial, and suppose that at each trial the bettor has the free choice of whether or not to bet. A theorem on gambling systems asserts that under any non-anticipative system the successive bets form a sequence of Bernoulli trial with unchanged probability for success. The importance of this statement was recognized by von Mises, who introduced the impossibility of a successful gambling system as a fundamental axiom (see [14], [15]). This topic was discussed still further by Kolmogrov (see[16]) and Liu and Wang (see [17] and [18]).

In this paper, we study the generalized Shannon-McMillan theorems for nonhomogeneous Markov chains on the generalized gambling system indexed by an infinite tree with the uniformly bounded degree by using the tools of the consistent distribution functions and constructing a nonnegative super-martingale. As corollaries, two Shannon-McMillan theorems for Markov chains indexed by a homogeneous tree and the general nonhomogeneous Markov chain are obtained. Liu and Yang's (see [1, 8]) results are extended.

2. Main results

Theorem 1. *Let T be an infinite tree with a uniformly bounded degree. Let $X = \{X_t, t \in T\}$ be a T-indexed Markov chain with the state space S defined as before, $S_n(\omega)$ be defined as (10). Denote by $H_t(\omega)$ the random conditional entropy of X_t relative to X_{1_t}, that is*

$$H_t(\omega) = -\sum_{x_t \in S} P_t(x_t|X_{1_t}) \log P_t(x_t|X_{1_t}),\ t \in T^{(n)}\backslash\{o\}.$$

Then

$$\lim_{n\to\infty} \sum_{t\in T^{(n)}\backslash\{o\}} \frac{Y_t^{b_t}[\log P_t(X_t|X_{1_t}) + H_t(\omega)]}{\sum_{k=1}^{t} Y_k^{b_k}} < \infty.\ a.s. \tag{11}$$

$$\lim_{n\to\infty}[S_n(\omega) - \frac{1}{\sum\limits_{t\in T^{(n)}\backslash\{o\}} Y_t^{b_t}} \sum_{t\in T^{(n)}\backslash\{o\}} Y_t^{b_t} H_t(\omega)] = 0.\ a.s. \tag{12}$$

Proof. On the probability space (Ω, F, P), let $\lambda = 1$ or $\lambda = -1$. Denote

$$\mu_Q(\lambda; x^{T^{(n)}}) = \frac{p(x_0) \prod\limits_{t\in T^{(n)}\backslash\{o\}} P_t(x_t|x_{1_t}) \exp\left\{\frac{\lambda Y_t^{b_t}(\log P_t(x_t|x_{1_t}) + H_t(\omega))}{\sum_{k=1}^{t} Y_k^{b_k}}\right\}}{\prod\limits_{t\in T^{(n)}\backslash\{o\}} U_t(\lambda; x_t)}. \tag{13}$$

where

$$U_t(\lambda; x_t) = E\left\{\exp\left\{\frac{\lambda Y_t^{b_t}(\log P_t(X_t|X_{1_t}) + H_t(\omega))}{\sum_{k=1}^{t} Y_k^{b_k}}\right\}|X_{1_t} = x_{1_t}\right\}$$
$$= \sum_{x_t\in S} \exp\left\{\frac{\lambda Y_t^{b_t}(\log P_t(x_t|x_{1_t}) + H_t(\omega))}{\sum_{k=1}^{t} Y_k^{b_k}}\right\} \cdot P_t(x_t|x_{1_t}). \tag{14}$$

By (14) and (13),

$$\sum_{x^{L_n}\in S^{L_n}} \mu_Q(\lambda; x^{T^{(n)}})$$

$$
\begin{aligned}
&= \sum_{x^{L_n}\in S^{L_n}} \frac{p(x_0)\prod\limits_{t\in T^{(n)}\setminus\{o\}} P_t(x_t|x_{1_t})\exp\left\{\frac{\lambda Y_t^{b_t}(\log P_t(x_t|x_{1_t})+H_t(\omega))}{\sum_{k=1}^t Y_k^{b_k}}\right\}}{\prod\limits_{t\in T^{(n)}\setminus\{o\}} U_t(\lambda;x_t)} \\
&= \mu_Q(\lambda;x^{T^{(n-1)}})\frac{\sum\limits_{x^{L_n}\in S^{L_n}}\prod\limits_{t\in L_n} P_t(x_t|x_{1_t})\exp\left\{\frac{\lambda Y_t^{b_t}(\log P_t(x_t|x_{1_t})+H_t(\omega))}{\sum_{k=1}^t Y_k^{b_k}}\right\}}{\prod\limits_{t\in L_n} U_t(\lambda;x_t)} \\
&= \mu_Q(\lambda;x^{T^{(n-1)}})\frac{\prod\limits_{t\in L_n}\sum\limits_{x_t\in S} P_t(x_t|x_{1_t})\exp\left\{\frac{\lambda Y_t^{b_t}(\log P_t(x_t|x_{1_t})+H_t(\omega))}{\sum_{k=1}^t Y_k^{b_k}}\right\}}{\prod\limits_{t\in L_n} U_t(\lambda;x_t)} \\
&= \mu_Q(\lambda;x^{T^{(n-1)}})\frac{\prod_{t\in L_n} U_t(\lambda;x_t)}{\prod_{t\in L_n} U_t(\lambda;x_t)} = \mu_Q(\lambda;x^{T^{(n-1)}}).\ a.s. \qquad (15)
\end{aligned}
$$

Therefore, $\mu_Q(\lambda;x^{T^{(n)}})$, $n=1,2,\cdots$ are a family of consistent distribution functions on $S^{T^{(n)}}$. Let

$$T_n(\lambda,\omega)=\frac{\mu_Q(\lambda;X^{T^{(n)}})}{P(X^{T^{(n)}})}. \qquad (16)$$

By (5), (13) and (16), we obtain

$$T_n(\lambda,\omega)=\frac{\exp\left\{\sum\limits_{t\in T^{(n)}\setminus\{o\}}\frac{\lambda Y_t^{b_t}(\log P_t(X_t|X_{1_t})+H_t(\omega))}{\sum_{k=1}^t Y_k^{b_k}}\right\}}{\prod\limits_{t\in T^{(n)}\setminus\{o\}} U_t(\lambda;X_t)},\ n\ge 0. \qquad (17)$$

Since $\mu_Q(\lambda;X^{T^{(n)}})$ and $P(X^{T^{(n)}})$ are two distribution functions, it is easy to see that $T_n(\lambda,\omega)$ is a nonnegative sup-martingale from Doob's martingale convergence theorem(see[12]). Therefore,

$$\lim_{n\to\infty} T_n(\lambda,\omega)=T_\infty(\lambda,\omega)<\infty.\ a.s.\ \omega\in D(\omega) \qquad (18)$$

Denote $P_t(x_t|X_{1_t})$ by P_t in brief, by the definition of $H_t(\omega)$, we have

$$\sum_{x_t\in S}\frac{\lambda Y_t^{b_t}[\log P_t(x_t|X_{1_t})+H_t(\omega)]}{\sum_{k=1}^t Y_k^{b_k}}\cdot P_t(x_t|X_{1_t})=\frac{\lambda Y_t^{b_t}[H_t(\omega)-H_t(\omega)]}{\sum_{k=1}^t Y_k^{b_k}}=0. \qquad (19)$$

By (14), (19) and the inequality $0\le e^x-1-x\le(1/2)x^2e^{|x|}$, the entropy density inequality $H_t(\omega)\le\log N$, noticing that $\lambda=\pm1$, $0\le Y_t^{b_t}\le a^{b_t}\le a^b$, we have

$$
\begin{aligned}
&0\le U_t(\lambda;X_t)-1 \\
&= \sum_{x_t\in S}\left\{\exp\{\frac{\lambda Y_t^{b_t}(\log P_t+H_t(\omega))}{\sum_{k=1}^t Y_k^{b_k}}\}-1-\frac{\lambda Y_t^{b_t}(\log P_t+H_t(\omega))}{\sum_{k=1}^t Y_k^{b_k}}\right\}P_t
\end{aligned}
$$

$$\le \frac{1}{2(\sum_{k=1}^t Y_k^{b_k})^2}\sum_{x_t\in S} Y_t^{2b_t}(\log P_t+H_t(\omega))^2\exp\{\frac{|Y_t^{b_t}||\log P_t+H_t(\omega)|}{\sum_{k=1}^t Y_k^{b_k}}\}P_t$$
$$\le \frac{1}{2(\sum_{k=1}^t Y_k^{b_k})^2}\sum_{x_t\in S} a^{2b}(\log P_t+H_t(\omega))^2\exp\{\frac{Y_t^{b_t}(|\log P_t|+|H_t(\omega)|)}{\sum_{k=1}^t Y_k^{b_k}}\}P_t$$
$$\le \frac{1}{2(\sum_{k=1}^t Y_k^{b_k})^2}\sum_{x_t\in S} a^{2b}(\log P_t+H_t(\omega))^2\exp\{\frac{a^b(-\log P_t+\log N)}{\sum_{k=1}^t Y_k^{b_k}}\}P_t. \quad (20)$$

It is easy to see $\sum_{k=1}^{t} Y_k^{b_k} \to \infty,\ as\ t\to\infty$ from Definition 2 of $\{Y_k^{b_k}, k\ge 0\}$. Therefore, there exists a positive integer m such that $\sum_{k=1}^t Y_k^{b_k}\ge 2a^b$ as $t\ge m$. Hence as $t\ge m$, by (20) and the entropy density inequality, we obtain

$$0\le U_t(\lambda;X_t)-1$$
$$\le \frac{1}{2(\sum_{k=1}^t Y_k^{b_k})^2}\sum_{x_t\in S} a^{2b}(\log P_t+H_t(\omega))^2\exp\{\frac{-\log P_t+\log N}{2}\}P_t$$
$$\le \frac{1}{2(\sum_{k=1}^t Y_k^{b_k})^2}\sum_{x_t\in S} a^{2b}(\log P_t+H_t(\omega))^2\exp\{\log(N/P_t)^{1/2}\}P_t$$
$$\le \frac{N}{2(\sum_{k=1}^t Y_k^{b_k})^2}\sum_{x_t\in S} a^{2b}(\log P_t+H_t(\omega))^2 P_t^{1/2}$$
$$\le \frac{N}{2(\sum_{k=1}^t Y_k^{b_k})^2}\sum_{x_t\in S} a^{2b}[(\log P_t)^2P_t^{1/2}+2H_t(\omega)\cdot P_t^{1/2}\log P_t+(H_t(\omega))^2]$$
$$\le \frac{N}{2(\sum_{k=1}^t Y_k^{b_k})^2}\sum_{x_t\in S} a^{2b}[(\log P_t)^2P_t^{1/2}-2\log N\cdot P_t^{1/2}\log P_t+(\log N)^2]. \quad (21)$$

We can calculate

$$\max\{x^{1/2}(\log x)^2,\ 0<x\le 1\}=16e^{-2};$$
$$\max\{-x^{1/2}\log x,\ 0<x\le 1\}=2e^{-1}.$$

By (21), noticing $Y_t^{b_t}\ge s^{b_t}\ge G=\min\{1,s^b\}$, we have

$$\sum_{t\in T^{(n)}\setminus\{o\}}(U_t(\lambda;X_t)-1)$$
$$\le \sum_{t\in T^{(n)}\setminus\{o\}}\frac{N}{2(\sum_{k=1}^t Y_k^{b_k})^2}\sum_{x_t\in S} a^{2b}[(\log P_t)^2P_t^{1/2}-2\log N\cdot P_t^{1/2}\log P_t+(\log N)^2]$$
$$\le \sum_{t\in T^{(n)}\setminus\{o\}}\sum_{x_t\in S}\frac{Na^{2b}}{2(\sum_{k=1}^t Y_k^{b_k})^2}[16e^{-2}+2(\log N)2e^{-1}+(\log N)^2]$$
$$= \sum_{t\in T^{(n)}\setminus\{o\}}\frac{N^2a^{2b}}{2(\sum_{k=1}^t Y_k^{b_k})^2}[16e^{-2}+2(\log N)2e^{-1}+(\log N)^2]$$
$$\le \sum_{t\in T^{(n)}\setminus\{o\}}\frac{N^2a^{2b}}{2(\sum_{k=1}^t G)^2}[16e^{-2}+2(\log N)2e^{-1}+(\log N)^2]$$
$$\le \sum_{t\in T^{(n)}\setminus\{o\}}\frac{N^2a^{2b}}{2G^2t^2}[16e^{-2}+4(\log N)e^{-1}+(\log N)^2]<\infty\ a.s. \quad (22)$$

where $\sum\limits_{t\in T^{(n)}\backslash\{o\}} \frac{1}{t^2} < \infty$ as $n \to \infty$ is obvious.

By the convergence theorem of infinite production, (22) implies that

$$\lim_{n\to\infty} \prod_{t\in T^{(n)}\backslash\{o\}} U_t(\lambda; X_t) \text{ converges. a.s. } \omega \in D(\omega) \tag{23}$$

By (17), (18) and (23), we obtain

$$\lim_{n\to\infty} \exp\left\{ \sum_{t\in T^{(n)}\backslash\{o\}} \frac{\lambda Y_t^{b_t}(\log P_t(X_t|X_{1_t}) + H_t(\omega))}{\sum_{k=1}^{t} Y_k^{b_k}} \right\} = \text{ a finite number. a.s.} \tag{24}$$

Letting $\lambda = 1$ and $\lambda = -1$ in (24), respectively, we have

$$\lim_{n\to\infty} \exp\left\{ \sum_{t\in T^{(n)}\backslash\{o\}} \frac{Y_t^{b_t}[\log P_t(X_t|X_{1_t}) + H_t(\omega)]}{\sum_{k=1}^{t} Y_k^{b_k}} \right\} = \text{ a finite number. a.s.} \tag{25}$$

$$\lim_{n\to\infty} \exp\left\{ \sum_{t\in T^{(n)}\backslash\{o\}} \frac{-Y_t^{b_t}[\log P_t(X_t|X_{1_t}) + H_t(\omega)]}{\sum_{k=1}^{t} Y_k^{b_k}} \right\} = \text{ a finite number. a.s.} \tag{26}$$

(25) and (26) imply that

$$\lim_{n\to\infty} \sum_{t\in T^{(n)}\backslash\{o\}} \frac{Y_t^{b_t}[\log P_t(X_t|X_{1_t}) + H_t(\omega)]}{\sum_{k=1}^{t} Y_k^{b_k}} \text{ converges. a.s.} \tag{27}$$

Hence (11) holds. By (27) and Kronecker's lemma, we have

$$\lim_{n\to\infty} \frac{1}{\sum_{k=1}^{|T^{(n)}|} Y_k^{b_k}} \sum_{t\in T^{(n)}\backslash\{o\}} Y_t^{b_t}[\log P_t(X_t|X_{1_t}) + H_t(\omega)] = 0. \text{ a.s.} \tag{28}$$

Moreover, from (10) and (28) we obtain

$$\begin{aligned} & \lim_{n\to\infty} \frac{1}{\sum\limits_{t\in T^{(n)}\backslash\{o\}} Y_t^{b_t}} \sum_{t\in T^{(n)}\backslash\{o\}} Y_t^{b_t}[\log P_t(X_t|X_{1_t}) + H_t(\omega)] \\ = & -\lim_{n\to\infty} \frac{1}{\sum\limits_{t\in T^{(n)}\backslash\{o\}} Y_t^{b_t}} \sum_{t\in T^{(n)}\backslash\{o\}} [-Y_t^{b_t} \log P_t(X_t|X_{1_t}) - Y_t^{b_t} H_t(\omega)] \\ = & -\lim_{n\to\infty} [S_n(\omega) - \frac{1}{\sum\limits_{t\in T^{(n)}\backslash\{o\}} Y_t^{b_t}} \sum_{t\in T^{(n)}\backslash\{o\}} Y_t^{b_t} H_t(\omega)] = 0 \ (29) \end{aligned}$$

(12) follows from (29) immediately. □

3. Some Corollaries

Corollary 1. *Let* $X = \{X_t, t \in T\}$ *be a nonhomogeneous Markov chain indexed by a homogeneous tree,* $f_n(\omega)$ *and* $H_t(\omega)$ *be defined as (7) and Theorem 1. Then*

$$\lim_{n\to\infty} \sum_{t\in T^{(n)}\setminus\{o\}} \frac{[\log P_t(X_t|X_{1_t}) + H_t(\omega)]}{t} < \infty,\ a.s. \tag{30}$$

$$\lim_{n\to\infty} [f_n(\omega) - \frac{1}{|T^{(n)}|} \sum_{t\in T^{(n)}\setminus\{o\}} H_t(\omega)] = 0.\ a.s. \tag{31}$$

Proof. Let T be a homogeneous tree, that is on the tree each vertex has M neighboring vertices. Let $Y_t \equiv 1,\ b_t = b,\ t \in T^{(n)}$, we obtain $\sum_{k=1}^{t} Y_k^{b_t} = t,\ \sum_{t\in T^{(n)}\setminus\{o\}} Y_t^{b_t} = |T^{(n)}| - 1$. Hence (30) and (31) follow from (11) and (12) directly. □

Remark. Equation (31) is a result of Yang and Ye (see[8]).
When the successor of each vertex on the infinite tree with the uniformly bounded degree has only one vertex, the nonhomogeneous Markov chain on the tree degenerates into the general nonhomogeneous Markov chain.

Corollary 2. *Let* $\{X_n, n \geq 0\}$ *be a nonhomogeneous Markov chain with the initial distribution and the transition probabilities as follows:*

$$p(i) > 0,\ i \in S.$$

$$P_t(j|i) > 0,\ i,j \in S,\ t = 1, 2, \cdots.$$

We denote

$$f_n(\omega) = -\frac{1}{n+1}[\log P(X_0) + \sum_{t=1}^{n} \log P_t(X_t|X_{t-1})], \tag{32}$$

$$H_t(\omega) = -\sum_{x_t\in S} P_t(x_t|X_{t-1}) \log P_t(x_t|X_{t-1}). \tag{33}$$

Then

$$\sum_{t=1}^{\infty} \frac{\log P_t(X_t|X_{t-1}) + H_t(\omega)}{t} < \infty,\ a.s. \tag{34}$$

$$\lim_{n\to\infty} [f_n(\omega) - \frac{1}{n+1} \sum_{t=1}^{n} H_t(\omega)] = 0.\ a.s. \tag{35}$$

Proof. At this time the nonhomogeneous Markov chain $X = \{X_t, t \in T\}$ indexed by the infinite tree is changed into the general nonhomogeneous Markov chain $\{X_n, n \geq 0\}$, we obtain $P_t(X_t|X_{1_t}) = P_t(X_t|X_{t-1})$, $|T^{(n)}| = n + 1$. (32)-(35) follow from (7), (10), (30) and (31), respectively. □

Remark. Equation (35) is just Theorem 2 of Liu and Yang (see [1]).

References

1. W. Liu and W.G. Yang, *An extension of Shannon-McMillan theorem and some limit properties for nonhomogeneous Markov chains*, Stochastic Process. Appl **61**(1996), 129-145.
2. W. Liu and W.G. Yang, *Some extension of Shannon-McMillan theorem*, J. of Combinatorics Information and System Science **21**(1996), 211-223.
3. Z. Ye and T. Berger, *Information Measure for Discrete Random Fields*, Science Press, Beijing, New York. 1998.
4. T. Berger and Z. Ye, *Entropic aspects of random fields on trees*, IEEE Trans. Inform. Theory **36**(1990), 1006-1018.
5. Z. Ye and T. Berger, *Ergodic regularity and asymptotic equipartition property of random fields on trees*, Combin. Inform. System. Sci **21**(1996), 157-184.
6. W.G. Yang, *Some limit properties for Markov chains indexed by homogeneous tree* , Stat. Probab. Letts. **65**(2003), 241-250.
7. W. Liu and W.G. Yang, *Some strong limit theorems for Markov chain fields on trees*, Probability in the Engineering and Informational Science **18**(2004), 411-422.
8. W.G. Yang and Z. Ye, *The asymptotic equipartition property for nonhomogeneous Markov chains indexed by a homogeneous tree*, IEEE Trans. Inform. Theory **53**(2007), 3275-3280
9. K.L. Chung, *A Course in Probability Theory*, Academic Press, New York. 1974.
10. J.G. Kemeny, J.L. Snell and A.W. Knapp, *Denumerabl Markov chains*, Springer, New York. 1974
11. F. Spitzer, *Markov random fields on an infinite tree*, Ann. Probab. **3**(1975), 387-398
12. J.L. Doob. *Stochastic Process*, Wiley, New York. 1953.
13. K.K. Wang, *Some research on Shannon-McMillan theorem for mth-Order nonhomogeneous Markov information source*, Stochastic Analysis and Applications **27** (2009), 1117-1128.
14. P. Billingsley, *Probability and Measure*, Wiley, New York. 1986.
15. R.V. Mises, *Mathematical Theory of Probability and Statistics*, Academic Press, New York. 1964.
16. A.N. Kolmogorov, *On the logical foundation of probability theory*, Springer-Verlag. New York, 1982.
17. W. Liu and Z. Wang, *An extension of a theorem on gambling systems to arbitrary binary random variables*, Statistics and Probability Letters **28**(1996), 51-58.
18. W. Liu, *A limit property of arbitrary discrete information sources*, Taiwanese J. Math. **3**(1999), 539-546.

Kangkang Wang received M.Sc. from Jiangsu University and Ph.D at Nanjing University of Aeronautics and Astronautics. Since 2005 he has been at Jiangsu University of Science and Technology. His research interests include limit theory in probability theory and information theory.

School of Mathematics and Physics, Jiangsu University of Science and Technology, Zhenjiang 212003, China.
e-mail: wkk.cn@126.com

Zurun Xu received M.Sc. from Jiangsu University. Since 2001 he has been at Jiangsu University of Science and Technology. His research interests include queueing theory.

School of Mathematics and Physics, Jiangsu University of Science and Technology, Zhenjiang 212003, China.
e-mail: marlboro-xu@163.com