Journal of Contemporary Applied Mathematics
V. 4, No 1, 2014, September
ISSN 2222-5498

The generalized Shannon-McMillan theorems for second-order nonhomogeneous Markov chains on the gambling systems indexed by a double rooted tree

Kangkang Wang

Abstract. In this paper, our aim is to establish a generalized Shannon-McMillan theorem for the second-order nonhomogeneous Markov chains indexed by a double rooted tree is discussed by constructing a nonnegative martingale and using analytical methods. As corollaries, we achieve some Shannon-Mcmillan theorems for the second-order nonhomogeneous Markov chains indexed by a homogeneous tree and a second-order nonhomogeneous Markov chain. Some results which have been gained are extended in a sense.

Key Words and Phrases: Shannon-McMillan theorem, double rooted tree, second-order Markov chains, the entropy density, generalized gambling system.

2000 Mathematics Subject Classifications: 60F15

1. Introduction

We specify that a tree is a graph $G = \{T, E\}$ which is connected and contains no circuits. Given any two vertices $\sigma, t(\ \sigma \neq t \in T)$, let $\overline{\sigma t}$ be the unique path connecting σ and t. Define the graph distance $d(\sigma, t)$ to be the number of edges contained in the path $\overline{\sigma t}$.

Let T_o be an arbitrary infinite tree that is partially finite (i.e. it has infinite vertices, and each vertex connects with finite vertices) and has a root o. Meanwhile, we consider another kind of double root tree T, that is, it is formed with the root o of T_o connecting with an arbitrary point denoted by the root -1. For a better explanation of the double root tree T, we take Cayley tree $T_{C,N}$ for example. It's a special case of the tree T_o, the root o of Cayley tree has N neighbors and all the other vertices of it have $N+1$ neighbors each. The double root tree $T'_{C,N}$ (see Fig.1) is formed with root o of tree $T_{C,N}$ connecting with another root -1.

Let σ, t be vertices of the double root tree T. Write $t \leq \sigma$ $(\sigma, t \neq -1)$ if t is on the unique path connecting o to σ, and $|\sigma|$ for the number of edges on this path. For any two vertices σ, t ($\sigma, t \neq -1$) of the tree T, denote by $\sigma \wedge t$ the vertex farthest from o satisfying $\sigma \wedge t \leq \sigma$ and $\sigma \wedge t \leq t$.

The set of all vertices with distance n from root o is called the n-th generation of T, which is denoted by L_n. We say that L_n is the set of all vertices on level n and especially root -1 is on the -1st level on tree T. We denote by $T^{(n)}$ the subtree of the tree T containing the vertices from level -1 (the root -1) to level n and denote by $T_o^{(n)}$ the subtree of the tree T_o containing the vertices from level 0 (the root o) to level n. Denote by $t(\neq o, -1)$ the t-th vertex from the root 0 to the upper part, from the left side to the right side on the tree. 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. Let $X^A = \{X_t, t \in A\}$, and let x^A be a realization of X^A and denote by $|A|$ the number of vertices of A .

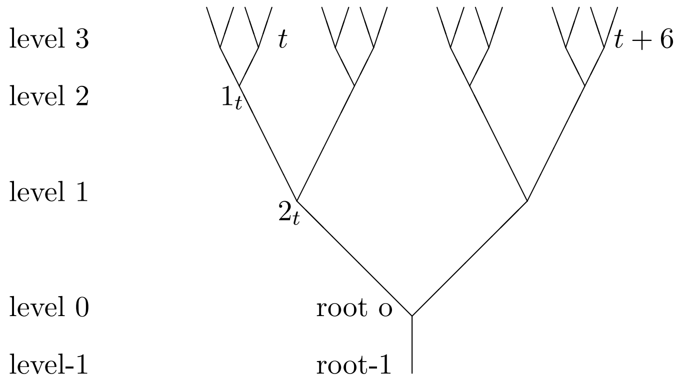

Fig.1 Double root tree $T'_{C,2}$

Definition 1. *Let $S = \{s_1, s_2, \cdots, s_N\}$ and $P(z|y,x)$ be nonnegative functions on S^3. Let*

$$P = (P(z|y,x)), \quad P(z|y,x) \geq 0, \quad x, y, z \in S.$$

If

$$\sum_{z \in S} P(z|y,x) = 1,$$

then P is called a second-order transition matrix.

Definition 2. *Suppose that T is a double rooted tree and $S = \{s_1, s_2, \cdots, s_N\}$ is a finite state space, and $\{X_t, t \in T\}$ is a collection of S-valued random variables defined on the probability space $(\Omega, \mathcal{F}, P)$. Let*

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

be a distribution on S^2, and

$$P_t = (P_t(z|y,x)), \quad x,y,z \in S, t \in T\backslash\{o\}\{-1\} \tag{2}$$

be a collection of second-order transition matrices. For any vertex t ($t \neq o, -1$), if

$$P(X_t = z|X_{1_t} = y, X_{2_t} = x, \textit{and } X_\sigma \textit{ for } \sigma \wedge t \leq 1_t)$$

$$= P(X_t = z|X_{1_t} = y, X_{2_t} = x) = P_t(z|y,x) \quad \forall x,y,z \in S \tag{3}$$

and

$$P(X_{-1} = x, X_o = y) = p(x,y), \quad x,y \in S, \tag{4}$$

then $\{X_t, t \in T\}$ is called a S-valued second-order nonhomogeneous Markov chain indexed by a tree T with the initial distribution (1) and second-order transition matrices (2), or called a T-indexed second-order nonhomogeneous Markov chain.

Remark 1. *Benjamini and Peres [9] have given the definition of the tree-indexed homogeneous Markov chains. Here we improve their definition and give the definition of the tree-indexed second-order nonhomogeneous Markov chains in a similar way.*

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

$$P(x^{T^{(n)}}) = P(X^{T^{(n)}} = x^{T^{(n)}}) = P(X_{-1} = x_{-1}, X_o = x_o) \prod_{t \in T^{(n)}\backslash\{o\}\{-1\}} P_t(x_t|x_{1_t}, x_{2_t}). \tag{5}$$

Let T be a tree, $\{X_t, t \in T\}$ be a stochastic process indexed by the tree T with the state space S. Denote

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

Let

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

$f_n(\omega)$ 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 2, we get by (5)

$$f_n(\omega) = -\frac{1}{|T^{(n)}|}[\log P(X_o, X_{-1}) + \sum_{t \in T^{(n)}\backslash\{o\}\{-1\}} \log P_t(X_t|X_{1_t}, X_{2_t})], \tag{7}$$

where $|T^{(n)}|$ represents the number of all the vertices from level -1 to level n.

Definition 3. *Suppose that* $\{f_n(x_0, \cdots, x_n), n \geq 0\}$ *is a set of real-valued functions defined on* $S^{n+1}(n = 1, 2, \cdots)$*, which will be called the generalized random selection functions if they take values in an interval of* $[0, b]$*. We let*

$$Y_0 = y(y \text{ is an arbitrary real number})$$

$$Y_{t+1} = f_{|t|}(X_{1_t}, X_{2_t} \cdots, X_o, X_{-1}), \quad |t| \geq 1, \tag{8}$$

where $|t|$ *stands for the number of the edges on the path from the root* -1 *to the vertex* t*.* $\{Y_t,\ t \in T^{(n)}\}$ *is called the generalized gambling system or the generalized random selection system indexed by a double rooted tree. The traditional random selection system*[17] *takes values in the set of* $\{0, 1\}$*.*

We first explain the conception of the traditional random selection, which is the crucial part of the gambling system. We give a set of real-valued functions $f_n(x_0, \cdots, x_n)$ defined on $S^n(n = 0, 1, 2, \cdots)$, which will be called the random selection function if they take values in a two-valued set $\{0, 1\}$. Then let

$$Y_0 = y(y \text{ is an arbitrary real number}),$$

$$Y_{n+1} = f_n(X_0 \cdots, X_n), \quad n \geq 0.$$

where $\{Y_n, n \geq 0\}$ be called the gambling system or the random selection system.

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 $\{g_n(x, y, z), n \geq 2\}$ be a real-valued function sequence defined on S^3. Interpret X_n as the result of the nth trial, the type of which may change at each step. Let $\mu_n = Y_n g_n(X_{n-2}, X_{n-1}, X_n)$ denote the gain of the bettor at the nth trial, where Y_n represents the bet size, $g_n(X_{n-2}, X_{n-1}, X_n)$ 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 two trials. 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} and X_{n-2} as $n \geq 2$, and b_0, b_1 are constants. Thus $\sum_{k=2}^n Y_k g_k(X_{k-2}, X_{k-1}, X_k)$ represents the total gain in the first n trials, $\sum_{k=2}^n b_k$ the accumulated entrance fees, and $\sum_{k=2}^n [Y_k g_k(X_{k-2}, X_{k-1}, X_k) - b_k]$ the accumulated net gain. Motivated by the classical definition of "fairness" of game of chance (see Kolmogorov[16]), we introduce the following definition:

Definition 4. *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*

$$\lim_{n\to\infty} \frac{1}{\sum_{k=2}^n Y_k} \sum_{k=2}^n [Y_k g_k(X_{k-2}, X_{k-1}, X_k) - b_k] = 0 \qquad a.s. \quad on \ \ \{\omega : \sum_{k=2}^\infty Y_k = \infty\}.$$

Definition 5. *Suppose that $\{Y_t, t \in T^{(n)}\}$ is a generalized gambling system defined by (8), we call*

$$S_n(\omega) = -\frac{1}{\sum_{t\in T^{(n)}} Y_t}[Y_0 \log P(X_o, X_{-1}) + \sum_{t\in T^{(n)}\setminus\{o\}\{-1\}} Y_t \log P_t(X_t|X_{1_t}, X_{2_t})] \quad (9)$$

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 tree models have drawn increasing interests from specialists in probability, physics and information theory in recent years. There have been some works on limit theorems for tree-indexed stochastic process, among them the convergence of $f_n(\omega)$ in a sense (L_1 convergence, convergence in probability, or almost sure convergence) indexed by a tree is called the tree-indexed Shannon-McMillan theorem or the asymptotic equipartition property(AEP) in information theory. Shannon-McMillan theorems on the Markov chain have been studied extensively(see[1],[2]). Liu has introduced a type of new analytical methods for Shannon-McMillan theorems in his work (see[12]). 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]). 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 Liu and Yang's works (see[1, 8]), they firstly construct a superior martingale, then they scale inequalities to attain the Shannon-McMillan theorems by virtue of superior limit and inferior limit properties. In this paper, we investigate Shannon-McMillan theorems for second-order nonhomogeneous Markov chains field on the tree by constructing the consistent distribution functions. We first give a series theorem by controlling the convergence of a series. Correspondingly, a limit theorem for the random conditional entropy is obtained by Kronecker's lemma. The limit theorem shows that the difference of the logarithm $\log P_t(X_t|X_{1_t}, X_{2_t})$ relative to its expectation in the first n levels on the tree is to be of smaller order of magnitude than the stochastic sequence $\sum\limits_{t\in T^{(n)}\setminus\{o\}\{-1\}} Y_t$ as n tends to infinity. In particular, if $Y_t \equiv 1,\ t \in T^{(n)}\setminus\{o\}\{-1\}$, the limit theorem for the random conditional entropy on the gambling system becomes the Shannon-McMillan theorem for second-order nonhomogeneous Markov chains field on the tree. Finally, when the tree model degenerates into the chain model, the Shannon-McMillan theorem for second-order nonhomogeneous Markov chains field on the tree is changed into the Shannon-McMillan theorem for the general second-order nonhomogeneous Markov chain. Liu and Yang's (see [1, 8]) results are extended in fact.

2. Main result and its proof

Theorem 1. *Suppose that T is a double rooted tree, $X = \{X_t, t \in T\}$ is a T-indexed Markov chain with the state space S defined as before, $S_n(\omega)$ is defined as (9). Denote by $H_t(\omega)$ the random conditional entropy of X_t relative to X_{1_t}, that is*

$$H_t(\omega) = -\sum_{x_t\in S} P_t(x_t|X_{1_t}, X_{2_t}) \log P_t(x_t|X_{1_t}, X_{2_t}), \quad t \in T^{(n)}\setminus\{o\}\{-1\}. \tag{10}$$

If $\sum_{k=1}^{n} Y_k = O(n)$, we have

$$\lim_{n\to\infty} \sum_{t\in T^{(n)}\setminus\{o\}\{-1\}} \frac{Y_t[\log P_t(X_t|X_{1_t}, X_{2_t}) + H_t(\omega)]}{\sum_{k=1}^{t} Y_k} < \infty. \qquad a.s. \tag{11}$$

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

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

$$\mu_Q(\lambda; x^{T^{(n)}}) = \frac{p(x_0, x_{-1}) \prod\limits_{t\in T^{(n)}\setminus\{o\}\{-1\}} P_t(x_t|x_{1_t}, x_{2_t}) \exp\left\{\frac{\lambda Y_t(\log P_t(x_t|x_{1_t}, x_{2_t}) + H_t(\omega))}{\sum_{k=1}^{t} Y_k}\right\}}{\prod\limits_{t\in T^{(n)}\setminus\{o\}\{-1\}} U_t(\lambda; x_t)}, \tag{13}$$

where

$$\begin{aligned} U_t(\lambda; x_t) &= E\left\{\exp\left\{\frac{\lambda Y_t(\log P_t(X_t|X_{1_t}, X_{2_t}) + H_t(\omega))}{\sum_{k=1}^{t} Y_k}\right\} |X_{1_t} = x_{1_t}, X_{2_t} = x_{2_t}\right\} \\ &= \sum_{x_t\in S} \exp\left\{\frac{\lambda Y_t(\log P_t(x_t|x_{1_t}, x_{2_t}) + H_t(\omega))}{\sum_{k=1}^{t} Y_k}\right\} \cdot P_t(x_t|x_{1_t}, x_{2_t}). \end{aligned}$$

$$t \in T^{(n)} \backslash \{o\}\{-1\}. \tag{14}$$

By (14) and (13), in the case $n \geq 1$ we can write

$$\begin{aligned}
&\sum_{x^{L_n} \in S^{L_n}} \mu_Q(\lambda; x^{T^{(n)}}) \\
= &\sum_{x^{L_n} \in S^{L_n}} \frac{p(x_0, x_{-1}) \prod_{t \in T^{(n)} \backslash \{o\}\{-1\}} P_t(x_t|x_{1_t}, x_{2_t}) \exp\left\{\frac{\lambda Y_t(\log P_t(x_t|x_{1_t}, x_{2_t}) + H_t(\omega))}{\sum_{k=1}^t Y_k}\right\}}{\prod_{t \in T^{(n)} \backslash \{o\}\{-1\}} U_t(\lambda; x_t)} \\
= &\mu_Q(\lambda; x^{T^{(n-1)}}) \frac{\sum_{x^{L_n} \in S^{L_n}} \prod_{t \in L_n} P_t(x_t|x_{1_t}, x_{2_t}) \exp\left\{\frac{\lambda Y_t(\log P_t(x_t|x_{1_t}, x_{2_t}) + H_t(\omega))}{\sum_{k=1}^t Y_k}\right\}}{\prod_{t \in L_n} U_t(\lambda; x_t)} \\
= &\mu_Q(\lambda; x^{T^{(n-1)}}) \frac{\prod_{t \in L_n} \sum_{x_t \in S} P_t(x_t|x_{1_t}, x_{2_t}) \exp\left\{\frac{\lambda Y_t(\log P_t(x_t|x_{1_t}, x_{2_t}) + H_t(\omega))}{\sum_{k=1}^t Y_k}\right\}}{\prod_{t \in L_n} U_t(\lambda; x_t)} \\
= &\mu_Q(\lambda; x^{T^{(n-1)}}) \frac{\prod_{t \in L_n} U_t(\lambda; x_t)}{\prod_{t \in L_n} U_t(\lambda; x_t)} = \mu_Q(\lambda; x^{T^{(n-1)}}). \qquad a.s.
\end{aligned} \tag{15}$$

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

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

Combining (5) with (13), we can rewrite (16) as

$$T_n(\lambda, \omega) = \frac{\exp\left\{\sum_{t \in T^{(n)} \backslash \{o\}\{-1\}} \frac{\lambda Y_t(\log P_t(X_t|X_{1_t}, X_{2_t}) + H_t(\omega))}{\sum_{k=1}^t Y_k}\right\}}{\prod_{t \in T^{(n)} \backslash \{o\}\{-1\}} U_t(\lambda; X_t)}, \quad n \geq 0. \tag{17}$$

Since $\mu_Q(\lambda; X^{T^{(n)}})$ and $P(X^{T^{(n)}})$ are two distribution functions, we immediately know that $T_n(\lambda, \omega)$ is a nonnegative sup-martingale from Doob's martingale convergence theorem. Therefore,

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

Denote $P_t(x_t|X_{1_t}, X_{2_t})$ by P_t in brief, by (10) we can conclude

$$\sum_{x_t \in S} \frac{\lambda Y_t[\log P_t(x_t|X_{1_t}, X_{2_t}) + H_t(\omega)]}{\sum_{k=1}^t Y_k} \cdot P_t(x_t|X_{1_t}, X_{2_t}) = \frac{\lambda Y_t[H_t(\omega) - H_t(\omega)]}{\sum_{k=1}^t Y_k} = 0. \tag{19}$$

Taking into account (14), (19) and the inequality $0 \leq e^x - 1 - x \leq (1/2)x^2 e^{|x|}$, the entropy density inequality $H_t(\omega) \leq \log N$, noticing that $\lambda = \pm 1$, $0 \leq Y_t \leq b$, we can write

$$0 \leq U_t(\lambda; X_t) - 1$$

$$
\begin{aligned}
&= \sum_{x_t\in S}\left\{\exp\{\frac{\lambda Y_t(\log P_t+H_t(\omega))}{\sum_{k=1}^t Y_k}\}-1-\frac{\lambda Y_t(\log P_t+H_t(\omega))}{\sum_{k=1}^t Y_k}\right\}P_t\\
&\le \frac{1}{2(\sum_{k=1}^t Y_k)^2}\sum_{x_t\in S} Y_t^2(\log P_t+H_t(\omega))^2\exp\{\frac{|Y_t||\log P_t+H_t(\omega)|}{\sum_{k=1}^t Y_k}\}P_t\\
&\le \frac{1}{2(\sum_{k=1}^t Y_k)^2}\sum_{x_t\in S} b^2(\log P_t+H_t(\omega))^2\exp\{\frac{b(-\log P_t+\log N)}{\sum_{k=1}^t Y_k}\}P_t. \qquad (20)
\end{aligned}
$$

It is easy to know $\sum\limits_{k=1}^{t} Y_k \to \infty,\ as\ t\to\infty$ from $\sum\limits_{k=1}^{n} Y_k = O(n)$. Moreover, there exists a positive integer m such that $\sum_{k=1}^t Y_k \ge 2b$ as $t \ge m$. Hence as $t\ge m$, according to (20) and the entropy density inequality, we can reach

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

We can easily calculate that

$$M_1=\max\{x^{1/2}(\log x)^2,\quad 0<x\le 1\}=16e^{-2};$$

$$M_2=\max\{-x^{1/2}\log x,\quad 0<x\le 1\}=2e^{-1}.$$

By use of (21), we get

$$
\begin{aligned}
&\sum_{t\in T^{(n)}\setminus\{o\}\{-1\}}(U_t(\lambda;X_t)-1)\\
&\le \sum_{t\in T^{(n)}\setminus\{o\}\{-1\}}\frac{N}{2(\sum_{k=1}^t Y_k)^2}\sum_{x_t\in S} b^2[(\log P_t)^2P_t^{1/2}-2\log N\cdot P_t^{1/2}\log P_t+(\log N)^2]\\
&\le \sum_{t\in T^{(n)}\setminus\{o\}\{-1\}}\sum_{x_t\in S}\frac{Nb^2}{2(\sum_{k=1}^t Y_k)^2}[M_1+2(\log N)M_2+(\log N)^2]
\end{aligned}
$$

$$= \sum_{t\in T^{(n)}\backslash\{o\}\{-1\}} \frac{b^2N^2}{2(\sum_{k=1}^t Y_k)^2}[16e^{-2}+4e^{-1}(\log N)+(\log N)^2] < \infty, \qquad a.s. \qquad (22)$$

where we easily see $\sum\limits_{t\in T^{(n)}\backslash\{o\}\{-1\}} \frac{1}{(\sum_{k=1}^t Y_k)^2} < \infty$ as $n\to\infty$ in virtue of $\sum\limits_{k=1}^t Y_k = O(t)$.

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

$$\lim_{n\to\infty} \prod_{t\in T^{(n)}\backslash\{o\}\{-1\}} U_t(\lambda;X_t) \quad converges. \qquad a.s. \qquad (23)$$

In virtue of (17), (18) and (23), we obtain

$$\lim_{n\to\infty} \exp\left\{ \sum_{t\in T^{(n)}\backslash\{o\}\{-1\}} \frac{\lambda Y_t(\log P_t(X_t|X_{1_t},X_{2_t})+H_t(\omega))}{\sum_{k=1}^t Y_k} \right\} = \ a\ finite\ number. \quad a.s. \qquad (24)$$

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

$$\lim_{n\to\infty} \exp\left\{ \sum_{t\in T^{(n)}\backslash\{o\}\{-1\}} \frac{Y_t[\log P_t(X_t|X_{1_t},X_{2_t})+H_t(\omega)]}{\sum_{k=1}^t Y_k} \right\} = \ a\ finite\ number. \quad a.s. \qquad (25)$$

$$\lim_{n\to\infty} \exp\left\{ \sum_{t\in T^{(n)}\backslash\{o\}\{-1\}} \frac{-Y_t[\log P_t(X_t|X_{1_t},X_{2_t})+H_t(\omega)]}{\sum_{k=1}^t Y_k} \right\} = \ a\ finite\ number. \quad a.s. \qquad (26)$$

(25) and (26) imply that

$$\lim_{n\to\infty} \sum_{t\in T^{(n)}\backslash\{o\}\{-1\}} \frac{Y_t[\log P_t(X_t|X_{1_t},X_{2_t})+H_t(\omega)]}{\sum_{k=1}^t Y_k} \quad converges. \quad a.s. \qquad (27)$$

Hence (11) holds. By virtue of (27) and Kronecker's lemma, we achieve

$$\lim_{n\to\infty} \frac{1}{\sum_{k=1}^{|T^{(n)}|} Y_k} \sum_{t\in T^{(n)}\backslash\{o\}\{-1\}} Y_t[\log P_t(X_t|X_{1_t},X_{2_t})+H_t(\omega)] = 0. \qquad a.s. \qquad (28)$$

Moreover, from (9) and (28) we gain

$$\lim_{n\to\infty} \frac{1}{\sum\limits_{t\in T^{(n)}\backslash\{o\}\{-1\}} Y_t} \sum_{t\in T^{(n)}\backslash\{o\}\{-1\}} Y_t[\log P_t(X_t|X_{1_t},X_{2_t})+H_t(\omega)]$$

$$= \ -\lim_{n\to\infty} \frac{1}{\sum\limits_{t\in T^{(n)}\backslash\{o\}\{-1\}} Y_t} \sum_{t\in T^{(n)}\backslash\{o\}\{-1\}} [-Y_t\log P_t(X_t|X_{1_t},X_{2_t}) - Y_tH_t(\omega)]$$

$$= \quad -\lim_{n\to\infty}[S_n(\omega) - \frac{1}{\sum\limits_{t\in T^{(n)}\setminus\{o\}\{-1\}} Y_t} \sum_{t\in T^{(n)}\setminus\{o\}\{-1\}} Y_t H_t(\omega)] = 0. \qquad a.s. \tag{29}$$

(12) follows from (29) immediately.

3. Some Shannon-McMillan theorems for second-order nonhomogeneous Markov chains on the homogeneous tree

Corollary 1. *Suppose that $X = \{X_t, t \in T\}$ is a second-order nonhomogeneous Markov chain indexed by a homogeneous tree, $f_n(\omega)$ and $H_t(\omega)$ be defined as (7) and (10). Then*

$$\lim_{n\to\infty} \sum_{t\in T^{(n)}\setminus\{o\}\{-1\}} \frac{[\log P_t(X_t|X_{1_t}, X_{2_t}) + H_t(\omega)]}{t} < \infty, \qquad a.s. \tag{30}$$

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

Proof. We let T be a homogeneous tree, that is on the tree each vertex has M neighboring vertices. Let $Y_t \equiv 1,\ t \in T^{(n)}$, we reach $\sum_{k=1}^{t} Y_k = t$, $\sum\limits_{t\in T^{(n)}\setminus\{o\}\{-1\}} Y_t = |T^{(n)}| - 2$, $S_n(\omega) = f_n(\omega)$. Hence $\sum_{k=1}^{n} Y_k = O(n)$ holds obviously. (30) and (31) follow from (11) and (12) directly.

Remark 2. *When the second-order nonhomogeneous Markov chain indexed by a tree degenerates into a common nonhomogeneous Markov chain indexed by the tree, we obtain $P_t(X_t|X_{1_t}, X_{2_t}) = P_t(X_t|X_{1_t})$. Equation (31) is a result of Yang and Ye (see[8]).*

When the successor of each vertex on the tree has only one vertex, the second-order nonhomogeneous Markov chain on the tree degenerates into the general second-order nonhomogeneous Markov chain.

Corollary 2. *Suppose that $\{X_n, n \geq 0\}$ is a second-order nonhomogeneous Markov chain with the initial distribution and the transition probabilities as follows:*

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

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

We denote

$$f_n(\omega) = -\frac{1}{n+2}[\log P(X_0, X_{-1}) + \sum_{t=1}^{n} \log P_t(X_t|X_{t-1}, X_{t-2})], \tag{32}$$

$$H_t(\omega) = -\sum_{x_t\in S} P_t(x_t|X_{t-1}, X_{t-2}) \log P_t(x_t|X_{t-1}, X_{t-2}). \tag{33}$$

Then

$$\sum_{t=1}^{\infty} \frac{\log P_t(X_t|X_{t-1}, X_{t-2}) + H_t(\omega)}{t} < \infty, \qquad a.s. \tag{34}$$

$$\lim_{n\to\infty} [f_n(\omega) - \frac{1}{n+2} \sum_{t=1}^{n} H_t(\omega)] = 0. \qquad a.s. \tag{35}$$

Proof. At present the second-order nonhomogeneous Markov chain $X = \{X_t, t \in T\}$ indexed by a tree is changed into the general second-order nonhomogeneous Markov chain $\{X_n, n \geq 0\}$, we reach $P_t(X_t|X_{1_t}, X_{2_t}) = P_t(X_t|X_{t-1}, X_{t-2})$, $|T^{(n)}| = n + 2$. (32)-(35) follow from (7), (10), (30) and (31), respectively.

Remark 3. *When the second-order nonhomogeneous Markov chain degenerates into an ordinary nonhomogeneous Markov chain, we obtain* $P_t(X_t|X_{t-1}, X_{t-2}) = P_t(X_t|X_{t-1})$. *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. (1998) Science Press, Beijing, New York.

[4] T. Berger, Z. Ye, Entropic aspects of random fields on trees. IEEE Trans.Inform.Theory. 36(1990),1006-1018.

[5] Z. Ye, 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, 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, Z. Ye, The asymptotic equipartition property for nonhomogeneous Markov chains indexed by a homogeneous tree. IEEE Trans.Inform.Theory. 53(2007), 3275-3280.

[9] I. Benjamini, and Y. Peres, Markov chains indexed by trees. Ann. Probab., 22(1994), 219-243.

[10] J.G. Kemeny, ,J.L. Snell and A.W. Knapp, Denumerabl Markov chains. (1976), Springer, New York.

[11] F. Spitzer, Markov random fields on an infinite tree. Ann.Probab.3(1975),387-398

[12] W. Liu, Strong Deviation Theorems and Analytic Method. (2003), Sciece Publication, Beijing.

[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. (1986) Wiley, New York.

[15] R.V. Mises, Mathematical Theory of Probability and Statistics. (1964) Academic Press. New York.

[16] A.N. Kolmogorov, On the logical foundation of probability theory. Lecture Notes in Mathematics. Springer-Verlag. New York,1982.1021:1-2.

[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(4)(1999), 539-546.

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

Received 12 July 2014
Accepted 15 June 2014