Azerbaijan Journal of Mathematics
V. 4, No 1, 2014, January
ISSN 2218-6816

Some Shannon-McMillan Theorems for Nonhomogeneous Markov chains Indexed by a Tree on Generalized Gambling Systems

Wang Kangkang*, Zong Decai

Abstract. In this paper, a class of generalized Shannon-McMillan theorems for the nonhomogeneous Markov chains field on an infinite tree with respect to the generalized random selection system is discussed by constructing a nonnegative martingale. As corollaries, some Shannon-Mcmillan theorems for the homogeneous Markov chains field on an infinite tree and the nonhomogeneous Markov chain are obtained. Some results which have been obtained are extended.
Key Words and Phrases: Shannon-McMillan theorem, the infinite tree, Markov chains field, generalized random selection system, relative entropy density.
2000 Mathematics Subject Classifications: 60F15

1. Introduction.

A tree is a graph $S = \{T, E\}$ 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 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. For a better explanation of the infinite root 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 σ, t be vertices of the infinite 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 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. We denote by

*Corresponding author.

$T^{(n)}$ the subtree of the tree T containing the vertices from level 0 (the root o) to level n. Let $t(\neq o)$be a vertex of the tree T. We denote the first predecessor of t by 1_t, the second predecessor of t by 2_t, and the n-th predecessor of t by n_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 .

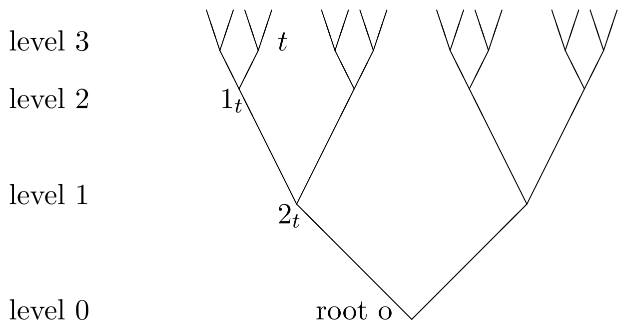

Fig.1 An infinite tree $T_{C,2}$

Definition 1 *Let* $S = \{s_0, s_1, s_2, \cdots\}$ *and* $P(y|x)$ *be a nonnegative function on* S^2. *Let*

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

If

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

then P *is called a transition matrix.*

Definition 2 *Let* T *be an infinite tree,* $S = \{s_0, s_1, s_2, \cdots\}$ *be a countable state space, and* $\{X_t, t \in T\}$ *be a collection of* S*-valued random variables defined on the probability space* $(\Omega, \mathcal{F}, P)$. *Let*

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

be a distribution on S*, and*

$$P_n = (P_n(y|x)), \quad x, y \in S, \tag{2}$$

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

$$\begin{aligned} &P(X_t = y | X_{1_t} = x, \textit{and } X_\sigma \textit{ for } \sigma \wedge t \leq 1_t) \\ &= P(X_t = y | X_{1_t} = x) = P_n(y|x) \quad t \in L_n, \quad \forall x, y \in S \end{aligned} \tag{3}$$

and

$$P(X_o = x) = P(x), \quad x \in S, \tag{4}$$

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

Definition 3. *Let $P_n = P_n(j|i)$ and $P = (P(s_0), P(s_1), P(s_2), \cdots)$ be defined as before, μ_P be a nonhomogeneous Markov measure on $(\Omega, \mathcal{F})$. If*

$$\mu_P(x_0) = P(x_0) \tag{5}$$

$$\mu_P(x^{T^{(n)}}) = P(x_0) \prod_{k=1}^{n} \prod_{t \in L_k} P_k(x_t|x_{1_t}) \quad n \geq 1, \tag{6}$$

then μ_P will be called a Markov chains field on an infinite tree T determined by the stochastic matrices P_n and the distribution P.

Let μ be an arbitrary probability measure, *log* is the natural logarithmic. Let

$$f_n(\omega) = -\frac{1}{|T^{(n)}|} \log \mu(X^{T^{(n)}}). \tag{7}$$

$f_n(\omega)$ is called the entropy density on subgraph $T^{(n)}$ with respect to μ. If $\mu = \mu_P$, then by (6),(7) we have

$$f_n(\omega) = -\frac{1}{|T^{(n)}|}[\log P(X_0) + \sum_{k=1}^{n} \sum_{t \in L_k} \log P_k(X_t|X_{1_t})]. \tag{8}$$

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 information theory. There have been some works on limit theorems for tree-indexed stochastic processes. Benjamini and Peres [1] have given the notion of the tree-indexed Markov chains and studied the recurrence and ray-recurrence for them. Berger and Ye [2] have studied the existence of entropy rate for some stationary random fields on a homogeneous tree. Ye and Berger (see [4],[5]), by using Pemantle's result [3] and a combinatorial approach, have studied the Shannon-McMillan theorem with convergence in probability for a PPS-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 field on a homogeneous tree (a particular case of tree-indexed Markov chains field and PPS-invariant random fields). Yang (see [6]) has studied the strong law of large numbers for frequency of occurrence of state and Shannon-McMillan theorem for homogeneous Markov chains indexed by a homogeneous tree. Recently, Yang (see [13]) has studied the strong law of large numbers and Shannon-McMillan theorem for nonhomogeneous Markov chains indexed by a homogeneous tree. Huang and Yang (see [11]) have also studied the strong law of large numbers for Markov chains indexed by an infinite tree with uniformly bounded degree. Peng and Yang have studied a class of small deviation theorems for functionals for arbitrary random field on a homogeneous trees (see[14]). Wang has also studied some Shannon-McMillan approximation theorems for arbitrary random field on the generalized Bethe tree (see[9]). Afterward, some scholars have investigated all kinds of applications of Shannon-McMillan theorems in the economic management and optimization controls (see[15-18]).

Definition 4. *Let $\{f_n(x_1, \cdots, x_n), n \geq 1\}$ be a sequence of real-valued functions defined on $S^n (n = 1, 2, \cdots)$, which will be called the generalized selection functions if $\{f_n, n \geq 1\}$ take values in an arbitrary interval of $[a, b]$ $(a, b \in R)$. We let*

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

$$Y_t = f_{|t|}(X_{1_t}, X_{2_t}, \cdots, X_0), \quad |t| \geq 1, \tag{9}$$

where $|t|$ stands for the number of the edges on the path from the root o to t. Then $\{Y_t,\ t \in T^{(n)}\}$ is called the generalized gambling system or the generalized random selection system indexed by an infinite tree with uniformly bounded degree. The traditional random selection system $\{Y_n,\ n \geq 0\}$ [10] 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 $f_n(x_1, \cdots, x_n)$ defined on $S^n (n = 1, 2, \cdots)$, which will be called the random selection functions if they take values in a two-valued set $\{0, 1\}$. Then let

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

$$Y_{n+1} = f_n(X_1, \cdots, X_n), \quad n \geq 0.$$

$\{Y_n, n \geq 1\}$ is called the gambling system (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 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 $\{Y_n, n \geq 1\}$ 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} as $n \geq 1$, and b_0 is a constant. Thus $\sum_{k=1}^{n} Y_k g_k(X_{k-1}, X_k)$ 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[10]), we introduce the following definition:

Definition 5. *The game is said to be fair, if for almost all $\omega \in \{\omega : \sum_{k=1}^{\infty} Y_k = \infty\}$, the accumulated net gain in the first n trial is to be of smaller order of magnitude than* *the accumulated stake* $\sum_{k=1}^{n} Y_k$ *as* n *tends to infinity, that is*

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

Definition 6. *Let* $\{Y_t, t \in T^{(n)}\}$ *be a generalized random selection system indexed by an infinite tree defined as (9). We call*

$$S_n(\omega) = -\frac{1}{\sum\limits_{k=1}^{n} \sum\limits_{t\in L_k} |Y_t|} [Y_0 \log P(X_0) + \sum_{k=1}^{n} \sum_{t\in L_k} Y_t \log P_k(X_t|X_{1_t})] \tag{10}$$

the generalized relative entropy density of nonhomogeneous Markov chain field $\{X_t, t \in T^{(n)}\}$ *on the generalized random selection system. Obviously, the generalized relative entropy density* $S_n(\omega)$ *is just the general relative entropy density* $f_n(\omega)$ *if* $Y_t \equiv 1,\ t \in T^{(n)}$.

In this paper, we study a class of generalized Shannon-McMillan theorems for nonhomogeneous Markov chains field on the generalized random selection system which takes values in a countable alphabet set on the infinite tree by constructing the consistent distribution functions and a nonnegative martingale. As corollaries, some Shannon-McMillan theorems for nonhomogeneous, homogeneous Markov chains field on an infinite tree and the general nonhomogeneous Markov chain are obtained. Liu and Yang's main results (see [7], [13]) which relate to the tree-indexed nonhomogeneous Markov chain field and the general nonhomogeneous Markov chain are extended.

2. Main result and its proof.

Theorem 1. *Let* $X = \{X_t, t \in T\}$ *be a nonhomogeneous Markov chains field on a homogeneous tree,* $\{Y_t, t \in T\}$, $S_n(\omega)$ *be defined as (9), (10). Denote by* $H(P_k(s_1|X_{1_t}), P_k(s_2|X_{1_t}), \cdots)$ *the random conditional entropy of* X_t *relative to* X_{1_t} *on the measure* μ_P*, that is*

$$H(P_k(s_1|X_{1_t}), P_k(s_2|X_{1_t}), \cdots) = -\sum_{x_t\in S} P_k(x_t|X_{1_t}) \log P_k(x_t|X_{1_t}) \quad t \in L_k, \quad k \geq 1.$$

Denote $\alpha > 0$, $G = max\{|a|, |b|\}$,

$$D(\omega) = \{\omega : \lim_{n} \sum_{k=1}^{n} \sum_{t\in L_k} |Y_t| = \infty\}. \tag{11}$$

We set

$$B_\alpha = \limsup_{n\to\infty} \frac{1}{\sum_{k=1}^{n} \sum_{t\in L_k} |Y_t|} \sum_{k=1}^{n} \sum_{t\in L_k} E[|Y_t| P_k(X_t|X_{1_t})^{-\alpha G} |X_{1_t}] < \infty. \tag{12}$$

Then

$$\lim_{n\to\infty}[S_n(\omega)-\frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}Y_tH(P_k(s_1|X_{1_t}),P_k(s_2|X_{1_t}),\cdots)]=0.$$

$$\mu_P-a.s.\qquad \omega\in D(\omega) \tag{13}$$

Proof. On the probability space $(\Omega,\mathcal{F},\mu_{\mathcal{P}})$, let $\lambda>0$ be a constant. Denote

$$Q_k(\lambda)=E[P_k(X_t|X_{1_t})^{-\lambda Y_t}|X_{1_t}=x_{1_t}]=\sum_{x_t\in S}P_k(x_t|x_{1_t})^{1-\lambda Y_t}, \tag{14}$$

$$q_k(\lambda;x_{1_t},x_t)=\frac{P_k(x_t|x_{1_t})^{1-\lambda Y_t}}{Q_k(\lambda)},\quad x_{1_t},x_t\in S. \tag{15}$$

$$g(\lambda;x^{T^{(n)}})=P(x_0)\prod_{k=1}^{n}\prod_{t\in L_k}q_k(\lambda;x_{1_t},x_t). \tag{16}$$

By (14-16) we can write that

$$\begin{aligned}
&\sum_{x^{L_n}\in S}g(\lambda,x^{T^{(n)}})\\
=&\sum_{x^{L_n}\in S}P(x_0)\prod_{k=1}^{n}\prod_{t\in L_k}\frac{P_k(x_t|x_{1_t})^{1-\lambda Y_t}}{Q_k(\lambda)}\\
=&\ g(\lambda,x^{T^{(n-1)}})\sum_{x^{L_n}\in S}\prod_{t\in L_n}\frac{P_n(x_t|x_{1_t})^{1-\lambda Y_t}}{E[P_n(X_t|X_{1_t})^{-\lambda Y_t}|X_{1_t}=x_{1_t}]}\\
=&\ g(\lambda,x^{T^{(n-1)}})\prod_{t\in L_n}\sum_{x_t\in S}\frac{P_n(x_t|x_{1_t})^{1-\lambda Y_t}}{E[P_n(X_t|X_{1_t})^{-\lambda Y_t}|X_{1_t}=x_{1_t}]}\\
=&\ g(\lambda,x^{T^{(n-1)}})\prod_{t\in L_n}\frac{E[P_n(X_t|X_{1_t})^{-\lambda Y_t}|X_{1_t}=x_{1_t}]}{E[P_n(X_t|X_{1_t})^{-\lambda Y_t}|X_{1_t}=x_{1_t}]}=g(\lambda,x^{T^{(n-1)}}).
\end{aligned}$$

Hence $g(\lambda;x^{T^{(n)}})$, $n=1,2,\cdots$ are a set of consistent distribution functions. Set

$$U_n(\lambda,\omega)=\frac{g(\lambda;X^{T^{(n)}})}{\mu_P(X^{T^{(n)}})}. \tag{17}$$

Since g and μ_P are two probability measures, $\{U_n(\lambda,\omega),\mathcal{F}_n,n\geq 1\}$ ($\mathcal{F}_n=\sigma(X^{T^{(n)}})$) is a nonnegative martingale which converges almost surely(see[12]). Thus, by Doob's martingale convergence theorem we get

$$\lim_{n\to\infty}U_n(\lambda,\omega)=U_\infty(\lambda,\omega)<\infty.\qquad \mu_P-a.s. \tag{18}$$

By (11) and (18), we have

$$\limsup_{n\to\infty}\frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\log U_n(\lambda,\omega)\le 0. \qquad \mu_P-a.s. \qquad \omega\in D(\omega) \tag{19}$$

By (6), (14)-(17), we can rewrite (19) as

$$\limsup_{n\to\infty}\frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}[-\lambda Y_t\log P_k(X_t|X_{1_t})-\log E(P_k(X_t|X_{1_t})^{-\lambda Y_t}|X_{1_t})]\le 0.$$

$$\mu_P-a.s. \qquad \omega\in D(\omega) \tag{20}$$

By the inequality $e^x-1-x\le(1/2)x^2e^{|x|}$, we have

$$x^{-\lambda}-1-(-\lambda)\log x\le(1/2)\lambda^2(\log x)^2x^{-|\lambda|}, \quad 0\le x\le 1. \tag{21}$$

Taking into account (12), (20), (21) and the inequality $\log x\le x-1,\ (x>0)$, noticing that $Y_t\in[a,b],\ |Y_t|\le max\{|a|,|b|\}=G,\ t\in L_k,\ k\ge 1$,

$$\max\{(\log x)^2x^h,0\le x\le 1,h>0\}=\frac{4e^{-2}}{h^2},$$

in the case of $0<|\lambda|<t<\alpha$, we can write

$$\begin{aligned}
&\limsup_{n\to\infty}\frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}[-\lambda Y_t\log P_k(X_t|X_{1_t})-E(-\lambda Y_t\log P_k(X_t|X_{1_t})|X_{1_t})]\\
\le\ &\limsup_{n\to\infty}\frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}[\log E(P_k(X_t|X_{1_t})^{-\lambda Y_t}|X_{1_t})-E(-\lambda Y_t\log P_k(X_t|X_{1_t})|X_{1_t})]\\
\le\ &\limsup_{n\to\infty}\frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}[E(P_k(X_t|X_{1_t})^{-\lambda Y_t}|X_{1_t})-1-E(-\lambda Y_t\log P_k(X_t|X_{1_t})|X_{1_t})]\\
\le\ &\limsup_{n\to\infty}\frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}E[\frac{1}{2}\lambda^2Y_t^2(\log P_k(X_t|X_{1_t}))^2P_k(X_t|X_{1_t})^{-|\lambda Y_t|}|X_{1_t}]\\
\le\ &\limsup_{n\to\infty}\frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}E[\frac{\lambda^2G}{2}|Y_t|(\log P_k(X_t|X_{1_t}))^2P_k(X_t|X_{1_t})^{-|\lambda|G}|X_{1_t}]\\
=\ &\frac{\lambda^2G}{2}\limsup_{n\to\infty}\frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}E[|Y_t|\log^2P_k(X_t|X_{1_t})
\end{aligned}$$

$$
\begin{aligned}
& \cdot P_k(X_t|X_{1_t})^{(\alpha-|\lambda|)G} P_k(X_t|X_{1_t})^{-\alpha G}|X_{1_t}] \\
\leq \quad & \frac{\lambda^2 G}{2} \limsup_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|} \sum_{k=1}^{n}\sum_{t\in L_k} E[\frac{4e^{-2}}{(\alpha-|\lambda|)^2 G^2}\cdot |Y_t| P_k(X_t|X_{1_t})^{-\alpha G}|X_{1_t}] \\
\leq \quad & \frac{2\lambda^2 e^{-2}}{(\alpha-t)^2 G} \limsup_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|} \sum_{k=1}^{n}\sum_{t\in L_k} E[|Y_t| P_k(X_k|X_{1_t})^{-\alpha G}|X_{1_t}] \\
= \quad & \frac{2\lambda^2 e^{-2}}{(\alpha-t)^2 G} B_\alpha < \infty. \qquad \mu_P - a.s. \qquad \omega\in D(\omega) \qquad (22)
\end{aligned}
$$

In the case of $0<\lambda<t<\alpha$, dividing both sides of (22) by λ, we obtain

$$
\limsup_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|} \sum_{k=1}^{n}\sum_{t\in L_k} Y_t[-\log P_k(X_t|X_{1_t}) - E(-\log P_k(X_t|X_{1_t})|X_{1_t})] \leq \frac{2\lambda e^{-2}B_\alpha}{(\alpha-t)^2 G}
$$

$$
\mu_P - a.s. \qquad \omega\in D(\omega) \qquad (23)
$$

Choose $0<\lambda_i<\alpha$, $(i=1,2,\cdots)$ such that $\lambda_i\to 0^+$ $(i\to\infty)$. Then for all i we have by (23) that

$$
\limsup_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|} \sum_{k=1}^{n}\sum_{t\in L_k} Y_t[-\log P_k(X_t|X_{1_t}) - E(-\log P_k(X_t|X_{1_t})|X_{1_t})] \leq 0.
$$

$$
\mu_P - a.s. \qquad \omega\in D(\omega) \qquad (24)
$$

When $-\alpha<-t<\lambda<0$, dividing two sides of (22) by λ, we attain

$$
\liminf_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|} \sum_{k=1}^{n}\sum_{t\in L_k} Y_t[-\log P_k(X_t|X_{1_t}) - E(-\log P_k(X_t|X_{1_t})|X_{1_t})] \geq \frac{2\lambda e^{-2}B_\alpha}{(\alpha-t)^2 G}.
$$

$$
\mu_P - a.s. \qquad \omega\in D(\omega) \qquad (25)
$$

Choose $-\alpha<-t<\lambda_i<0$, $(i=1,2,\cdots)$ such that $\lambda_i\to 0^-$ $(i\to\infty)$. Then for all i we have by (25) that

$$
\liminf_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|} \sum_{k=1}^{n}\sum_{t\in L_k} Y_t[-\log P_k(X_t|X_{1_t}) - E(-\log P_k(X_t|X_{1_t})|X_{1_t})] \geq 0.
$$

$$
\mu_P - a.s. \qquad \omega\in D(\omega) \qquad (26)
$$

It follows from (24) and (26) that

$$\lim_{n\to\infty}\frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}Y_t[-\log P_k(X_t|X_{1_t})-E(-\log P_k(X_t|X_{1_t})|X_{1_t})]=0.$$

$$\mu_P-a.s. \qquad \omega\in D(\omega). \tag{27}$$

Noticing that

$$\begin{aligned}&H(P_k(s_1|X_{1_t}),P_k(s_2|X_{1_t}),\cdots)\\ =\;&-\sum_{x_t\in S}P_k(x_t|X_{1_t})\log P_k(x_t|X_{1_t})=E(-\log P_k(X_t|X_{1_t})|X_{1_t}),\end{aligned}$$

it follows from (10) and (27) that

$$\begin{aligned}&\lim_{n\to\infty}[S_n(\omega)-\frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}Y_tH(P_k(s_1|X_{1_t}),P_k(s_2|X_{1_t}),\cdots)]\\ =\;&\lim_{n\to\infty}\frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}Y_t[-\log P_k(X_t|X_{1_t})-E(-\log P_k(X_t|X_{1_t})|X_{1_t})]=0.\end{aligned} \tag{28}$$

We complete the proof of the theorem.◀

Corollary 1. *Let $X=\{X_t,t\in T\}$ be a nonhomogeneous Markov chains field on an infinite tree, $f_n(\omega)$ be defined as (8). Denote $\alpha>0$. We set*

$$b_\alpha=\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{k=1}^{n}\sum_{t\in L_k}E[P_k(X_t|X_{1_t})^{-\alpha}|X_{1_t}]<\infty. \tag{29}$$

Then

$$\lim_{n\to\infty}[f_n(\omega)-\frac{1}{|T^{(n)}|}\sum_{k=1}^{n}\sum_{t\in L_k}H(P_k(s_1|X_{1_t}),P_k(s_2|X_{1_t}),\cdots)]=0. \qquad \mu_P-a.s. \tag{30}$$

Proof. Letting $a=0,\ b=1,\ Y_t\equiv 1,\ t\in T^{(n)},\ n\geq 0$, we have $\lim\limits_n\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|=\lim\limits_n|T^{(n)}|=+\infty,\ G=max\{0,1\}=1$. Hence $S_n(\omega)=f_n(\omega)$, $D(\omega)=\Omega$. (29), (30) follow from (12) and (13) immediately. ◀

$\{X_t,t\in T\}$ will be called $S-$valued homogeneous Markov chains field indexed by an infinite tree if for all $n\geq 0$,

$$P_n=P=(P(y|x)), \qquad \forall x,y\in S. \tag{31}$$

Corollary 2. *Let $\{X_t, t \in T\}$ be a homogeneous Markov chains field indexed by an infinite tree, $f_n(\omega)$ and $H(P_k(s_1|X_{1_t}), P_k(s_2|X_{1_t}), \cdots)$ be defined as above. Denote $0 < \alpha < 1/G$, if*

$$\sum_{i\in S}\sum_{j\in S} P(j|i)^{1-\alpha G} < \infty. \tag{32}$$

Then

$$\lim_{n\to\infty}[S_n(\omega) - \frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k} Y_t H(P_k(s_1|X_{1_t}), P_k(s_2|X_{1_t}), \cdots)] = 0. \qquad \mu_P\text{–}a.s. \tag{33}$$

Proof. By (31) and (32), we can write

$$\begin{aligned}
B_\alpha &= \limsup_{n\to\infty}\frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k} E[|Y_t|P_k(X_t|X_{1_t})^{-\alpha G}|X_{1_t}] \\
&= \limsup_{n\to\infty}\frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}\sum_{x_t\in S}|Y_t|P(x_t|X_{1_t})^{1-\alpha G} \\
&= \limsup_{n\to\infty}\frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}\sum_{i\in S}\sum_{j\in S}|Y_t|\delta_i(X_{1_t})P(j|i)^{1-\alpha G} \\
&\le \limsup_{n\to\infty}\frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}\sum_{i\in S}\sum_{j\in S}|Y_t|P(j|i)^{1-\alpha G} \\
&\le \sum_{i\in S}\sum_{j\in S}P(j|i)^{1-\alpha G}\limsup_{n\to\infty}\frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t| \\
&= \sum_{i\in S}\sum_{j\in S}P(j|i)^{1-\alpha G} < \infty.
\end{aligned} \tag{34}$$

It follows that (12) holds. Therefore, (33) follows from (13).◀

3. Some Shannon-McMillan theorems on a finite states space.

Corollary 3. *Let $X = \{X_t, t \in T\}$ be a nonhomogeneous Markov chains field on an infinite tree which takes values in the finite alphabet set $S = \{s_1, s_2, \cdots, s_N\}$, $f_n(\omega)$ be defined as (8). Denote by $H(P_k(s_1|X_{1_t}), \cdots, P_k(s_N|X_{1_t}))$ the random conditional entropy* *of X_t relative to X_{1_t} on the measure μ_P, that is*

$$H(P_k(s_1|X_{1_t}),\cdots,P_k(s_N|X_{1_t}))=-\sum_{x_t=s_1}^{s_N}P_k(x_t|X_{1_t})\log P_k(x_t|X_{1_t}),\quad t\in L_k,\quad k\geq 1.$$

Then

$$\lim_{n\to\infty}[S_n(\omega)-\frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}Y_tH(P_k(s_1|X_{1_t}),\cdots,P_k(s_N|X_{1_t}))]=0.$$

$$\mu_P-a.s.\quad \omega\in D(\omega) \tag{35}$$

Proof. Let $0<\alpha<1/G$. By (12) we can conclude

$$\begin{aligned}
B_\alpha &= \limsup_{n\to\infty}\frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}E[|Y_t|P_k(X_t|X_{1_t})^{-\alpha G}|X_0^{k-1}]\\
&= \limsup_{n\to\infty}\frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}\sum_{x_t=s_1}^{s_N}|Y_t|P_k(x_t|X_{1_t})^{1-\alpha G}\\
&\leq \limsup_{n\to\infty}\frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}\sum_{x_t=s_1}^{s_N}|Y_t|\\
&\leq \limsup_{n\to\infty}\frac{N}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|=N<\infty. \qquad \mu_P-a.s.
\end{aligned}\tag{36}$$

Hence (12) holds naturally. (35) follows from (13).◀

Corollary 4[13]. *Let $X=\{X_t,t\in T\}$ be a nonhomogeneous Markov chains field on an infinite tree which takes values in the finite alphabet set $S=\{s_1,s_2,\cdots,s_N\}$, $f_n(\omega)$ and $H(P_k(s_1|X_{1_t}),\cdots,\ P_k(s_N|X_{1_t}))$ be defined as above. Then*

$$\lim_{n\to\infty}[f_n(\omega)-\frac{1}{|T^{(n)}|}\sum_{k=1}^{n}\sum_{t\in L_k}H(P_k(s_1|X_{1_t}),\cdots,P_k(s_N|X_{1_t}))]=0.\qquad \mu_P-a.s. \tag{37}$$

Proof. Letting $Y_t\equiv 1,\ t\in T^{(n)},\ n\geq 1$, we obtain $\lim\limits_{n}\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|=\lim\limits_{n}|T^{(n)}|=+\infty$. Hence $S_n(\omega)=f_n(\omega),\ D(\omega)=\Omega$. (37) follows from (35) immediately.◀

Corollary 5[7]. *Let $\{X_n, n \geq 0\}$ be a nonhomogeneous Markov chain with the initial distribution and the transition probabilities as follows:*

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

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

Set

$$f_n(\omega) = -\frac{1}{n+1}[\log P(X_0) + \sum_{k=1}^{n} \log P_k(X_k|X_{k-1})],$$

$$H_k(X_k|X_{k-1}) = -\sum_{x_k=1}^{N} P_k(x_k|X_{k-1}) \log P_k(x_k|X_{k-1}).$$

Then

$$\lim_{n\to\infty}[f_n(\omega) - \frac{1}{n+1}\sum_{k=1}^{n} H_k(X_k|X_{k-1})] = 0. \qquad a.s. \qquad (38)$$

Proof. When the successor of each vertex of the tree T has only one vertex, the nonhomogeneous Markov chains field on the tree degenerates into the general nonhomogeneous Markov chain. Hence we easily get $|T^{(n)}| = n + 1$, $P_k(x_t|x_{1_t}) = P_k(x_k|x_{k-1})$. (38) follows from (37) naturally.◀

4. Derivation results.

Theorem 2. *Let $X = \{X_t, t \in T\}$ be a nonhomogeneous Markov chains field on an infinite tree which takes values in the countable alphabet set $S = \{s_1, s_2, \cdots\}$, $S_n(\omega)$ be defined as (10). Denote $\alpha \geq 0$, $0 < C < 1$. Set*

$$C_\alpha = \limsup_{n\to\infty} \frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k} E[|Y_t|P_k(X_t|X_{1_t})^{-(2+\alpha G)} I_{\{P_k(X_t|X_{1_t})\leq C\}}|X_{1_t}] < \infty.$$

$$\mu_P - a.s. \qquad (39)$$

Then

$$\lim_{n\to\infty}[S_n(\omega) - \frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k} Y_t H(P_k(s_1|X_{1_t}), P_k(s_2|X_{1_t}), \cdots)] = 0. \qquad \mu_P - a.s.$$

$$(40)$$

Proof. Let us denote $P_k(X_t|X_{1_t}) = P_k$ in brief. Taking into account (39) and the inequality $1-\frac{1}{x} \le \log x \le 0, (0<x<1)$, from the fourth inequality of (22) in the proof of Theorem 1, in the case of $0<|\lambda|<\alpha$, we can write

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}[-\lambda Y_t\log P_k(X_t|X_{1_t}) - E(-\lambda Y_t \log P_k(X_t|X_{1_t})|X_{1_t})]\\
\le\ &\limsup_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}E[\frac{\lambda^2G}{2}|Y_t|\log^2 P_k(X_t|X_{1_t})P_k(X_t|X_{1_t})^{-\alpha G}|X_{1_t}]\\
=\ &\frac{\lambda^2G}{2}\limsup_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}E[|Y_t|(\log P_k)^2P_k^{-\alpha G}(I_{\{P_k(X_t|X_{1_t})\le C\}}+I_{\{P_k(X_t|X_{1_t})>C\}})|X_{1_t}]\\
\le\ &\frac{\lambda^2G}{2}\Bigg\{\limsup_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}E[|Y_t|(\log P_k)^2P_k^{-\alpha G}I_{\{P_k(X_t|X_{1_t})\le C\}}|X_{1_t}]\\
&\quad+\limsup_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}E[|Y_t|(\log P_k)^2P_k^{-\alpha G}I_{\{P_k(X_t|X_{1_t})> C\}}|X_{1_t}]\Bigg\}\\
\le\ &\frac{\lambda^2G}{2}\Bigg\{\limsup_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}E[|Y_t|(\log P_k)^2P_k^{-\alpha G}I_{\{P_k(X_t|X_{1_t})\le C\}}|X_{1_t}]\\
&\quad+\limsup_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}E[|Y_t|C^{-\alpha G}\cdot(\log C)^2|X_{1_t}]\Bigg\}\\
\le\ &\frac{\lambda^2G}{2}\Bigg\{\limsup_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}E[|Y_t|(\log P_k)^2P_k^{-\alpha G}I_{\{P_k(X_t|X_{1_t})\le C\}}|X_{1_t}]\\
&\quad+C^{-\alpha G}\cdot(\log C)^2\Bigg\}\\
\le\ &\frac{\lambda^2G}{2}\Bigg\{\limsup_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}E[|Y_t|(1-\frac{1}{P_k})^2P_k^{-\alpha G}I_{\{P_k(X_t|X_{1_t})\le C\}}|X_{1_t}]\\
&\quad+C^{-\alpha G}\cdot(\log C)^2\Bigg\}\\
=\ &\frac{\lambda^2G}{2}\Bigg\{\limsup_{n\to\infty} \frac{1}{\sum\limits_{k=1}^{n}\sum\limits_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}E[|Y_t|(1-P_k)^2P_k^{-(2+\alpha G)}I_{\{P_k(X_t|X_{1_t})\le C\}}|X_{1_t}]\\
&\quad+C^{-\alpha G}\cdot(\log C)^2\Bigg\}
\end{aligned}$$

$$\leq \frac{\lambda^2 G}{2}\left\{\limsup_{n\to\infty}\frac{1}{\sum_{k=1}^{n}\sum_{t\in L_k}|Y_t|}\sum_{k=1}^{n}\sum_{t\in L_k}E[|Y_t|P_k^{-(2+\alpha G)}I_{\{P_k(X_t|X_{1_t})\leq C\}}|X_{1_t}] + C^{-\alpha G}\cdot(\log C)^2\right\}$$

$$= \frac{\lambda^2 G}{2}\left\{C_\alpha + C^{-\alpha G}\cdot(\log C)^2\right\} < \infty.$$

Imitating the proof of (23)-(28), Theorem 2 follows from Theorem 1.◄

5. Acknowledgements.

Authors would like to thank the referee for his valuable suggestions. The work is supported by the Project of Higher Schools' Natural Science Basic Research of Jiangsu Province of China (13KJB110006). Wang Kangkang is the corresponding author.

References

[1] Benjammini, I. and Peres, Y. 1994. Markov chains indexed bu trees. Ann. Probab. 22(2): 219-243.

[2] Berger, T. and Ye, Z. 1990. Entropic aspects of random fields on trees. IEEE Trans. Inform. Theory. 36(3): 1006-1018.

[3] Pemantle, R. 1992 Antomorphism invariant measure on trees. Ann. Probab. 20(3):1549-1566.

[4] Ye, Z.X. and Berger, T. 1996. Ergodic regularity and asymptotic equipartition property of random fields on trees. J.Combin.Inform.System.Sci, 21(1):157-184.

[5] Ye, Z.X. and Berger, T. 1998. Information Measures for Discrete Random Fields. Science Press, New York.

[6] Yang, W.G. 2003. Some limit properties for Markov chains indexed by homogeneous tree . Stat. Probab. Letts. 65(2): 241-250.

[7] Liu, W. and Yang, W.G. 1996 An extension of Shannon-McMillan theorem and some limit properties for nonhomogeneous Markov chains. Stochastic Process. Appl, 61(1), 129-145.

[8] Yang, W.G. and Liu, W. 2002. Strong law of large numbers and Shannon-McMillan theorem for Markov chains fields on trees. IEEE Trans.Inform.Theory. 48(1): 313-318

[9] Wang, K.K. and Zong, D.C. 2011, Some Shannon-McMillan approximation theoems for Markov chain field on the generalized Bethe tree. Journal of Inequalities and Applications. Article ID 470910, 18 pages doi:10.1155/2011/470910

[10] Kolmogorov, A.N. 1982 On the logical foundation of probability theory. Lecture Notes in Mathematics. Springer-Verlag. New York, 1021:1-2

[11] Huang, H.L. and Yang, W.G. 2008. Strong law of large numbers for Markov chains indexed by an infinite tree with uniformly bounded degree, Science in China, 51(2): 195-202,

[12] Doob, J.L. 1953. Stochastic Process. Wiley, New York.

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

[14] Peng, W.C., Yang, W.G. and Wang, B. 2010, A class of small deviation theorems for functionals of random fields on a homogeneous tree. Journal of Mathematical Analysis and Applications. 361, 293-301.

[15] S.H. Wang, Geometric entropy of group actions on rational and nonrational curves. J. Jiangsu Univ. Sci-tech. Nat. Sci. 26(4): 401-405, 2012.

[16] K.K. Wang, D.C. Zong, Li Fang, A class of Shannon-McMillan theorems for nonhomogeneous Markov information source on random selection systems. J. Jiangsu Univ. Sci-tech. Nat. Sci. 26(4): 406-410, 2012.

[17] K.K. Wang, and D.C. Zong, A class of strong deviation theorems on generalized gambling system for the sequence of arbitrary continuous random variables. J. Jiangsu Univ. Sci-tech. Nat. Sci. 25 (2): 195-199, 2011.

[18] K.K. Wang, H. Ye and Y. Ma, A class of strong deviation theorems for multivariate function sequence of mth-order countable nonhomogeneous Markov chains. J. Jiangsu Univ. Sci-tech. Nat. Sci. 25 (1): 93-96, 2011

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

Zong Decai
Department of Computer Science and Engineering, Changshu Institute of Technology, Changshu 215500, China

Received 10 May 2012
Accepted 29 October 2013