*Vietnam Journal of Mathematics* **39**:1(2011) 1-17

Vietnam Journal of MATHEMATICS



# A Type of Markov Approximation of Random Fields on a Homogeneous Tree and a Class of Small Deviation Theorems

**Kangkang Wang and Mingxing Zhu**

*School of Mathematics and Physics, Jiangsu University of Science and Technology, Zhenjiang 212003, China.*

Received August 13, 2009
Revised February 11, 2011

**Abstract.** In this paper, a class of small deviation theorems for an arbitrary bivariate function are introduced by introducing the sample relative entropy rate as a measure of deviation between the arbitrary random field and the Markov chains field on the homogeneous tree. As corollaries, a class of small deviation theorems for the frequencies of states ordered couples and a Shannon-McMillan approximation theorem for arbitrary random fields on the homogeneous tree are obtained.

2000 Mathematics Subject Classification: 60F15.

*Key words:* Shannon-McMillan theorem, the homogeneous tree, arbitrary random field, Markov random field, sample relative entropy density.

## 1. Introduction

Let $T$ be a homogeneous tree on which each vertex has $N+1$ neighboring vertices. We first fix any vertex as the "root" and label it by 0. Let $\sigma$, $\tau$ be vertices of a tree. Write $\tau \leq \sigma$ if $\tau$ is on the unique path connecting 0 to $\sigma$, $|\sigma|$ for the number of edges on this path. For any two vertices $\sigma$, $\tau$, denote by $\sigma \wedge \tau$ the vertex farthest from 0 satisfying

$$\sigma \wedge \tau \leq \sigma, \quad \text{and} \quad \sigma \wedge \tau \leq \tau.$$

If $\sigma \neq 0$, then we let $\bar{\sigma}$ stand for the vertex satisfying $\bar{\sigma} \leq \sigma$ and $|\bar{\sigma}| = |\sigma| - 1$ (we refer to $\sigma$ as a son of $\bar{\sigma}$). It is easy to see that the root has $N+1$ sons and all other vertices have $N$ sons.

**Definition 1.1.** ([7]) Let $T$ be a homogeneous tree, $S = \{s_0, s_1, s_2 \cdots, s_M\}$ be a finite state space, $\{X_\sigma, \sigma \in T\}$ be a collection of $S-$valued random variables defined on the measurable space $\{\Omega, \mathcal{F}\}$. Let

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

be a distribution on $S$, and

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

be a strictly positive stochastic matrix on $S^2$. If for any vertices $\sigma$, $\tau$,

$$\begin{aligned} &Q(X_\sigma = y | X_{\bar{\sigma}} = x, \text{ and } X_\tau \text{ for } \sigma \wedge \tau \leq \bar{\sigma}) \\ &= Q(X_\sigma = y | X_{\bar{\sigma}} = x) \\ &= Q(y|x) \quad \forall x, y \in S \end{aligned} \tag{3}$$

and

$$Q(X_0 = x) = q(x), \quad \forall x \in S, \tag{4}$$

then $\{X_\sigma, \sigma \in T\}$ will be called $S-$valued Markov chains indexed by a homogeneous tree with the initial distribution (1) and transition matrix (2).

Two special finite tree-indexed Markov chains are introduced in Kemeny et al. ([28]), Spitzer ([6]), and there the finite transition matrix is assumed to be positive and reversible to its stationary distribution, and these 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.

If $|\sigma| = n$, it is said to be on the $n$th level on a tree $T$. We denote by $T^{(n)}$ the subtree of $T$ containing the vertices from level 0 (the root) to level $n$, and by $L_n$ the set of all vertices on the level $n$. Let $B$ be a subgraph of $T$. Denote $X^B = \{X_\sigma, \sigma \in B\}$, and denote by $|B|$ the number of vertices of $B$. Let $S(\sigma)$ be the set of all sons of vertices $\sigma$. It is easy to see that $|S(0)| = N+1$ and $|S(\sigma)| = N$, where $\sigma \neq 0$.

Let $\Omega = S^T$, $\omega = \omega(\cdot) \in \Omega$, where $\omega(\cdot)$ is a function defined on $T$ and taking values in $S$, and $\mathcal{F}$ be the smallest Borel field containing all cylinder sets in $\Omega$, $\mu$ be the probability measure on $(\Omega, \mathcal{F})$. Let $X = \{X_\sigma, \sigma \in T\}$ be the coordinate stochastic process defined on the measurable space $(\Omega, \mathcal{F})$; that is, for any $\omega = \{\omega(t), t \in T\}$, define

$$\begin{aligned} &X_t(\omega) = \omega(t), \quad t \in T^{(n)}, \\ &X^{T^{(n)}} \triangleq \{X_t, t \in T^{(n)}\}, \\ &\mu(X^{T^{(n)}} = x^{T^{(n)}}) = \mu(x^{T^{(n)}}), \quad i = 1, 2. \end{aligned} \tag{5}$$

Now we give a definition of Markov chain fields on the tree $T$ by using the cylinder distribution directly, which is a natural extension of the classical definition of Markov chains (see [4]).

**Definition 1.2.** Let $Q = Q(j|i)$ and $q = (q(s_1), q(s_2) \cdots, q(s_M))$ be defined as before, $\mu_Q$ be another probability measure on $(\Omega, \mathcal{F})$. If

$$\mu_Q(x_0) = q(x_0), \tag{6}$$

$$\mu_Q(x^{T^{(n)}}) = q(x_0) \prod_{k=0}^{n-1} \prod_{\sigma \in L_k} \prod_{\tau \in S(\sigma)} Q(x_\tau | x_\sigma), \quad n \geq 1, \tag{7}$$

then $\mu_Q$ will be called a Markov chain field on the homogeneous tree $T$ determined by the stochastic matrix $Q$ and the distribution $q$.

Let $\mu$ be an arbitrary probability measure defined as (5), *log* is the natural logarithmic. Let

$$f_n(\omega) = -\frac{1}{|T^{(n)}|} \log \mu(X^{T^{(n)}}). \tag{8}$$

$f_n(\omega)$ is called the entropy density on subgraph $T^{(n)}$ with respect to $\mu$. If $\mu = \mu_Q$, then by (7), (8) we have

$$f_n(\omega) = -\frac{1}{|T^{(n)}|}[\log q(X_0) + \sum_{k=0}^{n-1} \sum_{\sigma \in L_k} \sum_{\tau \in S(\sigma)} \log Q(X_\tau | X_\sigma)]. \tag{9}$$

The convergence of $f_n(\omega)$ to a constant in a sense ($L_1$ convergence, convergence in probability, or almost sure convergence) is an important subject in information theory. Many scholars have investigated the Shannon-McMillan theorem in the stationary ergodic process $\{X_n, n \geq 0\}$. Shannon (see [13]) first proved the asymptotic equipartition property (AEP for short) for convergence in probability for stationary ergodic information source with a finite alphabet set. McMillan (see [14]) and Breiman(see [15]) proved the AEP in $L_1$ and a.s. convergence, respectively, for stationary ergodic information source. If $p(x_0, \cdots, x_n)$ and $p(x_n | x_0, \cdots, x_{n-1})$ denote the joint and conditional probability mass functions of a stationary ergodic $\{X_n, n \geq 0\}$ taking values in a set $S = \{s_1, \cdots, s_N\}$, then the Shannon-McMillan theorem asserts that

$$-\frac{1}{n} \log p(X_0, \cdots, X_n) = -\frac{1}{n}[\log p(X_0) + \sum_{t=1}^{n} \log p(X_t | X_0, \cdots, X_{t-1})] \to H, \ a.s.$$

where $H = \lim_{k \to \infty} E\{-\log p(X_k | X_0, \cdots, X_{k-1})\}$ is the entropy rate of $\{X_n, n \geq 0\}$. This individual ergodic theorem of information theory was proved first by Breiman (see [15]) for a finite state set, and later by Chung for a countable state set (see [16]). Convergence in probability already implies the existence of a set of roughly (to first order in the exponent) $\exp(nH)$ typical sequences of length $n$ all having roughly equal probability $\exp(nH)$.

The strong AEP was proved by Barren ([17]) and Orey ([18]), after Moy ([19], [20]), Perez ([21], [22]) and Kieffer ([23]) suggested its validity and proved $L^1$ convergence. Barren and Orey invoke Breiman's ([15]) extended ergodic theorem when they observe that $g_t(\omega) = \log p(X_0|X_{-1}, \cdots, X_{-t})$ almost surely converges to $g(\omega) = \log p(X_0|X_{-1}, X_{-t}, \cdots)$. Algoet and Cover ([24]) studied Shannon-McMillan theorem by arguing that $p(X_0, \cdots, X_n)$ is sandwiched in asymptotic growth rate between the $k$th Markov approximation and the infinite order approximation with no gap as $k \to \infty$. Liu and Yang (see [25]) have provided a class of small deviation theorems for an arbitrary sequence of random variables relative to nonhomogeneous Markov information sources. Afterward, Yang and Liu (see [26]) have moreover investigated the asymptotic equipartition property for $m$th-order nonhomogeneous Markov information sources. Wang and Li (see [27]) have studied the Shannon-McMillan approximation theorems for Markov information sources indexed by a homogeneous tree.

There have been some works on limit theorems for tree-indexed stochastic processes. Benjamini and Peres have given the notion of the tree-indexed homogeneous Markov chains and studied the recurrence and ray-recurrence for them (see [1]). Berger and Ye have studied the existence of entropy rate for some stationary random fields on a homogeneous tree (see [2]). Pemantle proved a mixing property and a weak law of large numbers for a PPG-invariant and ergodic random field on a homogeneous tree (see [5]). Ye and Berger, by using Pemantle's result and a combinatorial approach, have studied the asymptotic equipartition property in the sense of convergence in probability for a PPG-invariant and ergodic random field on a homogeneous tree(see [9], [10]). Peng and Yang have studied a class of small deviation theorems for functionals of random field and asymptotic equipartition property for arbitrary random field on a homogeneous trees (see [8]). Recently, Yang have studied some limit theorems for countable homogeneous Markov chains indexed by a homogeneous tree and strong law of large numbers and the asymptotic equipartition property for finite homogeneous Markov chains indexed by a homogeneous tree (see [7], [11]). But their results only concern the case of strong limit theorems for Markov chains field, they do not discuss the case of strong deviation theorems (the strong limit theorems which are represented by the inequalities) for arbitrary random fields.

In this paper, our aim is to establish a class of small deviation theorems (also called strong deviation theorems) for the arbitrary bivariate function by introducing the sample relative entropy rate as a measure of deviation between the arbitrary random fields and the Markov chains field on the homogeneous tree. We apply a new type of techniques distinct from that of Peng and Yang (see [8]) to the study of the small deviation theorems for arbitrary random field on a homogeneous tree and obtain the one which is more concise and accurate than Peng and Yang's result. As corollaries, a class of small deviation theorems for the frequencies of states ordered couples and a Shannon-McMillan approximation theorem for arbitrary random fields on the homogeneous tree are obtained.

## 2. Main Results and Their Proofs

**Lemma 2.1.** ([3]) *Let $\mu_1$ and $\mu_2$ be two probability measures on $(\Omega, \mathcal{F})$, $D \in \mathcal{F}$, $\{\tau_n, n \geq 0\}$ be a sequence of positive-valued random variables such that*

$$\liminf_n \frac{\tau_n}{|T^{(n)}|} > 0 \qquad \mu_1 - a.s.\ D, \tag{10}$$

*then*

$$\limsup_{n\to\infty} \frac{1}{\tau_n} \log \frac{\mu_2(X^{T^{(n)}})}{\mu_1(X^{T^{(n)}})} \leq 0 \qquad \mu_1 - a.s.\ D. \tag{11}$$

*Particularly, let $\tau_n = |T^{(n)}|$, then*

$$\limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \log \frac{\mu_2(X^{T^{(n)}})}{\mu_1(X^{T^{(n)}})} \leq 0 \qquad \mu_1 - a.s. \tag{12}$$

*Proof.* See reference [3]. ∎

Let

$$\varphi(\mu|\mu_Q) = \limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \log \frac{\mu(X^{T^{(n)}})}{\mu_Q(X^{T^{(n)}})}. \tag{13}$$

$\varphi(\mu|\mu_Q)$ is called the sample relative entropy rate with respect to $\mu$ and $\mu_Q$. $\varphi(\mu|\mu_Q)$ is also called asymptotic logarithmic likelihood ratio. By (12) and (13),

$$\varphi(\mu|\mu_Q) \geq \liminf_{n\to\infty} \frac{1}{|T^{(n)}|} \log \frac{\mu(X^{T^{(n)}})}{\mu_Q(X^{T^{(n)}})} \geq 0 \qquad \mu - a.s. \tag{14}$$

Hence $\varphi(\mu|\mu_Q)$ can be looked on as a type of a measure of the deviation between the arbitrary random field and the Markov chains field on the homogeneous tree.

Although $\varphi(\mu|\mu_Q)$ is not a proper metric between two probability measures, we nevertheless think of it as a measure of "dissimilarity" between their joint distribution $\mu$ and Markov distribution $\mu_Q$. Obviously, $\varphi(\mu|\mu_Q) = 0$ if and only if $\mu = \mu_Q$. It has been shown in (14) that $\varphi(\mu|\mu_Q) \geq 0$ a.s. in any case. Hence, $\varphi(\mu|\mu_Q)$ can be used as a random measure of the deviation between the true joint distribution $\mu(x^{T^{(n)}})$ and the Markov distribution $\mu_Q(x^{T^{(n)}})$. Roughly speaking, this deviation may be regarded as the one between coordinate stochastic process $x^{T^{(n)}}$ and the Markov case. The smaller $\varphi(\mu|\mu_Q)$ is, the smaller the deviation is.

**Theorem 2.2.** *Let $X = \{X_\sigma, \sigma \in T\}$ be an arbitrary random field on the homogeneous tree $T$. $\varphi(\mu|\mu_Q)$ is defined as (13). Let $f(x,y)$ be an arbitrary real bivariate function defined on $S^2$, $0 \leq c \leq \alpha^2 e^\alpha/2$, $\alpha > 0$. Denote*

$$D(c) = \{\omega : \varphi(\mu|\mu_Q) \leq c\}, \tag{15}$$

*then*

$$\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\{f(X_\sigma,X_\tau)-E_Q[f(X_\sigma,X_\tau)|X_\sigma]\}$$
$$\le\sqrt{2ce^{\alpha}}\sum_{i\in S}\sum_{j\in S}|f(i,j)| \qquad \mu-a.s.\ \omega\in D(c), \tag{16}$$

$$\liminf_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\{f(X_\sigma,X_\tau)-E_Q[f(X_\sigma,X_\tau)|X_\sigma]\}$$
$$\ge-\sqrt{2ce^{\alpha}}\sum_{i\in S}\sum_{j\in S}|f(i,j)| \qquad \mu-a.s.\omega\ \in D(c), \tag{17}$$

*where $E_Q(\cdot)$ represents the expectation with respect to the Markov measure $\mu_Q$.*

**Remark 2.3.** In Theorem 2.2, $\varphi(\mu|\mu_Q)$ is regarded as a measure of "dissimilarity" between their joint distribution $\mu$ and Markov distribution $\mu_Q$. When the difference between the joint distribution $\mu$ and Markov distribution $\mu_Q$ is controlled in a certain range, the difference between the functions $f(X_\sigma,X_\tau)$ and the conditional expectation of $f(X_\sigma,X_\tau)$ under the Markov measure $\mu_Q$ can also be controlled in a certain range determined by the bound of $\varphi(\mu|\mu_Q)$. The smaller the bound $c$ of $\varphi(\mu|\mu_Q)$ is, the smaller the deviation of $f(X_\sigma,X_\tau)$ relative to $E_Q[f(X_\sigma,X_\tau)|X_\sigma]$ is.

*Proof of Theorem 2.2.* Consider the probability space $(\Omega,\mathcal{F},\mu)$, let $\lambda$ be an arbitrary real number, $\delta_i(j)$ be the Kronecker function. We construct the following product distribution:

$$\mu_Q(x^{T^{(n)}};\lambda)=\exp\{\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\lambda\delta_i(x_\sigma)\delta_j(x_\tau)\}\times$$
$$\times\prod_{k=0}^{n-1}\prod_{\sigma\in L_k}\prod_{\tau\in S(\sigma)}[\frac{1}{1+(e^\lambda-1)Q(j|i)}]^{\delta_i(x_\sigma)}\times$$
$$\times q(x_0)\prod_{k=0}^{n-1}\prod_{\sigma\in L_k}\prod_{\tau\in S(\sigma)}Q(x_\tau|x_\sigma). \tag{18}$$

By (18) we have

$$\sum_{x^{L_n}\in S}\mu_Q(x^{T^{(n)}};\lambda)$$
$$=\sum_{x^{L_n}\in S}q(x_0)\prod_{k=0}^{n-1}\prod_{\sigma\in L_k}\prod_{\tau\in S(\sigma)}\exp\{\lambda\delta_i(x_\sigma)\delta_j(x_\tau)\}\times$$
$$\times[\frac{1}{1+(e^\lambda-1)Q(j|i)}]^{\delta_i(x_\sigma)}\cdot Q(x_\tau|x_\sigma)$$

$$= \mu_Q(x^{T^{(n-1)}};\lambda) \sum_{x^{L_n}\in S} \prod_{\sigma\in L_{n-1}} \prod_{\tau\in S(\sigma)} \exp\{\lambda\delta_i(x_\sigma)\delta_j(x_\tau)\}\times$$
$$\times [\frac{1}{1+(e^\lambda-1)Q(j|i)}]^{\delta_i(x_\sigma)} \cdot Q(x_\tau|x_\sigma)$$
$$= \mu_Q(x^{T^{(n-1)}};\lambda) \prod_{\sigma\in L_{n-1}} \prod_{\tau\in S(\sigma)} \frac{1}{[1+(e^\lambda-1)Q(j|i)]^{\delta_i(x_\sigma)}}\times$$
$$\times \sum_{x_\tau\in S} \exp\{\lambda\delta_i(x_\sigma)\delta_j(x_\tau)\}.Q(x_\tau|x_\sigma)$$
$$= \mu_Q(x^{T^{(n-1)}};\lambda) \prod_{\sigma\in L_{n-1}} \prod_{\tau\in S(\sigma)} \frac{1}{[1+(e^\lambda-1)Q(j|i)]^{\delta_i(x_\sigma)}}[\sum_{x_\tau\neq j} + \sum_{x_\tau=j}]$$
$$= \mu_Q(x^{T^{(n-1)}};\lambda) \prod_{\sigma\in L_{n-1}} \prod_{\tau\in S(\sigma)} \frac{e^{\lambda\delta_i(x_\sigma)}Q(j|x_\sigma)+1-Q(j|x_\sigma)}{[1+(e^\lambda-1)Q(j|i)]^{\delta_i(x_\sigma)}}. \tag{19}$$

When $x_\sigma = i$, we obtain from (19) that

$$\sum_{x^{L_n}\in S} \mu_Q(x^{T^{(n)}};\lambda)$$
$$= \mu_Q(x^{T^{(n-1)}};\lambda) \prod_{\sigma\in L_{n-1}} \prod_{\tau\in S(\sigma)} \frac{1+(e^\lambda-1)Q(j|i)}{1+(e^\lambda-1)Q(j|i)} = \mu_Q(x^{T^{(n-1)}};\lambda). \tag{20}$$

When $x_\sigma \neq i$, we acquire from (19) that

$$\sum_{x^{L_n}\in S} \mu_Q(x^{T^{(n)}};\lambda) = \mu_Q(x^{T^{(n-1)}};\lambda) \prod_{\sigma\in L_{n-1}} \prod_{\tau\in S(\sigma)} (Q(j|x_\sigma)+1-Q(j|x_\sigma))$$
$$= \mu_Q(x^{T^{(n-1)}};\lambda). \tag{21}$$

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

$$U_n(\lambda,\omega) = \frac{\mu_Q(X^{T^{(n)}};\lambda)}{\mu(X^{T^{(n)}})}. \tag{22}$$

By (18) and (22), we have

$$U_n(\lambda,\omega) = \exp\{\sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \lambda\delta_i(X_\sigma)\delta_j(X_\tau)\}$$
$$\cdot \prod_{k=0}^{n-1} \prod_{\sigma\in L_k} \prod_{\tau\in S(\sigma)} [\frac{1}{1+(e^\lambda-1)Q(j|i)}]^{\delta_i(X_\sigma)}$$
$$\cdot\, q(X_0) \prod_{k=0}^{n-1} \prod_{\sigma\in L_k} \prod_{\tau\in S(\sigma)} Q(X_\tau|X_\sigma) \Big/ \mu(X^{T^{(n)}}). \tag{23}$$

It is easy to see that $U_n(\lambda,\omega)$ is a nonnegative sup-martingale from Doob's martingale convergence theorem (see [12]). Moreover,

$$\lim_{n\to\infty} U_n(\lambda,\omega) = U_\infty(\lambda,\omega) < \infty \qquad \mu - a.s. \tag{24}$$

By (12) and (22) we have

$$\limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \log U_n(\lambda,\omega) \le 0 \qquad \mu - a.s. \tag{25}$$

By (7), (23) and (25) we have

$$\begin{aligned} &\limsup_{n\to\infty}\{ \frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)} \lambda\delta_i(X_\sigma)\delta_j(X_\tau) - \\ &\frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)} \delta_i(X_\sigma)\log[1+(e^\lambda-1)Q(j|i)] + \frac{1}{|T^{(n)}|}\log\frac{\mu_Q(X^{T^{(n)}})}{\mu(X^{T^{(n)}})}\} \\ &\le 0 \qquad \mu - a.s. \end{aligned} \tag{26}$$

Letting $\lambda = 0$ in (26), we have

$$\varphi(\mu|\mu_Q) \ge \liminf_{n\to\infty} \frac{1}{|T^{(n)}|}\log\frac{\mu(X^{T^{(n)}})}{\mu_Q(X^{T^{(n)}})} \ge 0 \qquad \mu - a.s. \tag{27}$$

By (15) and (26) we have

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)} \{\lambda\delta_i(X_\sigma)\delta_j(X_\tau) - \delta_i(X_\sigma)\log[1+(e^\lambda-1)Q(j|i)]\} \\ &\le c \qquad \mu - a.s. \quad \omega\in D(c). \end{aligned} \tag{28}$$

By (28), the inequalities $1-\frac{1}{x} \le \ln x \le x-1$, $(x>0)$, $e^x - 1 - x \le \frac{1}{2}x^2 e^{|x|}$ and the properties of superior limit

$$\limsup_{n\to\infty}(a_n - b_n) \le d \Rightarrow \limsup_{n\to\infty}(a_n - c_n) \le \limsup_{n\to\infty}(b_n - c_n) + d,$$

we have

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)} \lambda\{\delta_i(X_\sigma)\delta_j(X_\tau) - \delta_i(X_\sigma)Q(j|i)\} \\ \le\ &\limsup_{n\to\infty} \frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)} \{\delta_i(X_\sigma)\log[1+(e^\lambda-1)Q(j|i)] - \lambda\delta_i(X_\sigma)Q(j|i)\} \\ &+ c \end{aligned}$$

$$
\begin{aligned}
&\leq \limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \{\delta_i(X_\sigma)(e^{\lambda}-1)Q(j|i) - \lambda\delta_i(X_\sigma)Q(j|i)\} + c \\
&= \limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \delta_i(X_\sigma)(e^{\lambda}-1-\lambda)Q(j|i) + c \\
&\leq \limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \frac{1}{2}\lambda^2 e^{|\lambda|}\delta_i(X_\sigma)Q(j|i) + c \quad \mu - a.s.\ \omega \in D(c).
\end{aligned} \tag{29}
$$

Let $0 < \lambda \leq \alpha$, dividing the two sides of (29) by $\lambda$, we have

$$
\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \delta_i(X_\sigma)\{\delta_j(X_\tau) - Q(j|i)\} \\
\leq\ &\limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \frac{1}{2}\lambda e^{\alpha}\delta_i(X_\sigma)Q(j|i) + \frac{c}{\lambda} \\
\leq\ &\frac{1}{2}\lambda e^{\alpha} \limsup_{n\to\infty} \frac{|T^{(n)}|-1}{|T^{(n)}|} + \frac{c}{\lambda} = \frac{1}{2}\lambda e^{\alpha} + \frac{c}{\lambda} \qquad \mu - a.s. \quad \omega \in D(c).
\end{aligned} \tag{30}
$$

It is easy to show that in the case $0 < c \leq \frac{\alpha^2 e^{\alpha}}{2}$, the function $g(\lambda) = \frac{\lambda e^{\alpha}}{2} + \frac{c}{\lambda}$ attains its smallest value $g\left(\sqrt{\frac{2c}{e^{\alpha}}}\right) = \sqrt{2ce^{\alpha}}$ at $\lambda = \sqrt{\frac{2c}{e^{\alpha}}}$. Hence letting $\lambda = \sqrt{\frac{2c}{e^{\alpha}}}$ in (30), we have

$$
\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \delta_i(X_\sigma)\{\delta_j(X_\tau) - Q(j|i)\} \\
&\leq \sqrt{2e^{\alpha}c} \qquad \mu - a.s. \quad \omega \in D(c).
\end{aligned} \tag{31}
$$

In the case $c = 0$, (31) also follows from (30) by choosing $\lambda_i \to 0^+ (i \to \infty)$.

In the case $-\alpha \leq \lambda < 0$, dividing two sides of (29) by $\lambda$, we have

$$
\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \lambda\{\delta_i(X_\sigma)\delta_j(X_\tau) - \delta_i(X_\sigma)Q(j|i)\} \\
\geq\ &\frac{1}{2}\lambda \limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} e^{-\lambda}\delta_i(X_\sigma)Q(j|i) + \frac{c}{\lambda} \\
\geq\ &\frac{1}{2}\lambda e^{\alpha} + \frac{c}{\lambda} \qquad \mu - a.s. \quad \omega \in D(c).
\end{aligned} \tag{32}
$$

It is easy to show that in the case $0 < c \le \dfrac{\alpha^2 e^\alpha}{2}$, the function $h(\lambda) = \dfrac{\lambda e^\alpha}{2} + \dfrac{c}{\lambda}$ attains its largest value $h\left(-\sqrt{\frac{2c}{e^\alpha}}\right) = -\sqrt{2ce^\alpha}$ at $\lambda = -\sqrt{\frac{2c}{e^\alpha}}$. Hence letting $\lambda = -\sqrt{\frac{2c}{e^\alpha}}$ in (32), we obtain

$$\begin{aligned} &\liminf_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \delta_i(X_\sigma)\{\delta_j(X_\tau) - Q(j|i)\} \\ &\ge -\sqrt{2e^\alpha c} \qquad \mu - a.s. \quad \omega \in D(c). \end{aligned} \tag{33}$$

In the case $c = 0$, (33) also follows from (32) by choosing $\lambda_i \to 0^-(i \to \infty)$.
It follows from (31) and (33) that for any $f(i,j)$,

$$\limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \delta_i(X_\sigma) f(i,j)\{\delta_j(X_\tau) - Q(j|i)\} \le \sqrt{2e^\alpha c}|f(i,j)| \\ \mu - a.s. \quad \omega \in D(c), \tag{34}$$

$$\liminf_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \delta_i(X_\sigma) f(i,j)\{\delta_j(X_\tau) - Q(j|i)\} \ge -\sqrt{2e^\alpha c}|f(i,j)| \\ \mu - a.s. \quad \omega \in D(c). \tag{35}$$

Notice that

$$\begin{aligned} &f(X_\sigma, X_\tau) - E_Q[f(X_\sigma, X_\tau)|X_\sigma] \\ =&\sum_{i\in S}\sum_{j\in S} \delta_i(X_\sigma)\delta_j(X_\tau) f(i,j) - \sum_{j\in S} f(X_\sigma, j) Q(j|X_\sigma) \\ =&\sum_{i\in S}\sum_{j\in S} \delta_i(X_\sigma) f(i,j)\{\delta_j(X_\tau) - Q(j|i)\} \quad \mu - a.s.\ \omega \in D(c). \end{aligned} \tag{36}$$

By virtue of the properties of superior limit and inferior limit, (16), (17) follow from (34), (35) and (36). The proof is finished. ■

**Corollary 2.4.** *Let $X = \{X_\sigma, \sigma \in T\}$ be the Markov chains field determined by the measure $\mu_Q$ on the homogeneous tree with the initial distribution (6) and joint distribution (7). $f(x,y)$ is defined as above. Then*

$$\lim_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \{f(X_\sigma, X_\tau) - E_Q[f(X_\sigma, X_\tau)|X_\sigma]\} = 0 \quad \mu_Q - a.s. \tag{37}$$

*Proof.* In this case, $\mu \equiv \mu_Q$. It is obvious that $\varphi(\mu|\mu_Q) \equiv 0$. Hence letting $c = 0$ in Theorem 2.2, we obtain $D(0) = \Omega$, (37) follows from (16) and (17) immediately. ■

**Corollary 2.5.** *Let $X=\{X_\sigma,\sigma\in T\}$ be an arbitrary random field on the homogeneous tree $T$. $\varphi(\mu|\mu_Q)$ is defined as (13). Denote $0\le c\le\frac{\alpha^2e^\alpha}{2}$, $\alpha>0$. Let $i,j\in S$, $S_n(i,j)$ be the number of couples $(i,j)$ in the couples of random variables*

$$(X_\sigma,X_\tau),\quad 0\le k\le n-1,\ \sigma\in L_k,\ \tau\in S(\sigma),\ n\ge 1.$$

*That is*

$$S_n(i,j)=\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\delta_i(X_\sigma)\delta_j(X_\tau).$$

*Then*

$$\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\{S_n(i,j)-\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}N\delta_i(X_\sigma)Q(j|i)\}\le\sqrt{2e^\alpha c}$$
$$\mu_Q-a.s.\quad \omega\in D(c), \tag{38}$$
$$\liminf_{n\to\infty}\frac{1}{|T^{(n)}|}\{S_n(i,j)-\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}N\delta_i(X_\sigma)Q(j|i)\}\ge-\sqrt{2e^\alpha c}$$
$$\mu_Q-a.s.\quad \omega\in D(c). \tag{39}$$

*Proof.* Letting $f(x,y)=\delta_i(x)\delta_j(y)$ in Theorem 2.2, we have

$$\begin{aligned}
&\frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\{f(X_\sigma,X_\tau)-E_Q[f(X_\sigma,X_\tau)|X_\sigma]\}\\
=&\frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\{\delta_i(X_\sigma)\delta_j(X_\tau)-E_Q[\delta_i(X_\sigma)\delta_j(X_\tau)|X_\sigma]\}\\
=&\frac{1}{|T^{(n)}|}\{S_n(i,j)-\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\sum_{x_\tau\in S}\delta_i(X_\sigma)\delta_j(x_\tau)Q(x_\tau|X_\sigma)\}\\
=&\frac{1}{|T^{(n)}|}\{S_n(i,j)-\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\delta_i(X_\sigma)Q(j|X_\sigma)\}\\
=&\frac{1}{|T^{(n)}|}\{S_n(i,j)-\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}N\delta_i(X_\sigma)Q(j|i)\}
\end{aligned}\tag{40}$$

and

$$\sum_{u\in S}\sum_{v\in S}|f(u,v)|=\sum_{u\in S}\sum_{v\in S}\delta_i(u)\delta_j(v)=\sum_{u\in S}\delta_i(u)=1. \tag{41}$$

Therefore, (38) and (39) follow from (16), (17), (40) and (41). ■

**Corollary 2.6.** *Let $X=\{X_\sigma,\sigma\in T\}$ be an arbitrary random field on the homogeneous tree $T$. $\varphi(\mu|\mu_Q)$ is defined as (13). Denote $0\le c\le\frac{\alpha^2e^\alpha}{2}$, $\alpha>0$. Let*

*$S_n(j)$ be the number of $j$ in the set of random variables $X^{T^{(n)}} = \{X_\sigma, \sigma \in T^{(n)}\}$. That is*

$$S_n(j) = \sum_{k=0}^{n} \sum_{\sigma \in L_k} \delta_j(X_\sigma) = \delta_j(X_0) + \sum_{k=0}^{n-1} \sum_{\sigma \in L_k} \sum_{\tau \in S(\sigma)} \delta_j(X_\tau).$$

*Then*

$$\limsup_{n\to\infty} \frac{1}{|T^{(n)}|}\{S_n(j) - \sum_{i \in S} N S_{n-1}(i) Q(j|i)\} \le \sqrt{2e^\alpha c}(M+1)$$

$$\mu_Q - a.s. \quad \omega \in D(c), \tag{42}$$

$$\liminf_{n\to\infty} \frac{1}{|T^{(n)}|}\{S_n(j) - \sum_{i \in S} N S_{n-1}(i) Q(j|i)\} \ge -\sqrt{2e^\alpha c}(M+1)$$

$$\mu_Q - a.s. \ \omega \in D(c). \tag{43}$$

*Proof.* Letting $f(x,y) = \delta_j(y)$ in Theorem 2.2, we have by the definition of $S_n(j)$ that

$$\begin{aligned}
&\frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma \in L_k} \sum_{\tau \in S(\sigma)} \{f(X_\sigma, X_\tau) - E_Q[f(X_\sigma, X_\tau)|X_\sigma]\} \\
&= \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma \in L_k} \sum_{\tau \in S(\sigma)} \{\delta_j(X_\tau) - E_Q[\delta_j(X_\tau)|X_\sigma]\} \\
&= \frac{1}{|T^{(n)}|}\{S_n(j) - \delta_j(X_0) - \sum_{k=0}^{n-1} \sum_{\sigma \in L_k} \sum_{\tau \in S(\sigma)} \sum_{x_\tau \in S} \delta_j(x_\tau) Q(x_\tau|X_\sigma)\} \\
&= \frac{1}{|T^{(n)}|}\{S_n(j) - \delta_j(X_0) - \sum_{k=0}^{n-1} \sum_{\sigma \in L_k} \sum_{\tau \in S(\sigma)} Q(j|X_\sigma)\} \\
&= \frac{1}{|T^{(n)}|}\{S_n(j) - \delta_j(X_0) - \sum_{k=0}^{n-1} \sum_{\sigma \in L_k} \sum_{i \in S} N\delta_i(X_\sigma) Q(j|i)\} \\
&= \frac{1}{|T^{(n)}|}\{S_n(j) - \delta_j(X_0) - \sum_{i \in S} N S_{n-1}(i) Q(j|i)\} \qquad (44)
\end{aligned}$$

and

$$\sum_{u \in S} \sum_{v \in S} |f(u,v)| = \sum_{u \in S} \sum_{v \in S} \delta_j(v) = M+1. \tag{45}$$

Therefore, (42) and (43) follow from (16), (17), (44) and (45). ■

## 3. Shannon-McMillan Approximation Theorem for Arbitrary Random Fields on a Homogeneous Tree

The convergence of $f_n(\omega)$ to a constant in a sense ($L_1$ convergence, convergence in probability, or almost sure convergence) is called the Shannon-McMillan theorem or the asymptotic equipartition property in information theory. As we have mentioned in section 1, Yang and Liu have studied the asymptotic equipartition property for Markov chains field on the homogeneous tree. In the following, we will discuss the asymptotic equipartition property for arbitrary random field by comparing the deviation between an arbitrary probability measure and the Markov measure on a homogeneous tree.

**Corollary 3.1.** *Let $X = \{X_\sigma, \sigma \in T\}$ be an arbitrary random field on the homogeneous tree. $f_n(\omega)$ and $D(c)$ are defined as (8) and (15), respectively. $0 \le c \le \frac{\alpha^2 e^\alpha}{2}$, $\alpha > 0$. Let $H_k^Q(X_\tau|X_\sigma)$ be the random conditional entropy with respect to $X_\tau$ and $X_\sigma$ on the measure $\mu_Q$, that is*

$$H_k^Q(X_\tau|X_\sigma) = -\sum_{x_\tau \in S} Q(x_\tau|X_\sigma)\log Q(x_\tau|X_\sigma) \quad \sigma \in L_k,\ \tau \in S(\sigma),\ k \ge 0. \tag{46}$$

*Then*

$$\limsup_{n\to\infty}[f_n(\omega) - \frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)} H_k^Q(X_\tau|X_\sigma)] \le \sqrt{2e^\alpha c}\sum_{i\in S}\sum_{j\in S}|\log Q(j|i)| \qquad \mu - a.s. \quad \omega \in D(c), \tag{47}$$

$$\liminf_{n\to\infty}[f_n(\omega) - \frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)} H_k^Q(X_\tau|X_\sigma)] \ge -\sqrt{2e^\alpha c}\sum_{i\in S}\sum_{j\in S}|\log Q(j|i)| - c \quad \mu - a.s. \quad \omega \in D(c). \tag{48}$$

**Remark 3.2.** The corollary also reflects the relation between the relative entropy density $f_n(\omega)$ and the random conditional entropy under Markov measure $\mu_Q$. The deviation of $f_n(\omega)$ under arbitrary probability measure $\mu$ relative to its random conditional entropy under Markov measure $\mu_Q$ is determined by deviation between arbitrary probability measure $\mu$ and Markov measure $\mu_Q$. The smaller the sample relative entropy rate $\varphi(\mu|\mu_Q)$ with respect to $\mu$ and $\mu_Q$ is, the smaller the deviation between $f_n(\omega)$ and its random conditional entropy relative to Markov measure $\mu_Q$ is. In the case of $\mu = \mu_Q$, the deviation between them will disappear. We can see it from Corollary 3.3. Our result is more succinct and the boundary is more accurate than the one of Wang and Li (see [27]).

*Proof of Corollary 3.1.* Letting $f(x,y) = -\log Q(y|x)$ in Theorem 2.2, we have by (46) that

$$\begin{aligned}
&\frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\{f(X_\sigma,X_\tau)-E_Q[f(X_\sigma,X_\tau)|X_\sigma]\}\\
&=\frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\{-\log Q(X_\tau|X_\sigma)-E_Q[-\log Q(X_\tau|X_\sigma)|X_\sigma]\}\\
&=\frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\{-\log Q(X_\tau|X_\sigma)+\sum_{x_\tau\in S}\log Q(x_\tau|X_\sigma)Q(x_\tau|X_\sigma)\}\\
&=\frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\{-\log Q(X_\tau|X_\sigma)-H_k^Q(X_\tau|X_\sigma)\} \qquad (49)
\end{aligned}$$

and

$$\sum_{i\in S}\sum_{j\in S}|f(i,j)| = \sum_{i\in S}\sum_{j\in S}|-\log Q(j|i)| = \sum_{i\in S}\sum_{j\in S}|\log Q(j|i)|.$$

It follows from (16) and (17) that

$$\begin{aligned}
&\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\{-\log Q(X_\tau|X_\sigma)-H_k^Q(X_\tau|X_\sigma)\}\\
&\qquad\le\sqrt{2e^\alpha c}\sum_{i\in S}\sum_{j\in S}|\log Q(j|i)| \qquad \mu-a.s. \quad \omega\in D(c), \qquad (50)
\end{aligned}$$

$$\begin{aligned}
&\liminf_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\{-\log Q(X_\tau|X_\sigma)-H_k^Q(X_\tau|X_\sigma)\}\\
&\qquad\ge-\sqrt{2e^\alpha c}\sum_{i\in S}\sum_{j\in S}|\log Q(j|i)|, \quad \mu-a.s. \quad \omega\in D(c). \qquad (51)
\end{aligned}$$

Therefore, (8), (7), (27) and (50) imply

$$\begin{aligned}
&\limsup_{n\to\infty}[f_n(\omega)-\frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}H_k^Q(X_\tau|X_\sigma)]\\
&\le\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\{-\log Q(X_\tau|X_\sigma)-H_k^Q(X_\tau|X_\sigma)\}\\
&\quad+\limsup_{n\to\infty}\frac{1}{|T^{(n)}|}[\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)}\log Q(X_\tau|X_\sigma)-\log\mu(X^{T^{(n)}})]
\end{aligned}$$

$$\le \limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \{-\log Q(X_\tau|X_\sigma) - H_k^Q(X_\tau|X_\sigma)\}$$

$$+ \limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \log \frac{\mu_Q(X^{T^{(n)}})}{\mu(X^{T^{(n)}})}$$

$$\le \limsup_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \{-\log Q(X_\tau|X_\sigma) - H_k^Q(X_\tau|X_\sigma)\}$$

$$\le \sqrt{2e^\alpha c} \sum_{i\in S} \sum_{j\in S} |\log Q(j|i)| \qquad \mu - a.s.\ \omega \in D(c). \tag{52}$$

In a similar way, it follows from (8), (13), (15) and (51) that

$$\liminf_{n\to\infty} [f_n(\omega) - \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} H_k^Q(X_\tau|X_\sigma)]$$

$$\ge \liminf_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \{-\log Q(X_\tau|X_\sigma) - H_k^Q(X_\tau|X_\sigma)\}$$

$$+ \liminf_{n\to\infty} \frac{1}{|T^{(n)}|} [\sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \log Q(X_\tau|X_\sigma) - \log \mu(X^{T^{(n)}})]$$

$$\ge \liminf_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \{-\log Q(X_\tau|X_\sigma) - H_k^Q(X_\tau|X_\sigma)\}$$

$$+ \liminf_{n\to\infty} \frac{1}{|T^{(n)}|} \log \frac{\mu_Q(X^{T^{(n)}})}{\mu(X^{T^{(n)}})}$$

$$= \liminf_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \{-\log Q(X_\tau|X_\sigma) - H_k^Q(X_\tau|X_\sigma)\} - \varphi(\mu|\mu_Q)$$

$$\ge \liminf_{n\to\infty} \frac{1}{|T^{(n)}|} \sum_{k=0}^{n-1} \sum_{\sigma\in L_k} \sum_{\tau\in S(\sigma)} \{-\log Q(X_\tau|X_\sigma) - H_k^Q(X_\tau|X_\sigma)\} - c$$

$$\ge -\sqrt{2e^\alpha c} \sum_{i\in S} \sum_{j\in S} |\log Q(j|i)| - c \qquad \mu - a.s.\ \omega \in D(c). \tag{53}$$

The proof is accomplished. ■

**Corollary 3.3.** ([11]) *Let* $X = \{X_\sigma, \sigma \in T\}$ *be a Markov chains field on the homogeneous tree determined by the measure* $\mu_Q$ *with the initial distribution* (6) *and joint distribution* (7). $f_n(\omega)$ *and* $H_k^Q(X_\tau|X_\sigma)$ *are defined as* (9) *and* (46). *Then*

$$\lim_{n\to\infty}[f_n(\omega) - \frac{1}{|T^{(n)}|}\sum_{k=0}^{n-1}\sum_{\sigma\in L_k}\sum_{\tau\in S(\sigma)} H_k^Q(X_\tau|X_\sigma)] = 0 \qquad \mu_Q - a.s. \tag{54}$$

*Proof.* In this case, $\mu \equiv \mu_Q$. It is obvious that $\varphi(\mu|\mu_Q) \equiv 0$. Hence letting $c = 0$ in Corollary 3.1, we obtain $D(0) = \Omega$, (54) follows from (47) and (48) immediately. ∎

## References


1. I. Benjammini and Y. Peres, Markov chains indexed bu trees, *Ann. Probab.* **22** (1994), 219-243.
2. T. Berger and Z. Ye, Entropic aspects of random fields on trees, *IEEE Trans.Inform.Theory.* **36** (1990), 1006-1018.
3. 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.
4. 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.
5. R. Pemantle, Automorphism invariant measure on trees, *Ann. Probab.* **20** (1992), 1549-1566.
6. F. Spitzer, Markov random fields on an infinite tree, *Ann.Probab.* **3** (1975), 387-398.
7. W.G. Yang, Some limit properties for Markov chains indexed by a homogeneous tree, *Statist. Probab. Letts.* **65** (2003), 241-250.
8. W.C. Peng, W.G. Yang and B. Wang, A class of small deviation theorems for functionals of random fields on a homogeneous tree, *J. Math. Anal. Appl.* **361**(2010), 293-301.
9. Z.Ye and T. Berger, Ergodic regularity and asymptotic equipartition property of random fields on trees, *J. Combin. Inform. System. Sci.* **21** (1996), 157-184.
10. Z. Ye and T. Berger, *Information Measures for Discrete Random Fields*, Science Press, New York, 1998.
11. 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.
12. J.L. Doob, *Stochastic Process*, Wiley New York, 1953.
13. C. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.* **27** (1948), 379-423.
14. B. McMillan, The Basic Theorem of information theory, *Ann.Math.Statist.* **24** (1953), 196-216.
15. L. Breiman, The individual ergodic theorem of information theory, *Ann.Math.Statist.* **28** (1957), 809-811.
16. K.L. Chung, *A Course in Probability Theory*, Academic Press, New York, 1974.
17. A.R. Barron, The strong ergodic theorem for densities:Generalized Shannon-McMillan-Breiman theorem, *Ann.Probab.* **13** (1985), 1292-1303.
18. A. Orey, On the Shannon-Perez-Moy theorem, *Contemp. Math.* **41** (1985), 319-327.
19. S.C. Moy, Asymptotic properties of derivatives of stationary measures, *Pacific J. Math.* **10** (1960), 1371-1383.
20. S.C. Moy, Generalizations of Shannon-McMillan theorem, *Pacific J. Math.* **11**(1961), 705-714.

21. A. Perez, Extensions of Shannon-McMillan's limit theorem to more general stochastic processes, *In Trans. Third Prague Conf. Inform. Theory, Statist. Decision Functions and Random Processes,* (1964), 545-547.
22. A. Perez, Generalization of Chernoff's result on the asymptotic discernibility of two random processes, *Colloq. Math. Soc. Janos Bolyai,* **9** (1974), 619-632.
23. J.C. Kieffer, A counterexample to Perez's generalization of Shannon-McMillan theorem, *Ann.Probab.* **1** (1973), 362-364.
24. P.H. Algoet and T.M. Cover, A sandwich proof of Shannon-McMillan theorem, *Ann. Probab.* **16** (1988), 899-909.
25. W. Liu and W.G. Yang, The Markov approximation of the sequence of N-valued variables and a class of small deviation theorems, *Stochastic. Process. Appl.* **89** (2000), 117-130.
26. W.G. Yang and W. Liu, The asymptotic equipartition property for $m$th-order nonhomogeneous Markov information source, *IEEE Trans. Inform. Theory.* **50** (2004), 3325-3330.
27. K.K. Wang and F. Li, A class of Shannon-McMillan random deviation theorems for arbitrary random field indexed by a homogeneous tree, *Chinese Journal of Engineering Mathematics* **26** (2009), 260-266.
28. J.G. Kemeny, J.L. Snell and A.W. Knapp, *Denumerabl Markov chains,* Springer - New York, 1976.