Turk J Math
35 (2011) , 729 – 736.

doi:10.3906/mat-0903-3

A class of generalized Shannon-McMillan theorems for arbitrary discrete information source*

Kangkang Wang

Abstract

In this study, a class of strong limit theorems for the relative entropy densities of random sum of arbitrary information source are discussed by constructing the joint distribution and nonnegative super martingales. As corollaries, some Shannon-McMillan theorems for arbitrary information source, mth-order Markov information source and non-memory information source are obtained and some results for the discrete information source which have been obtained by authors are extended.

Key Words: Shannon-McMillan theorem, the consistent distribution, arbitrary information source, relative entropy density, m-order Markov information source, non-memory information source.

1. Introduction

Suppose $\{X_n, n \geq 0\}$ is an arbitrary information source defined on any probability space $(\Omega, \mathcal{F}, P)$ taking values in the alphabet set $S = \{s_1, s_2, \cdots\}$. Also let us denote the joint distribution of $\{X_n, n \geq 0\}$ as

$$P(X_0 = x_0, \cdots, X_n = x_n) = p(x_0, \cdots, x_n) > 0, \quad x_i \in S, \ 0 \leq i \leq n. \tag{1}$$

Denote

$$f_n(\omega) = -\frac{1}{n+1} \log p(X_0, \cdots, X_n), \tag{2}$$

where log is the natural logarithmic, $f_n(\omega)$ is called the relative entropy density of $\{X_i, 0 \leq i \leq n\}$.

Denote the conditional probability as follows:

$$p(X_n = x_n | X_0 = x_0, \cdots, X_{n-1} = x_{n-1}) = p_n(x_n | x_0, \cdots, x_{n-1}). \tag{3}$$

Then

$$P(X_0, \cdots, X_n) = p(X_0) \prod_{k=1}^{n} p_k(X_k | X_0, \cdots, X_{k-1}), \tag{4}$$

$$f_n(\omega) = -\frac{1}{n+1}[\log p(X_0) + \sum_{k=1}^{n} \log p_k(X_k | X_0, \cdots, X_{k-1})]. \tag{5}$$

2000 *AMS Mathematics Subject Classification:* 60F15.

*The author would like to thank the anonymous referee for his careful and valuable suggestions. This work is supported by Natural Science Foundation of High University of Jiangsu Province (09KJD110002).

Definition 1 *Suppose $\sigma_n(\omega)$ is an increasing nonnegative stochastic sequence, and $\sigma_n(\omega)\uparrow\infty$, we call*

$$f_{[\sigma_n(\omega)]}(\omega) = -\frac{1}{\sigma_n(\omega)+1}[\log p(X_0) + \sum_{k=1}^{[\sigma_n(\omega)]} \log p_k(X_k|X_0,\cdots,X_{k-1})] \tag{6}$$

the generalized relative entropy density of the arbitrary discrete information source $\{X_i, 0\le i\le [\sigma_n]\}$, where $[c]$ represents the integral part of c. We let $\sigma_n(\omega)=n$, the generalized relative entropy density (6) is just the general relative entropy density (5).

The convergence of $f_n(\omega)$ to a constant in a sense (L_1 convergence, convergence in probability, a.s. convergence) is called Shannon-McMillan theorem or the entropy theorem or the asymptotic equipartition property (AEP) in information theory. Shannon [7] first proved the AEP for the convergence in probability for stationary ergodic information sources with a finite alphabet set. McMillan [6] and Breiman [3] proved the AEP in L_1 and a.s. convergence, respectively, for stationary ergodic information sources. Chung [4] considered the case of the countable alphabet set. The AEP for general stochastic processes can be found, for example, in Barron [2] and Algoet and Cover [1]. Liu and Yang [5,8] have proved the AEP for a class of the nonhomogeneous Markov information source and Markov chains fields on Cayley trees. Many practical information sources, such as language and the image information, are often mth-order Markov information sources, and always nonhomogeneous. Hence it is of importance to study the AEP for the mth-order nonhomogeneous Markov information source in information theory. In this correspondence, we establish several Shannon-McMillan theorems for the generalized entropy density of the arbitrary information source with the countable and the finite alphabet sets. As corollaries, Shannon-McMillan theorems for the entropy density of the mth-order nonhomogeneous Markov information source and the non-memory information source are obtained.

Liu and Yang have studied Shannon-McMillan theorems for the nonhomogeneous Markov chain with a finite state space in Ref [5]. Analogously, Yang and Ye have given a Shannon-McMillan theorem for the arbitrary information source with a finite state space in Ref [9]. Moreover, the results of Ref [5] and Ref [9] are on the general entropy density with the natural sum n. This paper focuses on the generalized Shannon-McMillan theorems for the arbitrary information source with a countable state space, and the conclusion is about the generalized entropy density with the random sum σ_n. Therefore, the results of Liu and Yang(see [5]) and Yang, Ye(see [9]) are extended.

Definition 2 *From now on, we denote*

$$h_k(x_0,\cdots,x_{k-1}) = -\sum\nolimits_{x_k\in S} p_k(x_k|x_0,\cdots,x_{k-1})\log p_k(x_k|x_0,\cdots,x_{k-1}), \tag{7}$$

$$H_k(\omega) = h_k(X_0,\cdots,X_{k-1}),\quad k\ge 1. \tag{8}$$

$H_k(\omega)$ is called the random conditional entropy of X_k relative to $X_0,\cdots,X_{k-1}$.

We denote $X_0^n=\{X_0,\cdots,X_n\}$, $X_m^n=\{X_m,\cdots,X_n\}$. Denote by x_0^n, x_m^n the realizations of X_0^n and X_m^n, respectively.

2. Main results and the proof

In this section we investigate the asymptotic equipartition property (AEP) of the arbitrary information source $\{X_n, n\ge 0\}$. Let's now state the main result of this section.

Theorem 1 *Let $\{X_n, n \geq 0\}$ be an arbitrary information source and $\{\sigma_n(\omega), n \geq 0\}$ be a nonnegative stochastic sequence that increases to infinity as n goes to infinity. Take an $\alpha \in (0, +\infty)$, $0 < C \leq 1$, denote*

$$D = \{\omega : \lim_n \sigma_n(\omega) = \infty\}, \tag{9}$$

we assume

$$B_\alpha = \limsup_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} E[p_k(X_k|X_0^{k-1})^{-\alpha} I_{\{p_k(X_k|X_0^{k-1})\leq C\}}|X_0^{k-1}] < \infty. \quad a.s. \tag{10}$$

Then, the following holds:

$$\lim_{n\to\infty} [f_{[\sigma_n(\omega)]}(\omega) - \frac{1}{\sigma_n(\omega)} \sum_{k=1}^{[\sigma_n(\omega)]} H_k(\omega)] = 0. \quad a.s. \quad \omega \in D. \tag{11}$$

where, $[c]$ represents the integral part of c.

Proof. Let λ be an arbitrary constant. Denote

$$Q_k(\lambda) = E[p_k(X_k|X_0^{k-1})^{-\lambda}|X_0^{k-1} = x_0^{k-1}] = \sum_{x_k\in S} p_k(x_k|x_0^{k-1})^{1-\lambda}, \tag{12}$$

$$q_k(\lambda, x_k) = \frac{p_k(x_k|x_0^{k-1})^{1-\lambda}}{Q_k(\lambda)}, \quad x_k \in S. \tag{13}$$

$$g(\lambda, x_0, \cdots, x_n) = p(x_0) \prod_{k=1}^{n} q_k(\lambda, x_k). \tag{14}$$

It is easy to see that $g(\lambda, x_0, \cdots, x_n), n = 1, 2, \cdots$ are a family of consistent distribution functions defined on S^n. We set

$$T_n(\lambda, \omega) = \frac{g(\lambda, X_0, \cdots, X_n)}{p(X_0, \cdots, X_n)}. \tag{15}$$

Then $\{T_n(\lambda, \omega), n \geq 1\}$ is a non-negative super-martingale(see[4]). Applying Doob's Martingale Convergence Theorem to (15), we have

$$\lim_{n\to\infty} T_n(\lambda, \omega) = T_\infty(\lambda, \omega) < \infty. \tag{16}$$

Hence by using (9) and (16), we obtain

$$\limsup_{n\to\infty} \frac{1}{\sigma_n} \log T_{[\sigma_n(\omega)]}(\lambda, \omega) \leq 0, \quad a.s. \quad \omega \in D. \tag{17}$$

Taking into account (4) and (12)–(15), the left side of (17) can be rewritten as

$$\frac{1}{\sigma_n} \log T_{[\sigma_n]}(\lambda, \omega) = \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} [(-\lambda \log p_k(X_k|X_0^{k-1})) - \log E(p_k(X_k|X_0^{k-1})^{-\lambda}|X_0^{k-1})]. \tag{18}$$

On the other hand, by the inequality $e^x - 1 - x \leq (1/2)x^2 e^{|x|}$, we have

$$x^{-\lambda} - 1 + \lambda \log x \leq (1/2)\lambda^2 (\log x)^2 x^{-|\lambda|}, \quad 0 \leq x \leq 1. \tag{19}$$

By means of (10), (18), (19) and the inequality $\log x \le x-1 (x\ge 0)$, noticing that

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

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

$$\begin{aligned}
&\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=1}^{[\sigma_n]}[(-\lambda\log p_k(X_k|X_0^{k-1})) - E(-\lambda\log p_k(X_k|X_0^{k-1})|X_0^{k-1})]\\
\le\ &\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=1}^{[\sigma_n]}[\log E(p_k(X_k|X_0^{k-1})^{-\lambda}|X_0^{k-1}) - E(-\lambda\log p_k(X_k|X_0^{k-1})|X_0^{k-1})]\\
\le\ &\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=1}^{[\sigma_n]}[E(p_k(X_k|X_0^{k-1})^{-\lambda}|X_0^{k-1}) - 1 - E(-\lambda\log p_k(X_k|X_0^{k-1})|X_0^{k-1})]\\
\le\ &\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=1}^{[\sigma_n]}E[(1/2)\lambda^2(\log(p_k(X_k|X_0^{k-1}))^2 p_k(X_k|X_0^{k-1})^{-|\lambda|}|X_0^{k-1}]\\
=\ &\limsup_{n\to\infty}\frac{1}{\sigma_n(\omega)}\sum_{k=1}^{[\sigma_n]}E[\frac{\lambda^2}{2}(\log(p_k(X_k|X_0^{k-1}))^2 p_k(X_k|X_0^{k-1})^{\alpha-|\lambda|}p_k(X_k|X_0^{k-1})^{-\alpha}|X_0^{k-1}]\\
\le\ &\frac{\lambda^2}{2}\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=1}^{[\sigma_n]}E[\frac{4e^{-2}}{(\alpha-|\lambda|)^2}\cdot p_k(X_k|X_0^{k-1})^{-\alpha}|X_0^{k-1}]\\
\le\ &\frac{2\lambda^2e^{-2}}{(\alpha-t)^2}\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=1}^{[\sigma_n]}E[p_k(X_k|X_0^{k-1})^{-\alpha}(I_{\{p_k(X_k|X_0^{k-1})\le C\}}+I_{\{p_k(X_k|X_0^{k-1})>C\}})|X_0^{k-1}]\\
\le\ &\frac{2\lambda^2e^{-2}}{(\alpha-t)^2}\{\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=1}^{[\sigma_n]}E[p_k(X_k|X_0^{k-1})^{-\alpha}I_{\{p_k(X_k|X_0^{k-1})\le C\}}|X_0^{k-1}] + \limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=1}^{[\sigma_n]}C^{-\alpha}\}\\
\le\ &\frac{2\lambda^2e^{-2}}{(\alpha-t)^2}\{B_\alpha + C^{-\alpha}\}<\infty. \qquad a.s. \qquad \omega\in D. \qquad (20)
\end{aligned}$$

We now consider the case $0<\lambda<t<\alpha$, dividing both sides of (20) by λ, we have

$$\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=1}^{[\sigma_n]}[(-\log p_k(X_k|X_0^{k-1})) - E(-\log p_k(X_k|X_0^{k-1})|X_0^{k-1})] \le \frac{2\lambda e^{-2}}{(\alpha-t)^2}(B_\alpha + C^{-\alpha}).$$

$$a.s. \qquad \omega\in D. \qquad (21)$$

We can take the limit in (21) as $\lambda\to 0$ and we get

$$\limsup_{n\to\infty}\frac{1}{\sigma_n}\sum_{k=1}^{[\sigma_n]}[(-\log p_k(X_k|X_0^{k-1})) - E(-\log p_k(X_k|X_0^{k-1})|X_0^{k-1})] \le 0.$$

$$a.s. \qquad \omega\in D. \qquad (22)$$

In the case of $-\alpha < -t < \lambda < 0$, analogously, it follows from (20) that

$$\liminf_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} [(-\log p_k(X_k|X_0^{k-1})) - E(-\log p_k(X_k|X_0^{k-1})|X_0^{k-1})] \geq 0.$$

$$a.s. \qquad \omega \in D. \tag{23}$$

Moreover, (22) and (23) imply that

$$\lim_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} [(-\log p_k(X_k|X_0^{k-1})) - E(-\log p_k(X_k|X_0^{k-1})|X_0^{k-1})] = 0. \quad a.s. \quad \omega \in D. \tag{24}$$

By the definition of $H_k(\omega)$ we can write

$$H_k(\omega) = -\sum_{x_k \in S} p_k(x_k|X_0^{k-1}) \log p_k(x_k|X_0^{k-1}) = E(-\log p_k(X_k|X_0^{k-1})|X_0^{k-1}).$$

From (6) and (24) we conclude that

$$\begin{aligned} &\lim_{n\to\infty} [f_{[\sigma_n]}(\omega) - \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} H_k(\omega)] \\ = \quad &\lim_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} [(-\log p_k(X_k|X_0^{k-1})) - E(-\log p_k(X_k|X_0^{k-1})|X_0^{k-1})] \\ &- \lim_{n\to\infty} \frac{1}{\sigma_n(\omega)} \log p(X_0) = 0. \qquad a.s. \quad \omega \in D. \end{aligned} \tag{25}$$

Consequently, (11) follows from (25). This completes the proof of Theorem 1. □

Remark 1. When $C = 1$, B_α reduces to

$$\limsup_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} E[p_k(X_k|X_0^{k-1})^{-\alpha}|X_0^{k-1}] < \infty. \quad a.s. \tag{26}$$

Then (11) also holds.

Remark 2. In Remark 1, (26) means that $\sum_{k=1}^{[\sigma_n(\omega)]} E[p_k(X_k|X_0^{k-1})^{-\alpha}|X_0^{k-1}]$ has to be infinite of the same order of $\sigma_n(\omega)$ or infinite of the lower order of $\sigma_n(\omega)$, otherwise, Theorem 1 can not hold.

Remark 3. When $C = 1$, replace $S = \{1, 2, \cdots\}$ with $S = \{1, 2, \cdots, N\}$, (10) holds naturally. Therefore, (11) is also valid.

Remark 4. When $C = 1$, $S = \{1, 2, \cdots, N\}$, letting $\sigma_n(\omega) = n$, $n \geq 0$, (11) holds still and $D(\omega) = \Omega$.

Suppose $\{X_n, n \geq 0\}$ is an mth-order nonhomogeneous Markov information source, then as $n \geq m$,

$$P(X_n = x_n | X_0 = x_0, \cdots, X_{n-1} = x_{n-1}) = P(X_n = x_n | X_{n-m} = x_{n-m}, \cdots, X_{n-1} = x_{n-1}). \tag{27}$$

Denote

$$q(i_0, \cdots, i_{m-1}) = P(X_0 = i_0, \cdots, X_{m-1} = i_{m-1}), \tag{28}$$

$$p_n(j|i_1, \cdots, i_m) = P(X_n = j | X_{n-m} = i_1, \cdots, X_{n-1} = i_m). \tag{29}$$

We define $q(i_0, \cdots, i_{m-1})$ as the m dimensional initial distribution, define $p_n(j|i_1, \cdots, i_m), n \geq m$ as the mth-order transition probabilities, and

$$P_n = (p_n(j|i_1, \cdots, i_m)) \tag{30}$$

are called the mth-order transition matrices. In this case,

$$p(x_0, \cdots, x_n) = q(x_0, \cdots, x_{m-1}) \prod_{k=m}^{n} p_k(x_k | x_{k-m}, \cdots, x_{k-1}), \tag{31}$$

$$f_n(\omega) = -\frac{1}{n+1}[\log q(X_0, \cdots, X_{m-1}) + \sum_{k=m}^{n} \log p_k(X_k | X_{k-m}, \cdots, X_{k-1})]. \tag{32}$$

Corollary 1 *Let $\{X_n, n \geq 0\}$ be an mth-order nonhomogeneous Markov information source with the m dimensional initial distribution (28) and the mth-order transition probabilities (29). Let $f_{[\sigma_n(\omega)]}(\omega)$ be defined as (6), denote*

$$H(p_k(X_{k-m}^{k-1}, 1), \cdots, p_k(X_{k-m}^{k-1}, N)) = -\sum_{x_k \in S} p_k(x_k | X_{k-m}^{k-1}) \log p_k(x_k | X_{k-m}^{k-1}).$$

Then

$$\lim_{n\to\infty} [f_{[\sigma_n(\omega)]}(\omega) - \frac{1}{\sigma_n(\omega)} \sum_{k=m}^{[\sigma_n(\omega)]} H(p_k(X_{k-m}^{k-1}, 1), \cdots, p_k(X_{k-m}^{k-1}, N))] = 0. \quad a.s. \quad \omega \in D. \tag{33}$$

Proof. At this moment, $p_k(x_k|x_0^{k-1}) = p_k(x_k|x_{k-m}^{k-1}), \quad k \geq m$. (33) follows from (11) accordingly for $H_k(\omega) = H(p_k(X_{k-m}^{k-1}, 1), \cdots, p_k(X_{k-m}^{k-1}, N))$. □

Remark 5. Let $m = 1$, $\sigma_n(\omega) = n$, $n \geq 0$ in Corollary 1, the corollary is Theorem 2. of Liu and Yang(see[5]).

Corollary 2 *Let $\{X_n, n \geq 0\}$ be a non-memory information source, $f_{[\sigma_n(\omega)]}(\omega)$ be defined as before, denote*

$$H(p_k(1), \cdots, p_k(N)) = -\sum_{x_k \in S} p_k(x_k) \log p_k(x_k). \tag{34}$$

Then

$$\lim_{n\to\infty} [f_{[\sigma_n(\omega)]}(\omega) - \frac{1}{\sigma_n(\omega)} \sum_{k=1}^{[\sigma_n(\omega)]} H(p_k(1), \cdots, p_k(N))] = 0. \qquad a.s. \quad \omega \in D. \tag{35}$$

Proof. At this moment, we have $p_k(X_k|X_{k-m}^{k-1}) = p_k(X_k)$, thus, $H_k(\omega) = H(p_k(1), \cdots, p_k(N))$. (35) follows from (11). □

3. A weaker Shannon-McMillan Theorem for the arbitrary information source

Taking into account the theoretical and practical importance of Shannon-McMillan Theorems in information theory, in this section we will extract a weaker condition and present another kind of proof technique. For this purpose, we introduce the following theorem.

Theorem 2 *Let* $\{X_n, n \geq 0\}$ *be an arbitrary information source taking values in a countable alphabet set* $S = \{s_1, s_2, \cdots\}$. *Let* $f_{[\sigma_n(\omega)]}(\omega)$ *and* $H_k(\omega)$ *be defined by (6) and (8). Take an* $\alpha \in (0, +\infty)$, $0 < C < 1$, *we assume that*

$$\limsup_{n\to\infty} \frac{1}{\sigma_n(\omega)} \sum_{k=1}^{[\sigma_n(\omega)]} E[(p_k(X_k|X_0^{k-1}) - 1)^2 p_k(X_k|X_0^{k-1})^{-\alpha} I_{\{p_k(X_k|X_0^{k-1}) \leq C\}} |X_0^{k-1}] < \infty. \quad a.s. \tag{36}$$

Then, (11) holds also.

Proof. Let us denote $p_k(X_k|X_0^{k-1}) = p_k$ in brief, taking into account (36) and the inequality $\log(1+x) \geq x - \frac{x^2}{2}$, $(x > -1)$, from the second inequality of (20) in the proof of Theorem 1 we can conclude that in the case of $0 < |\lambda| < \alpha$,

$$\begin{aligned}
& \limsup_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} [(-\lambda \log p_k(X_k|X_0^{k-1})) - E(-\lambda \log p_k(X_k|X_0^{k-1})|X_0^{k-1})] \\
\leq\ & \limsup_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} E[(1/2)\lambda^2 (\log(p_k(X_k|X_0^{k-1}))^2 p_k(X_k|X_0^{k-1})^{-|\lambda|} |X_0^{k-1}] \\
\leq\ & \frac{\lambda^2}{2} \limsup_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} E[(\log(p_k(X_k|X_0^{k-1}))^2 p_k(X_k|X_0^{k-1})^{-\alpha} |X_0^{k-1}] \\
=\ & \frac{\lambda^2}{2} \limsup_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} E[(\log p_k)^2 p_k^{-\alpha} (I_{\{p_k(X_k|X_0^{k-1}) \leq C\}} + I_{\{p_k(X_k|X_0^{k-1}) > C\}}) |X_0^{k-1}] \\
\leq\ & \frac{\lambda^2}{2} \{\limsup_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} E[(\log p_k)^2 p_k^{-\alpha} I_{\{p_k(X_k|X_0^{k-1}) \leq C\}} |X_0^{k-1}] + \limsup_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} C^{-\alpha} (\log C)^2\} \\
\leq\ & \frac{\lambda^2}{2} \{\limsup_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} E[\{p_k - 1 - \frac{(p_k-1)^2}{2}\}^2 p_k^{-\alpha} I_{\{p_k(X_k|X_0^{k-1}) \leq C\}} |X_0^{k-1}] + C^{-\alpha} (\log C)^2\} \\
=\ & \frac{\lambda^2}{2} \{\limsup_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} E[\{(p_k - 1)(\frac{3-p_k}{2})\}^2 p_k^{-\alpha} I_{\{p_k(X_k|X_0^{k-1}) \leq C\}} |X_0^{k-1}] + C^{-\alpha} (\log C)^2\} \\
\leq\ & \frac{9\lambda^2}{8} \{\limsup_{n\to\infty} \frac{1}{\sigma_n} \sum_{k=1}^{[\sigma_n]} E[(p_k - 1)^2 p_k^{-\alpha} I_{\{p_k(X_k|X_0^{k-1}) \leq C\}} |X_0^{k-1}] + \frac{4}{9} C^{-\alpha} (\log C)^2\} < \infty. \\
& a.s. \quad \omega \in D.
\end{aligned} \tag{37}$$

Imitating the proof of (21)–(25), we can obtain (11). This completes the proof of Theorem 2. □

References

[1] Algoet, P.H. and Cover, T.M.: A sandwich proof of Shannon-Mcmillan theorem. Ann.Probab. 16, 899-909 (1988).

[2] Barron, A.R.: The strong ergodic theorem for densities:Generalized Shannon-McMillan-Breiman theorem. Ann.Probab. 13,1292-1303 (1985)

[3] Breiman, L.: the individual ergodic theorem of information theory. Ann.Math.Statist. 28, 809-811 (1957)

[4] Chung K.L.: A Course in Probability Theory. Academic Press, New York. (1974)

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

[6] Mcmillan, B.: The Basic Theorem of information theory. Ann.Math.Statist. 24,196-216 (1953)

[7] Shannon, C.: A mathematical theory of communication. Bell Syst. Tech. J. 27,379-423 (1948)

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

[9] Yang W.G., Ye Z.X. and Liu W.: A local convergence theorem for partial sums of stochastic adapt sequences. Czechoslovak Mathematical Journal, 2, 86-91 (2006)

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

Received: 03.03.2009