Probability in the Engineering and Informational Sciences, **35**, 2021, 297–315.
doi:10.1017/S0269964819000354

SOME LIMIT THEOREMS OF DELAYED AVERAGES FOR COUNTABLE NONHOMOGENEOUS MARKOV CHAINS

PINGPING ZHONG, WEIGUO YANG, ZHIYAN SHI, AND YAN ZHANG
Faculty of Science, Jiangsu University, Zhenjiang 212013, China
E-mail: shizhiyan1984@126.com
Jingjiang College of Jiangsu University, Zhenjiang 212013, China

The purpose of this paper is to establish some limit theorems of delayed averages for countable nonhomogeneous Markov chains. The definition of the generalized C-strong ergodicity and the generalized uniformly C-strong ergodicity for countable nonhomogeneous Markov chains is introduced first. Then a theorem about the generalized C-strong ergodicity and the generalized uniformly C-strong ergodicity for the nonhomogeneous Markov chains is established, and its applications to the information theory are given. Finally, the strong law of large numbers of delayed averages of bivariate functions for countable nonhomogeneous Markov chains is proved.

Keywords: countable nonhomogeneous markov chains, generalized c-strong ergodicity, strong law of large numbers of delayed averages

Jel Classification: 60G60, 60F15

1. INTRODUCTION

We begin with introducing some notations that will be used throughout the paper.

Let $\{X_n,\ n \geq 0\}$ be a nonhomogeneous Markov chain taking values in state space $S = \{1, 2, \ldots\}$ with the transition matrices

$$P_n = (p_n(i,j)), \quad i,j \in S,\ n \geq 1, \tag{1.1}$$

where $p_n(i,j) = P(X_n = j \,|\, X_{n-1} = i)$. Let

$$P^{(m,n)} = P_{m+1}P_{m+2}\cdots P_n, \tag{1.2}$$

and let $p^{(m,n)}(i,j)$ be the (i,j) element of $P^{(m,n)}$. It is easy to see that

$$p^{(m,n)}(i,j) = P(X_n = j \,|\, X_m = i). \tag{1.3}$$

Let $f^{(0)}$ be a distributional row vector, called initial distribution. Let

$$f^{(k)} = f^{(0)}P_1\cdots P_k = f^{(0)}P^{(0,k)}, \quad f^{(k)}(j) = P(X_k = j).$$

If the Markov chain is homogeneous, $\{P_n,\ n \geq 1\}$ will be denoted simply by P and $P^{(m,m+k)}$ be denoted as P^k.

 0269-9648/19 $25.00

If $A = (a_{ij})$ is a real matrix defined on $S \times S$, where $S = \{1, 2, \ldots\}$. The norm $\|\cdot\|$ of A is defined as

$$\|A\| = \sup_{i \in S} \sum_{j \in S} |a_{ij}|.$$

It is obvious that the definition of $\|A\|$ is the norm of A as an operator on ℓ_∞ (space of bounded sequences with sup-norm). If $f = (f_1, f_2, \ldots)$ is a row vector, define $\|f\| = \sum_{j \in S} |f_j|$. If $g = (g_1, g_2, \ldots)'$ is a column vector, define $\|g\| = \sup_{i \in S} |mg_i|$. The norm defined above satisfies the following two properties (see [1]), which will be used repeatedly in this paper.

(a) $\|AB\| \le \|A\| \cdot \|B\|$ for all matrices A and B;
(b) $\|P\| = 1$ for all transition matrix P.

Let Q be a constant transition matrix, that is, each row of this transition matrix Q is the same. Let $P = (p_{ij})$ be a transition matrix. The delta coefficient of P, denoted by $\delta(P)$, is defined as

$$\delta(P) = \sup_{i,k} \sum_{j=1}^{\infty} [p_{ij} - p_{kj}]^+, \tag{1.4}$$

where $[p_{ij} - p_{kj}]^+ = \max\{0, p_{ij} - p_{kj}\}$ (see [5, p. 144]).

Let $\{X_n,\ n \ge 0\}$ be a nonhomogeneous Markov chain taking values in state space $S = \{1, 2, \ldots\}$ with the transition matrices $\{P_n,\ n \ge 1\}$. The sequence $\{P_n,\ n \ge 1\}$ is said to be weakly ergodic if for each $m \ge 0$,

$$\delta(P^{(m,k)}) \to 0 \quad \text{as} \quad k \to \infty. \tag{1.5}$$

We also call the Markov chain $\{X_n,\ n \ge 0\}$ weakly ergodic (see [5, p. 149]).

If for each $m \ge 0$

$$\lim_n \|P^{(m,m+n)} - Q\| = 0, \tag{1.6}$$

then the sequence $\{P_n,\ n \ge 1\}$ is said to be strongly ergodic (with a constant transition matrix Q) (see [5, p. 157]). We also call the Markov chain $\{X_n,\ n \ge 0\}$ strongly ergodic with respect to Q. If the Markov chain is homogeneous with the transition matrix P, then (1.6) becomes

$$\lim_n \|P^n - Q\| = 0. \tag{1.7}$$

If for each $m \ge 0$,

$$\lim_n \left\| \frac{1}{n} \sum_{t=1}^{n} P^{(m,m+t)} - Q \right\| = 0, \tag{1.8}$$

then $\{P_n,\ n \ge 1\}$ will be called C-strongly ergodic with respect to Q (see [5, p. 184]), and the Markov chain will also be called C-strongly ergodic with respect to Q. If the Markov chain is homogeneous with the transition matrix P, then (1.8) becomes

$$\lim_n \left\| \frac{1}{n} \sum_{t=1}^{n} P^t - Q \right\| = 0. \tag{1.9}$$

The sequence $\{P_n, n \geq 1\}$ will be called uniformly C-strongly ergodic (see [17]) if

$$\lim_n \sup_{m \geq 0} \left\| \frac{1}{n} \sum_{k=1}^{n} P^{(m,m+k)} - Q \right\| = 0. \tag{1.10}$$

Let $\{a_n,\ n \geq 0\}$ and $\{\phi(n),\ n \geq 0\}$ be two sequences of nonnegative integers such that $\phi(n)$ converges to infinite as $n \to \infty$. If for all $m \geq 0$,

$$\lim_n \left\| \frac{1}{\phi(n)} \sum_{t=a_n+1}^{a_n+\phi(n)} P^{(m,m+t)} - Q \right\| = 0, \tag{1.11}$$

then $\{P_n,\ n \geq 1\}$ will be called generalized C-strongly ergodic with respect to Q. We also call the Markov chain $\{X_n,\ n \geq 0\}$ generalized C-strongly ergodic with respect to Q.

The sequence $\{P_n, n \geq 1\}$ is said to be generalized uniformly C-strongly ergodic with respect to Q if

$$\lim_n \sup_{m \geq 0} \left\| \frac{1}{\phi(n)} \sum_{t=a_n+1}^{a_n+\phi(n)} P^{(m,m+t)} - Q \right\| = 0. \tag{1.12}$$

It is apparent that both C-strong ergodicity and uniform C-strong ergodicity of Markov chains are the special cases of generalized C-strong ergodicity and generalized uniform C-strong ergodicity of Markov chains, respectively.

An irreducible transition matrix P, of period d ($d \geq 1$) partitions the state space S into d disjoint subspaces $C_0, C_1, \ldots, C_{d-1}$, and P^d yields d transition matrices $\{T_l,\ 0 \leq l \leq d-1\}$, where T_l is defined on C_l. If the irreducible periodic transition matrix P is finite, then each T_l is automatically strongly ergodic. But if P is infinite, the strong ergodicity of T_l is not guaranteed. As in [1], we shall consider an irreducible transition matrix P, of period d, in which T_l is strongly ergodic for $l = 0, 1, \ldots, d-1$ only. As in [1], such a transition matrix will be called periodic strongly ergodic. It has been proved that periodic strongly ergodic of P implies C-strongly ergodic of P (see [1, Lem 1.1]), but the inverse is false (see [9]).

There have been some works on the weak ergodicity, the strong ergodicity, and the C-strongly ergodicity for nonhomogeneous Markov chains. It has been proved that the strong ergodicity implies the weak ergodicity and strong ergodicity and weak ergodicity are equivalent when the Markov chain is homogenous (see [5, Chapter 5]). Isaacson and Madson (see [5, Thm. V.5.1]) proved that if $\lim_n \|P_n - P\| = 0$, where P is periodic strongly ergodic, then the nonhomogeneous Markov chain is C-strongly ergodic. Bowerman, David and Isaacson (see [1]) obtained the convergence rates of Cesaro averages $(1/n)\sum_{t=1}^{n} P^{(m,m+t)}$ under the condition of $\|P_t - P\| \leq G/t^{\alpha},\ t \geq 1$, where P is periodic strongly ergodic. Yang (see [16]) studied the C-strong ergodicity and uniform C-strong ergodicity for countable nonhomogeneous Markov chains (see following Corollary 1). He provided an example (see [17, Example 1]) to show that (3.6) can not imply (3.10) (see below). Also from this example, one can find that the uniform C-strong ergodicity can not be implied by the C-strong ergodicity. Isaacson and Senta (see [6]) studied the necessary and sufficient conditions for the strongly ergodic of nonhomogeneous Markov chains. Tan (see [12]) studied the weak ergodicity of nonhomogeneous Markov chains. Mukhamedov (see [10]) studied the Dobrushin ergodicity coefficient and the ergodicity of nonhomogeneous Markov chains. Recently, Waqas and Yang (see [14]) have studied the generalized absolute mean strong ergodicity for the countable nonhomogeneous Markov chains.

Some works on the limit theorems for nonhomogeneous Markov chains also have been studied. Rosenblatt (see [11]) studied some theorems concerning the strong law of large numbers for nonhomogeneous Markov chains. Chiang and Chow (see [2]) proved a limit theorem for a class of nonhomogeneous Markov processes. Liu and Liu (see [7]) studied the strong law of large numbers for functionals of countable nonhomogeneous Markov chains. Liu and Yang (see [8]) studied the strong law of large numbers and Shannon–McMillan theorem for finite nonhomogeneous Markov chains. Yang (see [15]) studied the Shannon–McMillan theorem for a nonhomogeneous Markov information source. Dietz and Sethuraman (see [4]) studied large deviations for a class of nonhomogeneous Markov chains. Yang (see [16]) studied the strong law of large numbers of bivariate functions for countable nonhomogeneous Markov chains under the condition $\lim_n (1/n) \sum_{k=1}^{n} \|P_k - P\| = 0$ where P is periodic strongly ergodic. Yang (see [17]) studied another strong law of large numbers for countable nonhomogeneous Markov chains under the condition of uniformly C-strong ergodicity. He also provided two examples in [17] to show that the results in [15,16] are not overlapping. Zhong, Yang and Liang (see [19]) studied the Shannon–McMillan theorem for finite asymptotic circular Markov chains. Yang, Wang and Shi (see [18]) studied strong law of large numbers for countable asymptotic circular Markov chains. Wang and Yang (see [13]) studied the generalized entropy ergodic theorem for finite nonhomogeneous Markov chains.

As mentioned before, Yang (see [16]) studied the C-strong ergodicity, the uniform C-strong ergodicity, and the strong law of large numbers of bivariate functions for countable nonhomogeneous Markov chains, respectively. He also (see [15]) studied the Shannon–McMillan theorem for finite nonhomogeneous Markov information source. In this paper, we first give the definition of the generalized C-strong ergodicity and the generalized uniform C-strong ergodicity for countable nonhomogeneous Markov chains. Then, we obtain a theorem about the generalized C-strong ergodicity and generalized uniform C-strong ergodicity for this Markov chains. As a corollary, we get the same theorem with Yang's work (see [16]) about the C-strong ergodicity and the uniform C-strong ergodicity for countable nonhomogeneous Markov chains. We also obtain a corollary showing that our result of generalized C-strong ergodicity and generalized uniform C-strong ergodicity for countable nonhomogeneous Markov chains can imply more results. Moreover, we give some applications to the information theory. Finally, we establish the strong law of large numbers of the delayed averages of bivariate functions for countable nonhomogeneous Markov chains. As corollaries, the main result of [15] is deduced. We also get the strong law of large numbers for the frequencies of occurrence of states of delayed averages for countable nonhomogeneous Markov chains, which can not be implied by previous known results.

The approach used in this paper is different from that used in [15–19], where the strong law of large numbers for martingale is applied. The strong law of large numbers of the delayed averages of bivariate functions for countable nonhomogeneous Markov chains in this paper follows from Lemma 3, which is similar to Lemma 1 in [13]. The essence of the approach, used to prove Lemma 1 in [13], is first to construct a one-parameter class of random variables with means 1 and then to prove the existence of a.e. convergence of certain random variables by using Borel–Cantelli Lemma.

This article is organized as follows. In Section 2, we state some lemmas. In Section 3, we present a theorem about the generalized C-strong ergodicity and the generalized uniform C-strong ergodicity for the nonhomogeneous Markov chains, and give its applications to the information theory. In Section 4, we state a strong law of large numbers of delayed averages of bivariate functions for countable nonhomogeneous Markov chains. In Section 5, we provide the proofs of Lemma 4, Theorem 1, and Theorem 2.

2. SOME LEMMAS

LEMMA 1 (see [5, Lem V.5.2]): *Let P be a transition matrix. Assume that P is periodic strongly ergodic with period d, and let Q be the constant transition matrix each row of which is the left eigenvector $\pi = (\pi_1, \pi_2, \ldots)$ of P, which solves uniquely the system of equations $\pi P = \pi$ and $\sum_i \pi_i = 1$. Then*

$$\lim_n \left\| \frac{1}{d} \sum_{l=0}^{d-1} P^{nd+l} - Q \right\| = 0. \tag{2.1}$$

LEMMA 2: *Let $\{a_n,\ n \geq 0\}$ and $\{\phi(n),\ n \geq 0\}$ be two sequences of nonnegative integers such that for any positive integers n, m,*

$$\phi(m+n) - \phi(n) \geq m, \quad \frac{\phi(m+n)}{\phi(n)} \longrightarrow 1 \quad (n \to \infty). \tag{2.2}$$

Let $\{b_n,\ n \geq 0\}$ be a sequence of real numbers, and let b be a real number. If

$$\lim_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} |b_k - b| = 0, \tag{2.3}$$

then for any positive integer m, we have

$$\lim_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} |b_{k+m} - b| = 0. \tag{2.4}$$

Remark 1: It is easy to see that if $\phi(n) = n^\alpha$ $(\alpha \geq 1)$, then (2.2) holds.

PROOF: By (2.2), for any $m \geq 0$,

$$\begin{aligned}
\frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} |b_{k+m} - b| &= \frac{1}{\phi(n)} \sum_{v=a_n+m+1}^{a_n+m+\phi(n)} |b_v - b| \\
&\leq \frac{1}{\phi(n)} \sum_{v=a_n+1}^{a_n+m+\phi(n)} |b_v - b| \leq \frac{1}{\phi(n)} \sum_{v=a_n+1}^{a_n+\phi(m+n)} |b_v - b| \\
&= \frac{\phi(m+n)}{\phi(n)} \frac{1}{\phi(m+n)} \sum_{v=a_n+1}^{a_n+\phi(m+n)} |b_v - b|,
\end{aligned} \tag{2.5}$$

(2.4) follows from (2.2), (2.3), and (2.5). ■

LEMMA 3: *Suppose $\{X_n,\ n \geq 0\}$ is a nonhomogeneous Markov chain taking values in state space $S = \{1, 2, \ldots\}$ with the transition matrices (1.1). Suppose $\{a_n,\ n \geq 0\}$ and $\{\phi(n),\ n \geq 0\}$ are two sequences of nonnegative integers such that $\phi(n)$ converges to infinite* *as $n \to \infty$. Let $\{f_n(x,y),\ n \geq 1\}$ be a sequence of real functions defined on S^2. If for all $\varepsilon > 0$*

$$\sum_{n=1}^{\infty} \exp[-\varepsilon\phi(n)] < \infty, \tag{2.6}$$

and there exists a real number $0 < \gamma < \infty$ such that

$$\limsup_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} E[|f_k(X_{k-1},X_k)|^2 e^{\gamma|f_k(X_{k-1},X_k)|}|X_{k-1}] = c(\gamma;\omega) < \infty \text{ a.e.}, \tag{2.7}$$

then, we have

$$\lim_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \{f_k(X_{k-1},X_k) - E[f_k(X_{k-1},X_k)|X_{k-1}]\} = 0 \text{ a.e.} \tag{2.8}$$

Proof: By going through the proof of Lemma 1 in [13], we observe that this lemma still holds for the countable nonhomogeneous Markov chain. ∎

Lemma 4: *Let $\{X_n,\ n \geq 0\}$ be a nonhomogeneous Markov chain taking values in state space $S = \{1, 2, \ldots\}$ with the transition matrices (1.1). Let $\{f_n(x,y),\ n \geq 1\}$ be a sequence of real functions defined on S^2. Let $\{a_n,\ n \geq 0\}$ and $\{\phi(n),\ n \geq 0\}$ be defined as in Lemma 3 such that (2.6) holds. If there exists a real number $0 < \gamma < \infty$ such that*

$$M = \sup_{i,k} \sum_{j \in S} f_k^2(i,j) e^{\gamma|f_k(i,j)|} p_k(i,j) < \infty, \tag{2.9}$$

then for any positive integer t, we have

$$\lim_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \{f_k(X_{k-1},X_k) - E[f_{k+t}(X_{k+t-1},X_{k+t})|X_{k-1}]\} = 0 \text{ a.e.} \tag{2.10}$$

The proof of this lemma can be found in Section 5.

Remark 2: It is easy to see that if $\{f_n(x,y),\ n \geq 1\}$ is bounded, then the condition (2.9) holds for any $\gamma > 0$.

3. THE GENERALIZED C-STRONG ERGODICITY

In this section, we will state a theorem about the generalized C-strong ergodicity and the generalized uniform C-strong ergodicity for countable nonhomogeneous Markov chains, and give its applications.

Theorem 1: *Let $\{X_n,\ n \geq 0\}$ be a nonhomogeneous Markov chain with transition matrices (1.1). Let $\{a_n,\ n \geq 0\}$ and $\{\phi(n),\ n \geq 0\}$ be two sequences of nonnegative integers such that (2.2) holds. Let P be a periodic strongly ergodic transition matrix with period d, and let Q be the constant transition matrix each row of which is the left eigenvector $\pi = (\pi_1, \pi_2, \ldots)$ of P, which solves uniquely the system of equations $\pi P = \pi$ and $\sum_i \pi_i = 1$.*

(a) *If*

$$\lim_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \|P_k - P\| = 0, \tag{3.1}$$

then for any positive integers m and v, we have

$$\lim_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \|P^{(m+k,m+k+v)} - P^v\| = 0, \tag{3.2}$$

and

$$\lim_n \left\| \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} P^{(m,m+k)} - Q \right\| = 0. \tag{3.3}$$

(b) *If*

$$\lim_n \sup_{m\geq 0} \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \|P_{m+k} - P\| = 0, \tag{3.4}$$

then

$$\lim_n \sup_{m\geq 0} \left\| \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} P^{(m,m+k)} - Q \right\| = 0. \tag{3.5}$$

The proof of this theorem can be found in Section 5.

Remark 3: Obviously if $\|P_n - P\| \to 0$ $(n \to \infty)$, then (3.1) and (3.4) hold.

COROLLARY 1 (see Theorem 1 of [16]): *Let $\{X_n,\ n \geq 0\}$ be a nonhomogeneous Markov chain with the transition matrices (1.1). Let P be a periodic strongly ergodic transiton matrix with period d, and let Q be the constant transition matrix each row of which is the left eigenvector $\pi = (\pi_1, \pi_2, \ldots)$ of P, which is the unique solution of equations $\pi P = \pi$ and $\sum_i \pi_i = 1$.*

(a) *If*

$$\lim_n \frac{1}{n} \sum_{k=1}^{n} \|P_k - P\| = 0, \tag{3.6}$$

then for any positive integer m and v, we have

$$\lim_n \frac{1}{n} \sum_{k=1}^{n} \|P^{(m+k,m+k+v)} - P^v\| = 0, \tag{3.7}$$

and

$$\lim_n \left\| \frac{1}{n} \sum_{k=1}^{n} P^{(m,m+k)} - Q \right\| = 0. \tag{3.8}$$

(b) *If*

$$\lim_n \sup_{m\geq 0} \frac{1}{n}\sum_{k=1}^{n} \|P_{m+k} - P\| = 0, \tag{3.9}$$

then

$$\lim_n \sup_{m\geq 0} \left\| \frac{1}{n}\sum_{k=1}^{n} P^{(m,m+k)} - Q \right\| = 0. \tag{3.10}$$

PROOF: Letting $a_n = 0$ and $\phi(n) = n$ in Theorem 1, this corollary follows. ■

COROLLARY 2: *Let $\{X_n,\ n \geq 0\}$ be a nonhomogeneous Markov chain with transition matrices $\{P_n,\ n \geq 1\}$. Let P be a periodic strongly ergodic transition matrix, and let Q be the constant transition matrix each row of which is the left eigenvector $\pi = (\pi_1, \pi_2, \ldots)$ of P, which solves uniquely the system of equations $\pi P = \pi$ and $\sum_i \pi_i = 1$. If*

$$\lim_n \|P_n - P\| = 0, \tag{3.11}$$

then for any positive integer m,

$$\lim_n \sup_{m\geq 0} \left\| \frac{1}{n}\sum_{k=2^n+1}^{2^n+n} P^{(m,m+k)} - Q \right\| = 0. \tag{3.12}$$

PROOF: Letting $a_n = 2^n$ and $\phi(n) = n$ in Theorem 1, it is easy to see that (3.11) implies (3.4) and $\phi(n) = n$ satisfies (2.2), thus this corollary follows.

Next we give the applications of Theorem 1 to the information theory.

Let $\{X_n,\ n \geq 0\}$ be a sequence of random variables taking values in S. Let $\{a_n,\ n \geq 0\}$ and $\{\phi(n),\ n \geq 0\}$ be two sequences of nonnegative integers such that $\phi(n)$ converges to infinite as $n \to \infty$. Let $H(X_{a_n}, X_{a_n+1}, \ldots, X_{a_n+\phi(n)})$ be the joint entropy of random vector $(X_{a_n}, X_{a_n+1}, \ldots, X_{a_n+\phi(n)})$ and assume that it is finite. If the limit

$$\lim_n \frac{1}{\phi(n)} H(X_{a_n}, X_{a_n+1}, \ldots, X_{a_n+\phi(n)})$$

exists, it will be called the generalized entropy rate of the sequence of random variables $\{X_n,\ n \geq 0\}$. Denote it by $H^{\infty}_{a_n,\phi(n)}$. If $a_n = 0, \phi(n) = n$, $H^{\infty}_{a_n,\phi(n)}$ will be called the entropy rate of $\{X_n,\ n \geq 0\}$(see [3, p. 63]). Denote it by H^{∞}.

Since

$$\begin{aligned} H(X_{a_n}, \ldots, X_{a_n+\phi(n)}) = {} & H(X_{a_n}) + H(X_{a_n+1} \mid X_{a_n}) + \cdots \\ & + H(X_{a_n+\phi(n)} \mid X_{a_n}, \ldots, X_{a_n+\phi(n)-1}). \end{aligned} \tag{3.13}$$

If $\{X_n,\ n \geq 0\}$ is a nonhomogeneous Markov chain, then

$$\begin{aligned} & H(X_{a_n}, \ldots, X_{a_n+\phi(n)}) \\ & \quad = H(X_{a_n}) + H(X_{a_n+1} \mid X_{a_n}) + \cdots + H(X_{a_n+\phi(n)} \mid X_{a_n+\phi(n)-1}) \\ & \quad = H(X_{a_n}) + \sum_{k=a_n+1}^{a_n+\phi(n)} H(X_k \mid X_{k-1}). \end{aligned} \tag{3.14}$$

https://doi.org/10.1017/S0269964819000354 Published online by Cambridge University Press

It is easy to see that

$$H(X_k \mid X_{k-1}) = -\sum_i P(X_{k-1} = i) \sum_j p_k(i,j) \log p_k(i,j), \tag{3.15}$$

where $\{P_n = p_n(i,j),\ n \geq 1\}$ are the transition matrices of the nonhomogeneous Markov chain, and log is the natural logarithm. ■

Corollary 3: *Let $\{X_n,\ n \geq 0\}$ be a nonhomogeneous Markov chain with the initial distribution $f^{(0)}$ and the transition matrices (1.1). Let $\{a_n,\ n \geq 0\}$ and $\{\phi(n),\ n \geq 0\}$ be two sequences of nonnegative integers such that (2.2) holds. Suppose that $\{P_n,\ n \geq 1\}$ is generalized C-strongly ergodic with respect to the constant transition matrix Q, that is (1.11) holds. Let*

$$g_n(i) = -\sum_j p_n(i,j) \log p_n(i,j), \quad g(i) = -\sum_j p(i,j) \log p(i,j), \tag{3.16}$$

g_n and g be all column vectors with ith elements $g_n(i)$ and $g(i)$, respectively. If

$$\lim_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \|g_k - g\| = 0, \tag{3.17}$$

and $\{H(X_{a_n}),\ n \geq 0\}$ is uniform bound, where $\|g\|$ is finite, then the generalized entropy rate of the nonhomogeneous Markov chain exists, and

$$\lim_n \frac{1}{\phi(n)} H(X_{a_n}, \ldots, X_{a_n+\phi(n)}) = -\sum_i \pi_i \sum_j p(i,j) \log p(i,j), \tag{3.18}$$

where $\pi = (\pi_1, \pi_2, \ldots)$ is the row vector of the constant transition matrix Q.

Remark 4: If the nonhomogeneous Markov chain is finite, then $\{H(X_{a_n}),\ n \geq 0\}$ is uniform bound.

Remark 5: By this corollary, we know the generalized entropy rate of the countable nonhomogeneous Markov chain exists and all of them are equal under some conditions.

Proof of Corollary 3: Since $\{X_n,\ n \geq 0\}$ is a nonhomogeneous Markov chain, then (3.14) holds.

Since

$$\begin{aligned}
&\left| \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} H(X_k|X_{k-1}) + \sum_i \pi_i \sum_j p(i,j) \log p(i,j) \right| \\
&= \left| \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \left[-\sum_i P(X_{k-1} = i) \sum_j p_k(i,j) \log p_k(i,j) \right] \right. \\
&\quad \left. + \sum_i \pi_i \sum_j p(i,j) \log p(i,j) \right|
\end{aligned}$$

https://doi.org/10.1017/S0269964819000354 Published online by Cambridge University Press

$$= \left| \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} f^{(0)} P^{(0,k-1)} g_k - \pi g \right|$$
$$\le \left| \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} f^{(0)} P^{(0,k-1)} g_k - \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} f^{(0)} P^{(0,k-1)} g \right|$$
$$+ \left| \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} f^{(0)} P^{(0,k-1)} g - f^{(0)} Q g \right|$$
$$\le \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \|g_k - g\| + \|g\| \cdot \left\| \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} P^{(0,k-1)} - Q \right\|. \tag{3.19}$$

Since $\{H(X_{a_n}),\ n \ge 0\}$ is uniform bound, then $H(X_{a_n})/\phi(n) \to 0 (n \to 0)$. By (3.14), (3.19), (3.17), and (1.11), this corollary follows. ■

Corollary 4: *Let $\{X_n,\ n \ge 0\}$ be a nonhomogeneous Markov chain with the transition matrices (1.1). Suppose that $\{P_n,\ n \ge 1\}$ is C-strongly ergodic with respect to the constant transition matrix Q, that is (1.8) holds. Let g_n and g be defined as in Corollary 3. If*

$$\lim_n \frac{1}{n} \sum_{k=1}^{n} \|g_k - g\| = 0, \tag{3.20}$$

where $\|g\|$ is finite, then the entropy rate of the nonhomogeneous Markov chain exists, and

$$\lim_n \frac{1}{n} H(X_0, \ldots, X_n) = -\sum_i \pi_i \sum_j p(i,j) \log p(i,j), \tag{3.21}$$

where $\pi = (\pi_1, \pi_2, \ldots)$ is the row vector of the constant transition matrix Q.

Proof: Letting $a_n = 0$ and $\phi(n) = n$ in Corollary 3, this corollary follows. ■

4. STRONG LAW OF LARGE NUMBERS

In this section, we will establish the strong law of large numbers of delayed averages of bivariate functions for countable nonhomogeneous Markov chains.

Theorem 2: *Let $\{X_n,\ n \ge 0\}$ be a nonhomogeneous Markov chain taking values in state space $S = \{1, 2, \ldots\}$ with the transition matrices (1.1). Let $\{f_n(x,y),\ n \ge 1\}$ be a sequence of real functions defined on S^2. Let $\{a_n,\ n \ge 0\}$ and $\{\phi(n),\ n \ge 0\}$ be two sequences of nonnegative integers such that (2.2) holds. Let P be a C-strongly ergodic transition matrix with respect to the constant transition matrix Q. Let*

$$g_n(i) = \sum_j f_n(i,j) p_n(i,j), \tag{4.1}$$

https://doi.org/10.1017/S0269964819000354 Published online by Cambridge University Press

and $g(i)$, $i \in S$, be another function defined on S. Let g_n and g be column vectors with ith elements $g_n(i)$ and $g(i)$, respectively. If (2.9), (3.1), and (3.17) hold, where $\|g\|$ is finite, then

$$\lim_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} f_k(X_{k-1}, X_k) = \sum_i g(i)\pi_i \text{ a.e.,} \tag{4.2}$$

where $\pi = (\pi_1, \pi_2, \ldots)$ is the row vector of the constant transition matrix Q.

The proof of this theorem can be found in Section 5.

Corollary 5 (see Theorem 4 of [15]): *Let $\{X_n,\ n \geq 0\}$ be a nonhomogeneous Markov chain taking values in state space $S = \{1, 2, \ldots, N\}$ with the initial distribution $(q(1), q(2), \ldots, q(N))$ and the transition matrices $P_n = (p_n(i,j))_{N\times N}$. Let $f_n(\omega)$ be the relative entropy density of $\{X_k,\ 0 \leq k \leq n\}$, that is*

$$f_n(\omega) = -(1/n)\left[\log q(X_0) + \sum_{k=1}^{n} \log p_k(X_{k-1}, X_k)\right]. \tag{4.3}$$

Let $P = (p(i,j))_{N\times N}$ be another finite transition matrix, and assume P is irreducible. If

$$\lim_n \frac{1}{n} \sum_{k=1}^{n} |p_k(i,j) - p(i,j)| = 0, \quad \forall\ i, j \in S, \tag{4.4}$$

then

$$\lim_n f_n(\omega) = -\sum_{i=1}^{N} \pi_i \sum_{j=1}^{N} p(i,j) \log p(i,j) \text{ a.e.,} \tag{4.5}$$

where $\{\pi_1, \pi_2, \ldots, \pi_N\}$ is the unique stationary distribution determined by the transition matrix P.

Proof: Let $\phi(n) = n$, $a_n = 0$, $f_n(x,y) = -\log p_n(x,y)$, and $\gamma = \frac{1}{2}$ in Theorem 2. Since

$$\begin{aligned}
&\sup_{i,k} \sum_{j\in S} f_k^2(i,j) e^{\gamma|f_k(i,j)|} p_k(i,j) \\
&\quad = \sup_{i,k} \sum_{j\in S} [-\log p_k(i,j)]^2 e^{(1/2)|-\log p_k(i,j)|} p_k(i,j) \\
&\quad = \sup_{i,k} \sum_{j\in S} [\log p_k(i,j)]^2 \sqrt{p_k(i,j)} < \infty,
\end{aligned}$$

so (2.9) holds. Let $g(i) = -\sum_j p(i,j) \log p(i,j)$ and $g_k(i) = -\sum_j p_k(i,j) \log p_k(i,j)$ in Theorem 2. By (4.4) and Lemma 2 of [13], we have

$$\lim_n \frac{1}{n} \sum_{k=1}^{n} |g_k(i) - g(i)| = 0, \quad \forall\ i, j \in S. \tag{4.6}$$

If the matrices, row vectors, and column vectors are all finite, then the convergence in norm is equivalent to the point-wise convergence. Hence, when $a_n = 0$ and $\phi(n) = n$, (4.4)

https://doi.org/10.1017/S0269964819000354 Published online by Cambridge University Press

is equivalent to (3.1) and (4.6) is equivalent to (3.17). Since P is finite and irreducible, so it is C-strongly ergodic. This corollary follows from Theorem 2. ■

Let $\{X_n,\ n \geq 0\}$ be a sequence of random variables taking values in state space $S = \{1, 2, \ldots\}$. Let $\{a_n,\ n \geq 0\}$ and $\{\phi(n),\ n \geq 0\}$ be two sequences of nonnegative integers such that $\phi(n)$ converges to infinite as $n \to \infty$. Let

$$S_{a_n,\phi(n)}(i, \omega) = \sum_{k=a_n+1}^{a_n+\phi(n)} I_i(X_k), \tag{4.7}$$

where $I_i(j) = \delta_{ji}$ is the Kronecker δ function. It is easy to see that $S_{a_n,\phi(n)}(i, \omega)$ is the number of occurrence of states i in $\{X_{a_n+1}, \ldots, X_{a_n+\phi(n)}\}$.

What follows is the strong law of large numbers for frequency of occurrence of states in $\{X_{a_n+1}, \ldots, X_{a_n+\phi(n)}\}$ for countable nonhomogeneous Markov chains and can not be implied in previous known results.

Corollary 6: *Let $\{X_n,\ n \geq 0\}$ be a nonhomogeneous Markov chain taking values in state space $S = \{1, 2, \ldots\}$ with the transition matrices (1.1). Let $\{a_n,\ n \geq 0\}$ and $\{\phi(n),\ n \geq 0\}$ be two sequences of nonnegative integers such that (2.2) holds. Let $P = (p(i, j))$ be a C-strongly ergodic transition matrix with respect to the constant transition matrix Q. Let $S_{a_n,\phi(n)}(i, \omega)$ be defined by (4.7). If (3.1) holds, then*

$$\lim_n \frac{S_{a_n,\phi(n)}(i, \omega)}{\phi(n)} = \pi_i \text{ a.e.}, \tag{4.8}$$

where $\pi = (\pi_1, \pi_2, \ldots)$ is the row vector of the constant transition matrix Q.

Proof: Letting $f_n(x, y) = I_i(y)$ ($\forall\ n$) in Theorem 2, it is easy to see that (2.9) holds for any $\gamma > 0$, and we have

$$\sum_{k=a_n+1}^{a_n+\phi(n)} f_k(X_{k-1}, X_k) = \sum_{k=a_n+1}^{a_n+\phi(n)} I_i(X_k) = S_{a_n,\phi(n)}(i, \omega), \tag{4.9}$$

and $g_n(l) = \sum_j I_i(j) p_n(l, j) = p_n(l, i)$. Let $g(l) = p(l, i)$. It is obvious that (3.1) implies (3.17) in this case. Since $P = (p(i, j))$ is a C-strongly ergodic transition matrix with respect to the constant transition matrix Q, we can easily prove that each row vector of the constant transition matrix Q is the unique stationary distribution determined by P. Hence we have

$$\sum_l g(l)\pi_l = \sum_l \pi_l p(l, i) = \pi_i. \tag{4.10}$$

By Theorem 2, (4.9) and (4.10), this corollary follows. ■

5. THE PROOFS

In this section, we will prove Lemma 4, Theorem 1, and Theorem 2.

Proof of Lemma 4: By (2.9), we have

$$\begin{aligned}
&\limsup_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} E[f_k^2(X_{k-1}, X_k) e^{\gamma |f_k(X_{k-1}, X_k)|} | X_{k-1}] \\
&= \limsup_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \sum_j f_k^2(X_{k-1}, j) e^{\gamma |f_k(X_{k-1}, j)|} p_k(X_{k-1}, j) \\
&\le M < \infty \text{ a.e.}
\end{aligned}$$

By Lemma 3, we have

$$\lim_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \{f_k(X_{k-1}, X_k) - E[f_k(X_{k-1}, X_k)|X_{k-1}]\} = 0 \text{ a.e.} \tag{5.1}$$

Let $g_k(x, y) = E[f_{k+1}(X_k, X_{k+1})|X_k = y]$ in Lemma 3 and $h(x) = x^2 e^{\gamma|x|}$. It is easy to prove that $h(x)$ is a convex function. By (2.9), Jensen inequality and smooth property for conditional expectation, we have

$$\begin{aligned}
&\limsup_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} E[g_k^2(X_{k-1}, X_k) e^{\gamma |g_k(X_{k-1}, X_k)|} | X_{k-1}] \\
&= \limsup_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} E\{(E[f_{k+1}(X_k, X_{k+1})|X_k])^2 e^{\gamma |E[f_{k+1}(X_k, X_{k+1})|X_k]|} | X_{k-1}\} \\
&= \limsup_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} E\{h(E[f_{k+1}(X_k, X_{k+1})|X_k])|X_{k-1}\} \\
&\le \limsup_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} E\{E[h(f_{k+1}(X_k, X_{k+1}))|X_k]|X_{k-1}\} \\
&= \limsup_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} E\{E[h(f_{k+1}(X_k, X_{k+1}))|X_0, \ldots, X_k]|X_0, \ldots, X_{k-1}\} \\
&= \limsup_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} E\{E[h(f_{k+1}(X_k, X_{k+1}))|X_0, \ldots, X_{k-1}]|X_0, \ldots, X_k\} \\
&= \limsup_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} E[h(f_{k+1}(X_k, X_{k+1}))|X_{k-1}] \\
&= \limsup_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} E[f_{k+1}^2(X_k, X_{k+1}) e^{\gamma |f_{k+1}(X_k, X_{k+1})|} | X_{k-1}] \\
&= \limsup_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \sum_i \sum_j f_{k+1}^2(i, j) e^{\gamma |f_{k+1}(i, j)|} p_{k+1}(i, j) p_k(X_{k-1}, i) \\
&\le M < \infty \text{ a.e.}
\end{aligned}$$

https://doi.org/10.1017/S0269964819000354 Published online by Cambridge University Press

By Lemma 3, we have

$$\lim_{n} \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \{E[f_{k+1}(X_k, X_{k+1})|X_k] - E[E[f_{k+1}(X_k, X_{k+1})|X_k]|X_{k-1}]\} = 0 \text{ a.e.} \tag{5.2}$$

Using (5.2) and smooth property for conditional expectation, we have

$$\lim_{n} \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \{E[f_{k+1}(X_k, X_{k+1})|X_k] - E[f_{k+1}(X_k, X_{k+1})|X_{k-1}]\} = 0 \text{ a.e.} \tag{5.3}$$

By (2.9) and Jensen inequality for the conditional expectation, for any m we have

$$\begin{aligned}
&(E[f_{a_n+m}(X_{a_n+m-1}, X_{a_n+m}) \mid X_{a_n+m-1}])^2 \\
&\quad \le E[f^2_{a_n+m}(X_{a_n+m-1}, X_{a_n+m}) \mid X_{a_n+m-1}] \\
&\quad = \sum_j f^2_{a_n+m}(X_{a_n+m-1}, j) \cdot p_{a_n+m-1}(X_{a_n+m-1}, j) \\
&\quad \le \sum_j f^2_{a_n+m}(X_{a_n+m-1}, j) \cdot e^{\gamma|f_{a_n+m}(X_{a_n+m-1}, j)|} \cdot p_{a_n+m-1}(X_{a_n+m-1}, j) \\
&\quad \le M < \infty.
\end{aligned}$$

So

$$\lim_{n} \frac{1}{\phi(n)} E[f_{a_n+m}(X_{a_n+m-1}, X_{a_n+m}) \mid X_{a_n+m-1}] = 0 \text{ a.e.} \tag{5.4}$$

Based on (5.1) and (5.4), we have

$$\lim_{n} \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \{f_k(X_{k-1}, X_k) - E[f_{k+1}(X_k, X_{k+1})|X_k]\} = 0 \text{ a.e.} \tag{5.5}$$

By (5.5) and (5.3), we have

$$\lim_{n} \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \{f_k(X_{k-1}, X_k) - E[f_{k+1}(X_k, X_{k+1})|X_{k-1}]\} = 0 \text{ a.e.} \tag{5.6}$$

By induction, (2.10) holds for all positive integer t. ∎

Proof of Theorem 1: Since $\{\phi(n),\ n \ge 0\}$ satisfies (2.2), using (3.1) and Lemma 2, for any positive integer m we have

$$\lim_{n} \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \|P_{m+k} - P\| = 0. \tag{5.7}$$

Since

$$\begin{aligned}
&\|P^{(m+k,m+k+2)} - P^2\| \\
&\quad \le \|P_{m+k+1}P_{m+k+2} - P_{m+k+1}P\| + \|P_{m+k+1}P - P^2\| \\
&\quad \le \|P_{m+k+2} - P\| + \|P_{m+k+1} - P\|,
\end{aligned} \tag{5.8}$$

https://doi.org/10.1017/S0269964819000354 Published online by Cambridge University Press

by (5.7) and (5.8) we have for any positive integer m,

$$\lim_n \frac{1}{\phi(n)} \sum_{k=a_n+1}^{a_n+\phi(n)} \|P^{(m+k,\ m+k+2)} - P^2\| = 0. \tag{5.9}$$

By induction, (3.2) holds. Next we will prove (3.3).

Let $L = [\phi(n)/d]$, (i.e., $\phi(n) = Ld + r,\ 0 \le r < d$), where $[x]$ is the largest integer less than x. Since

$$\begin{aligned}
&\sum_{t=a_n+1}^{a_n+\phi(n)} P^{(m,m+t)} - \phi(n)Q \\
&= \sum_{t=a_n+1}^{a_n+Jd+r} (P^{(m,m+t)} - Q) + \sum_{t=a_n+Jd+r+1}^{a_n+\phi(n)} (P^{(m,\ m+t)} - Q) \\
&= \sum_{t=a_n+1}^{a_n+Jd+r} (P^{(m,m+t)} - Q) + \sum_{j=J}^{L-1} \sum_{k=1}^{d} (P^{(m,\ m+a_n+jd+k+r)} - Q),
\end{aligned} \tag{5.10}$$

where $J<L$, then

$$\begin{aligned}
\left\| \sum_{t=a_n+1}^{a_n+\phi(n)} P^{(m,m+t)} - \phi(n)Q \right\| &\le \left\| \sum_{t=a_n+1}^{a_n+Jd+r} (P^{(m,\ m+t)} - Q) \right\| \\
&\quad + \sum_{j=J}^{L-1} \left\| \sum_{k=1}^{d} P^{(m,\ m+a_n+jd+k+r)} - dQ \right\|.
\end{aligned} \tag{5.11}$$

Let $M \le Jd$,

$$\begin{aligned}
&\sum_{j=J}^{L-1} \left\| \sum_{k=1}^{d} P^{(m,\ m+a_n+jd+k+r)} - dQ \right\| \\
&= \sum_{j=J}^{L-1} \left\| P^{(m,\ m+a_n+jd+r+1-M)} \left(\sum_{k=1}^{d} P^{(m+a_n+jd+r+1-M,\ m+a_n+jd+r+k)} - dQ \right) \right\| \\
&= \sum_{j=J}^{L-1} \left\| P^{(m,\ m+a_n+jd+r+1-M)} \left[\sum_{k=1}^{d} (P^{(m+a_n+jd+r+1-M,\ m+a_n+jd+r+k)} - P^{k+M-1}) \right] \right. \\
&\quad \left. + P^{(m,\ m+a_n+jd+r+1-M)} \left[\sum_{k=1}^{d} P^{k+M-1} - dQ \right] \right\| \\
&\le \sum_{j=J}^{L-1} \left[\left\| \sum_{k=1}^{d} (P^{(m+a_n+jd+r+1-M,\ m+a_n+jd+r+k)} - P^{k+M-1}) \right\| + \left\| \sum_{k=1}^{d} P^{k+M-1} - dQ \right\| \right] \\
&\le \sum_{j=J}^{L-1} \sum_{k=1}^{d} \|P^{(m+a_n+jd+r+1-M,\ m+a_n+jd+r+k)} - P^{k+M-1}\| \\
&\quad + (L-J) \left\| \sum_{k=1}^{d} P^{M+k-1} - dQ \right\|.
\end{aligned} \tag{5.12}$$

https://doi.org/10.1017/S0269964819000354 Published online by Cambridge University Press

So

$$
\begin{aligned}
&\left\| \frac{1}{\phi(n)} \sum_{t=a_n+1}^{a_n+\phi(n)} P^{(m,\ m+t)} - Q \right\| \\
&\quad \le \frac{1}{\phi(n)} \left\| \sum_{t=a_n+1}^{a_n+Jd+r} (P^{(m,\ m+t)} - Q) \right\| \\
&\qquad + \frac{1}{\phi(n)} \sum_{j=J}^{L-1} \sum_{k=1}^{d} \| P^{(m+a_n+jd+r+1-M,\ m+a_n+jd+r+k)} - P^{k+M-1} \| \\
&\qquad + \frac{L-J}{\phi(n)} \left\| \sum_{k=1}^{d} P^{M+k-1} - dQ \right\| \\
&\quad \le \frac{2(Jd+r)}{\phi(n)} + \sum_{k=1}^{d} \frac{1}{\phi(n)} \sum_{v=a_n+1}^{a_n+\phi(n)} \| P^{(m+v+1-M,\ m+v+k)} - P^{k+M-1} \| \\
&\qquad + \frac{L-J}{\phi(n)} \left\| \sum_{k=1}^{d} P^{M+k-1} - dQ \right\|. \qquad \textbf{(5.13)}
\end{aligned}
$$

From Lemma 1, $\forall\, \varepsilon > 0$, and choosing sufficient large M, we have

$$
\frac{L-J}{\phi(n)} \left\| \sum_{k=1}^{d} P^{M+k-1} - dQ \right\| < \varepsilon. \qquad \textbf{(5.14)}
$$

As $n \to \infty$, the first term of (5.13) converges to zero. By (3.2), the second term of (5.13) also converges to zero. By (3.2), (5.13), (5.14) and the arbitrariness of ϵ, (3.3) follows. Similarly, (3.5) follows from (3.4). ■

Proof of Theorem 2: Obviously (2.2) implies (2.6). By Lemma 4, (2.10) holds. Now by (4.1)

$$
\begin{aligned}
&E[f_{k+t}(X_{k+t-1}, X_{k+t}) \mid X_{k-1}] \\
&\quad = \sum_i \sum_j f_{k+t}(i,j) P(X_{k+t-1} = i, X_{k+t} = j \mid X_{k-1}) \\
&\quad = \sum_i \sum_j f_{k+t}(i,j) P(X_{k+t} = j \mid X_{k+t-1} = i) P(X_{k+t-1} = i \mid X_{k-1}) \\
&\quad = \sum_i \sum_j f_{k+t}(i,j) p_{k+t}(i,j) p^{(k-1,k+t-1)}(X_{k-1}, i) \\
&\quad = \sum_i g_{k+t}(i) p^{(k-1,k+t-1)}(X_{k-1}, i). \qquad \textbf{(5.15)}
\end{aligned}
$$

https://doi.org/10.1017/S0269964819000354 Published online by Cambridge University Press

Letting the elements of P^t be $p^{(t)}(i,j)$, we obtain

$$\begin{aligned}
&\left|\frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)} E[f_{k+t}(X_{k+t-1},X_{k+t}) \mid X_{k-1}] - \frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\sum_i g(i)p^{(t)}(X_{k-1},i)\right| \\
&= \left|\frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\sum_i g_{k+t}(i)p^{(k-1,k+t-1)}(X_{k-1},i) - \frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\sum_i g(i)p^{(t)}(X_{k-1},i)\right| \\
&\le \frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\left|\sum_i g_{k+t}(i)p^{(k-1,k+t-1)}(X_{k-1},i) - \sum_i g(i)p^{(k-1,k+t-1)}(X_{k-1},i)\right| \\
&\quad + \frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\left|\sum_i g(i)p^{(k-1,k+t-1)}(X_{k-1},i) - \sum_i g(i)p^{(t)}(X_{k-1},i)\right| \\
&\le \frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\left|\sum_i p^{(k-1,k+t-1)}(X_{k-1},i)(g_{k+t}(i)-g(i))\right| \\
&\quad + \frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\sup_i |g(i)|\cdot\sup_l\sum_i |p^{(k-1,k+t-1)}(l,i)-p^{(t)}(l,i)| \\
&\le \frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\| g_{k+t}-g \| + \frac{\|g\|}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\|P^{(k-1,k+t-1)}-P^t\|. \qquad \textbf{(5.16)}
\end{aligned}$$

By (3.1) and Theorem 1, we have

$$\lim_n \frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\|P^{(k-1,k+t-1)}-P^t\| = 0. \qquad \textbf{(5.17)}$$

By (3.17) and Lemma 2, we have

$$\lim_n \frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\|g_{k+t}-g\| = 0. \qquad \textbf{(5.18)}$$

Combining (5.16), (5.17), (5.18), and (2.10), for any t we have

$$\lim_n \frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\left\{f_k(X_{k-1},X_k) - \sum_i g(i)p^{(t)}(X_{k-1},i)\right\} = 0 \text{ a.e.} \qquad \textbf{(5.19)}$$

By (5.19), for any N we have

$$\lim_n \left\{\frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)} f_k(X_{k-1},X_k) - \frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\frac{1}{N}\sum_{t=1}^{N}\sum_i g(i)p^{(t)}(X_{k-1},i)\right\} = 0 \text{ a.e.} \qquad \textbf{(5.20)}$$

Since

$$\begin{aligned}&\left|\frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\frac{1}{N}\sum_{t=1}^{N}\sum_i g(i)p^{(t)}(X_{k-1},i)-\sum_i g(i)\pi_i\right|\\&\le\frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\sum_i\left|\frac{1}{N}\sum_{t=1}^{N}g(i)(p^{(t)}(X_{k-1},i)-\pi_i)\right|\\&\le\frac{1}{\phi(n)}\sum_{k=a_n+1}^{a_n+\phi(n)}\sup_i|g(i)|\cdot\sup_l\sum_i\left|\frac{1}{N}\sum_{t=1}^{N}p^{(t)}(l,i)-\pi_i\right|\\&=\|g\|\cdot\left\|\frac{1}{N}\sum_{t=1}^{N}P^t-Q\right\|,\end{aligned}\tag{5.21}$$

and P is a C-strongly ergodic transition matrix with the constant transition matrix Q, thus we have

$$\lim_N\left\|\frac{1}{N}\sum_{t=1}^{N}P^t-Q\right\|=0.\tag{5.22}$$

(4.2) follows from (5.20), (5.21), and (5.22). ■

Acknowledgments

This work was supported by the National Natural Science Foundations of China 11571142,11601191. Youth talent cultivation project of Jiangsu University, Young science and technology talents lifting project of Jiangsu association for science and technology.

References

1. Bowerman, B., David, H.T., & Isaacson, D. (1977). The convergence of Cesàro averages for certain nonstationary Markov chains. *Stochastic Processes and their Applications* 5: 221–230.
2. Chiang, T.S. & Chow, Y.S. (1989). A limit theorem for a class of inhomogeneous Markov processes. *Annals of Probability* 17: 1483–1502.
3. Cover, T.M. & Thomas, J.A. (1991). *Elements of Information Theory*. New York: Wiley.
4. Dietz, Z. & Sethuraman, S. (2005). Large deviations for a class of nonhomogeneous Markov chains. *Annals of Applied Probability* 15: 421–486.
5. Isaacson, D. & Madsen, R. (1976). *Markov Chains Theory and Application*. New York: Wiley.
6. Isaacson, D. & Senta, E. (1982). Ergodicity for countable inhomogeneous Markov chains. *Linear Algebra and its Applications* 48: 37–44.
7. Liu, G.X. & Liu, W. (1994). On the strong law of large numbers for functionals of countable nonhomogeneous Markov chains. *Stochastic Processes and their Applications* 50: 375–391.
8. Liu, W. & Yang, W.G. (1996). An extension of Shannon-McMillan theorem and some limit properties for nonhomogeneous Markov chains. *Stochastic Processes and their Applications* 61: 129–145.
9. Ma, H.L. & Yang, W.G. (2011). Erratum to “The asymptotic equipartition property for asymptotic circular Markov chains”. *Probability in the Engineering and Informational Sciences* 25(2): 265–267.
10. Mukhamedov, F. (2013). The Dobrushin ergodicity coefficient and the ergodicity of noncommutative Markov chains. *Journal of Mathematical Analysis and Applications* 408: 364–373.
11. Rosenblatt-Roth, M. (1964). Some theorems concerning the strong law of large numbers for non-homogeneous Markov chains. *The Annals of Mathematical Statistics* 35: 566–576.
12. Tan, C.P. (1996). On the weak ergodicity of nonhomogeneous Markov chains. *Statistics & Probability Letters* 26: 293–295.
13. Wang, Z.Z. & Yang, W.G. (2016). The generalized entropy ergodic theorem for nonhomogeneous Markov Chains. *Journal of Theoretical Probability* 29: 761–775.

14. Waqas, A. & Yang, W.G. (2019). Generalized absolute mean strong ergodicity of nonhomogeneous Markov Chains. *Pakistan Journal of Statistics* 35: 53–63.
15. Yang, W.G. (1998). The asymptotic equipartition property for a nonhomogeneous Markov information source. *Probability in the Engineering and Informational Sciences* 12: 509–518.
16. Yang, W.G. (2002). Convergence in the Cesáro sense and strong law of large numbers for countable nonhomogeneous Markov chains. *Linear Algebra and its Applications* 354: 275–286.
17. Yang, W.G. (2009). Strong law of large numbers for countable nonhomogeneous Markov chains. *Linear Algebra and its Applications* 430: 3008–3018.
18. Yang, W.G., Wang, B., & Shi, Z.Y. (2014). Strong law of large numbers for countable asymptotic circular Markov Chains. *Communications in Statistics – Theory and Methods* 43(18): 3943–3954.
19. Zhong, P.P., Yang, W.G., & Liang, P.P. (2010). The asymptotic equipartition property for asymptotic circular Markov chains. *Probability in the Engineering and Informational Sciences* 24: 279–288.