frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 17 March 2023
DOI 10.3389/fphy.2023.1156610

OPEN ACCESS

EDITED BY
Song Zheng,
Zhejiang University of Finance and Economics, China

REVIEWED BY
Weigang Sun,
Hangzhou Dianzi University, China
Zhiyan Shi,
Jiangsu University, China
Huilin Huang,
Wenzhou University, China

*CORRESPONDENCE
Zhaobiao Rui,
ruizhaobiao@chu.edu.cn

SPECIALTY SECTION
This article was submitted to Statistical and Computational Physics, a section of the journal Frontiers in Physics

RECEIVED 01 February 2023
ACCEPTED 27 February 2023
PUBLISHED 17 March 2023

CITATION
Qin X, Rui Z, Chen S and Peng W (2023), Small deviation properties concerning arrays of non-homogeneous Markov information sources.
Front. Phys. 11:1156610.
doi: 10.3389/fphy.2023.1156610

Small deviation properties concerning arrays of non-homogeneous Markov information sources

Ximei Qin[1], Zhaobiao Rui[1]*, Shu Chen[1,2] and Weicai Peng[1]

[1]School of Mathematics and Big Data, Chaohu University, Hefei, China, [2]School of Microelectronics and Data Science, Anhui University of Technology, Ma'anshan, China

In this study, we first define the logarithmic likelihood ratio as a measure between arbitrary generalized information sources and non-homogeneous Markov sources and then establish a class of generalized information sources for small deviation theorems, strong limit theorems, and the asymptotic equipartition property. The present outcomes generalize some existing results.

KEYWORDS

non-homogeneous Markov chains, generalized information sources, small deviation properties, general relative entropy, asymptotic equipartition property

1 Introduction

In information theory, the asymptotic equipartition property (abbreviated as AEP) is a type of property of random sources. It is the basis of the typical set concept used in data compression theory. The AEP is the constant convergence of certain random processes in some types of convergence, such as $\mathcal{L}_1$ convergence, probability convergence, and almost surely convergence. In some circumstances, it is also known as the Shannon–McMillan–Breiman theorem or entropy ergodic theorem. Small deviation properties are types of strong limit theorems (i.e., in the sense of almost everywhere convergence) presented by inequalities for information sources, and they usually involve the generalization of the strong law of large numbers (SLLNs). In this paper, following the research [1-3], we mainly consider the strong limiting properties and small deviation properties of generalized information entropy (a type of array of non-homogeneous Markov chains), which is an important issue in the study of limit theory.

In 1948, Shannon first explored the AEP of i.i.d. sequences (i.e., independent identically distributed sequences) and the entropy ergodic theorem of convergence in the sense of probability (see [4]). Then, in the 1950s, McMillan and Breiman established the AEP for certain types of information sources in the sense of $\mathcal{L}_1$ and almost everywhere (abbreviated as a.e.) convergence, respectively (see [5-6]). In 1960, Chung relaxed the conditions and found that the AEP still holds for random sources equipped with countable state (see [7]). From the 1970s to the early stages of the 21st century, the AEP for various general stochastic processes was investigated by many studies, such as [8–13]. Recently, many scholars, such as Yang (e.g., [14–17]), Shi (e.g., [3,18–21]), Huang [22,23], and Peng [24–26], by generalizing the method proposed by [27], [11], and Wang [2,28,29], studied the AEP and the limit properties (including AEP and SLLNs) of some types of Markov chains (such as homogeneous and non-homogeneous; finite state space and infinite state space; and Markov chains indexed by the set of positive integers and tree-indexed Markov chains).

However, most of the aforementioned results do not consider arrays of information sources, which play significant roles in information science. In recent works, [30] explored the conditions and SLLNs for almost certain convergence of double random variable arrays, and [31] established several kinds of convergences for row negatively correlated random variable arrays under certain conditions. More related studies can be found in their references. Therefore, the limit behavior and the AEP, as well as the small deviation properties of the arrays of information sources, aroused our interest. This paper, in line with [3], [30], and [31], first introduces the logarithmic likelihood ratio as a measure between arbitrary generalized information sources and non-homogeneous Markov sources and then establishes a class of generalized information sources for small deviation theorems and strong limit theorems. The outcomes generalize some existing results.

The rest of the content is arranged as follows: Section 2, the preliminaries, gives some notations and establishes some definitions and lemmas. Section 3 states the main results and presents the strong limit behaviors and strong deviation properties of non-homogeneous Markov sources.

2 Preliminaries

In this section, we first introduce several notation and then establish the definition of the generalized divergence rate distance of the arbitrary measure μ with respect to the Markov measure $\tilde{\mu}$. In the rest of this section, the probability space $(\Omega, \mathcal{F}, \mu)$ that we explore in our main results is fixed. Let $\xi = \{\xi_i^{(n)}, v_n \leqslant i \leqslant u_n\}_{n\in\mathbb{N}^+}$ be a general information source, where $\{\xi_i^{(n)}, v_n \leqslant i \leqslant u_n\}$ are an array of non-negative integer-valued random variables over the $(u_n - v_n + 1)$th Cartesian product $\mathcal{X}_{v_n} \times \mathcal{X}_{v_n+1} \times\cdots\times \mathcal{X}_{u_n}$ of an arbitrary discrete source alphabet $\mathcal{X}\,(\mathcal{X} = \{s_{n1}, s_{n2}, \cdots\}, n \in \mathbb{N}^+)$ with the distribution

$$p_n\left(x_{v_n}^{(n)}, \ldots, x_{u_n}^{(n)}\right) = \mu\left(\xi_{v_n}^{(n)} = x_{v_n}^{(n)}, \xi_{v_n+1}^{(n)} = x_{v_n+1}^{(n)}, \xi_{v_n+2}^{(n)} = x_{v_n+2}^{(n)}, \ \xi_{v_n+3}^{(n)} = x_{v_n+3}^{(n)}, \ldots, \xi_{u_n}^{(n)} = x_{u_n}^{(n)}\right) > 0,$$

where $x_i^{(n)} \in \mathcal{X}, \ v_n \leqslant i \leqslant u_n, \ n \in \mathbb{N}^+$ and $\{(v_n, u_n): v_n, u_n \in \mathbb{Z}, -\infty < v_n < u_n < +\infty\}$, $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$, $\mathbb{N}^+ = \{1, 2, \cdots\}$.

For any arbitrary information source, $\xi = \{\xi^n = (\xi_{v_n}^{(n)}, \ldots, \xi_{u_n}^{(n)})\}_{n\in\mathbb{N}^+}$ denote $p^{(n)}\left(x_{v_n}^{(n)}, \ldots, x_{u_n}^{(n)}\right) = \mu\left(\xi_{v_n}^{(n)} = x_{v_n}^{(n)}, \ldots, \xi_{u_n}^{(n)} = x_{u_n}^{(n)}\right)$. Let

$$f_n(\omega) = -\frac{1}{n}\log p^{(n)}\left(\xi_{v_n}^{(n)}, \ldots, \xi_{u_n}^{(n)}\right),$$

which is called the entropy density of $p^{(n)}\left(\xi_{v_n}^{(n)}, \ldots, \xi_{u_n}^{(n)}\right)$.

Supposing that $\tilde{\mu}$ is a non-homogeneous Markov information source, then there exists a distribution $\{q^{(n)}(1), q^{(n)}(2)\cdots q^{(n)}(n), q^{(n)}(i) > 0, v_n \leqslant i \leqslant u_n\}$ and a transition probability density $\left\{p_i^{(n)}(x, y), v_n \leqslant i \leqslant u_n\right\}$, which is called the nth step transition probability density, such that

$$q^{(n)}\left(x_{v_n}^{(n)}, \ldots, x_{u_n}^{(n)}\right) = q^{(n)}\left(x_{v_n}^{(n)}\right) \prod_{i=v_n+1}^{u_n} p_i^{(n)}\left(x_{i-1}^{(n)}, x_i^{(n)}\right),$$
$$x_i^{(n)} \in \mathcal{X}, \ v_n \leqslant i \leqslant u_n, \ n \in \mathbb{N}^+$$

and

$$-\frac{1}{n}\log q^{(n)}\left(\xi_{v_n}^{(n)}, \ldots, \xi_{u_n}^{(n)}\right) = -\left[\frac{1}{n}\log q^{(n)}\left(\xi_{v_n}^{(n)}\right) + \frac{1}{n}\sum_{i=v_n+1}^{u_n} \log p_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)\right].$$

Definition 2.1: Defining

$$H(\mu\|\tilde{\mu}) := -\liminf_n \frac{1}{n}\log\left[\frac{q^{(n)}\left(\xi_{v_n}^{(n)}\right)\prod_{i=v_n+1}^{u_n} p_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)}{p^{(n)}\left(\xi_{v_n}^{(n)}, \ldots, \xi_{u_n}^{(n)}\right)}\right].$$

Here, $H(\mu\|\tilde{\mu})$ is called the generalized divergence rate distance of the arbitrary measure μ with respect to the Markov measure $\tilde{\mu}$.

We use log to represent the logarithm operator. Let $0\log\ 0 = 0$, which can be verified since $x\log\ x \to 0$ as $x \to 0$.

Lemma 2.1: [27] Let $\{\xi_n\}_{n\in\mathbb{N}^+}$ be a sequence of non-negative random variables with $\mathbb{E}[\xi_n] \leqslant 1$, then

$$\limsup_n \frac{1}{n}\log\xi_n \leqslant 0 \qquad a.s.$$

The proof of Lemma 2.1 can be found in [27], which is omitted in this study.

3 Main results and proofs

In this section, we first derive the strong deviation theorem (Theorem 3.1) for a sequence of measurable functions defined on $\mathbb{N}^2$ under certain conditions. Then, by considering the special case with $c = 0$ in Theorem 3.1, we derive the strong law of large numbers for strong ergodic information sources (Theorem 3.2). Finally, we obtain the small deviation behavior (Theorem 3.3) and the asymptotic property of the entropy density $f_n(\omega)$ (Corollary 3.1).

Theorem 3.1: Let $f_n(\omega)$ and $H(\mu\|\tilde{\mu})$ be as given in Definition 2.1, $f_i^{(n)}(x, y)$ be a sequence of measurable functions defined on $\mathbb{N}^2$, and $\tilde{\xi}_{(i-1,i)}^{(n)} = f_i^{(n)}(\xi_{i-1}^{(n)}, \xi_i^{(n)})$. Let $c > 0$ be a real-valued constant and

$$\mathcal{D}(\omega) = \{\omega: H(\mu\|\tilde{\mu}) \leqslant c\}. \tag{3.1}$$

Supposing that there exists $\alpha > 0$, for each $v_n \leqslant i \leqslant u_n$

$$\mathbb{E}_{\tilde{\mu}}\left[e^{\alpha|f_i^{(n)}(\xi_{i-1}^{(n)}, \xi_i^{(n)})|}\right] < \infty, \tag{3.2}$$

and for arbitrary $v_n \leqslant i \leqslant u_n$,

$$b_\alpha = \mathbb{E}_{\tilde{\mu}}\left\{\left[f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)\right]^2 e^{\alpha|f_i^{(n)}(\xi_{i-1}^{(n)}, \xi_i^{(n)})|} | \xi_{i-1}^{(n)} = k\right\} \leqslant \tau. \tag{3.3}$$

Let

$$H_t(\alpha, \tau) = \frac{2\tau}{e^2(t-\alpha)^2}, \tag{3.4}$$

where $t \in (0, \alpha)$. Then, in the case of $0 \leqslant c \leqslant t^2 H_t(\alpha, \tau)$, it can be found that

$$\limsup_n \frac{1}{n}\sum_{i=v_n+1}^{u_n}\left\{f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) - \mathbb{E}_{\tilde{\mu}}\left[f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) | \xi_{i-1}^{(n)}\right]\right\} \leqslant 2\sqrt{cH_t(\alpha, \tau)} \quad a.s. \tag{3.5}$$

and

$$\liminf_{n} \frac{1}{n} \sum_{i=v_n+1}^{u_n} \left\{ f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) - \mathbb{E}_{\tilde{\mu}}\left[f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) | \xi_{i-1}^{(n)} \right] \right\} \geqslant -2\sqrt{cH_t(\alpha,\tau)} \quad a.s. \tag{3.6}$$

Note: In Eq. 3.1 of Theorem 3.1, $\mathcal{D}(\omega)$ defines the range of the generalized divergence rate distance of the arbitrary measure μ with respect to the Markov measure $\tilde{\mu}$. It measures the difference between arbitrary generalized information sources and non-homogeneous Markov sources. In the rest of the content, we omit ω for notation explicitly. Equation 3.2 states the restriction that the array $f_i^{(n)}(\xi_{i-1}^{(n)}, \xi_i^{(n)})$ is integrable in the exponential sense. Equation 3.3 gives the moment condition for the conditional mathematical expectation of the array.

Proof: Let λ be a negative constant and

$$g^{(n)}\left(x_{v_n}^{(n)}, \ldots, x_{u_n}^{(n)}\right) = q^{(n)}\left(x_{v_n}^{(n)}\right) \prod_{i=v_n+1}^{u_n} \frac{e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} \cdot p_i^{(n)}\left(x_{i-1}^{(n)}, x_i^{(n)}\right)}{\mathbb{E}_{\tilde{\mu}}\left[e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} | \xi_{i-1}^{(n)} = x_{i-1}^{(n)} \right]}.$$

Let

$$\Lambda_n(\lambda,\omega) = \frac{g^{(n)}\left(x_{v_n}^{(n)}, \ldots, x_{u_n}^{(n)}\right)}{p^{(n)}\left(x_{v_n}^{(n)}, \ldots, x_{u_n}^{(n)}\right)} = \frac{q^{(n)}\left(x_{v_n}^{(n)}\right) \prod\limits_{i=v_n+1}^{u_n} \frac{e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} \cdot p_i^{(n)}\left(x_{i-1}^{(n)}, x_i^{(n)}\right)}{\mathbb{E}_{\tilde{\mu}}\left[e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} | \xi_{i-1}^{(n)} = x_{i-1}^{(n)} \right]}}{p^{(n)}\left(x_{v_n}^{(n)}, \ldots, x_{u_n}^{(n)}\right)}, \tag{3.7}$$

then

$$\begin{aligned}
\mathbb{E}[\Lambda_n(\lambda,\omega)] &= \sum_{x_{v_n}^{(n)} \cdots x_{u_n}^{(n)}} \frac{q^{(n)}\left(x_{v_n}^{(n)}\right) \prod\limits_{i=v_n+1}^{u_n} \frac{e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} \cdot p_i^{(n)}\left(x_{i-1}^{(n)}, x_i^{(n)}\right)}{\mathbb{E}_{\tilde{\mu}}\left[e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} | \xi_{i-1}^{(n)} = x_{i-1}^{(n)} \right]}}{p^{(n)}\left(x_{v_n}^{(n)}, \ldots, x_{u_n}^{(n)}\right)} \cdot p^{(n)}\left(x_{v_n}^{(n)}, \ldots, x_{u_n}^{(n)}\right) \\
&= \sum_{x_{v_n}^{(n)} \cdots x_{u_n}^{(n)}} q^{(n)}\left(x_{v_n}^{(n)}\right) \prod_{i=v_n+1}^{u_n} \frac{e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} \cdot p_i^{(n)}\left(x_{i-1}^{(n)}, x_i^{(n)}\right)}{\mathbb{E}_{\tilde{\mu}}\left[e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} | \xi_{i-1}^{(n)} = x_{i-1}^{(n)} \right]} \\
&= \sum_{x_{v_n}^{(n)}} q^{(n)}\left(x_{v_n}^{(n)}\right) \sum_{x_{v_n+1}^{(n)} \cdots x_{u_n}^{(n)}} \prod_{i=v_n+1}^{u_n} \frac{e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} \cdot p_i^{(n)}\left(x_{i-1}^{(n)}, x_i^{(n)}\right)}{\mathbb{E}_{\tilde{\mu}}\left[e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} | \xi_{i-1}^{(n)} = x_{i-1}^{(n)} \right]} \\
&= \sum_{x_{v_n}^{(n)}} q^{(n)}\left(x_{v_n}^{(n)}\right) \sum_{x_{v_n+1}^{(n)} \cdots x_{u_n}^{(n)}} \frac{e^{\lambda f_{v_n+1}^{(n)}\left(\xi_{v_n}^{(n)}, \xi_{v_n+1}^{(n)}\right)} \cdot p_{v_n+1}^{(n)}\left(x_{v_n}^{(n)}, x_{v_n+1}^{(n)}\right)}{\mathbb{E}_{\tilde{\mu}}\left[e^{\lambda f_{v_n+1}^{(n)}\left(\xi_{v_n}^{(n)}, \xi_{v_n+1}^{(n)}\right)} | \xi_{v_n}^{(n)} = x_{v_n}^{(n)} \right]} \cdots \frac{e^{\lambda f_{u_n}^{(n)}\left(\xi_{u_n-1}^{(n)}, \xi_{u_n}^{(n)}\right)} \cdot p_{u_n}^{(n)}\left(x_{u_n-1}^{(n)}, x_{u_n}^{(n)}\right)}{\mathbb{E}_{\tilde{\mu}}\left[e^{\lambda f_{u_n}^{(n)}\left(\xi_{u_n-1}^{(n)}, \xi_{u_n}^{(n)}\right)} | \xi_{u_n-1}^{(n)} = x_{u_n-1}^{(n)} \right]} \\
&= \sum_{x_{v_n}^{(n)}} q^{(n)}\left(x_{v_n}^{(n)}\right) \sum_{x_{v_n+1}^{(n)} \cdots x_{u_n}^{(n)}} \frac{e^{\lambda f_{v_n+1}^{(n)}\left(\xi_{v_n}^{(n)}, \xi_{v_n+1}^{(n)}\right)} \cdot p_{v_n+1}^{(n)}\left(x_{v_n}^{(n)}, x_{v_n+1}^{(n)}\right)}{e^{\lambda f_{v_n+1}^{(n)}\left(\xi_{v_n}^{(n)}, \xi_{v_n+1}^{(n)}\right)} \cdot p_{v_n+1}^{(n)}\left(x_{v_n}^{(n)}, x_{v_n+1}^{(n)}\right)} \cdots \frac{e^{\lambda f_{u_n}^{(n)}\left(\xi_{u_n-1}^{(n)}, \xi_{u_n}^{(n)}\right)} \cdot p_{u_n}^{(n)}\left(x_{u_n-1}^{(n)}, x_{u_n}^{(n)}\right)}{e^{\lambda f_{u_n}^{(n)}\left(\xi_{u_n-1}^{(n)}, \xi_{u_n}^{(n)}\right)} \cdot p_{u_n}^{(n)}\left(x_{u_n-1}^{(n)}, x_{u_n}^{(n)}\right)} \\
&= \sum_{x_{v_n}^{(n)}} q^{(n)}\left(x_{v_n}^{(n)}\right) = 1.
\end{aligned}$$

Combining Lemma 2.1, we can obtain

$$\limsup_{n} \frac{1}{n} \log \Lambda_n(\lambda,\omega) \leqslant 0 \quad a.s. \tag{3.8}$$

With Eqs 3.1, 3.7, we have

$$\begin{aligned}
\frac{1}{n}\log\Lambda_n(\lambda,\omega) = \frac{1}{n}\sum_{i=v_n+1}^{u_n} \lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) &- \frac{1}{n}\sum_{i=v_n+1}^{u_n} \log \mathbb{E}_{\tilde{\mu}}\left[e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} | \xi_{i-1}^{(n)} \right] \\
&+ \frac{1}{n}\log \frac{q^{(n)}\left(x_{v_n}^{(n)}\right) \prod\limits_{i=v_n+1}^{u_n} p_i^{(n)}\left(x_{i-1}^{(n)}, x_i^{(n)}\right)}{p^{(n)}\left(x_{v_n}^{(n)}, \ldots, x_{u_n}^{(n)}\right)}.
\end{aligned} \tag{3.9}$$

With Eqs 3.8, 3.9, we have

$$\begin{aligned}
&\limsup_{n} \left\{ \frac{1}{n}\sum_{i=v_n+1}^{u_n} \lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) - \frac{1}{n}\sum_{i=v_n+1}^{u_n} \log \mathbb{E}_{\tilde{\mu}}\left[e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} | \xi_{i-1}^{(n)} \right] \right\} \\
&\leqslant -\liminf_{n} \frac{1}{n} \log \frac{q^{(n)}\left(x_{v_n}^{(n)}\right) \prod\limits_{i=v_n+1}^{u_n} p_i^{(n)}\left(x_{i-1}^{(n)}, x_i^{(n)}\right)}{p^{(n)}\left(x_{v_n}^{(n)}, \ldots, x_{u_n}^{(n)}\right)} \leqslant H(\mu \| \tilde{\mu}).
\end{aligned} \tag{3.10}$$

Hence, with Eq. 3.10, we have

$$\begin{aligned}
&\limsup_{n} \frac{\lambda}{n} \sum_{i=v_n+1}^{u_n} \left\{ f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) - \mathbb{E}_{\tilde{\mu}}\left[f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) | \xi_{i-1}^{(n)} \right] \right\} \\
&\leqslant \limsup_{n} \frac{1}{n} \sum_{i=v_n+1}^{u_n} \left\{ \log \mathbb{E}_{\tilde{\mu}}\left[e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} | \xi_{i-1}^{(n)} \right] - \mathbb{E}_{\tilde{\mu}}\left[\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) | \xi_{i-1}^{(n)} \right] \right\} + H(\mu \| \tilde{\mu}).
\end{aligned} \tag{3.11}$$

We consider that the maximum of $x^2 e^{-hx}$ is $\frac{4e^{-2}}{h^2}$ with $(h > 0)$. Hereafter, we restrict the analysis to $0 < \lambda < t$ and $0 < t < \alpha$. According to the inequality $2 - 1/x \leqslant 1 + \log x \leqslant x$ and $e^x - 1 \leqslant x + x^2 e^{|x|}/2$ with $x > 0$, the properties of the superior limit, and Eq. 3.4, we have

$$\begin{aligned}
&\limsup_{n} \frac{1}{n} \sum_{i=v_n+1}^{u_n} \left\{ \log\left[\mathbb{E}_{\tilde{\mu}} e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} | \xi_{i-1}^{(n)} - \mathbb{E}_{\tilde{\mu}}\left[\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) | \xi_{i-1}^{(n)} \right] \right] \right\} \\
&\leqslant \limsup_{n} \frac{1}{n} \sum_{i=v_n+1}^{u_n} \left\{ \mathbb{E}_{\tilde{\mu}}\left[e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} | \xi_{i-1}^{(n)} \right] - 1 - \mathbb{E}_{\tilde{\mu}}\left[\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) | \xi_{i-1}^{(n)} \right] \right\} \\
&= \limsup_{n} \frac{1}{n} \sum_{i=v_n+1}^{u_n} \mathbb{E}_{\tilde{\mu}}\left\{ \left[e^{\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)} - 1 - \lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) \right] | \xi_{i-1}^{(n)} \right\} \\
&\leqslant \limsup_{n} \frac{1}{n} \sum_{i=v_n+1}^{u_n} \mathbb{E}_{\tilde{\mu}}\left\{ \frac{\lambda^2}{2}\left[f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) \right]^2 e^{|\lambda f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)|} | \xi_{i-1}^{(n)} \right\} \\
&= \limsup_{n} \frac{1}{n} \sum_{i=v_n+1}^{u_n} \frac{\lambda^2}{2} \mathbb{E}_{\tilde{\mu}}\left\{ e^{\alpha |f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)|} \left[f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) \right]^2 e^{(\lambda-\alpha)|f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)|} | \xi_{i-1}^{(n)} \right\} \\
&\leqslant \frac{\lambda^2}{2} \limsup_{n} \frac{1}{n} \sum_{i=v_n+1}^{u_n} \mathbb{E}_{\tilde{\mu}}\left[\frac{4e^{-2}}{(\lambda-\alpha)^2} e^{\alpha |f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right)|} | \xi_{i-1}^{(n)} \right] \\
&\leqslant \lambda^2 H_t(\alpha,\tau) \quad a.s.
\end{aligned} \tag{3.12}$$

From Eqs 3.11, 3.12, we can obtain

$$\limsup_{n} \frac{\lambda}{n} \sum_{i=v_n+1}^{u_n} \left\{ f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) - \mathbb{E}_{\tilde{\mu}}\left[f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) | \xi_{i-1}^{(n)} \right] \right\}, \tag{3.13}$$

$$\leqslant \lambda^2 H_t(\alpha,\tau) + H(\mu \| \tilde{\mu}) \ a.s. \tag{3.14}$$

Considering $0 < \lambda < t < \alpha$ and $0 \leqslant c \leqslant t^2 H_t(\alpha, \tau)$, with Eqs 3.2, 3.13, we have

$$\begin{aligned}
\limsup_{n} \frac{1}{n} \sum_{i=v_n+1}^{u_n} \left\{ f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) - \mathbb{E}_{\tilde{\mu}}\left[f_i^{(n)}\left(\xi_{i-1}^{(n)}, \xi_i^{(n)}\right) | \xi_{i-1}^{(n)} \right] \right\} &\leqslant \lambda H_t(\alpha,\tau) + \frac{H(\mu \| \tilde{\mu})}{\lambda} \\
&\leqslant \lambda H_t(\alpha,\tau) + \frac{c}{\lambda}.
\end{aligned}$$

Defining function $g(\lambda) = \lambda H_t(\alpha, \tau) + \frac{c}{\lambda}$ and $0 \leqslant c \leqslant t^2 H_t(\alpha, \tau)$, we can arrive at

$$\inf_{\lambda\in(0,t)} g(\lambda) = g\left(\sqrt{\frac{c}{H_t(\alpha,\tau)}}\right) = 2\sqrt{cH_t(\alpha,\tau)}.$$

Considering $0\leqslant c\leqslant t^2H_t(\alpha,\tau)$, it can be found that

$$\limsup_n \frac{1}{n}\sum_{i=v_n+1}^{u_n}\left\{f_i^{(n)}\left(\xi_{i-1}^{(n)},\xi_i^{(n)}\right)-\mathbb{E}_{\tilde\mu}\left[f_i^{(n)}\left(\xi_{i-1}^{(n)},\xi_i^{(n)}\right)|\xi_{i-1}^{(n)}\right]\right\} \leqslant 2\sqrt{cH_t(\alpha,\tau)} \quad a.s.$$

Similarly, supposing $-\alpha<-t<\lambda<0$, we have

$$\liminf_n \frac{\lambda}{n}\sum_{i=v_n+1}^{u_n}\left\{f_i^{(n)}\left(\xi_{i-1}^{(n)},\xi_i^{(n)}\right)-\mathbb{E}_{\tilde\mu}\left[f_i^{(n)}\left(\xi_{i-1}^{(n)},\xi_i^{(n)}\right)|\xi_{i-1}^{(n)}\right]\right\} \geqslant -2\sqrt{cH_t(\alpha,\tau)} \quad a.s.$$

In particular, and only if $c=0$

$$\limsup_n \frac{1}{n}\sum_{i=v_n+1}^{u_n}\left\{f_i^{(n)}\left(\xi_{i-1}^{(n)},\xi_i^{(n)}\right)-\mathbb{E}_{\tilde\mu}\left[f_i^{(n)}\left(\xi_{i-1}^{(n)},\xi_i^{(n)}\right)|\xi_{i-1}^{(n)}\right]\right\}=0.$$

The proof is completed. □

In the following content, we assume that $\mathcal{P}$ is a strongly ergodic matrix and the vector π is the unique invariant measure determined by $\mathcal{P}$.

Theorem 3.2:: Supposing the conditions of Theorem 3.1 hold, if

$$\sup_k\sum_j \tilde\xi_{(k,j)}^{(n)} p_i^{(n)}(k,j)<\infty\,. \tag{3.15}$$

for any positive integer k,

$$\mathcal{C}_\alpha(i)=\mathbb{E}_{\tilde\mu}\left\{\left[f_i^{(n)}\left(\xi_{i-1}^{(n)},\xi_i^{(n)}\right)\right]^2 e^{\alpha|f_i^{(n)}(\xi_{i-1}^{(n)},\xi_i^{(n)})|}|\xi_{i-1}^{(n)}=k\right\}\leqslant\tau \tag{3.16}$$

and for $v_n\leqslant i\leqslant u_n$,

$$\mathbb{E}_{\tilde\mu}\left[e^{\alpha|f_i^{(n)}(\xi_{i-1}^{(n)},\xi_i^{(n)})|}\right]<\infty, \tag{3.17}$$

then

$$\lim_n \frac{1}{n}\sum_{i=v_n+1}^{u_n} f_i^{(n)}\left(\xi_{i-1}^{(n)},\xi_i^{(n)}\right)=\sum_k\pi_k\sum_j\tilde\xi_{(k,j)}^{(n)}p_k^{(n)}\left(\xi_k^{(n)},\xi_j^{(n)}\right) \quad a.s.$$

Proof: According to Theorem 3.1, we consider, under the condition of $c=0$, that

$$\lim_n \frac{1}{n}\sum_{i=v_n+1}^{u_n}\left\{f_i^{(n)}\left(\xi_{i-1}^{(n)},\xi_i^{(n)}\right)-\mathbb{E}_{\tilde\mu}\left[f_i^{(n)}\left(\xi_{i-1}^{(n)},\xi_i^{(n)}\right)|\xi_{i-1}^{(n)}\right]\right\}=0 \quad a.s. \tag{3.18}$$

With Eqs 3.14, 3.16, for arbitrary k, we have $\sum\tilde\xi_{(k,j)}^{(n)}p_i^{(n)}(\xi_k^{(n)},\xi_j^{(n)})<\infty$, and under the condition of $v_n^j\leqslant i\leqslant u_n$, we have $\mathbb{E}_{\tilde\mu}[f_i^{(n)}(\xi_{i-1}^{(n)},\xi_i^{(n)})|\xi_{i-1}^{(n)}]<\infty$ and

$$\lim_n \frac{\mathbb{E}_{\tilde\mu}\left[f_{i+1}^{(n)}\left(\xi_i^{(n)},\xi_{i+1}^{(n)}\right)|\xi_i^{(n)}\right]}{n}=0, \tag{3.19}$$

$$\lim_n \frac{\mathbb{E}_{\tilde\mu}\left[f_{v_{n+1}}^{(n)}\left(\xi_{v_n}^{(n)},\xi_{v_{n+1}}^{(n)}\right)|\xi_{v_n}^{(n)}\right]}{n}=0. \tag{3.20}$$

With Eqs 3.17–3.19, we can arrive at

$$\lim_n\left\{\frac{1}{n}\sum_{i=v_n+1}^{u_n} f_i^{(n)}\left(\xi_{i-1}^{(n)},\xi_i^{(n)}\right)-\frac{1}{n}\sum_{i=v_n+1}^{u_n}\mathbb{E}_{\tilde\mu}\left[f_i^{(n)}\left(\xi_{i-1}^{(n)},\xi_i^{(n)}\right)|\xi_i^{(n)}\right]\right\}=0 \quad a.s. \tag{3.21}$$

Since $e^{\alpha|x|}$ is convex, according to the Jensen's inequality of conditional expectation, we arrive at

$$\mathbb{E}_{\tilde\mu}\left[e^{\alpha|\mathbb{E}_{\tilde\mu}(f_{i+1}^{(n)}(\xi_i^{(n)},\xi_{i+1}^{(n)})|\xi_i^{(n)})|}\right]\leqslant\mathbb{E}_{\tilde\mu}\left\{\mathbb{E}_{\tilde\mu}\left[e^{\alpha|f_{i+1}^{(n)}(\xi_i^{(n)},\xi_{i+1}^{(n)})|\xi_i^{(n)}|}\right]\right\} =\mathbb{E}_{\tilde\mu}\left[e^{\alpha|f_{i+1}^{(n)}(\xi_i^{(n)},\xi_{i+1}^{(n)})|}\right]<\infty.$$

It is easy to obtain the conclusion that $g(x)=x^2e^{\alpha|x|}$ is a convex function. With Eq. 3.15, we have

$$\begin{aligned}
&\lim_n\frac{1}{n}\sum_{i=v_n+1}^{u_n}\mathbb{E}_{\tilde\mu}\left\{\left[\mathbb{E}_{\tilde\mu}\left(\tilde\xi_{(i,i+1)}^{(n)}|\xi_i^{(n)}\right)\right]^2\cdot e^{\alpha|\mathbb{E}_{\tilde\mu}(\tilde\xi_{(i,i+1)}^{(n)}|\xi_i^{(n)})|}|\xi_{i-1}^{(n)}\right\}\\
&=\lim_n\frac{1}{n}\sum_{i=v_n+1}^{u_n}\mathbb{E}_{\tilde\mu}\left\{g\left[\mathbb{E}_{\tilde\mu}\left(\tilde\xi_{(i,i+1)}^{(n)}|\xi_i^{(n)}\right)\right]|\xi_{i-1}^{(n)}\right\}\\
&\leqslant\lim_n\frac{1}{n}\sum_{i=v_n+1}^{u_n}\mathbb{E}_{\tilde\mu}\left\{\mathbb{E}_{\tilde\mu}\left[g\left(\tilde\xi_{(i,i+1)}^{(n)}\right)|\xi_i^{(n)}\right]|\xi_{i-1}^{(n)}\right\}\\
&=\lim_n\frac{1}{n}\sum_{i=v_n+1}^{u_n}\mathbb{E}_{\tilde\mu}\left\{g\left[\tilde\xi_{(i,i+1)}^{(n)}\right]|\xi_{i-1}^{(n)}\right\}\\
&=\lim_n\frac{1}{n}\sum_{i=v_n+1}^{u_n}\mathbb{E}_{\tilde\mu}\left\{\left[\tilde\xi_{(i,i+1)}^{(n)}\right]^2\cdot e^{\alpha|\tilde\xi_{(i,i+1)}^{(n)}|}|\xi_{i-1}^{(n)}\right\}\leqslant\tau.
\end{aligned}$$

Let $\mathbb{E}_{\tilde\mu}(\tilde\xi_{(i,i+1)}^{(n)}|\xi_i^{(n)})$, satisfying the condition of Eqs 3.2 and 3.3, then

$$\lim_n\frac{1}{n}\sum_{i=v_n+1}^{u_n}\left\{\mathbb{E}_{\tilde\mu}\left[\tilde\xi_{(i,i+1)}^{(n)}|\xi_i^{(n)}\right]-\mathbb{E}_{\tilde\mu}\left[\mathbb{E}_{\tilde\mu}\left(\tilde\xi_{(i,i+1)}^{(n)}|\xi_i^{(n)}\right)\Big|\xi_{i-1}^{(n)}\right]\right\}=0 \quad a.s.$$

Then,

$$\mathbb{E}_{\tilde\mu}\left\{\mathbb{E}_{\tilde\mu}\left[\tilde\xi_{(i,i+1)}^{(n)}|\xi_i^{(n)}\right]\Big|\xi_{i-1}^{(n)}\right\}=\mathbb{E}_{\tilde\mu}\left[\tilde\xi_{(i,i+1)}^{(n)}|\xi_{i-1}^{(n)}\right].$$

Therefore,

$$\lim_n\frac{1}{n}\sum_{i=v_n+1}^{u_n}\left\{\mathbb{E}_{\tilde\mu}\left[\tilde\xi_{(i,i+1)}^{(n)}|\xi_i^{(n)}\right]-\mathbb{E}_{\tilde\mu}\left[\tilde\xi_{(i,i+1)}^{(n)}|\xi_{i-1}^{(n)}\right]\right\}=0 \quad a.s.$$

With Eqs 3.18, 3.19, we can find

$$\lim_n\frac{1}{n}\sum_{i=v_n+1}^{u_n}\left\{\mathbb{E}_{\tilde\mu}\left[\tilde\xi_{(i,i+1)}^{(n)}|\xi_i^{(n)}\right]-\mathbb{E}_{\tilde\mu}\left[\tilde\xi_{(i+1,i+2)}^{(n)}|\xi_i^{(n)}\right]\right\}=0 \quad a.s. \tag{3.22}$$

With Eqs 3.20, 3.21, we have

$$\lim_n\frac{1}{n}\sum_{i=v_n+1}^{u_n}\left\{\tilde\xi_{(i-1,i)}^{(n)}-\mathbb{E}_{\tilde\mu}\left[\tilde\xi_{(i+1,i+2)}^{(n)}|\xi_i^{(n)}\right]\right\}=0 \quad a.s.$$

For positive integer h, calculating by induction, we arrive at

$$\lim_n\frac{1}{n}\sum_{i=v_n+1}^{u_n}\left\{\tilde\xi_{(i-1,i)}^{(n)}-\mathbb{E}_{\tilde\mu}\left[\tilde\xi_{(i+h,i+h+1)}^{(n)}|\xi_i^{(n)}\right]\right\}=0 \quad a.s.$$

With the strong ergodicity of $\mathcal{P}$ and the invariant of π, we can find

$$
\begin{aligned}
&\left|\frac{1}{n}\sum_{i=v_n+1}^{u_n}\mathbb{E}_{\tilde{\mu}}\left[\tilde{\xi}^{(n)}_{(i+h,i+h+1)}|\xi^{(n)}_i\right]-\sum_k \pi_k\tilde{\xi}^{(n)}_{(k,j)}p^{(n)}_k\left(\xi^{(n)}_k,\xi^{(n)}_j\right)\right|\\
&\quad=\left|\frac{1}{n}\sum_{i=v_n+1}^{u_n}\tilde{\xi}^{(n)}_{(i+h,i+h+1)}\cdot p^{(n)}_{i+h}\left(\xi^{(n)}_{i+h},\xi^{(n)}_{i+h+1}\right)-\sum_k \pi_k\tilde{\xi}^{(n)}_{(k,j)}p^{(n)}_k\left(\xi^{(n)}_k,\xi^{(n)}_j\right)\right|\\
&\quad=\left|\frac{1}{n}\sum_{i=v_n+1}^{u_n}\tilde{\xi}^{(n)}_{(i,i+1)}\left[p^{(n)}_{i+1}\left(\xi^{(n)}_{i+1},\xi^{(n)}_{i+2}\right)\cdots p^{(n)}_{i+h+1}\left(\xi^{(n)}_{i+h},\xi^{(n)}_{i+h+1}\right)\right]p^{(n)}_{i+h+1}\left(\xi^{(n)}_{i+h},\xi^{(n)}_{i+h+1}\right)\right.\\
&\qquad\left.-\sum_k \pi_k\tilde{\xi}^{(n)}_{(k,j)}p^{(n)}_k\left(\xi^{(n)}_k,\xi^{(n)}_j\right)\right|\\
&\quad=\left|\frac{1}{n}\sum_{i=v_n+1}^{u_n}\sum_k\sum_j\tilde{\xi}^{(n)}_{(k,j)}\left[p^{(n)}_{i+1}\left(\xi^{(n)}_{i+1},\xi^{(n)}_k\right)\right]^h p^{(n)}_k\left(\xi^{(n)}_k,\xi^{(n)}_j\right)-\sum_k \pi_k\tilde{\xi}^{(n)}_{(k,j)}p^{(n)}_k\left(\xi^{(n)}_k,\xi^{(n)}_j\right)\right|\\
&\quad=\left|\frac{1}{n}\sum_{i=v_n+1}^{u_n}\sum_k\sum_j\tilde{\xi}^{(n)}_{(k,j)}p^{(n)}_k\left(\xi^{(n)}_k,\xi^{(n)}_j\right)\left[p^{(n)}_i\left(\xi^{(n)}_i,\xi^{(n)}_k\right)\right]^h-\sum_k \pi_k\tilde{\xi}^{(n)}_{(k,j)}p^{(n)}_k\left(\xi^{(n)}_k,\xi^{(n)}_j\right)\right|\\
&\quad=\left|\frac{1}{n}\sum_{i=v_n+1}^{u_n}\sum_{k,j,l}\delta^{(n)}_l\left(\xi^{(n)}_i\right)\tilde{\xi}^{(n)}_{(k,j)}p^{(n)}_k\left(\xi^{(n)}_k,\xi^{(n)}_j\right)\left[p^{(n)}_l\left(\xi^{(n)}_l,\xi^{(n)}_k\right)\right]^h-\sum_k \pi_k\tilde{\xi}^{(n)}_{(k,j)}p^{(n)}_k\left(\xi^{(n)}_k,\xi^{(n)}_j\right)\right|\\
&\quad=\left|\frac{1}{n}\sum_l\delta_n(l)\sum_k\sum_j\tilde{\xi}^{(n)}_{(k,j)}p^{(n)}_k\left(\xi^{(n)}_k,\xi^{(n)}_j\right)\left\{\left[p^{(n)}_l\left(\xi^{(n)}_l,\xi^{(n)}_k\right)\right]^h-\pi_k\right\}\right|\\
&\quad=\left|\frac{1}{n}S_n(l)\sum_k\sum_j\tilde{\xi}^{(n)}_{(k,j)}p^{(n)}_k\left(\xi^{(n)}_k,\xi^{(n)}_j\right)\left\{\left[p^{(n)}_l\left(\xi^{(n)}_l,\xi^{(n)}_k\right)\right]^h-\pi_k\right\}\right|\\
&\quad\leqslant\sup_k\sum_j\left|\tilde{\xi}^{(n)}_{(k,j)}p^{(n)}_k\left(\xi^{(n)}_k,\xi^{(n)}_j\right)\cdot\sup_l\sum_k\left\{\left[p^{(n)}_l\left(\xi^{(n)}_l,\xi^{(n)}_k\right)\right]^h-\pi_k\right\}\right|\\
&\quad\to 0\quad(h\to\infty).
\end{aligned}
$$

The proof is completed.

Theorem 3.3: Let $f_n(\omega)$ and $H(\mu\|\tilde{\mu})$ be as given in Definition 2.1. Let $0 < t < 1$ and

$$H_t = \frac{2N}{e^2(t-1)^2}.$$

Assume that $0 < c < t^2H_t$ and $j\ (1\leqslant j\leqslant N)$ are constant, then we have

$$\limsup_n \frac{1}{n}\sum_{i=v_n+1}^{u_n}\left\{f_n(\omega)-H\left[p^{(n)}_i\left(\xi^{(n)}_{i-1},j\right)\right]\right\}\leqslant 2\sqrt{cH_t}\quad a.s.$$

and

$$\liminf_n \frac{1}{n}\sum_{i=v_n+1}^{u_n}\left\{f_n(\omega)-H\left[p^{(n)}_i\left(\xi^{(n)}_{i-1},j\right)\right]\right\}\geqslant -2\sqrt{cH_t}-c\quad a.s.$$

Proof: Under the conditions of Theorem 3.1, let $\tilde{\xi}^{(n)}_{(x,y)} = -\log p^{(n)}_i(x,y)$ and $\alpha = 1$, and

$$
\begin{aligned}
\mathbb{E}_\mu\left[e^{|\tilde{\xi}^{(n)}_{(i-1,i)}|}|\xi^{(n)}_{i-1}=k\right] &= \sum_{y=1}^N e^{|-\log p^{(n)}_i(x,y)|}\cdot p^{(n)}_i(x,y)\\
&=\sum_{y=1}^N\frac{p^{(n)}_i(x,y)}{p^{(n)}_i(x,y)}\\
&=N,
\end{aligned}
$$

then

$$\limsup_n\frac{1}{n}\sum_{i=v_n+1}^{u_n}\mathbb{E}_\mu\left[e^{|\tilde{\xi}^{(n)}_{(i-1,i)}|}|\xi^{(n)}_{i-1}=k\right]\leqslant N.$$

Supposing $1\leqslant j\leqslant N$ is a constant, we have

$$
\begin{aligned}
\mathbb{E}_\mu\left[-\log p^{(n)}_i\left(\xi^{(n)}_{i-1},\xi^{(n)}_i\right)|\xi^{(n)}_{i-1}\right] &= \sum_{y=1}^N p^{(n)}_i\left(\xi^{(n)}_{i-1},y\right)\log p^{(n)}_i\left(\xi^{(n)}_{i-1},y\right)\\
&= H\left[p^{(n)}_i\left(\xi^{(n)}_{i-1},j\right)\right].
\end{aligned}
$$

Applying $0 < c < t^2H_t$, we have

$$
\begin{aligned}
&\limsup_n\left\{\frac{1}{n}\sum_{i=v_n+1}^{u_n}\left[-\log p^{(n)}_i\left(\xi^{(n)}_{i-1},\xi^{(n)}_i\right)\right]-\frac{1}{n}\sum_{i=v_n+1}^{u_n}H\left[p^{(n)}_i\left(\xi^{(n)}_{i-1},j\right)\right]\right\}\\
&\leqslant 2\sqrt{cH_t}\quad a.s.
\end{aligned}
$$

and

$$
\begin{aligned}
&\liminf_n\left\{\frac{1}{n}\sum_{i=v_n+1}^{u_n}\left[-\log p^{(n)}_i\left(\xi^{(n)}_{i-1},\xi^{(n)}_i\right)\right]-\frac{1}{n}\sum_{i=v_n+1}^{u_n}H\left[p^{(n)}_i\left(\xi^{(n)}_{i-1},j\right)\right]\right\}\\
&\geqslant -2\sqrt{cH_t}\quad a.s.
\end{aligned}
$$

With Eqs 3.5, 3.11, we have

$$
\begin{aligned}
&\limsup_n\left\{f_n(\omega)-\frac{1}{n}\sum_{i=v_n+1}^{u_n}H\left[p^{(n)}_i\left(\xi^{(n)}_{i-1},j\right)\right]\right\}\\
&\quad\leqslant\limsup_n\left\{f_n(\omega)-\frac{1}{n}\sum_{i=v_n+1}^{u_n}\left[-\log p^{(n)}_i\left(\xi^{(n)}_{i-1},\xi^{(n)}_i\right)\right]\right\}\\
&\quad+\limsup_n\left\{\frac{1}{n}\sum_{i=v_n+1}^{u_n}\left[-\log p^{(n)}_i\left(\xi^{(n)}_{i-1},\xi^{(n)}_i\right)\right]-\frac{1}{n}\sum_{i=v_n+1}^{u_n}H\left[p^{(n)}_i\left(\xi^{(n)}_{i-1},j\right)\right]\right\}\\
&\quad\leqslant-\liminf_n\frac{1}{n}\log\left[\frac{p_n\left(\xi^{(n)}_1,\ldots,\xi^{(n)}_n\right)}{q^{(n)}\left(x^{(n)}_1\right)\prod_{i=v_n+1}^{u_n}p^{(n)}_i\left(x^{(n)}_{i-1},x^{(n)}_i\right)}\right]+2\sqrt{cH_t}\\
&\quad\leqslant 2\sqrt{cH_t}\quad a.s.
\end{aligned}
$$

With Eqs 3.6, 3.11, we have

$$
\begin{aligned}
&\liminf_n\left\{f_n(\omega)-\frac{1}{n}\sum_{i=v_n+1}^{u_n}H\left[p^{(n)}_i\left(\xi^{(n)}_{i-1},j\right)\right]\right\}\\
&\geqslant\liminf_n\left\{f_n(\omega)-\frac{1}{n}\sum_{i=v_n+1}^{u_n}\left[-\log p^{(n)}_i\left(\xi^{(n)}_{i-1},\xi^{(n)}_i\right)\right]\right\}\\
&+\liminf_n\left\{\frac{1}{n}\sum_{i=v_n+1}^{u_n}\left[-\log p^{(n)}_i\left(\xi^{(n)}_{i-1},\xi^{(n)}_i\right)\right]-\frac{1}{n}\sum_{i=v_n+1}^{u_n}H\left[p^{(n)}_i\left(\xi^{(n)}_{i-1},j\right)\right]\right\}\\
&\geqslant-\limsup_n\frac{1}{n}\log\left[\frac{p_n\left(\xi^{(n)}_1,\ldots,\xi^{(n)}_n\right)}{q^{(n)}\left(x^{(n)}_1\right)\prod_{i=v_n+1}^{u_n}p^{(n)}_i\left(x^{(n)}_{i-1},x^{(n)}_i\right)}\right]-2\sqrt{cH_t}\\
&\geqslant -H(\mu\|\tilde{\mu})-2\sqrt{cH_t}\\
&\geqslant -2\sqrt{cH_t}-c\quad a.s.
\end{aligned}
$$

The proof is completed.

Corollary 3.1: Supposing the conditions of Theorem 3.1 hold, then

$$\lim_n\left\{f_n(\omega)-\frac{1}{n}\sum_{i=v_n+1}^{u_n}H\left[p^{(n)}_i\left(\xi^{(n)}_{i-1},j\right)\right]\right\}=0.$$

Proof: It is easy to obtain this conclusion regarding the strong limit theory of entropy when $c = 0$. □

We point out that Corollary 3.1 implies that our main outcomes generalize the known results, such as Liu and Yang [12].

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

Author contributions

Conceptualization: XQ, ZR, and WP; methodology: XQ; software: ZR; validation: XQ and SC; writing—original draft preparation: XQ; visualization: ZR. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by the University Key Project of the Natural Science Foundation of Anhui Province (grant nos KJ2021A1032, KJ2019A0683, and KJ2021A1031), Key Project of the Natural Science Foundation of Chaohu University (grant no. XLZ-202201), Key Construction Discipline of Chaohu University (grant no. kj22zdjsxk01, kj22yjzx05, and kj22xjzz01), Anhui Province Social Science Innovation Development Research Project (grant no. 2021CX077), and University Outstanding Young Talents Project of Anhui Province (grant no. gxyq2021018).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Peng W. Conditional entropy, entropy density, and strong law of large numbers for generalized controlled tree-indexed Markov chains. *Commun Statistics-Theory Methods* (2017) 46(23):11880–91. doi:10.1080/03610926.2017.1285935

2. Wang Z, Yang W. Markov approximation and the generalized entropy ergodic theorem for non-null stationary process. *Proc Math Sci* (2020) 130(1):13. doi:10.1007/s12044-019-0542-4

3. Shi Z, Wang Z, Zhong P, Fan Y. The generalized entropy ergodic theorem for nonhomogeneous bifurcating Markov chains indexed by a binary tree. *J Theor Probab* (2021) 35:1367–90. doi:10.1007/s10959-021-01117-1

4. Shannon C. A mathematical theory of communication. *Bell Syst Tech J* (1948) 27: 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x

5. McMillan B. The basic theorems of information theory. *Ann Math Stat* (1953) 24: 196–219. doi:10.1214/aoms/1177729028

6. Breiman L. The individual ergodic theorem of information theory. *Ann Math Stat* (1957) 28:809–11. doi:10.1214/aoms/1177706899

7. Chung K. A note on the ergodic theorem of information theory. *Ann Math Stat* (1961) 32(2):612–4. doi:10.1214/aoms/1177705069

8. Kieffer J. A simple proof of the Moy-Perez generalization of the Shannon-McMillan theorem. *Pac J. Math.* (1974) 51:203–6. doi:10.2140/pjm.1974.51.203

9. Barron A. The strong ergodic theorem for densities: Generalized shannnon-MaMillan-breiman theorem. *Ann Prob* (1985) 13:1292–303.

10. Han T. *Information-spectrum methods in information theory*. New York: Springer (1990).

11. Liu W. Relative entropy densities and a class of limit theorems of the sequence of m-valued random variables. *Ann Probab* (1990) 1990:829–39.

12. Liu W, Yang W. The Markov approximation of the sequences of N-valued random variables and a class of small deviation theorems, Stochastic Process. *Appl* (2000) 89: 117–30.

13. Liu W. *The strong deviation theorms and analytical methods (in Chinese)*. Beijing: Science Press (2003).

14. Yang W, Liu W. Strong law of large numbers and Shannon-McMillan theorem for Markov chain fields on trees. *IEEE Trans Inf Theor* (2002) 48(1):313–8. doi:10.1109/18.971762

15. Yang W, Ye Z. The asymptotic equipartition property for nonhomogeneous Markov chains indexed by a homogeneous tree. *IEEE Trans Inf Theor* (2007) 53(9): 3275–80. doi:10.1109/tit.2007.903134

16. Yang W. A class of strong deviation theorems for the random fields associated with nonhomogeneous Markov chains indexed by a Bethe tree. *Stochastic Anal Appl* (2012) 30(2):220–37. doi:10.1080/07362994.2012.649619

17. Yang W, Zhao Y, Pan H. Strong laws of large numbers for asymptotic even–odd Markov chains indexed by a homogeneous tree. *J Math Anal Appl* (2014) 410(1): 179–89. doi:10.1016/j.jmaa.2013.08.009

18. Shi Z, Ji J, Yang W. A class of small deviation theorem for the sequences of countable state random variables with respect to homogeneous Markov chains. *Commun Stat* (2016) 46:6823–30. doi:10.1080/03610926.2015.1137594

19. Shi Z, Bao D, Fan Y, Wu B. The asymptotic equipartition property of Markov chains in single infinite markovian environment on countable state space. *Stochastics Int J Probab Stochastic Process* (2019) 91:945–57. doi:10.1080/17442508.2019.1567730

20. Shi Z, Ding C. A class of small deviation theorems for functionals of random fields on a tree with uniformly bounded degree in random environment. *Probab Eng Informational Sci* (2020) 36:169–83. doi:10.1017/s026996482000042x

21. Zhao M, Shi Z, Yang W, Wang. A class of strong deviation theorems for the sequence of real valued random variables with respect to continuous-state non-homogeneous Markov chains. *Commun Statistics-Theory Methods* (2021) 50(23): 5475–87. doi:10.1080/03610926.2020.1734838

22. Huang H, Yang W. Strong law of large numbers for Markov chains indexed by an infinite tree with uniformly bounded degree. *Sci China Ser A: Math* (2008) 51(2): 195–202. doi:10.1007/s11425-008-0015-1

23. Huang H. The generalized entropy ergodic theorem for nonhomogeneous Markov chains indexed by a homogeneous tree. *Probab Eng Informational Sci* (2020) 34(2): 221–34. doi:10.1017/s0269964818000554

24. Peng W, Yang W, Wang B. A class of small deviation theorems for functionals of random fields on a homogeneous tree. *J Math Anal Appl* (2010) 361(2):293–301. doi:10.1016/j.jmaa.2009.06.079

25. Peng W, Yang W, Shi Z. Strong law of large numbers for Markov chains indexed by spherically symmetric trees. *Probab Eng Informational Sci*, (2015) 29, 473–81. doi:10.1017/s0269964815000108

26. Peng W, Xi X. Shannon-McMillan theorem and strong law of large numbers for Markov chains indexed by generalized spherically symmetric trees. *Commun Statistics-Theory Methods* (2021) 2021:1–12. doi:10.1080/03610926.2021.1955385

27. Algoet P, Cover T. A sandwich proof of the Shannon-MaMillan-Breiman theorem. *Ann.Prob.* (1998) 16(2):899–909.

28. Wang Z. A kind of asymptotic properties of moving averages for Markov chains in markovian environments. *Commun Statistics-Theory Methods* (2017) 46:10926–40. doi:10.1080/03610926.2016.1252404

29. Wang Z, Yang W. The generalized entropy ergodic theorem for nonhomogeneous Markov chains. *Theor Probab* (2017) 29:761–75. doi:10.1007/s10959-015-0597-9

30. Chen M, Zhu Y, Niu X, Peng W, Wang Z. On almost sure convergence for double arrays of dependent random variables. *Commun Statistics-Theory Methods* (2022) 2002: 1–15. doi:10.1080/03610926.2022.2095403

31. Wang M, Wang M, Wang R, Wang X. Complete convergence and complete/moment convergence for arrays of rowwise negatively dependent random variables under sub–linear expectations. *J Math inequalities* (2022) 16(4):1347–70. doi:10.7153/jmi-2022-16-89