Journal of Statistical Physics, Vol. 91, Nos. 1/2, 1998

A Level 1 Large-Deviation Principle for the Autocovariances of Uniquely Ergodic Transformations with Additive Noise

S. C. Carmona,[1] C. Landim,[2] A. Lopes,[1] and S. Lopes[1]

Received November 25, 1997

A large-deviation principle (LDP) at level 1 for random means of the type

$$M_n \equiv \frac{1}{n} \sum_{j=0}^{n-1} Z_j Z_{j+1}, \qquad n = 1, 2,...$$

is established. The random process $\{Z_n\}_{n \geqslant 0}$ is given by $Z_n = \Phi(X_n) + \xi_n$, $n = 0, 1, 2,...$, where $\{X_n\}_{n \geqslant 0}$ and $\{\xi_n\}_{n \geqslant 0}$ are independent random sequences: the former is a stationary process defined by $X_n = T^n(X_0)$, X_0 is uniformly distributed on the circle S^1, $T: S^1 \to S^1$ is a continuous, uniquely ergodic transformation preserving the Lebesgue measure on S^1, and $\{\xi_n\}_{n \geqslant 0}$ is a random sequence of independent and identically distributed random variables on S^1; Φ is a continuous real function. The LDP at level 1 for the means M_n is obtained by using the level 2 LDP for the Markov process $\{V_n = (X_n, \xi_n, \xi_{n+1})\}_{n \geqslant 0}$ and the contraction principle. For establishing this level 2 LDP, one can consider a more general setting: $T: [0, 1) \to [0, 1)$ is a measure-preserving Lebesgue measure, $\Phi: [0, 1) \to \mathbb{R}$ is a real measurable function, and ξ_n are independent and identically distributed random variables on $\mathbb{R}$ (for instance, they could have a Gaussian distribution with mean zero and variance σ^2). The analogous result for the case of autocovariance of order k is also true.

KEY WORDS: Large deviation; level 1 entropy function; level 2 entropy function; contraction principle; ergodic transformations; Markov process.

[1] Instituto de Matemática, UFRGS, Porto Alegre, RS, Brazil.
[2] IMPA, Rio de Janeiro, RJ, Brazil.

1. INTRODUCTION

Given a probability space $(\Omega, \mathscr{F}, P)$ and a measurable transformation $T: \Omega \to \Omega$, we say that T is *measure preserving* if

$$P(T^{-1}(A)) = P(A), \qquad \forall A \in \mathscr{F} \tag{1.1}$$

We say that T is *uniquely ergodic* if there exists only one invariant measure for T, in the sense of (1.1).

We will parametrize points $y = \exp\{2\pi i x\}$ in the circle S^1 by $x \in [0, 1)$ and we shall identify x and y according to the convenience. In the sequel we will consider $T: S^1 \to S^1$ a continuous map. With this identification on mind, we point out that the main motivation of this paper came from Time Series Analysis. In that context, the process

$$X_n = T^n(X_0) \tag{1.2}$$

where $T: [0, 1) \to [0, 1)$ is a continuous, measure preserving transformation and T^n is the composition of T, n times, is called *signal process*.

One of the examples of transformations we are interested in is T defined by

$$T(x) = (x + \alpha) \bmod 1, \qquad x \in [0, 1) \tag{1.3}$$

where α is irrational. It is well known that this transformation preserves the Lebesgue measure λ on $([0, 1), \mathscr{B}([0, 1)))$, where $\mathscr{B}(A)$ is the Borel σ-field of subsets of A; moreover, T is uniquely ergodic (Durrett, 1996). For this example, the process $\{X_n\}_{n \geqslant 0}$ in (1.2) is stationary if and only if X_0 is uniformly distributed on $[0, 1)$. We observe that this process may be viewed as a Markov process with transition function $p(x, A) = \delta_A(T(x))$, $A \in \mathscr{B}([0, 1))$,

$$\delta_A(v) = \begin{cases} 1, & v \in A \\ 0, & v \notin A \end{cases}$$

having the Lebesgue measure on $[0, 1)$ as its unique stationary distribution. Other examples of uniquely ergodic transformations appear in Lopes and Rocha (1994), Coelho *et al.* (1994), and Lopes and Lopes (1995, 1996).

Let Φ be any continuous real function on S^1 and $\{\xi_n\}_{n \geqslant 0}$ a sequence of independent and identically distributed random variables in the circle S^1, common distribution η, and independent of $\{X_n\}_{n \geqslant 0}$. We define

$$Z_n = \Phi(X_n) + \xi_n, \qquad n = 0, 1, \ldots \tag{1.4}$$

When $\Phi(x) = \cos(2\pi x)$ and T is given by (1.3), the process Z_n is called the *harmonic model.*

The main goal in this paper is to establish a level 1 large-deviation principle (LDP) for the random means

$$M_n \equiv \frac{1}{n} \sum_{j=0}^{n-1} Z_j Z_{j+1}, \qquad n = 1, 2,... \tag{1.5}$$

where Z_n is defined in (1.4) with $T: [0, 1) \to [0, 1)$ being a continuous, uniquely ergodic transformation preserving the Lebesgue measure λ (or a measure absolutely continuous with respect to the Lebesgue measure). For simplifying the exposition we shall assume that T preserves the Lebesgue measure. We refer the reader to Lopes and Lopes (1995, 1996) for the motivation of analyzing such process and where all results at level 2 LDP considered here are applied. The strategy we shall follow consists in firstly to get a level 2 LDP for the process

$$V_n = (X_n, \xi_n, \xi_{n+1}), \qquad n = 0, 1, 2,... \tag{1.6}$$

and then, using the Contraction Principle (see Ellis, 1985), to obtain the level 1 LDP for (1.5).

The level 2 LDP is considered in Sections 3 and 4. The assumptions on $\{\xi_n\}_{n \geqslant 0}$ and T may be weakened in this case: the random variables ξ_n, $n \geqslant 0$, may not have compact support (the random variable ξ_n can be Gaussian distributed) and the transformation T can be discontinuous. In Section 5 we obtain the level 1 LDP for (1.5) when ξ_n has compact support. In Section 6 we make some remarks about special situations and extension results. In Remark 6.5 we point out that similar results are also valid for the autocovariance of order k, that is, for sums of the form

$$\frac{1}{n} \sum_{j=0}^{n-k} Z_j Z_{j+k}$$

The level 1 LDP for the random means M_n as in (1.5) is not true when ξ_n is a Gaussian distributed random variable.

2. STATEMENT OF THE MAIN RESULTS

In what follows we introduce notations, definitions, and we state the main results of this paper.

The random process $\{V_n\}_{n\geq 0}$ in (1.6) is a Markov process with phase space $S=[0,1)\times\mathbb{R}^2$ and transition function

$$\Pi((x,y,z),d(x_1,y_1,z_1))=\delta_{T(x)}(dx_1)\,\delta_{\{z\}}(dy_1)\,\eta(dz_1),\qquad (x,y,z)\in S \tag{2.1}$$

where the random variables $\xi_1,\xi_2,\ldots$ are independent and identically distributed with common distribution η on $\mathbb{R}$.

It is worth to remark that, given a Markov process with phase space S, transition function Π, and initial distribution μ, the Kolmogorov Existence theorem (see Billingsley, 1995) allows one to construct a measure $\mathbb{P}_\mu$ on sequence space $(S^{\mathbb{N}},\sigma(\mathscr{C}))$ so that the sequence $Y_n(w)=w_n$, $w\in S^{\mathbb{N}}$, has the same distribution as the original Markov process.

From now on, let us assume $\Omega=S^{\mathbb{N}}$ as being the space of sequences of elements of S, $\sigma(\mathscr{C})$ be the σ-field generated by the cylinder sets, and $\mathbb{P}_\mu$ the probability measure on $(\Omega,\sigma(\mathscr{C}))$ given by

$$\mathbb{P}_\mu[V_0\in A_0,\ldots,V_n\in A_n]=\int_{A_0}\mu(dv_0)\int_{A_1}\Pi(v_0,dv_1)\cdots\int_{A_n}\Pi(v_{n-1},dv_n) \tag{2.2}$$

$\forall A_0,\cdots,A_n\in\mathscr{B}(S)$, where μ is a (initial) distribution on $(S,\mathscr{B}(S))$. If $\mu(\cdot)=\delta_v(\cdot)$, for $v\in S$, the above measure is denoted by $\mathbb{P}_v$ and the corresponding expectation by $\mathbb{E}_v$.

If η is the distribution of ξ_n, it is not difficult to see that the product measure $\lambda\times\eta\times\eta$ on $(S,\mathscr{B}(S))$ is the unique stationary distribution for the Markov process $\{V_n\}_{n\geq 0}$ (in the sense that the only initial distribution that makes $\{V_n\}_{n\geq 0}$ a stationary process is $\lambda\times\eta\times\eta$). By the ergodic theorem (see Durrett, 1996), for any $\lambda\times\eta\times\eta$-integrable function g,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}g(V_j(w))=\int_S g(v)(\lambda\times\eta\times\eta)(dv),\qquad \mathbb{P}_{\lambda\times\eta\times\eta}\text{-a.s.} \tag{2.3}$$

where $V_j(w)=(X_j,\xi_j,\xi_{j+1})(w)=w_j$, for all $w\in\Omega$. Moreover, the above convergence holds $\mathbb{P}_v$-a.s., $\forall v\in S$ (see Doob, 1953).

Let $\mathscr{M}_1(S)$ be the space of probability measures on $\mathscr{B}(S)$; it is a Polish space (complete, separable metric space) if we impose on it the weak topology (which is compatible with the Lévy metric) (see Appendix in Dembo and Zeitouni, 1993). For measures in $\mathscr{M}_1(S)$ we shall introduce some definitions. By writing $S=S_1\times S_2\times S_3$, for $i\in\{1,2,3\}$ let π_i be the projection of S onto S_i, and π_{ij} be the projection of S onto $S_i\times S_j$, for $i,j\in\{1,2,3\}$, defined by $\pi_i(s_1,s_2,s_3)=s_i$ and $\pi_{ij}(s_1,s_2,s_3)=(s_i,s_j)$. If ν is a measure in $\mathscr{M}_1(S)$, then define a probability measure $\pi_i\nu$ on $\mathscr{B}(S_i)$ by requiring that, for each $i \in \{1, 2, 3\}$,

$$\pi_i\nu(F) = \nu(\pi_i^{-1}(F)) = \nu\{(s_1, s_2, s_3) \in S : s_i \in F\}, \qquad \forall F \in \mathscr{B}(S_i)$$

The measure $\pi_i\nu$ is called the *i-dimensional marginal of* ν. Similarly, define $\pi_{ij}\nu$ as the probability measure on $\mathscr{B}(S_i \times S_j)$, for each $i, j \in \{1, 2, 3\}$, given by

$$\pi_{ij}\nu(F) = \nu(\pi_{ij}^{-1}(F)) = \nu\{(s_1, s_2, s_3) \in S : (s_i, s_j) \in F\}, \qquad \forall F \in \mathscr{B}(S_i \times S_j)$$

The measure $\pi_{ij}\nu$ is called the *(i, j)-dimensional marginal of* ν. We also define, for each $\nu \in \mathscr{M}_1(S)$, a new measure νT^{-1} in $\mathscr{M}_1(S)$ by requiring

$$\nu T^{-1}(A \times B \times C) = \nu(T^{-1}(A) \times B \times C)$$

for all measurable rectangle $A \times B \times C$.

Let us introduce the empirical means

$$L_n(w, \cdot) = \frac{1}{n} \sum_{j=0}^{n-1} \delta_{V_j(w)}(\cdot) = \frac{1}{n} \sum_{j=0}^{n-1} \delta_{(X_j(w), \xi_j(w), \xi_{j+1}(w))}(\cdot) \tag{2.4}$$

$w \in S^{\mathbb{N}}$, $n = 1, 2, \ldots$. Clearly, for each $w \in S^{\mathbb{N}}$, $L_n(w, \cdot) \in \mathscr{M}_1(S)$. Moreover, L_n is $\sigma(\mathscr{C})$-measurable:

$$L_n^{-1}(A) = \{w \in \Omega : L_n(w, \cdot) \in A\} \in \sigma(\mathscr{C}), \qquad \forall A \in \mathscr{B}(\mathscr{M}_1(S))$$

The distribution of L_n on $\mathscr{B}(\mathscr{M}_1(S))$ is $Q_{n,\mu}(\cdot)$ given by

$$Q_{n,\mu}(A) = \mathbb{P}_\mu[L_n^{-1}(A)], \qquad \forall A \in \mathscr{B}(\mathscr{M}_1(S)) \tag{2.5}$$

where μ is a distribution on $(S, \mathscr{B}(S))$. In particular, if $\mu(\cdot) = \delta_v(\cdot)$, for $v \in S$, we shall use the notation $Q_{n,v}(\cdot)$.

Since

$$\int_S g(v)\, L_n(w, dv) = \frac{1}{n} \sum_{j=0}^{n-1} g(V_j(w))$$

it follows from (2.3) that

$$L_n(w, \cdot) \Rightarrow \lambda \times \eta \times \eta, \qquad \mathbb{P}_v\text{-a.s.}, \quad \forall v \in S$$

and then

$$\lim_{n \to \infty} Q_{n,\nu}(A) = 0$$

if $\lambda \times \eta \times \eta \notin \bar{A}$, $A \in \mathscr{B}(\mathscr{M}_1(S))$, where $\bar{A}$ is the closure of A. Hence, the sequence $\{Q_{n,\nu}(\cdot) : n = 1, 2,...\}$ converges weakly, when n goes to infinity, to the unit point measure $\delta_{\lambda \times \eta \times \eta}$ on $\mathscr{M}_1(S)$. We shall show that the sequence V_n obeys a LDP at level 2 (see Ellis, 1985), with the entropy function $I(\nu)$, $\nu \in \mathscr{M}_1(S)$ (this statement is equivalent to that the family $\{Q_n, \nu(\cdot) : n \geqslant 1\}$ obeys a LDP with entropy function $I(\nu)$).

In Sections 3 and 4 we prove that $I(\nu)$ is given by

$$I(\nu) = \begin{cases} \displaystyle\int_S \ln \frac{m}{m_{12}}\, d\nu, & \text{if} \quad \nu \in \mathscr{M}_0 \quad \text{and} \quad \displaystyle\int_S \left| \ln \frac{m}{m_{12}} \right| d\nu < +\infty \\ +\infty & \text{otherwise} \end{cases} \tag{2.6}$$

where

$$\mathscr{M}_0 = \{\nu \in \mathscr{M}_1(S) : \pi_1 \nu = \lambda,\ \pi_{12} \nu = \pi_{13} \nu T^{-1},\ \nu \ll \lambda \times \eta \times \eta\} \tag{2.7}$$

$$m(x, y, z) = \frac{d\nu}{d\lambda \times \eta \times \eta}(x, y, z)$$

and

$$m_{12}(x, y) = \int_{\mathbb{R}} m(x, y, z)\, \eta(dz) \tag{2.8}$$

We may say that

$$\frac{m(x, y, z)}{m_{12}(x, y)} \equiv m(z/x, y)$$

is the conditional density of $\pi_3 \nu / \pi_{12} \nu$, with respect to the measure η.

Now we state the main result in this paper (the level 2 LDP) which will be proved in Sections 3 and 4.

Theorem 2.1. For $I(\nu)$ given in (2.6) and for any $(x, y, z) \in S$,

(a) Lower Bound: for all open set $G \subset \mathscr{M}_1(S)$,

$$\varliminf_{n \to \infty} \frac{1}{n} \ln Q_{n,(x,y,z)}(G) \geqslant -\inf_{\nu \in G} I(\nu) \tag{2.9}$$

(b) Upper Bound: for all closed set $F \subset \mathcal{M}_1(S)$,

$$\overline{\lim_{n \to \infty}} \frac{1}{n} \ln Q_{n,(x,y,z)}(F) \leqslant - \inf_{v \in F} I(v) \tag{2.10}$$

(c) Compactness of the Level Sets: $\forall s > 0$, $\{v \in \mathcal{M}_1(S) : I(v) \leqslant s\}$ is a compact set in the weak topology.

A corollary (see Theorem 4.3.1 in Dembo and Zeitouni, 1993) of this theorem is that if Ψ is a bounded real-valued weakly continuous functional on $\mathcal{M}_1(S)$, then

$$\lim_{n \to \infty} \frac{1}{n} \ln \mathbb{E}^{Q_{n,(x,y,z)}}\{e^{-n\Psi(v)}\} = - \inf_{v \in \mathcal{M}_1(S)} [\Psi(v) + I(v)]$$

To prove Theorem 2.1 we use the same approach of Donsker and Varadhan (1975a): starting with the functional

$$I(v) = - \inf_{\psi \in \mathcal{W}} \int_S \ln \frac{\Pi\psi}{\psi} \, dv \tag{2.11}$$

where

$$\Pi\psi(x, y, z) = \int_S \psi(x_1, y_1, z_1) \, \Pi((x, y, z), d(x_1, y_1, z_1)) \tag{2.12}$$

with Π defined in (2.1) and

$$\begin{aligned} \mathcal{W} = \{\psi : S \to \mathbb{R} : \psi \text{ is continuous, } \exists a, b \text{ such that} \\ 0 < a \leqslant \psi(x, y, z) \leqslant b < +\infty, \forall (x, y, z) \in S\} \end{aligned} \tag{2.13}$$

we prove that $I(v)$ in (2.11) coincides with $I(v)$ in (2.6) and then we show that $\{Q_{n,v}(\cdot) : n \geqslant 1\}$ obeys a Weak Large-Deviation Principle with entropy function $I(v)$ (i.e., Theorem 2.1 is valid but the upper bound holds only for compact subsets of $\mathcal{M}_1(S)$). To extend the upper bound to closed sets, it is enough that $\{Q_{n,v}(\cdot) : n \geqslant 1\}$ be exponentially tight, which is proved in Lemma 4.1 of Section 4. If the distribution of ξ_n has compact support then $\mathcal{M}_1(S)$ is compact which implies that the Weak LDP is in fact the LDP for the process.

It is important to observe that the functional $I(\cdot)$ in (2.11) is lower semicontinuous in the weak topology of $\mathcal{M}_1(S)$ and convex. Moreover, $I(v) = 0$ if and only if v is the invariant measure of Π (see Lemma 2.5 in Donsker and Varadhan, 1975a).

Now, returning to the means (1.5), we have

$$Z_j Z_{j+1} = [\Phi(X_j) + \xi_j][\Phi(T(X_j)) + \xi_{j+1}]$$

If $g: S \to \mathbb{R}$ is defined by

$$g(x, y, z) = [\Phi(x) + y][\Phi(T(x)) + z] \tag{2.14}$$

we may write

$$\frac{1}{n} \sum_{j=0}^{n-1} Z_j Z_{j+1}(w) = \frac{1}{n} \sum_{j=0}^{n-1} g(V_j)(w) = \int_S g(v)\, dL_n(w, v) \tag{2.15}$$

Assuming that T is continuous and the distribution η has compact support (which is true if the random variables ξ_n have distribution on S^1), the function g is continuous and bounded. Therefore, the operator $\mathscr{G}: \mathscr{M}_1(S) \to \mathbb{R}$, such that $\mathscr{G}(v) = \int_S g(v)\, v(dv)$, is weakly continuous. Using the Contraction Principle (see Ellis, 1985), the LDP at level 1 for (1.5) is obtained taking into account (2.15): the level 1 entropy function $I_Z(\cdot)$ is given by

$$I_Z(r) = \inf_{\langle v, g \rangle = r} I(v) = \inf \left\{ I(v) : v \in \mathscr{M}_1(S), \int_S g(v)\, v(dv) = r \right\}$$

where $I(\cdot)$ is the level 2 entropy function for $\{Q_{n,v}(\cdot) : n \geq 1\}$. In this way the LDP at level 1 follows from the LDP at level 2.

3. LEVEL 2 LARGE DEVIATIONS: LOWER BOUND

The goal here is to prove part (a) of Theorem 2.1. For proving it we need some lemmas.

First we consider the random process $\{X_n\}_{n \geq 0}$ introduced in (1.2); it can be seen as a Markov process with transition function $p(x, A) = \delta_A(T(x))$, $x \in [0, 1)$. Throughout this section and Section 4 T is not assumed to be a continuous transformation and the support of the random variables ξ_n is not compact, necessarily. As T is uniquely ergodic and preserves the Lebesgue measure, the uniform distribution on [0,1) is the unique stationary measure for the process $X_n = \phi \circ T^n$.

Let

$$L_n^{(1)}(x, \cdot) = \frac{1}{n} \sum_{j=0}^{n-1} \delta_{X_j(x)}(\cdot) = \frac{1}{n} \sum_{j=0}^{n-1} \delta_{T^j(x)}(\cdot), \qquad x \in [0, 1)$$

The ergodic theorem says that

$$L_n^{(1)}(x, \cdot) \underset{n \to \infty}{\Rightarrow} \lambda(\cdot), \qquad \forall x \in [0, 1)$$

where λ is the Lebesgue measure on $[0,1)$.

Let $Q_{n,x}^{(1)}(\cdot)$ be the distribution of $L_n^{(1)}(x, \cdot)$ on $\mathscr{B}(\mathcal{M}_1([0, 1)))$. Notice that, once the initial point x is fixed, the process $\{X_n\}_{n \geqslant 0}$ is deterministic as well as $L_n^{(1)}(x, \cdot)$. The next lemma follows from this observation.

Lemma 3.1. (a) For all open set $G \subset \mathcal{M}_1([0, 1))$,

$$\varliminf_{n \to \infty} \frac{1}{n} \ln Q_{n,x}^{(1)}(G) \geqslant -\inf_{v \in G} I^{(1)}(v)$$

and (b) for all closed set $F \subset \mathcal{M}_1([0, 1))$,

$$\varlimsup_{n \to +\infty} \frac{1}{n} Q_{n,x}^{(1)}(F) \leqslant -\inf_{v \in F} I^{(1)}(v)$$

where the entropy function at level 2 $I^{(1)}(v)$ for the process $\{X_n\}_{n \geqslant 0}$ is given by

$$I^{(1)}(v) = \begin{cases} 0, & \text{if } \ v = \lambda \\ +\infty, & \text{if } \ v \neq \lambda \end{cases}$$

Secondly, we consider the Markov process $\{V_n\}_{n \geqslant 0}$ introduced in (1.6). Its transition function Π is given in (2.1) and its phase space is $S = [0, 1) \times \mathbb{R}^2$. Let $I(\cdot)$ be the entropy function defined in (2.11).

Lemma 3.2. $I(v) < +\infty$ if and only if $v \in \mathcal{M}_0$ and the density $m(x, y, z)$ of v with respect to $\lambda \times \eta \times \eta$ satisfies

$$\int_S \left| \ln \frac{m(x, y, z)}{m_{12}(x, y)} \right| m(x, y, z)(\lambda \times \eta \times \eta)(d(x, y, z)) < +\infty \tag{3.1}$$

where $m_{12}(x, y)$ is given in (2.8) and $\mathcal{M}_0$ is the set introduced in (2.7). Moreover,

$$I(v) = \begin{cases} \displaystyle\int_S \ln \frac{m(x, y, z)}{m_{12}(x, y)} \, v(d(x, y, z)), & \text{if } \ v \in \mathcal{M} \text{ and (3.1) holds} \\ +\infty, & \text{otherwise} \end{cases} \tag{3.2}$$

Proof. Suppose that $\nu \in \mathscr{M}_0$ and that (3.1) holds. Let

$$m(x, y, z) = \frac{d\nu}{d\lambda \times \eta \times \eta}(x, y, z), \qquad (x, y, z) \in S$$

and

$$m_{13}(x, z) = \int_{\mathbb{R}} m(x, y, z)\, \eta(dy) \equiv \frac{d\pi_{13}\nu}{d\lambda \times \eta}(x, z)$$

Since $\pi_{13}\nu T^{-1}(A \times B) = \pi_{13}\nu(T^{-1}(A) \times B)$, for all $A \in \mathscr{B}([0, 1))$, $B \in \mathscr{B}(\mathbb{R})$, and T is λ-preserving, we get

$$\frac{d\pi_{13}\nu T^{-1}}{d\lambda \times \eta}(x, z) = m_{13}(T^{-1}(x), z)$$

Taking into account that $\pi_{12}\nu = \pi_{13}\nu T^{-1}$, we have $m_{12}(x, y) = m_{13}(T^{-1}(x), y)$. Let

$$l = \int_S m(x, y, z) \ln \frac{m(x, y, z)}{m_{12}(x, y)}\, \eta(dz)\, \eta(dy)\, dx$$

By hypothesis, $l < +\infty$. Notice that, for $\psi \in \mathscr{W}$,

$$\ln \frac{\Pi\psi}{\psi}(x, y, z) = \ln \int_{\mathbb{R}} \psi(T(x), z, u)\, \eta(du) - \ln \psi(x, y, z), \qquad \forall (x, y, z) \in S$$

So, if we show that, for all $\psi \in \mathscr{W}$,

$$\int_{[0,1)} \int_{\mathbb{R}} \left[\ln \int_{\mathbb{R}} \psi(T(x), z, u)\, \eta(du) \right] m_{13}(x, z)\, \eta(dz)\, dx$$
$$- \int_S m(x, y, z) \ln \psi(x, y, z)\, \eta(dz)\, \eta(dy)\, dx \geqslant -l \tag{3.3}$$

then $I(\nu) \leqslant l$, $I(\cdot)$ being the functional in (2.11).

Recall that the marginal density of $\pi_1\nu$ is $m_1(x) \equiv 1$, for all $x \in [0, 1)$, so that, for each $x \in [0, 1)$, $m(x, y, z)$ is a probability density (with respect to $\eta \times \eta$) of some measure μ_x on $\mathscr{B}(\mathbb{R}^2)$. For each $x \in [0, 1)$, let us define

$A_x = \{(y, z) \in \mathbb{R}^2 : m(x, y, z) > 0\}$. Clearly $\mu_x(A_x) = 1$. Let $B_x = \{y \in \mathbb{R} : m_{12}(x, y) > 0\}$. Since the first marginal $\mu_x^{(1)}$ of μ_x has density $m_{12}(x, y)$ with respect to η, $\mu_x^{(1)}(B_x) = 1$ which means that

$$1 = \int_{B_x} m_{12}(x, y)\, \eta(dy) = \int_{B_x} \left[\int_{\mathbb{R}} m(x, y, z)\, \eta(dz) \right] \eta(dy)$$

Hence, $\mu_x(B_x \times \mathbb{R}) = 1$ and we may identify A_x with $B_x \times \mathbb{R}$, in terms of integration. Then,

$$\begin{aligned} I &\equiv \int_S m(x, y, z) \ln \psi(x, y, z)\, \eta(dz)\, \eta(dy)\, dx \\ &= \int_{[0, 1)} \iint_{A_x} [\ln \psi(x, y, z)]\, m(x, y, z)\, \eta(dy)\, \eta(dz)\, dx \\ &= \int_{[0, 1)} \iint_{A_x} \ln \left[\frac{\psi(x, y, z)}{m(x, y, z)} m_{12}(x, y) \right] \frac{m(x, y, z)}{m_{12}(x, y)} \eta(dz)\, m_{12}(x, y)\, \eta(dy)\, dx \\ &\quad + \int_{[0, 1)} \iint_{A_x} \ln \left[\frac{m(x, y, z)}{m_{12}(x, y)} \right] m(x, y, z)\, \eta(dz)\, \eta(dy)\, dx \end{aligned}$$

But, for each $(x, y) \in [0, 1) \times \mathbb{R}$ with $m_{12}(x, y) > 0$, $m(x, y, z)/m_{12}(x, y)$ is a density with respect to η for some probability measure on $\mathscr{B}(\mathbb{R})$. Using Jensen's inequality in the first integral on the right hand side of the last equality (this is possible if one substitutes A_x by $B_x \times \mathbb{R}$), we obtain

$$I \leqslant \int_{[0, 1)} \int_{\mathbb{R}} \ln \left[\int_{\mathbb{R}} \psi(x, y, z)\, \eta(dz) \right] m_{12}(x, y)\, \eta(dy)\, dx + l$$

Since $\pi_{12} \nu = \pi_{13} \nu T^{-1}$ and $\lambda = \lambda T^{-1}$, we may write

$$\begin{aligned} I &\leqslant \int_{[0, 1)} \int_{\mathbb{R}} \ln \left[\int_{\mathbb{R}} \psi(x, y, z)\, \eta(dz) \right] m_{13}(T^{-1}(x), y)\, \eta(dy)\, dT^{-1}(x) + l \\ &= \int_{[0, 1)} \int_{\mathbb{R}} \ln \left[\int_{\mathbb{R}} \psi(T(v), y, z)\, \eta(dz) \right] m_{13}(v, y)\, \eta(dy)\, dv + l \end{aligned}$$

and we get (3.3).

Now suppose that $I(\nu) < +\infty$. Let $I(\nu) = l$ for $\nu \in \mathcal{M}_1(S)$. Then,

$$\int_{[0,1)} \int_{\mathbb{R}} \left[\ln \int_{\mathbb{R}} \psi(T(x), z, u)\, \eta(du) \right] \pi_{13} \nu(d(x, z))$$

$$- \int_S \ln \psi(x, y, z)\, \nu(d(x, y, z)) \geqslant -l, \qquad \forall \psi \in \mathcal{W} \tag{3.4}$$

From Lusin's theorem (see Rudin, 1974), (3.4) also holds for all non-negative measurable functions on S, bounded away from zero and infinity. We denote this set by $\mathcal{W}^*$.

Let $\psi \in \mathcal{W}^*$ be defined by $\psi(x, y, z) = \psi_1(x)\, \psi_2(y)\, \psi_3(z)$, where ψ_1 is any continuous function with $\psi_1(x) > 0$, $\forall x \in [0, 1)$, $\psi_2 \equiv 1$, and

$$\psi_3(z) = \begin{cases} k, & z \in A \\ 1, & z \in A^c \end{cases}$$

where $k > 1$ and $A \in \mathcal{B}(\mathbb{R})$. For such ψ, (3.4) implies that

$$\pi_3 \nu(A) \ln k \leqslant l + \ln[k\eta(A) + \eta(A^c)] + \int_{[0,1)} \ln \psi_1(T(x))\, \pi_1 \nu(dx)$$

$$- \int_{[0,1)} \ln \psi_1(x)\, \pi_1 \nu(dx) \tag{3.5}$$

Suppose that $\pi_1 \nu \neq \lambda$. From Lemma 3.1, we know that for all $M > 0$, there exists a positive continuous function ψ_1 on $[0,1)$ such that

$$\int_{[0,1)} \ln \frac{\psi_1(T(x))}{\psi_1(x)}\, \pi_1 \nu(dx) < -M$$

So, we may choose M, ψ_1, and k in such a way that (3.5) implies that

$$\pi_3 \nu(A) \ln k \leqslant l + \ln[k\eta(A) + \eta(A^c)] - M < 0$$

which is a contradiction, if M is large enough. Therefore, $\pi_1 \nu = \lambda$.

Now take $\psi(x, y, z) = \psi_1(x)\, \psi_2(y)\, \psi_3(z)$ with

$$\psi_1(x)\, \psi_2(y) = \begin{cases} k, & (x, y) \in A_1 \times A_2 \\ 1, & (x, y) \notin A_1 \times A_2 \end{cases}$$

where $A_1 \in \mathscr{B}([0, 1))$, $A_2 \in \mathscr{B}(\mathbb{R})$, and $\psi_3 \equiv 1$. Notice that

$$\psi_1(T(x))\,\psi_2(y) = k \Leftrightarrow (x, y) \in T^{-1}(A_1) \times A_2$$

Hence (3.4) implies that, for $k > 1$,

$$\pi_{12}\nu(A_1 \times A_2) - \pi_{13}\nu(T^{-1}(A_1) \times A_2) \leqslant \frac{l}{\ln k}$$

By making $k \to \infty$, we get

$$\pi_{12}\nu(A_1 \times A_2) \leqslant \pi_{13}\nu T^{-1}(A_1 \times A_2)$$

from what follows the equality of measures $\pi_{12}\nu$ and $\pi_{13}\nu T^{-1}$, if one takes the complement of the set.

To show that $\nu \ll \lambda \times \eta \times \eta$ first we show that $\pi_{13}\nu T^{-1} \ll \lambda \times \eta$. Choose $\psi \in \mathscr{W}^*$ such that $\psi(x, y, z) = \psi_1(x, z)\,\psi_2(y)$, $\psi_2 \equiv 1$, and for $A \in \mathscr{B}([0, 1) \times \mathbb{R})$,

$$\psi_1(x, z) = \begin{cases} k, & (x, z) \in A \\ 1, & (x, z) \in A^c \end{cases}$$

By Jensen's inequality and (3.4) we get

$$\ln \int_{[0,1)} \int_{\mathbb{R}} \psi_1(T(x), u)\,\eta(du)\,dx - \int_{[0,1)} \int_{\mathbb{R}} \ln \psi_1(x, z)\,\pi_{13}\nu(d(x, z)) \geqslant -l$$

Since $\lambda = \lambda T^{-1}$, the last inequality implies that

$$\pi_{13}\nu(A) \ln k \leqslant l + \ln[(k-1)(\lambda \times \eta)(A) + 1]$$

If $(\lambda \times \eta)(A) = 0$ we have

$$\pi_{13}\nu(A) \leqslant \frac{l}{\ln k} \to 0, \qquad \text{when} \quad k \to +\infty$$

Hence, $\pi_{13}\nu \ll \lambda \times \eta$. Consequently, $\pi_{13}\nu T^{-1} \ll \lambda \times \eta$ since

$$\pi_{13}\nu T^{-1}(A \times B) = \pi_{13}\nu(T^{-1}(A) \times B)$$

and we conclude that

$$\frac{d\pi_{13}\nu}{d\lambda \times \eta}(x, y) = m_{13}(x, y) \Leftrightarrow \frac{d\pi_{13}\nu T^{-1}}{d\lambda \times \eta}(x, y) = m_{13}(T^{-1}(x), y)$$

Using this fact and that T is λ-preserving, (3.4) may be written as

$$\int_{[0,1)}\int_{\mathbb{R}}\left[\ln\int_{\mathbb{R}}\psi(x,z,u)\,\eta(du)\right]\pi_{13}\nu T^{-1}(d(x,z))$$
$$-\int_S \ln\psi(x,y,z)\,\nu(d(x,y,z))\geqslant -l \tag{3.6}$$

Finally, for having $\nu\ll\lambda\times\eta\times\eta$ it suffices that $\nu\ll\pi_{13}\nu T^{-1}\times\eta$. To prove this last statement, choose $\psi\in\mathscr{W}^*$ as

$$\psi(x,y,z)=\begin{cases}k, & (x,y,z)\in A\\ 1, & (x,y,z)\in A^c\end{cases}$$

where $A\in\mathscr{B}(S)$. Jensen's inequality and (3.6) imply that

$$\nu(A)\leqslant\frac{l}{\ln k}+\frac{1}{\ln k}\ln\{k[\pi_{13}\nu T^{-1}\times\eta](A)+[\pi_{13}\nu T^{-1}\times\eta](A^c)\}$$

from what we conclude, by taking $k\to+\infty$, that $\nu\ll\pi_{13}\nu T^{-1}\times\eta$.

It remains to show that (3.1) holds. By defining

$$u_n(x,y,z)=\left(\frac{m(x,y,z)}{m_{12}(x,y)}\vee\frac{1}{n}\right)\wedge n\equiv\left(a(x,y,z)\vee\frac{1}{n}\right)\wedge n,\qquad n\geqslant 1$$

and following the same arguments as in the proof of Lemma 2.1 in Donsker and Varadhan (1975a), we get (3.1). In what follows we outline the main steps.

From the Dominated Convergence theorem,

$$\lim_{n\to\infty}\int_S|u_n(x,y,z)\,m_{12}(x,y)-m(x,y,z)|\,(\lambda\times\eta\times\eta)(d(x,y,z))=0 \tag{3.7}$$

But $I(\nu)=l<+\infty$ implies that (3.6) holds for all $\psi\in\mathscr{W}$ and then from Lusin's theorem it also holds for $\psi\in\mathscr{W}^*$. Hence, for $\psi=u_n$ and using Jensen's inequality, we get from (3.6),

$$\int_S\ln u_n(x,y,z)\,\nu(d(x,y,z))$$
$$\leqslant\ln\int_S u_n(x,y,z)\,m_{12}(x,y)\,\eta(dy)\,\eta(dz)\,dx+l \tag{3.8}$$

for obtaining the above inequality we also used the fact that $\pi_{13}\nu T^{-1} = \pi_{12}\nu$. Since $m(x, y, z)$ is a probability density with respect to $\lambda \times \eta \times \eta$, it follows from (3.7) and (3.8) that

$$\overline{\lim_{n\to\infty}} \int_S \ln u_n(x, y, z)\, \nu(d(x, y, z)) \leqslant l \tag{3.9}$$

By the Monotone Convergence Theorem

$$\int_S (\ln u_n)^- \, d\nu \underset{n\to\infty}{\uparrow} \int_S (\ln a)^- \, d\nu = \int_S a(\ln a)^- \, d(\pi_{12}\nu)\, d\eta < +\infty$$

and then (3.9) implies that

$$\overline{\lim_{n\to\infty}} \int_S (\ln u_n)^+ \, d\nu \leqslant l + \int_S a(\ln a)^- \, d(\pi_{12}\nu)\, d\eta < +\infty.$$

Hence

$$\int_S |\ln a| \, d\nu < +\infty$$

which is (3.1). Moreover,

$$\int_S m(x, y, z) \ln \frac{m(x, y, z)}{m_{12}(x, y)} (\lambda \times \eta \times \eta)(d(x, y, z)) \leqslant l = I(\nu)$$

From the whole proof we also conclude that, if $I(\nu) < +\infty$ then

$$I(\nu) = \int_S m(x, y, z) \ln \frac{m(x, y, z)}{m_{12}(x, y)} (\lambda \times \eta \times \eta)(d(x, y, z))$$

so we have (3.2); besides, (2.11) and (3.2) are equal. ∎

For proving the lower bound (2.9) in Theorem 2.1 we shall consider a new Markov process with transition function Π' absolutely continuous with respect to Π.

Let us introduce the set

$$\mathscr{M}_2 = \left\{ \nu \in \mathscr{M}_0 : \frac{d\nu}{d\lambda \times \eta \times \eta}(x, y, z) \equiv m(x, y, z) \text{ and } \exists c, d \right.$$
$$\left. \text{such that } 0 < c \leqslant m(x, y, z) \leqslant d < +\infty,\ \forall (x, y, z) \in S \right\}$$

Let ν be in $\mathcal{M}_2$ with density $m(x, y, z)$. Define

$$\Pi'((x, y, z), d(x_1, y_1, z_1)) = \frac{m(x_1, y_1, z_1)}{m_{12}(x_1, y_1)} \Pi((x, y, z), d(x_1, y_1, z_1)) \quad (3.10)$$

with Π as in (2.1).

Lemma 3.3. Under the above conditions, ν is the only invariant measure for Π'.

Proof. Clearly $I(\nu) < +\infty$ which implies, from Lemma 3.2, that $\nu \in \mathcal{M}_0$. It is not difficult to show that

$$\int_S \Pi'((x, y, z), A_1 \times A_2 \times A_3) \, \nu(d(x, y, z)) = \nu(A_1 \times A_2 \times A_3)$$

for any measurable rectangle $A_1 \times A_2 \times A_3$. ∎

Lemma 3.4. Let G be an open subset of $\mathcal{M}_1(S)$. Then

$$\inf_{\nu \in G} I(\nu) = \inf_{\nu \in G \cap \mathcal{M}_2} I(\nu)$$

Proof. This lemma can be proved as Lemma 2.9 in Donsker and Varadhan (1975a) so we omit it. ∎

Lemmas (3.2)–(3.4) allow one to prove the lower bound (2.9) by using the same arguments as in Donsker and Varadhan (1975a). In what follows, we outline the main steps of the proof.

Proof of the Lower Bound. Let $\nu \in \mathcal{M}_2$ and, for simplifying the notation,

$$W(x, y, z) \equiv \ln \frac{m(x, y, z)}{m_1 2(x, y)} = \ln a(x, y, z), \qquad (x, y, z) \in S$$

where m is the density of ν with respect to $\lambda \times \eta \times \eta$. Then $I(\nu) = \int_S W \, d\nu$.

Let $S(\nu; \varepsilon)$ be the sphere with center ν of radius $\varepsilon > 0$, in the weak topology on $\mathcal{M}_1(S)$. Define $E_{n,\nu,\varepsilon} = \{w : L_n(w, \cdot) \in S(\nu; \varepsilon)\}$. One may show that

$$Q_{n,\nu}[S(\nu; \varepsilon)] = \int_{E_{n,\nu,\varepsilon}} \prod_{j=0}^{n-1} \frac{m_1 2(x_j, y_j)}{m(x_j, y_j, z_j)} \, d\mathbb{P}'_\nu$$

where $\mathbb{P}'_v$ is the probability measure in $\Omega = S^{\mathbb{N}}$ induced by the transition function $\Pi'(v, du)$ defined in (3.10).

For each $\varepsilon' > 0$, define

$$F_{n,\nu,\varepsilon'} = \left\{ w : \left| \frac{W(V_0)(w) + \cdots + W(V_{n-1})(w)}{n} - \int_S W\,d\nu \right| < \varepsilon' \right\}$$

Then

$$Q_{n,v}[S(\nu; \varepsilon)] \geqslant \exp\{-n[I(\nu) + \varepsilon']\}\ \mathbb{P}'_v[E_{n,\nu,\varepsilon} \cap F_{n,\nu,\varepsilon'}]$$

By Lemma 3.3, ν is the unique invariant measure for Π'. From the ergodic theorem (see Doob, 1953),

$$L_n(w, \cdot) \Rightarrow \nu, \qquad \mathbb{P}'_v\text{-a.s.}, \quad \forall v \in S$$

so that, $\forall \varepsilon > 0$, $\forall \varepsilon' > 0$,

$$\lim_{n \to +\infty} \mathbb{P}'_v[E_{n,\nu,\varepsilon}] = 1 \qquad \text{and} \qquad \lim_{n \to +\infty} \mathbb{P}'_v[F_{n,\nu,\varepsilon'}] = 1$$

Hence,

$$\varliminf_{n \to +\infty} \frac{1}{n} \ln Q_{n,v}[S(\nu; \varepsilon)] \geqslant -I(\nu), \qquad \nu \in \mathcal{M}_2$$

Now, let G be an open subset of $\mathcal{M}_1(S)$ and take $\nu \in G \cap \mathcal{M}_2$. Since G is an open set, there exists $\varepsilon > 0$ such that $S(\nu; \varepsilon) \subset G$. By using the last inequality and Lemma 3.4 we get (2.9). ∎

4. LEVEL 2 LARGE DEVIATIONS: UPPER BOUND

Following the same ideas as in Donsker and Varadhan (1975a), one can prove the upper bound in (2.10) for compact sets. Since the measure η has not compact support, $\mathcal{M}_1(S)$ is not a compact set. So, the inequality for closed sets does not follow as a consequence.

Proof of the Upper Bound. Let $\psi \in \mathcal{W}$, $u = \Pi\psi$, and $e^{-W} = \psi/u$. Notice that $W = \ln \Pi\psi - \ln \psi$ is bounded and continuous. From the Markov property it follows that

$$\mathbb{E}_v\{\exp\{-[W(V_0) + \cdots + W(V_{n-1})]\}\ u(V_{n-1})\} = \psi(v), \qquad \forall v \in S, \quad n \geqslant 1$$

where $V_n = (X_n, \xi_n, \xi_n + 1)$ as before. Then

$$\mathbb{E}_v\{\exp\{-[W(V_0) + \cdots + W(V_{n-1})]\}\} \leqslant M$$

for some constant $M > 0$. This inequality may be written as

$$\mathbb{E}^{Q_{n,v}}\left\{\exp\left\{-n\int_S W\, d\mu\right\}\right\} \leqslant M$$

where $\mu \in \mathscr{M}_1(S)$ is the integration variable.

First, take $F \subset \mathscr{M}_1(S)$ as being any measurable set. From the above inequality we get

$$Q_{n,v}(F) \leqslant M \sup_{\mu \in F} \int_S \ln\left(\frac{\Pi\psi}{\psi}\right) d\mu, \qquad \forall \psi \in \mathscr{W}$$

and then

$$\overline{\lim_{n \to +\infty}} \frac{1}{n} \ln Q_{n,v}(F) \leqslant \inf_{\psi \in \mathscr{W}} \sup_{\mu \in F} \int_S \ln\left(\frac{\Pi\psi}{\psi}\right) d\mu$$

Secondly, for any $F \subset \bigcup_{i=1}^k F_i$, for F_i measurable sets,

$$\overline{\lim_{n \to +\infty}} \frac{1}{n} \ln Q_{n,v}(F) \leqslant \inf_{\substack{F_1, \ldots, F_k \\ F \subset \bigcup_{i=1}^k F_i}} \sup_{1 \leqslant i \leqslant k} \inf_{\psi \in \mathscr{W}} \sup_{\mu \in F_i} \int_S \ln\left(\frac{\Pi\psi}{\psi}\right) d\mu$$

Now, if F is a compact set it can be shown (see Donsker and Varadhan, 1975b) that the expression on the right hand side of the above inequality is equal to

$$\sup_{\mu \in F} \inf_{\psi \in \mathscr{W}} \int_S \ln\left(\frac{\Pi\psi}{\psi}\right) d\mu = -\inf_{\mu \in F} I(\mu)$$

Up to now, (2.10) holds for compact sets F. Therefore, $\{Q_{n,v}(\cdot) : n \geqslant 1\}$ satisfies a weak LDP with rate function $I(\cdot)$ given in (2.11). But Lemma 4.1, to be proved below, tells us that this family of measures is exponentially tight, so (2.10) holds for closed sets F as well (see Lemma 1.2.18 in Dembo and Zeitouni, 1993). ∎

Relying on Lemma 1.2.18 in Dembo and Zeitouni (1993), since the lower bound in (2.9) holds for all open sets and the family of measures

$\{Q_{n,\nu}(\cdot) : n \geqslant 1\}$ is exponentially tight, then $I(\cdot)$ in (3.2) is a *good rate function*, that is, the level sets $\{\nu : I(\nu) \leqslant s\}$ are compact in the weak topology. Moreover, this property is carried out to the rate function $I_Z(\cdot)$ for the process M_n in (1.5).

Lemma 4.1. The family of measures $\{Q_{n,(x,y,z)}(\cdot) : n \geqslant 1\}$ is exponentially tight.

Proof. We shall prove that $\forall L \geqslant 1$, there exists a compact set $C_L \subset [0, 1) \times \mathbb{R}^2$ such that

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \ln Q_{n,(x,y,z)}(C_L^c) \leqslant -L$$

For each $\psi \in \mathscr{W}$ and $t > 0$ define the functional

$$\Psi_{t\psi}(\nu) = \exp\left[nt \int_S \psi \, d\nu\right], \qquad \nu \in \mathscr{M}_1(S)$$

Then

$$\Psi_{t\psi}(L_n(w,\cdot)) = \exp\left[t \sum_{j=0}^{n-1} \psi(V_j(w))\right]$$

where $(V_n)_{n\geqslant 0}$ is the random process in (2.6) and $L_n(w,\cdot)$ is defined in (2.4). Besides,

$$\begin{aligned} \mathbb{E}^{Q_{n,(x,y,z)}} \Psi_{t\psi}(\cdot) &= \int_{\mathscr{M}_1(S)} \exp\left[nt \int_S \psi \, d\nu\right] Q_{n,(x,y,z)}(d\nu) \\ &= \mathbb{E}_{(x,y,z)} \exp\left[t \sum_{j=0}^{n-1} \psi(V_j)\right] \end{aligned} \tag{4.1}$$

where $\mathbb{E}^{Q_{n,(x,y,z)}}$ is the expectation corresponding to the measure $Q_{n,(x,y,z)}$.

For each $\delta > 0$, define

$$A_\delta = \left\{\nu \in \mathscr{M}_1(S) : \int_S \psi \, d\nu \geqslant \delta\right\}$$

Using (4.1), we get

$$Q_{n,(x,y,z)}(A_\delta) \leqslant \exp\{-nt\delta\} \, \mathbb{E}_{(x,y,z)} \exp\left[t \sum_{j=0}^{n-1} \psi(V_j)\right] \tag{4.2}$$

Choose $\{K_m\}$ as a sequence of compact subsets of $\mathbb{R}$ for which $\eta(K_m^c) \to 0$ as $m \to +\infty$. Define $\tilde{K}_m = [0, 1) \times K_m^2$, $m \geqslant 1$. Clearly $\tilde{K}_m$ is a compact subset of S and

$$\begin{aligned}\tilde{K}_m^c &= [0, 1) \times (K_m^2)^c = [0, 1) \times \{(K_m^c)^2 \cup (K_m^c \times K_m) \cup (K_m \times K_m^c)\}\\ &\equiv [0, 1) \times (B_m^1 \cup B_m^2 \cup B_m^3)\end{aligned}$$

Let us introduce the functions $\psi_m = \mathscr{X}_{\tilde{K}_m^c}$, $m \geqslant 1$, and the sets

$$A_\delta^m = \left\{ v : \int_S \psi_m \, dv \geqslant \delta \right\} = \{ v : v(\tilde{K}_m^c) \geqslant \delta \}$$

Then, by using (4.2),

$$\begin{aligned}&Q_{n, (x, y, z)}(A_\delta^m)\\ &\quad\leqslant \sum_{i=1}^{3} Q_{n, (x, y, z)}(\{ v : v([0, 1) \times B_m^i) \geqslant \delta/3 \})\\ &\quad\leqslant \exp\{-nt\delta/3\} \sum_{i=1}^{3} \mathbb{E} \exp\left\{ t \sum_{j=0}^{n-1} \mathscr{X}_{B_m^i}(\xi_j, \xi_{j+1}) \right\}\\ &\quad\equiv \exp\{-nt\delta/3\} \sum_{i=1}^{3} I_i \qquad (4.3)\end{aligned}$$

where $\mathbb{E}$ is the expectation corresponding to the independent and identically distributed random process ξ_n, $\xi_0 = y$ and $\xi_1 = z$ with probability one.

One can see that

$$I_1 = \int_{\mathbb{R}^{n-1}} \exp\left\{ t \sum_{j=0}^{n-1} \mathscr{X}_{B_m^1}(z_j, z_{j+1}) \right\} \eta(dz_n) \cdots \eta(dz_2)$$

where $z_0 = y$ and $z_1 = z$. Since

$$\sum_{j=0}^{n-1} \mathscr{X}_{B_m^1}(z_j, z_{j+1}) \leqslant \sum_{j=0}^{n-1} \mathscr{X}_{K_m^c}(z_{j+1})$$

we get, for any $0 < \varepsilon < 1$ and for m large enough such that $z \in K_m$,

$$\begin{aligned}I_1 &\leqslant \left(\int_{\mathbb{R}} \exp\{ t\mathscr{X}_{K_m^c}(v) \} \, \eta(dv) \right)^{n-1}\\ &\leqslant [e^t \eta(\{ v : \mathscr{X}_{K_m^c}(v) \geqslant \varepsilon \}) + e^{t\varepsilon} \eta(\{ v : \mathscr{X}_{K_m^c}(v) < \varepsilon \})]^{n-1}\end{aligned}$$

Now, for each $t > 0$ we choose m so large that

$$e^t \eta(\{v : \mathscr{X}_{K_m^c}(v) \geqslant \varepsilon\}) < 1$$

This is possible because $\mathscr{X}_{K_m^c}(\cdot)$ converges to zero as m goes to infinity in η-measure. Besides, $0 < \varepsilon < 1$ being arbitrary, we choose ε so small that $e^{t\varepsilon} < 2$. Hence, for m large enough and depending on t, $I_1 \leqslant 3^n$.

Similarly, there exists m sufficiently large and depending on t such that $I_2 \leqslant 3^n$ and $I_3 \leqslant 3^n$. Therefore, $\forall t > 0$, $\exists m \equiv m_t > 0$ such that $I_1 + I_2 + I_3 \leqslant 9^n$ which implies by (4.3) that

$$Q_{n, (x, y, z)}(A_\delta^{m_t}) \leqslant \exp\{-nt\delta/3\}\ 9^n$$

Let $L \geqslant 1$. For $l \geqslant L$, take $\delta = 1/l$ and $t = 3l(l + \ln 9 + 1)$. Then, by writing $m_t \equiv m_l$

$$Q_{n, (x, y, z)}(A_{1/l}^{m_l}) \leqslant e^{-n(1+l)}, \qquad \forall n \geqslant 1, \quad \forall l \geqslant L \tag{4.4}$$

Let

$$C_L = \bigcap_{l \geqslant L} \left\{v : v(\tilde{K}_{m_l}) \geqslant 1 - \frac{1}{l}\right\} \subseteq \mathscr{M}_1(S)$$

This set is relatively compact. By Prohorov's theorem (see Appendix of Dembo and Zeitouni, 1993, page 319), C_L is compact in $\mathscr{M}_1(S)$. Since

$$C_L^c = \bigcup_{l \geqslant L} \{v : lv(\tilde{K}_{m_l}^c) > 1\}$$

we get, from (4.4),

$$Q_{n, (x, y, z)}(C_L^c) \leqslant e^{-nL}, \qquad \forall L \geqslant 1. \quad \blacksquare$$

5. LEVEL 1 LARGE DEVIATIONS

In this section we assume that T is a continuous transformation and η has compact support.

The rate function that governs large deviations for the means

$$M_n = \frac{1}{n} \sum_{j=0}^{n-1} Z_j Z_{j+1}, \qquad n = 1, 2, \ldots$$

introduced in (1.5) is obtained by using the level 2 large deviations for $\{V_n\}_n \geqslant 0$ in (1.6).

Since $Z_j = \Phi(X_j) + \xi_j$, we have

$$Z_j Z_{j+1} = [\Phi(X_j) + \xi_j][\Phi(T(X_j)) + \xi_{j+1}]$$

If $g: S \to \mathbb{R}$ is defined by

$$g(x, y, z) = [\Phi(x) + y][\Phi(T(x)) + z]$$

and taking into account (2.15) the ergodic theorem implies that

$$\begin{aligned}\lim_{n \to +\infty} M_n &= \int_S g(v)(\lambda \times \eta \times \eta)(dv)\\ &= \int_{[0,1)} \Phi(x)\, \Phi(T(x))\, dx + \int_{[0,1)} \int_{\mathbb{R}} \Phi(x)\, z\eta(dz)\, dx\\ &\quad + \int_{[0,1)} \int_{\mathbb{R}} \Phi(T(x))\, y\eta(dy)\, dx + \int_{\mathbb{R}} \int_{\mathbb{R}} yz\eta(dy)\, \eta(dz),\\ &\qquad \mathbb{P}_{\lambda \times \eta \times \eta}\text{-a.s. and } \mathbb{P}_{x,y,z}\text{-a.s.}, \quad \forall (x, y, z) \in S.\end{aligned}$$

In particular, if ξ_n has zero mean, then

$$\lim_{n \to +\infty} M_n = \int_{[0,1)} \Phi(u)\, \Phi(T(u))\, du, \qquad \mathbb{P}_{(x,y,z)}\text{-a.s.}, \quad \forall (x, y, z) \in S$$

In Sections 3 and 4 of this paper we established a full LDP for the family of distributions $Q_{n,(x,y,z)}(\cdot)$ of $L_n(w, \cdot)$. Taking into account (2.15) and using the Contraction Principle (see Ellis, 1985), the entropy function for $\{M_n\}_n \geqslant 1$ is given by

$$I_Z(r) = \inf_{\langle v, g \rangle = r} I(v) = \inf \left\{ I(v) : v \in \mathscr{M}_1(S), \int_S g(v)\, v(dv) = r \right\} \tag{5.1}$$

where $I(\cdot)$ is the level 2 entropy function for the process $\{V_n\}_{n \geqslant 0}$. Clearly $I_Z(r) = 0$ if and only if $r = \int_S g(v)(\lambda \times \eta \times \eta)(dv)$ because $I(v) = 0$ if and only if $v = \lambda \times \eta \times \eta$. If ξ_n has zero mean then $I_Z(r) = 0$ if and only if $r = \int_{[0,1)} \Phi(u)\, \Phi(T(u))\, du$.

6. SOME REMARKS

Remark 6.1. Large deviations for the empirical pair measures

$$\frac{1}{n}\sum_{j=0}^{n-1}\delta_{(\xi_j,\,\xi_{j+1})}(\cdot)$$

may be studied similarly to what was done in Sections 3 and 4. One can prove that its entropy function is given by

$$I^{(2)}(\nu)=\begin{cases}\displaystyle\iint_{\mathbb{R}^2}\ln\frac{m(x,y)}{m_1(x)}\,\nu(d(x,y)), & \text{if}\quad \nu\in\mathcal{M}\\ +\infty, & \text{otherwise,}\end{cases}$$

where

$$\mathcal{M}=\left\{\nu\in\mathcal{M}_1(\mathbb{R}^2):\pi_1\nu=\pi_2\nu,\ \nu\ll\eta\times\eta,\ \iint_{\mathbb{R}^2}\left|\ln\frac{m(x,y)}{m_1(x)}\right|\nu(d(x,y))<+\infty\right\}$$

with $m(x,y)\equiv(d\nu/(d\eta\times\eta))(x,y)$ and $m_1(x)\equiv\int_{\mathbb{R}}m(x,y)\,\eta(dy)$.

Remark 6.2. Let $Y_n=(X_n,\xi_n)$, $n\geqslant 0$, and consider the empirical measures

$$\frac{1}{n}\sum_{j=0}^{n-1}\delta_{Y_n}(\cdot)$$

Large deviations for the family of distributions of the above empirical measures is governed by the entropy function

$$I^{(2)}(\nu)=\begin{cases}\displaystyle\iint_{[0,1)\times\mathbb{R}}\ln m(x,y)\,\nu(d(x,y)), & \text{if}\quad \nu\in\mathcal{M}\\ +\infty, & \text{otherwise}\end{cases}$$

where

$$\mathcal{M}=\Big\{\nu\in\mathcal{M}_1([0,1)\times\mathbb{R}):\pi_1\nu=\lambda,\ \nu\ll\lambda\times\eta,\ m(x,y)=\frac{d\nu}{d\lambda\times\eta}(x,y),$$

$$\iint_{[0,1)\times\mathbb{R}}|\ln m(x,y)|\,\nu(d(x,y))<+\infty\Big\}$$

This result may be obtained similarly to Sections 3 and 4 of this paper.

Remark 6.3. One can generalize level 2 large deviations by considering the empirical pair measures corresponding to $Y_n = (X_n, \xi_n)$, $n \geqslant 0$. Let

$$\begin{aligned} L_n(w, \cdot) &= \frac{1}{n} \sum_{j=0}^{n-1} \delta_{(Y_j(w),\, Y_{j+1}(w))}(\cdot) \\ &= \frac{1}{n} \sum_{j=0}^{n-1} \delta_{(X_j(w),\, X_{j+1}(w),\, \xi_j(w),\, \zeta_{j+1}(w))}(\cdot) \end{aligned} \tag{6.1}$$

Clearly, for each $w \in S^{\mathbb{N}}$, $L_n(w, \cdot) \in \mathscr{M}_1(S)$, where $S = [0, 1)^2 \times \mathbb{R}^2$.

The ergodic theorem implies that

$$L_n(w, \cdot) \Rightarrow \bar{\lambda} \times \eta \times \eta, \qquad \mathbb{P}_{(x, y, z, t)}\text{-a.s.}, \quad \forall (x, y, z, t) \in S$$

where $\bar{\lambda}$ is a measure on $\mathscr{B}([0, 1)^2)$ defined by

$$\bar{\lambda}(A_1 \times A_2) = \lambda(A_1 \cap T^{-1}(A_2)), \qquad \forall A_1, A_2 \in \mathscr{B}([0, 1))$$

For each $\nu \in \mathscr{M}_1(S)$, we define the measure $\nu T^{-1} \in \mathscr{M}_1(S)$ by

$$\nu T^{-1}(A \times B \times C \times D) = \nu(T^{-1}(A) \times T^{-1}(B) \times C \times D)$$

for any measurable rectangle.

Let us define

$$\mathscr{M}_0 = \{ \nu \in \mathscr{M}_1(S) : \pi_{12} \nu = \bar{\lambda},\ \nu \ll \bar{\lambda} \times \eta \times \eta,\ \pi_{123} \nu = \pi_{124} \nu T^{-1} \}$$

If $\nu \in \mathscr{M}_0$, let $m(x, y, z, t)$ be the density of ν with respect to $\bar{\lambda} \times \eta \times \eta$, $m_{123}(x, y, z)$ be the marginal density of $\pi_{123} \nu$ with respect to $\bar{\lambda} \times \eta$ and $m_{124}(x, y, t)$ be the marginal density of $\pi_{124} \nu$ with respect to $\bar{\lambda} \times \eta$. The definition of $\pi_{124} \nu T^{-1}$ tells us that

$$\frac{d\pi_{124} \nu T^{-1}}{d\bar{\lambda} \times \eta}(x, y, z) = m_{124}(T^{-1}(x), T^{-1}(y), z), \qquad (x, y, z) \in [0, 1)^2 \times \mathbb{R}$$

Moreover, from the condition $\pi_{123} \nu = \pi_{124} \nu T^{-1}$, we have

$$m_{123}(x, y, z) = m_{124}(T^{-1}(x), T^{-1}(y), z), \qquad (x, y, z) \in [0, 1)^2 \times \mathbb{R}$$

One can prove, as in Sections 3 and 4 of this paper, that the level 2 large deviations for $L_n(w, \cdot)$ in (6.1) is governed by the entropy function

$$I^{(2)}(\nu) = \begin{cases} \int_S \ln \dfrac{m}{m_{123}}\, d\nu, & \text{if} \quad \nu \in \mathscr{M}_0 \quad \text{and} \quad \int_S \left| \ln \dfrac{m}{m_{123}} \right| d\nu < +\infty \\ +\infty, & \text{otherwise} \end{cases}$$

Remark 6.4. Returning to the process $\{V_n\}_n \geqslant 0$ in (1.6), let us define

$$\Lambda(\psi) = \varlimsup_{n \to +\infty} \frac{1}{n} \ln \left(\sup_{v \in S} \mathbb{E}_v \exp \left\{ \sum_{j=0}^{n-1} \psi(V_j) \right\} \right)$$

where $\mathbb{E}_v$ is the expectation corresponding to the measure $\mathbb{P}_v$ on $(S^{\mathbb{N}}, \sigma(\mathscr{C}))$, introduced in (2.2). Let $B(S : \mathbb{R})$ be the set of bounded measurable real functions.

Let

$$\Lambda^*(\nu) = \sup \left\{ \int_S \psi \, d\nu - \Lambda(\psi) : \psi \in B(S : \mathbb{R}) \right\}$$

By Lemma 4.1.36 in Deuschel and Stroock (1989), $\Lambda^*(\nu) = I(\nu)$, $\nu \in \mathscr{M}_1(S)$, where $I(\nu)$ is defined in (2.11).

Remark 6.5. The results for random means $\{M_n\}_{n \geqslant 1}$ in (1.5) may be extended to

$$M_n \equiv \frac{1}{n} \sum_{j=0}^{n-k} Z_j Z_{j+k}, \qquad n \geqslant k, \quad k \geqslant 1 \tag{6.2}$$

When $\{Z_n\}_{n \geqslant 0}$ has zero mean they are called the *autocovariances of order k of the process Z_n*.

The level 1 LDP for the random means (6.2) follows from the level 1 LDP for the corresponding autocovariances of order 1, in (1.5). To see this, let us consider first the case $k = 2$. One can verify that the process $\{M_n\}_{n \geqslant 2}$ in (6.2) has the same distribution as the process

$$a_n^{(1)} M_n^{(1)} + a_n^{(2)} M_n^{(2)}, \qquad n \geqslant 1$$

where

$$M_n^{(1)} \equiv \frac{1}{n} \sum_{j=0}^{n-1} Y_j Y_{j+1} \qquad \text{and} \qquad M_n^{(2)} \equiv \frac{1}{n} \sum_{j=0}^{n-1} W_j W_{j+1}$$

$\{Y_n\}_{n\geqslant 0}$ and $\{W_n\}_{n\geqslant 0}$ are independent random sequences with the same distribution as the process $\{Z_n\}_{n\geqslant 0}$ given by (1.5), with T^2 instead of T, since T^2 is a uniquely ergodic transformation, where T is given by (1.1). The sequences $\{a_n^{(1)}\}_{n\geqslant 1}$ and $\{a_n^{(2)}\}_{n\geqslant 1}$ are real sequences converging to $\frac{1}{2}$ as n goes to infinity.

Since $\{a_n^{(1)}\}_{n\geqslant 1}$ and $\{a_n^{(2)}\}_{n\geqslant 1}$ are deterministic sequences their entropy function is

$$\tilde{I}(r)=\begin{cases}0, & \text{if} \quad r=\frac{1}{2}\\ +\infty, & \text{if} \quad r\neq\frac{1}{2}\end{cases}$$

The level 1 entropy functions for $M_n^{(1)}$ and $M_n^{(2)}$ are equal and coincide with $I_Z(r)$ in (5.1).

Relying on the independence of the sequences $a_n^{(1)}$, $a_n^{(2)}$, $M_n^{(1)}$, $M_n^{(2)}$ and using the Contraction Principle (see Dembo and Zeitouni, 1993), we obtain the level 1 entropy function for $a_n^{(1)}M_n^{(1)}$ (which is the same for $a_n^{(2)}M_n^{(2)}$):

$$I^{(1)}(u)=\inf_{s\in\mathbb{R}}\left\{\tilde{I}(1/2)+I_Z(s):\frac{s}{2}=u\right\}=I_Z(2u), \qquad u\in\mathbb{R}$$

Hence, the level 1 entropy function for $a_n^{(1)}M_n^{(1)}+a_n^{(2)}M_n^{(2)}$ is given by

$$\begin{aligned}I_Z^{(2)}(t)&=\inf_{u,\,v\in\mathbb{R}}\{I^{(1)}(u)+I^{(1)}(v):u+v=t\}\\&=\inf_{u,\,v\in\mathbb{R}}\{I_Z(2u)+I_Z(2v):v=t-u\}\\&=\inf_{u\in\mathbb{R}}\{I_Z(2u)+I_Z(2(t-u))\}, \qquad \text{for} \quad t\in\mathbb{R}\end{aligned}$$

Similarly, for each $k\geqslant 1$, the level 1 entropy function for $\{M_n\}_{n\geqslant k}$ in (6.2) is

$$I_Z^{(k)}(t)=\inf_{u_1,\dots,u_k}\left\{\sum_{i=1}^{k}I_Z(ku_i):\sum_{i=1}^{k}u_i=t\right\}, \qquad \text{for} \quad t\in\mathbb{R}$$

ACKNOWLEDGMENTS

This work was partially supported by CNPq and PRONEX: “Fenômenos Críticos em Probabilidade e Processos Estocásticos” and “Sistemas Dinâmicos.”

REFERENCES

P. Billingsley, *Probability and Measure* (John Wiley, 3rd ed., New York, 1995).

Z. Coelho, A. Lopes, and L. F. C. Rocha, Absolutely Continuous Invariant Measures for a Class of Affine Interval Exchange Maps, *Proceedings of the American Mathematical Society* **123**:3533–3542 (1994).

A. Dembo and O. Zeitouni, *Large Deviations Techniques* (Jones and Bartlett, Boston, 1993).

J.-D. Deuschel and D. W. Stroock, *Large Deviations* (Academic Press, Boston, 1989).

M. D. Donsker and S. R. S. Varadhan, Asymptotic Evaluation of Certain Markov Process Expectations for Large Time, I, *Communications on Pure and Applied Mathematics* **28**:1–47 (1975a).

M. D. Donsker and S. R. S. Varadhan, Asymptotic Evaluation of Certain Wiener Integrals for Large Time, in *Functional Integration and its Applications. Proc. of the International Conference*, A. M. Arthurs, ed. (Clarendon Press, Oxford, 1975b), pp. 15–33.

J. L. Doob, *Stochastic Processes* (John Wiley, New York, 1953).

R. Durrett, *Probability: Theory and Examples* (Duxbury Press, 2nd ed., Boston, 1996).

R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics* (Springer-Verlag, New York, 1985).

A. Lopes and S. Lopes, Parametric Estimation and Spectral Analysis of Piecewise Linear Maps of the Interval, preprint (1995). To appear in *Advances in Applied Probability*, December, 1998.

A. Lopes and S. Lopes, Unique Ergodicity, Large Deviations and Parametric Estimation, submitted (1996).

A. Lopes and L. F. C. Rocha, Invariant Measure for the Gauss Map Associated with Interval Exchange Maps, *Indiana University Mathematics Journal* **43**:1399–1438 (1994).

W. Rudin, *Real and Complex Analysis* (McGraw Hill, 2nd ed., New York, 1974).