Elect. Comm. in Probab. **14** (2009), 540–551

ELECTRONIC COMMUNICATIONS in PROBABILITY

IDENTIFICATION OF THE RATE FUNCTION FOR LARGE DEVIATIONS OF AN IRREDUCIBLE MARKOV CHAIN

WEI LIU
Department of Mathematics, Wuhan University, 430072-HUBEI, China
email: Liuwei830505@hotmail.com

LIMING WU
Laboratoire de Mathématiques Appliquées, CNRS-UMR 6620, Université Blaise Pascal, 63177 Aubière, France. And Institute of Applied Mathematics, Chinese Academy of Sciences, 100190 Beijing, China
email: Li-Ming.Wu@math.univ-bpclermont.fr

Submitted May 26, 2009, *accepted in final form* November 16, 2009

AMS 2000 Subject classification: 60F10, 60J05
Keywords: Large deviations, irreducible Markov processes, Feynman-Kac semigroups

Abstract

For an irreducible Markov chain $(X_n)_{n\geq 0}$ we identify the rate function governing the large deviation estimation of empirical mean $\frac{1}{n}\sum_{k=0}^{n-1} f(X_k)$ by means of the Donsker-Varadhan's entropy. That allows us to obtain the lower bound of large deviations for the empirical measure $\frac{1}{n}\sum_{k=0}^{n-1}\delta_{X_k}$ in full generality.

1 Introduction

Large deviations of Markov processes were opened by Donsker-Varadhan [8] (1975-1983) under strong regularity conditions. Generalizations of their fundamental works are very numerous and various. In this paper we are interested in general irreducible Markov processes. In this direction the first general results were obtained by Ney and Nummelin [14] (1987). de Acosta [5] (1988) derived a universal lower bound for bounded additive functionals valued in a separable Banach space, and the boundedness condition was finally removed by de Acosta-Ney [6] (1998) : that is a definite work for the lower bound. For related works see also Jain [10] (1990) and some recent works [17, 18, 20] by the second named author, and Kontoyiannis-Meyn [11] and Meyn [12]. However the rate function, a basic object describing the exact exponential rate in a large deviation estimation, is expressed by means of the *convergence parameter*, a quantity proper to the theory of irreducible Markov processes. It is not related to the Donsker-Varadhan entropy, unlike the work of Deuschel-Stroock [7] in the more classical strong mixing case.

The main purpose of this note is to identify the rate function appearing in the large deviation results in [14, 5, 6] etc by means of the Donsker-Varadhan entropy. The key is to prove the lower semi-continuity of the Cramer functional associated with the convergence parameter of the

Feynman-Kac semigroup. This result is based mainly on the increasing continuity of the convergence parameter due to de Acosta [5].

This paper is organized as follows. In the next section we present necessary backgrounds and calculate the Legendre transform of the Cramer functional. We prove in section 3 the lower semi-continuity of the Cramer functional, which gives us the desired identification of the rate function of de Acosta-Ney [6] and as corollary we obtain the lower bound of large deviations for the occupation measures. Finally in section 4 we present some remarks about the upper bound of Ney-Nummelin [14] and provide a very simple counter-example to the upper bound of large deviations for unbounded additive functionals.

2 Preliminaries and Legendre transform of the Cramer functional

2.1 Preliminaries

We recall some known facts about irreducible Markov chains from Nummelin [15] and Meyn-Tweedie [13].

Let $K(x,dy)$ be a nonnegative kernel on a measurable space $(E,\mathscr{B})$ where $\mathscr{B}$ is countably generated. It is said *irreducible* if there is some non-zero nonnegative σ-finite measure μ such that

$$\forall A \in \mathscr{B} \text{ with } \mu(A) > 0,\ \sum_{n=1}^{+\infty} K^n(x,A) > 0, \text{ for } \textbf{all} \ \ x \in E. \tag{1}$$

Such measure μ is said to be a *maximal* irreducible measure of K, if $\mu K \ll \mu$. All maximal irreducible measures of K are equivalent. Below μ is some fixed maximal irreducible measure of K.

A couple (s,ν) where $s \geq 0$ with $\mu(s) := \int_E s(x)d\mu(x) > 0$ and ν a probability measure on E is said *K-small*, if there is some $m_0 \geq 1$ and constant $c > 0$ such that

$$K^{m_0}(x,A) \geq cs \otimes \nu(x,A) = cs(x)\nu(A),\ x \in E, A \in \mathscr{B} \tag{2}$$

A real measurable function $s \geq 0$ with $\mu(s) > 0$ is said to be K-small, if there is some probability measure such that (s,ν) is K-small. A subset $A \in \mathscr{B}$ with $\mu(A) > 0$ is said K-small if 1_A does.

By Nummelin [15, Theorem 2.1], any irreducible kernel K has always such a small couple (s,ν).

A non-empty set $F \in \mathscr{B}$ is said *K-closed*, if $K(x,F^c) = 0$ for all $x \in F$. For every K-closed F, $\mu(F^c) = 0$ by the irreducibility of K.

According to Nummelin [15, Definition 3.2, Proposition 3.4 and 4.7], the *convergence parameter* of K, say $R(K)$, is given by : for every K-small couple (s,ν),

$$\begin{aligned} R(K) &= \sup\left\{ r \geq 0 : \sum_{n=0}^{+\infty} r^n \nu K^n s < +\infty \right\} \\ &= \sup\left\{ r \geq 0 : (\exists K\text{-closed } F)(\forall K\text{-small } s), \sum_{n=0}^{+\infty} r^n K^n s < +\infty, x \in F \right\}. \end{aligned} \tag{3}$$

It is well known that $0 \leq R(K) < +\infty$.

Lemma 2.1. *For every K-small couple (s,ν),*

$$\begin{aligned} -\log R(K) &= \limsup_{n\to\infty}\frac{1}{n}\log \nu K^n s \\ &= \inf_{K-closed\ F}\sup_{x\in F}\limsup_{n\to\infty}\frac{1}{n}\log K^n s(x) \\ &= \mathrm{esssup}_{x\in E}\limsup_{n\to\infty}\frac{1}{n}\log K^n s(x). \end{aligned} \tag{4}$$

Here and hereafter $\mathrm{esssup}_{x\in E}$ is taken always w.r.t μ.

Its proof is quite easy, so omitted.

2.2 Cramer functional

Let $(X_n)_{n\geq 0}$ be a Markov chain valued in E, defined on $(\Omega,\mathscr{F},(\mathscr{F}_n),(\mathbb{P}_x)_{x\in E})$. Throughout this paper we assume always that its transition kernel $P(x,dy)$ is irreducible, and μ is a fixed maximal irreducible measure of P.

For every $V\in r\mathscr{B}$ (the space of all real $\mathscr{B}$-measurable functions on E), consider the kernel $P^V(x,dy):=e^{V(x)}P(x,dy)$. We have the following Feynman-Kac formula,

$$(P^V)^n f(x) = \mathbb{E}^x f(X_n)\exp\left(\sum_{k=0}^{n-1}V(X_k)\right),\ 0\leq f\in r\mathscr{B}. \tag{5}$$

It is obvious that P^V is irreducible with the maximal irreducible measure μ.

Define now our Cramer functional

$$\Lambda(V) = -\log R(P^V). \tag{6}$$

Since $R(P^V)\in[0,+\infty)$ by [15, Theorem 3.2], $\Lambda(V)>-\infty$. By (5) and Lemma 2.1, we have

$$\begin{aligned} \Lambda(V) &= \limsup_{n\to\infty}\frac{1}{n}\log \mathbb{E}^\nu s(X_n)\exp\left(\sum_{k=0}^{n-1}V(X_k)\right) \\ &= \mathrm{esssup}_{x\in E}\limsup_{n\to\infty}\frac{1}{n}\log \mathbb{E}^x s(X_n)\exp\left(\sum_{k=0}^{n-1}V(X_k)\right) \end{aligned} \tag{7}$$

for every P^V-small couple (s,ν). From the above expression we see by Hölder's inequality that $\Lambda: r\mathscr{B}\to(-\infty,+\infty]$ is convex.

We can now recall the universal lower bound of large deviation in de Acosta [5] and de Acosta-Ney [6].

Theorem 2.2. ([5, 6]) *Let $f:E\to\mathbb{B}$ be a measurable function with values in a separable Banach space $(\mathbb{B},\|\cdot\|)$. Then for every open subset G of $\mathbb{B}$ and every initial measure ν,*

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}_\nu\left(\frac{1}{n}\sum_{k=0}^{n-1}f(X_k)\in G\right)\geq -\inf_{z\in G}\Lambda_f^*(z), \tag{8}$$

where $\Lambda_f(y):=\Lambda(\langle y,f\rangle),\ y\in\mathbb{B}'$ (the topological dual space of $\mathbb{B}$), and

$$\Lambda_f^*(z) = \sup_{y\in\mathbb{B}'}(\langle y,z\rangle - \Lambda_f(y)). \tag{9}$$

Here $\langle y,z\rangle$ denotes the duality bilinear relation between $\mathbb{B}$ and $\mathbb{B}'$.

The main objective is to identify Λ_f^* by means of the Donsker-Varadhan entropy.

2.3 Legendre transform of the Cramer functional on a weighted space

At first we recall the Donsker-Varadhan's entropy $J : \mathscr{M}_1(E) \to [0,+\infty]$, where $\mathscr{M}_1(E)$ is the space of all probability measures on $(E,\mathscr{B})$. For any $\nu \in \mathscr{M}_1(E)$,

$$J(\nu) = \sup_{u\in\mathscr{U}} \int_E \log\frac{u}{Pu} d\nu, \tag{10}$$

where $\mathscr{U} = \{u \in b\mathscr{B} : \inf_{x\in E} u(x) > 0\}$ ($b\mathscr{B}$ being the space of all real and bounded measurable functions on $(E,\mathscr{B})$). Consider the modified Donsker-Varadhan's entropy ([17])

$$J_\mu(\nu) = \begin{cases} J(\nu), & \nu \ll \mu; \\ +\infty, & \text{otherwise.}\end{cases} \tag{11}$$

Let us now fix some reference measurable function $\Phi : E \to \mathbb{R}$ such that $\Phi \ge 1$ everywhere. Introduce the weighted functions space

$$b_\Phi\mathscr{B} := \left\{V : E \to \mathbb{R}\ \mathscr{B}-measurable;\ \|V\|_\Phi := \sup_{x\in E}\frac{|V(x)|}{\Phi(x)} < +\infty\right\}.$$

Let

$$M_{b,\Phi}(E) := \left\{\text{(maybe signed) measure } \nu \text{ on } E;\ \int \Phi d|\nu| < +\infty\right\}$$

where $|\nu| = \nu^+ + \nu^-$ (ν^+, ν^- are respectively the positive and negative part of ν in Hahn-Jordan's decomposition). If $\Phi = 1$, we write simply $\mathscr{M}_b(E)$ for $M_{b,\Phi}(E)$. Obviously

$$\nu(V) = \int_E V d\nu =: \langle \nu, V\rangle$$

is a bilinear form on $b_\Phi\mathscr{B} \times M_{b,\Phi}(E)$. The main result of this section is

Theorem 2.3. *Let $\Phi \ge 1$ be some reference measurable function on E. Let $\Lambda^{*,\Phi} : M_{b,\Phi}(E) \to \mathbb{R}\bigcup\{+\infty\}$ be the Legendre transform of Λ on $b_\Phi\mathscr{B}$, i.e.,*

$$\Lambda^{*,\Phi}(\nu) := \sup\left\{\langle \nu, V\rangle - \Lambda(V);\ V \in b_\Phi\mathscr{B}\right\}. \tag{12}$$

Then we have for every $\nu \in M_{b,\Phi}(E)$,

$$\Lambda^{*,\Phi}(\nu) = \begin{cases} J_\mu(\nu), & \text{if } \nu \in \mathscr{M}_1(E) \\ +\infty, & \text{otherwise.}\end{cases} \tag{13}$$

We begin with

Lemma 2.4. (de Acosta [5, lemma 6.1]) *Define for* $\nu \in \mathscr{M}_1(E)$,

$$J'(\nu) = \sup\left\{\int \log \frac{u}{Pu} d\nu : u \in \mathscr{B},\ u > 0\ everywhere,\ \nu(\{u = \infty\}) = 0,\ (\log(\frac{u}{Pu}))^- \in L^1(\nu)\right\}.$$

Then $J'(\nu) = J(\nu)$ *for all* $\nu \in \mathscr{M}_1(E)$.

Proof of Theorem 2.3. **Step 1:** $\geq$ **in (13).** This can be derived from [17, Proposition B.9], but we will follow the argument below proposed by the referee. At first since $\Lambda(V) = 0$ if $V = 0, \mu - a.s.$, one deduces easily that $\Lambda^{*,\Phi}(\nu) = +\infty$ for all ν not absolutely continuous w.r.t. μ.
Fix now $\nu \ll \mu$. For each $u \in b\mathscr{B}$ let $V := \log \frac{u}{Pu}$. Since $P^V u = u$, we see that $\Lambda(V) \leq 0$ by Lemma 2.1. Hence

$$\int \log \frac{u}{Pu} d\nu \leq \int V d\nu - \Lambda(V) \leq \Lambda^{*,\Phi}(\nu).$$

Taking the supremum over all $1 \leq u \in b\mathscr{B}$ and recalling that $J(\nu) = \sup_{1\leq u \in b\mathscr{B}} \int \log \frac{u}{Pu} d\nu$ (Donsker-Varadhan's formula), we get the desired result.
Step 2: $\Lambda^{*,\Phi}(\nu) \leq J_\mu(\nu)$ **for** $\nu \in M_1(E)$**.** The argument below follows quite closely that of de Acosta [5, Proof of (6.3)].
We can assume that $J_\mu(\nu) < +\infty$ (trivial otherwise). Then $\nu \ll \mu$. For every $V \in r\mathscr{B}$ with $\int V^- d\nu < +\infty$, and $\lambda > \Lambda(V)$, we take

$$u := \sum_{k=0}^{+\infty} e^{-\lambda k}(P^V)^k s(x),$$

where s is a P^V-small function. Then we have $\nu([u = +\infty]) = 0$ and

$$u > s, \qquad e^{-\lambda} e^V Pu = u - s.$$

Noting that

$$\log \frac{u}{Pu} \geq \log e^{-\lambda + V} = -\lambda + V \geq -\lambda - V^- \in L^1(\nu),$$

by de Acosta's Lemma 2.4 we have

$$\begin{aligned}
\int V d\nu - J(\nu) &\leq \int V d\nu + \int \log \frac{Pu}{u} d\nu \\
&= \lambda + \int \log \frac{e^{-\lambda+V} Pu}{u} d\nu \\
&= \lambda + \int \log \frac{u-s}{u} d\nu \\
&< \lambda.
\end{aligned}$$

As $\lambda > \Lambda(V)$ is arbitrary we have so proved

$$\int V d\nu \leq J_\mu(\nu) + \Lambda(V), \ \text{ if } \ V^- \in L^1(\nu),\ \nu \in \mathscr{M}_1(E). \tag{14}$$

It yields the desired claim since $V \in b_\Phi \mathscr{B} \implies V \in L^1(\nu)$ for $\nu \in M_{b,\Phi}(E)$.
Combining Step 1 and Step 2, we obtain (13). □

As an application, we have the following result.

Corollary 2.5. *Let V be a nonnegative measurable function. If there exists $\delta \ge 0$ such that $\Lambda(\delta V) < +\infty$, then*

$$J_\mu(\nu) < +\infty \implies \int_E V d\nu < +\infty.$$

Proof. By Theorem 2.3, we have

$$J_\mu(\nu) \ge \int \delta(V \wedge n) d\nu - \Lambda(\delta(V \wedge n)), \ \forall n \ge 1.$$

Thus we have

$$\begin{aligned}\delta \int V \wedge n d\nu &\le J_\mu(\nu) + \Lambda(\delta(V \wedge n)) \\ &\le J_\mu(\nu) + \Lambda(\delta V) < +\infty\end{aligned}$$

Letting $n \to +\infty$, we get the result. □

3 Identification of the rate function

Let $L_n := \frac{1}{n}\sum_{k=0}^{n-1} \delta_{X_k}$ be the empirical measure. Assuming the existence of invariant probability measure μ, then for every $\nu \ll \mu$, $\mathbb{P}_\nu(L_n \in \cdot)$ satisfies the weak* LDP on $(\mathscr{M}_1(E), \tau)$ with rate function J_μ (by [17, Theorem B.1, Theorem B.5 and Proposition B.9]), where τ is the topology $\sigma(\mathscr{M}_b(E), b\mathscr{B})$ restricted to $\mathscr{M}_1(E)$. Then for a measurable function $f : E \to \mathbb{B}$, inspired by the contraction principle, the rate function governing the LDP of

$$\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) = \int_E f(x) dL_n(x)$$

should be

$$J_\mu^f(z) = \inf\{J_\mu(\nu) : \int_E \|f\| d\nu < +\infty, \ \nu(f) = z\}, \ z \in \mathbb{B}. \tag{15}$$

That is not completely exact, but not far. The main result of this paper is

Theorem 3.1. *For any measurable function $f : E \to \mathbb{B}$ where $\mathbb{B}$ is a separable Banach space, the rate function Λ_f^* given in Theorem 2.2 (due to de Acosta and Ney [6]) is exactly the lower semi-continuous regularization $\widetilde{J}_\mu^f$ of $z \to J_\mu^f(z)$.*

It is based on the following lower semi-continuity of Λ which is of independent interest.

Proposition 3.2. *Let $\Phi : E \to [1, +\infty)$ be a measurable function. Then $\Lambda : b_\Phi\mathscr{B} \to (-\infty, +\infty]$ is lower semi-continuous w.r.t. the weak topology $\sigma(b_\Phi\mathscr{B}, \mathscr{M}_{b,\Phi}(E))$, or equivalently for any $V \in b_\Phi\mathscr{B}$,*

$$\Lambda(V) = \sup\left\{\langle \nu, V\rangle - J_\mu(\nu); \ \nu \in \mathscr{M}_1(E), \nu(\Phi) < +\infty\right\}. \tag{16}$$

As $\Lambda(V)$ is the rate of the exponential growth of the Feynmann-Kac semigroup $(P^V)^n$, its identification is a fundamental subject in heat theory : formula of type (16) is the counterpart of the famous Rayleigh principle (for the maximal eigenvalue of $\mathscr{L}+V$ where $\mathscr{L}$ is the generator of a symmetric Markov semigroup) and it follows from the LDP of Donsker-Varadhan (when this last holds) by Varadhan's Laplace integral lemma.

Proof of Theorem 3.1 assuming Proposition 3.2. Let $\Phi(x) := \|f(x)\|+1$. For every $y \in \mathbb{B}'$ (the dual space of $\mathbb{B}$), $V(x) := \langle y, f(x)\rangle \in b_\Phi \mathscr{B}$. Then by Proposition 3.2,

$$\Lambda_f(y) = \Lambda(V) = \sup\left\{\int_E \langle y, f(x)\rangle d\nu - J_\mu(\nu);\ \nu \in \mathscr{M}_1(E), \nu(\|f\|) < +\infty\right\}$$
$$= \sup\left\{\langle y, z\rangle - J_\mu^f(z);\ z \in \mathbb{B}\right\} = (J_\mu^f)^*(y).$$

Let us prove that J_μ^f is convex on $\mathbb{B}$. We have only to prove that $J_\mu^f(z) \le [J_\mu^f(z_1) + J_\mu^f(z_2)]/2$ for any $z, z_1, z_2 \in \mathbb{B}$ such that $z = (z_1+z_2)/2$. We may assume that $J_\mu^f(z_k) < +\infty,\ k = 1,\ 2.$
Given any $\varepsilon > 0$, there exist ν_1 and ν_2 such that

$$\nu_k(\|f\|) < +\infty,\ \nu_k(f) = z_k,\ J_\mu(\nu_k) \le J_\mu^f(z_k) + \varepsilon \ \text{ for }\ k = 1,\ 2.$$

Let $\nu = \frac{\nu_1+\nu_2}{2}$, then $\nu(f) = z$. Because $J_\mu(\nu)$ is convex in $\nu \in \mathscr{M}_1(E)$, we have

$$J_\mu(\nu) \le \frac{J_\mu(\nu_1) + J_\mu(\nu_2)}{2} \le \frac{J_\mu^f(z_1) + J_\mu^f(z_2)}{2} + \varepsilon$$

Hence

$$J_\mu^f(z) \le \frac{J_\mu^f(z_1) + J_\mu^f(z_2)}{2} + \varepsilon$$

which completes the proof of the convexity of J_μ^f, for $\varepsilon > 0$ is arbitrary.
Consequently by the famous Fenchel's theorem in convex analysis,

$$\Lambda_f^* = (J_\mu^f)^{**}$$

is the lower semi-continuous regularization $\widetilde{J}_\mu^f$ of $z \to J_\mu^f$. □

Proposition 3.2 is mainly based on the following continuity of the convergence parameter due to de Acosta.

Lemma 3.3. (de Acosta [5, Theorem 2.1]) *For each $j \ge 1$ let $E_j \in \mathscr{B}$ and assume $E_j \uparrow E$. Let $\mathscr{B}_j = \{A \in \mathscr{B} : A \subset E_j\}$. Let K be an irreducible kernel on $E_j \in \mathscr{B}_j$ and for $j \ge 1$, let K_j be an irreducilbe kernel on $E_j \in \mathscr{B}_j$. Assume that for all $x \in E, A \in \mathscr{B}, K_j(x, A\cap E_j) \uparrow K(x,A)$. Then $R(K_j) \downarrow R(K)$.*

We also require the following general result.

Lemma 3.4. *Let μ be a σ-finite measure on $(E, \mathscr{B})$ and $\Lambda : b_\Phi\mathscr{B} \to (-\infty, +\infty]$ be a convex function such that*

(i) If $V_1 = V_2,\ \mu - a.e.$, then $\Lambda(V_1) = \Lambda(V_2)$;

(ii) If $V_1 \le V_2$, then $\Lambda(V_1) \le \Lambda(V_2)$.

Then Λ is lower semi-continuous (l.s.c. in short) w.r.t. the weak topology $\sigma(b_\Phi\mathscr{B}, \mathscr{M}_{b,\Phi}(E))$ iff for every non-decreasing sequence (V_n) in $b_\Phi\mathscr{B}$ such that $V(x) = \sup_n V_n(x) \in b_\Phi\mathscr{B}$,

$$\Lambda(V_n) \to \Lambda(V).$$

Proof. The necessity is obvious because if $V_n \uparrow V$ in $b_\Phi\mathscr{B}$, then $V_n \to V$ in $\sigma(b_\Phi\mathscr{B}, \mathscr{M}_{b,\Phi}(E))$ by dominated convergence. Let us prove the sufficiency.
Consider the isomorphism $V \to V/\Phi$ from $b_\Phi\mathscr{B}$ to $b\mathscr{B}$ and define

$$\tilde{\Lambda}(U) := \Lambda(U\Phi), \ \forall U \in b\mathscr{B}.$$

Then $\tilde{\Lambda}$ satisfies again (i) and (ii) on $b\mathscr{B}$. By the isomorphism above, the l.s.c. of Λ on $b_\Phi\mathscr{B}$ w.r.t. $\sigma(b_\Phi\mathscr{B}, \mathscr{M}_{b,\Phi}(E))$ is equivalent to the l.s.c. of $\tilde{\Lambda}$ on $b\mathscr{B}$ w.r.t. $\sigma(b\mathscr{B}, \mathscr{M}_b(E))$. In other words we may assume without loss of generality that $\Phi = 1$.
Because of condition (i), Λ is well defined on $L^\infty(\mu)$ (which is the dual of $L^1(\mu)$), and the l.s.c. of Λ on $b\mathscr{B}$ w.r.t. $\sigma(b\mathscr{B}, \mathscr{M}_b(E))$ is equivalent to that of Λ on L^∞ w.r.t. $\sigma(L^\infty, L^1)$.
Below we prove the l.s.c. of Λ on $L^\infty(\mu)$ w.r.t. $\sigma(L^\infty, L^1)$. By taking an equivalent measure if necessary we may assume that μ is a probability measure. For any $L \in \mathbb{R}$, since $[\Lambda \le L]$ is convex, by the Krein-Smulyan theorem(see [4], page 163), it is closed in L^∞ w.r.t. $\sigma(L^\infty, L^1)$ iff

$$[\Lambda \le L] \bigcap B(R)$$

is closed for every $R > 0$, where $B(R) := \{V \in L^\infty(\mu);\ \|V\|_\infty \le R\}$. Since $B(R)$ equipped with the weak*-topology $\sigma(L^\infty, L^1)$ is metrizable (see [2, Chap.IV, p.111]), for the desired l.s.c., it remains to prove that if $V_n \to V$ in $\sigma(L^\infty, L^1)$ and $\sup_n \|V_n\|_\infty \le R < +\infty$, then

$$\liminf_{n\to\infty} \Lambda(V_n) \ge \Lambda(V).$$

Taking a subsequence if necessary, we may assume that $l := \lim_{n\to\infty} \Lambda(V_n)$ exists in $[-\infty, +\infty]$.
As μ is a probability measure, $V_n \to V$ in the weak topology of $L^2(\mu)$. By the Mazur theorem ([22, Chap. V, §1, Theorem 2]), there exists a sequence (U_n), each U_n is a convex combination U_n of $\{V_k, k \ge n\}$ such that $U_n \to V$ in $L^2(\mu)$. By the convexity of Λ,

$$\limsup_{n\to\infty} \Lambda(U_n) \le \lim_{n\to\infty} \Lambda(V_n) = l.$$

Now taking a subsequence U_{n_k} which converges to V, $\mu - a.e.$ and set $W_k = \inf_{j\ge k} U_{n_j}$. Then $W_k \uparrow W$ and $W = V$, $\mu - a.e.$. By condition (ii), $\Lambda(W_k) \le \Lambda(U_{n_k})$. By the assumed increasing continuity of Λ we get finally

$$l \ge \limsup_{k\to\infty} \Lambda(U_{n_k}) \ge \lim_{k\to\infty} \Lambda(W_k) = \Lambda(W) = \Lambda(V).$$

□

Proof of Proposition 3.2. By Theorem 2.3, the right hand side (r.h.s.) of (16) is exactly Λ^{**}, the double Legendre transform of Λ basing on the duality between $b_\Phi\mathscr{B}$ and $\mathscr{M}_{b,\Phi}(E)$. By Fenchel's theorem, (16) is equivalent to the l.s.c. of Λ w.r.t. $\sigma(b_\Phi\mathscr{B}, \mathscr{M}_{b,\Phi}(E))$, which holds true by Lemma 3.4 and de Acosta's Lemma 3.3. □

We end this section by two corollaries. We begin with the rate function governing the lower bound of large deviation of the occupation measure $L_n := \frac{1}{n}\sum_{k=0}^{n-1} \delta_{X_k}$.
Let $\mathscr{A}$ be the σ-algebra on $\mathscr{M}_1(E)$ generated by $\nu \to \nu(V)$, $V \in b\mathscr{B}$.

Corollary 3.5. *For every initial measure ν and any τ-open subset $G \in \mathscr{A}$,*

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_\nu(L_n \in G) \geq - \inf_{\beta\in G} J_\mu(\beta). \tag{17}$$

This corollary was proved in de Acosta [5, Theorem 6.3] under an extra assumption guaranteeing $J_\mu = J$. This last assumption is also crucial in Jain's work [10], but difficult to verify out of the classical absolutely continuous framework of Donsker-Varadhan [8].
If one uses the Donsker-Varadhan's entropy J instead of J_μ in (17), the lower bound (17) would be wrong. A simple counter-example is: $E = \{0, 1\}$, $P(0, 1) = 1/2$, $P(1, 1) = 1$. Let $\nu = \delta_1$ and $G = \{\delta_0\}$. The l.h.s. of (17) is $-\infty$, but $J(\delta_0) = \log 2$.
Notice that the above lower bound was proved by the second named author for much more general essentially irreducible Markov processes but also with a mild technical condition (see [17, Theorem B.1]).

Proof. Let β_0 be an arbitrary element of G but fixed. By the definition of the τ-topology, there is a bounded and measurable function $f : E \to \mathbb{R}^d$ (for some $d \geq 1$) and $\delta > 0$ such that

$$\mathscr{N} := \{\beta \in \mathscr{M}_1(E);\ |\beta(f) - \beta_0(f)| < \delta\} \subset G.$$

Here $|\cdot|$ is the Euclidian norm on $\mathbb{R}^d$. Hence by Theorem 2.2,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_\nu(L_n \in \mathscr{N}) \geq - \inf_{|z-z_0|<\delta} \Lambda_f^*(z)$$

where $z_0 = \beta_0(f)$. By Theorem 3.1,

$$\inf_{|z-z_0|<\delta} \Lambda_f^*(z) = \inf_{|z-z_0|<\delta} J_\mu^f(z) \leq J_\mu(\beta_0).$$

Therefore we get

$$\text{the l.h.s. of (17)} \geq -J_\mu(\beta_0)$$

where (17) follows for $\beta_0 \in G$ is arbitrary. □

Corollary 3.6. *The following equality holds true:*

$$R(P) = \exp\left(\inf_{\nu\in\mathscr{M}_1(E)} J_\mu(\nu)\right). \tag{18}$$

Proof. It follows from Proposition 3.2.(16) with $V = 0$. □

A pathological phenomenon may happen : $R(P) > 1$.

Example 3.7. Let $(B_t)_{t\geq 0}$ be a Brownian Motion on a connected complete and stochastic complete Riemannian manifold M with sectional curvature less than $-K$ ($K > 0$) and let $P(x, dy) = \mathbb{P}_x(B_1 \in dy)$. P is irreducible with a maximal irreducible measure given by the Riemann volume measure $\mu = dx$. It is well known that $\|P\|_{L^2(M,\,dx)} < 1$ ([16]). It is obvious (from (3)) that $\frac{1}{R(P)} \leq \|P\|_{L^2(M,\,dx)}$, and the converse inequality holds by [20, Lemma 5.3] and the symmetry of P on $L^2(M, dx)$. Thus $R(P) = \dfrac{1}{\|P\|_{L^2(M,\,dx)}} > 1$.

4 Some remarks on the upper bound of large deviations

Theorem 4.1. (Ney and Nummelin [14]) *Let $f : E \to \mathbb{R}^d$ be measuable. The following upper bound of large deviation holds*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x \left(\frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \in F, X_n \in C \right) \le -\inf_{z\in F} \Lambda_f^*(z),\ \mu - a.e.\ x \in E \tag{19}$$

for all compact subsets F of $\mathbb{R}^d$, where

(a) either C is P-small set if f is bounded;

(b) or for all $\lambda > 0$, there are some constant $c(\lambda) > 0$ and $\nu \in \mathscr{M}_1(E)$ such that

$$e^{-\lambda|f(x)|} P(x, dy) > c(\lambda)\nu(dy),\ \forall x \in C \tag{20}$$

if f is unbounded.

Furthermore if $d = 1$ or $\Lambda(\delta|f|) < +\infty$, then (19) holds for all closed subsets F of $\mathbb{R}^d$.

Indeed for every $V = \langle f(x), y\rangle$ where $y \in \mathbb{R}^d$, if f is bounded then every P-small set C is also P^V-small; if f is unbounded and C satisfies (20), then C is again P^V-small. Thus by Lemma 2.1,

$$\Lambda_f(y) = \mathrm{esssup}_{x\in E} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}^x 1_C(X_n) \exp\left(\sum_{k=0}^{n-1} \langle f, y\rangle(X_k) \right),\ \forall y \in \mathbb{R}^d$$

where the Theorem 4.1 follows by Gärtner-Ellis theorem (see [9]).
Baxter et al. (1991) [1] found for the first time a Doeblin recurrent Markov chain which does not verify the level-2 LDP (but it satisfies the level-1 LDP for $L_n(f)$ for bounded f by Theorems 2.2 and 4.1 since the whole space E is P-small in such case). Furthermore Bryc and Smolenski (1993) [3] constructed an exponentially recurrent Markov chain for which even the level-1 LDP of $L_t(f)$ for some bounded and measurable f fails.
We now present a counter-example for which (19) does not hold for P-small C but unbounded f.

Example 4.2. (see [19]) Let $h(x)$ be a strictly increasing differentiable function on $\mathbb{R}^+$ such that

$$h(0) = 0,\ \sup_{x\in\mathbb{R}^+} |h^{'}(x)| < +\infty, a := \limsup_{x\to\infty} \frac{h(x)}{x} > \liminf_{x\to\infty} \frac{h(x)}{x} =: b > 0$$

Assume that $(\xi_k)_{k\ge0}$ is a sequence of independent and identically distributed nonnegative random variables and $\mathbb{P}(\xi_0 > t) = e^{-h(t)},\ t \ge 0$. Then $\{X_k = (\xi_k, \xi_{k+1})\}_{k\ge0}$ is a Markov chain valued in $E = (\mathbb{R}^+)^2$, which is Doeblin recurrent, i.e., E is P-small. Let us show that (19) may be wrong with $C = E$ for some unbounded f.
Indeed let $f(X_k) := \xi_{k+1} - \xi_k,\ \forall\ k \ge 1$, then $\frac{1}{n}S_n(f) := \frac{1}{n}\sum_{k=1}^n f(X_k) = \frac{1}{n}(\xi_n - \xi_0)$. We have for all $r > 0$,

$$\begin{aligned} e^{-h((n+1)r)}(1 - e^{-h(r)}) &= \mathbb{P}(\xi_{n+1} > (n+1)r, \xi_1 < r) \\ &\le \mathbb{P}(\frac{S_n(f)}{n} > r) \le \mathbb{P}(\xi_{n+1} > nr) = e^{-h(nr)} \end{aligned}$$

Consequently for all $r > 0$,

$$\limsup_{n\to+\infty} \frac{1}{n} \log \mathbb{P}(\frac{S_n(f)}{n}) > r) = -br,\ \liminf_{n\to+\infty} \frac{1}{n} \log \mathbb{P}(\frac{S_n(f)}{n}) > r) = -ar$$

Then $(\frac{1}{n}S_n(f))$ does not verify any LDP, which implies that the condition (20) is essential.

References

[1] Baxter, J.R., Jain, N.C., Varadhan, S.R.S., Some familiar examples for which the large deviation principle does not hold. *Comm. Pure Appl. Math.* **44** (1991), 911-923. MR1127039 MR1127039

[2] N. Bourbaki, "*Espaces Vectoriels Topologiques(Livre V)", Chaps. III-V.* Hermann, Paris, 1955.

[3] Bryc, W. and Smolenski, W., On the convergence of averages of mixing sequences. *J. Theoret. Probab.* **6** (1993), 473-483. MR1230342 MR1230342

[4] Conway, J.B., *A Course in Functional Analysis* . Springer, Berlin, 1985. MR0768926 MR0768926

[5] A. de Acosta, Large deviations for vector valued additive functionals of a Markov process: Lower bounds. *Ann. Proba.* **16** (1988), 925-960. MR0942748 MR0942748

[6] A. de Acosta and P. Ney, Large deviation lower bounds for arbitrary additive functionals of a Markov chain. *Ann. Proba.* **26** (1998), 1660-1682. MR1675055 MR1675055

[7] J.D. Deuschel and D.W. Stroock, *Large deviations*. Pure and Appl. Math. **137**, Academic Press, Inc., Boston, MA, 1989. MR0997938 MR0997938

[8] M.D. Donsker, S.R.S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time, I-IV. *Comm. Pur. Appl. Math.* **28** (1975), 1-47 and 279-301 ; **29** (1976), 389-461 ; **36** (1983), 183-212 . MR0386024

[9] A. Dembo and O. Zeitouni, *Large Deviations Techiniques and Applications.* **38**. Springer-Verlag, New York, 1998. MR1619036 MR1619036

[10] N.C. Jain, Large deviation lower bounds for additive functionals of Markov processes. *Ann. Proba.* **18** (1990), 1071-1098. MR1062059 MR1062059

[11] I. Kontoyiannis and S.P. Meyn, Spectral Theory and Limit Theory for Geometrically Ergodic Markov Processes. *Ann. Appl. Prob.*, **13** (2003), 304-362. MR1952001 MR1952001

[12] S.P. Meyn, Large Deviation Asymptotics and Control Variates for Simulating Large Functions. *Ann. Appl. Prob.*, **16** (2006), 310-339. MR2209344 MR2209344

[13] S.P. Meyn and R.L. Tweedie, *Markov Chains and Stochastic Stability.* Springer-Verlag, London, 1993. MR1287609 MR1287609

[14] P. Ney and E. Nummelin, Markov additive processes (I). Eigenvalue properties and limit theorems; (II). Large deviations. *Ann. Proba.* **15** (1987), 561-592, 593-609. MR0885131 and MR0885132 MR0885131

[15] E. Nummelin, *General Irreducible Markov Chains and Non-Negative Operators*. Cambridge Tracts in Math., **83**, Cambridge Univ. Press, 1984. MR0776608 MR0776608

[16] Shing-Tung Yau and Richard Schoen, *Lectures on Differential Geometry, in chinese*. Higher Education Press in china, Beijing 2004. MR1333601

[17] L. Wu, Uniformly integrable operators and large deviations for Markov proceeses. *J. Funct. Anal.* **172** (2000), 301-376. MR1753178 MR1753178

[18] L. Wu, Some notes on large deviations of Markov processes . *Acta Math. Sin. (Engl. Ser.)*, **16** (2000), 369-394. MR1787093 MR1787093

[19] L. Wu, On large deviations for moving average processes. In *Probability, Finance and Insurance, pp.15-49*, the proceeding of a Workshop at the University of Hong-Kong (15-17 July 2002), Eds: T.L. Lai, H.L. Yang and S.P. Yung. World Scientific 2004, Singapour. MR2189197 MR2189197

[20] L. Wu, Essential spectral radius for Markov semigroups (I) : discrete time case. *Probab. Th. Rel. Fields* **128** (2004), 255-321. MR2031227 MR2031227

[21] L. Wu, Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stochastic Process. Appl.* **91** (2001), 205-238. MR1807683 MR1807683

[22] K. Yosida, *Functional Analysis*, Third Version. Grundlehren der mathematischen Wissenschaften 123, Springer-Verlag 1971.