LARGE DEVIATIONS FOR OCCUPATION MEASURES FOR MARKOV PROCESSES
(discrete time, non compact case)

R. LIPTSER
Electrical Engineering-Systems, Tel Aviv University,
69978 - Ramat Aviv, Tel Aviv, ISRAEL
and
Institute for Problems of Information Transmission,
Academy of Science, Moscow, 101447, RUSSIA

Abstract

A simple proof of Donsker-Varadhan's large deviation principle (LDP) for occupation measure of Markov process, valued in **R**, with the discrete time is given. A proof is based on a new version of Dupui-Ellis's large deviation principle for two-dimensional occupation measures. In our setting, an existence of the invariant measure does not assumed. This condition is replaced (from point of view of applications) on more natural one. It is given an example of Markov process, defined by non linear recursion, for which sufficient conditions of existing the large deviation principle are easy verified.
Key words: Large Deviations, Exponential Tightness, Local Large Deviations.

1. Introduction. Main Result.

1. It is well known from Donsker and Varadhan [1], the LDP for occupation measures $(\pi_n, n \geq 1)$

$$\pi_n(A) = \frac{1}{n}\sum_{k=1}^{n} I(\xi_k \in A) \tag{1.1}$$

of Markov process $\xi = (\xi_k)_{k\geq 0}$, valued in R, with fixed initial point $\xi_0 = x$. It takes place under the assumptions:

(F) $\xi = (\xi_k)_{k\geq 0}$ is Feller's process;

(I) there exists the unique invariant measure $\alpha(dx)$: $\alpha(\Gamma) = \int_R \pi(x,\Gamma)\alpha(dx)$, $\Gamma \in R$;

($\mathbf{H}^*$) there exists non negative measurable function $v(x)$ such that for every $N > 0$ $\sup_{|x|\leq N} v(x) < \infty$ and for the function

$$w(x) = \log \frac{e^{v(x)}}{\int_R e^{v(y)}\pi(x,dy)}$$

the following properties hold $\inf_{x\in R} w(x) = w_\star > -\infty, \quad \lim_{i\to\infty} \inf_{|x|>i}[w(x) - w_\star] = \infty$;

(RM) there exists σ-finite measure $l = l(dx)$ such that $\pi(x, dy) = p(y|x)l(dx) \quad \alpha - a.s.$ and for every x $p(y|x) > 0, \quad l - a.s.$

Introduce a metric space $(\mathbf{S}, \rho)$ ($\mathbf{S}$ is space of probability measures on R, ρ is Levy-Prokhorov's metric.

Theorem. (Donsker, Varadhan). *Assume (*$\mathbf{F}$*), (*$\mathbf{I}$*), (*$\mathbf{H}^*$*), and (*$\mathbf{RM}$*). Then the family $\pi_n, n \geq 1$ obeys the LDP in $(\mathbf{S}, \rho)$ with the rate function*

$$J(\mu) = \sup_{u\in\mathcal{N}} \int_R \log \frac{e^{u(x)}}{\int_R e^{u(y)}\pi(x, dy)} \mu(dx), \tag{1.2}$$

where $\mathcal{N}$ is set of compactly supported continuous functions. Level sets of $J(\mu), \mu \in \mathbf{S}$ are compacts in $(\mathbf{S}, \rho)$.

2. Here and in the sequel the following notations are used

$$\lambda^\circ_\gamma(dx, dy) = \pi(x, dy)\gamma(dx), \quad \forall\gamma \in \mathbf{S}$$

and $\mathbf{S}_{\gamma\gamma}$ for designating of a set of probability measures on R^2 with the same marginals γ. Let $\lambda \in \mathbf{S}_{\mu\mu}$. Following Donsker and Varadhan, a value

$$H(\lambda|\lambda^\circ_\mu) = \begin{cases} \int_{R^2} \log \frac{d\lambda}{d\lambda^\circ_\mu}(x, y)\lambda(dx, dy), & \lambda \ll \lambda^\circ_\mu \\ \infty, & \text{otherwise} \end{cases} \tag{1.3}$$

is named the conditional entropy of λ with respect to λ°_μ.

A difficult part of Donsker-Varadhan's proof, concerning to the lower bound, is identity

$$J(\mu) \equiv \inf_{\lambda\in\mathbf{S}_{\mu\mu}} H(\lambda|\lambda^\circ_\mu)(= J'(\mu)). \tag{1.4}$$

It would be noted that the inequality $J(\mu) \leq J'(\mu)$ is obvious while the proof of the opposite one, even for the compact case, (see Donsker and Varadhan [2]), seems sufficiently complicated. Later, Donsker and Varadhan have establish the LDP, avoiding the identity $J(\mu) \equiv J'(\mu)$ and using Varadhan's contraction principle [3] and the LDP of occupation measures for, so called, "third level", [4].

3. The aim of this paper is to obtain the LDP not applying neither $J(\mu) \equiv J'(\mu)$ nor the result for the "third level". Our proof is based on a new version of Dupui-Ellis's large deviation principle for two-dimensional occupation measures. In the present paper, instead of (**I**) and (**RM**) we introduce assumptions:

(**I'**) there exists a probability measure on R^2 $\lambda' = \lambda'(dx, dy)$, obeying the same marginals, say, $\alpha = \alpha(dx)$, such that $\lambda' \sim \lambda^\circ_\alpha$ and $H(\lambda'|\lambda^\circ_\alpha) < \infty$ (it would be noted (**I**) implies (**I'**) with $\lambda' = \lambda^\circ_\alpha$);

(**RM'**) transition probabilities $\pi(x, dy)$ and $\pi'(x, dy)$ ($\lambda'(dx, dy) = \pi'(x, dy)\alpha(dx)$) obey conditional densities with respect to some σ-finite measure $l = l(dx)$, say, $p(y|x)$ and $p'(y|x)$ such that for every $x \in R$

$$\begin{aligned} \pi(x, dy) &= p(y|x)l(dy), \\ \pi'(x, dy) &= p'(y|x)l(dy), \quad \alpha - \text{a.s.} \\ p'(y|x) &> 0, \quad (l \times \alpha) - \text{a.s.} \end{aligned}$$

In this paper, ξ_0 is a random variable distributed with $\alpha_0 = \alpha_0(dx)$ for which the following condition is assumed

($\mathbf{H_0}$) Let $v(x)$ be from ($\mathbf{H}^*$) and $u(x) = v(x) + \frac{1}{2}[w(x) - w_*]$. There exists $b > 0$ such that

$$\begin{aligned} \int_R e^{bu(x)}\alpha_0(dx) &< \infty \\ \int_R p(y|x)\alpha_0(dx) &> 0 \ l - \text{a.s.} \end{aligned}$$

We give new proof of Donsker-Varadhan's type theorem.

Theorem 1.1 *Under* (**F**), (**I'**), ($\mathbf{H}^*$), (**RM'**) *and* ($\mathbf{H_0}$), *the family* $(\pi_n, n \geq 1)$ *obeys the LDP in the metric space* $(\mathbf{S}, \rho)$ *with rate function*

$$J'(\mu) = \inf_{\lambda \in \mathbf{S}_{\mu\mu}} H(\lambda|\lambda^\circ_\mu) \quad \forall \mu \in \mathbf{S}.$$

4. We derive the statement of this theorem by using Varadhan's contraction principle and LDP for two-dimensional occupation measures $\pi^2_n(dx, dy), n \geq 1$, where

$$\pi_n^2(A\times B)=\frac{1}{n}\sum_{k=1}^n I(\xi_{k-1}\in A,\xi_k\in B). \qquad (1.5)$$

Following Varadhan [3], the LDP for the family $(\pi_n^2, n\ge 1)$ in the metric space $(\mathbf{S}^2,\rho^2)$ ($\mathbf{S}^2$ is a space of probabilistic measures on R^2 and ρ^2 is Levy-Prokhorov's metric) is defined as:

(0) there exists (rate) function $J^2(\lambda), \lambda\in\mathbf{S}^2$ values in $[0,\infty]$ level sets of which are compacts;

(1) for every closed (open) in the metric ρ^2 set F^2 (G^2) from $\mathbf{S}^2$

$$\begin{aligned}\limsup_n \frac{1}{n}\log P(\pi_n^2\in F^2) &\le -\inf_{\lambda\in F^2} J^2(\lambda);\\ \liminf_n \frac{1}{n}\log P(\pi_n^2\in G^2) &\ge -\inf_{\lambda\in G^2} J^2(\lambda).\end{aligned}$$

Theorem 1.2 (comp. Ellis [5] and Dupui and Ellis [6], [19]) *Under (*__F__*), (*__I'__*), (*$\mathbf{H}^*$*), (*__RM'__*), and (*$\mathbf{H_0}$*), the family* $(\pi_n^2, n\ge 1)$ *obeys the LDP in* $(\mathbf{S}^2,\rho^2)$ *with the rate function*

$$J^2(\lambda)=\begin{cases}\infty, & \textit{marginals of } \lambda \textit{ are different}\\ H(\lambda|\lambda^\circ_\gamma), & \textit{marginals of } \lambda \textit{ are the same } (=\gamma).\end{cases}$$

5. Despite Theorem 1.2 is closed to corresponding results from [5] [6] its proof is essenially different. It requires the following notions.

Definition 1. ([7]) The family $(\pi_n^2, n\ge 1)$ is said to be exponentially tight in $(\mathbf{S}^2,\rho^2)$, if there exists a sequence of compacts $K_j^2\in\mathbf{S}^2, j\ge 1$ such that

$$\limsup_n \frac{1}{n}\log P(\pi_n^2\in\mathbf{S}^2\setminus K_j^2)=-\infty \qquad (1.6)$$

Definition 2. ([8]) The family $(\pi_n^2, n\ge 1)$ is said to be LD relatively compact in $(\mathbf{S}^2,\rho^2)$, if any of infinite subsequence $(\pi_{n'}^2)$ of (π_n^2) contains further subsequence $(\pi_{\tilde n}^2)$, which satisfies the LDP in $(\mathbf{S}^2,\rho^2)$ with some rate function $\tilde J^2=\tilde J^2(\lambda)$.

Definition 3. (comp. [9]) The family $(\pi_n^2, n \geq 1)$ is said to be satisfied the local LDP in $(\mathbf{S}^2, \rho^2)$ with local rate function $I = I(\lambda)$, if for every $\lambda \in \mathbf{S}^2$

$$\begin{aligned} & \liminf_{\delta \to 0} \liminf_n \frac{1}{n} \log P(\rho^2(\pi_n^2, \lambda) \leq \delta) \\ = \; & \limsup_{\delta \to 0} \limsup_n \frac{1}{n} \log P(\rho^2(\pi_n^2, \lambda) \leq \delta) \\ = \; & -I(\lambda)). \end{aligned} \tag{1.7}$$

The exponential tightness for the family $(\pi_n^2, n \geq 1)$, under $(\mathbf{H}^*)$ and $(\mathbf{H_0})$, is established in Theorem 2.1 (Section 2). Theorems 3.1 and 4.1 (Sections 3 and 4 respectively) bring the local LDP for this family with $I(\lambda) \equiv J^2(\lambda)$.

The next link in proving Theorem 2.1 is Puhalskii's theorem [8] or, more exactly, its straightforward statement:

$$\text{“exponential tightness”} \implies \text{“LD relative compactness”}. \tag{1.8}$$

Thus, a scheme of the proof for Theorem 1.2 is the following. By Theorem 2.1 the family $(\pi_n^2, n \geq 1)$ is exponentially tight. By (1.8) one can take a subsequence $(\pi_{\tilde{n}}^2)$, obeying the LDP with rate function $\tilde{J}^2(\lambda)$. All these facts imply the local LDP for $(\pi_{\tilde{n}}^2)$: for every $\lambda \in \mathbf{S}^2$

$$\begin{aligned} & \liminf_{\delta \to 0} \liminf_{\tilde{n}} \frac{1}{\tilde{n}} \log P(\rho^2(\pi_{\tilde{n}}^2, \lambda) \leq \delta) \\ = \; & \limsup_{\delta \to 0} \limsup_{\tilde{n}} \frac{1}{\tilde{n}} \log P(\rho^2(\pi_{\tilde{n}}^2, \lambda) \leq \delta) \\ = \; & -\tilde{J}^2(\lambda)). \end{aligned} \tag{1.9}$$

On the other hand by virtue of the local LDP for the original sequence with the local rate function $I(\lambda)$ from (1.7) and obvious inequalities $\liminf_n \leq \liminf_{\tilde{n}} \leq \limsup_{\tilde{n}} \leq \limsup_n$, we arrive to an identity $\tilde{J}^2(\lambda) \equiv J^2(\lambda)$, what in turn implies that $J^2(\lambda)$ is rate function (the same method has been used in [10]). The upper and lower bounds from (1) and (2) are checked also by using Puhalskii theorem (see (1.8)). Other approaches for proving LDP for (π_n) and (π_n^2) can be found in Akosta [11], Gärtner [12], Orey and Pelikan [13], Veretennikov [14].

6. As was mentioned above, (1.8) is implied by Theorems 3.1 and 4.1, which give us the upper and lower bounds in the local LDP for the family $(\pi_n^2, n \geq 1)$. A proof of the lower bound requires only conditions (**I'**) (**RM'**). It uses a change of probability measures, proposed by Donsker and Varadhan in [1], [2] and a regularization method borrowed from recent Wu's paper [15]. Other approach can be found in Jain [16].

A proof for the upper bound in the local LDP requires condition (**F**) and the exponential tightness for family $(\pi_n^2, n \ge 1)$.

2. Exponential tightness.

Due to Definition 1, (1.6) has to be verified. To this end, we need a few auxiliary results.

Letting $\gamma = \gamma(y), y \ge 0$ a positive decreasing function such that $\lim_{y\to\infty} \gamma(y) = 0$, put

$$K_j^2 = \bigcap_{i \ge j} \{\lambda \in \mathbf{S}^2 : \int_{(|x|>i)\cup(|y|>i)} \lambda(dx, dy) \le \gamma(i)\}. \tag{2.1}$$

The set K_j^2 is tight and, by virtue of Prokhorov's theorem (see e.g. [17]), is relatively compact. Since $\{x, y : (|x| > l) \cup (|y| > l)\}$ is open set, a limit of each converging sequence from K_j^2 belongs to K_j^2 too, i.e. K_j^2 is compact. Evidently $K_j^2 \subseteq K_{j+1}^2$.

For $j \ge 1$ $\lambda \in \mathbf{S}^2$, put

$$L(j, \lambda) = \min\,(i \ge j : \int_{(|x|>i)\cup(|y|>i)} \lambda(dx, dy) > \gamma(i)). \tag{2.2}$$

Lemma 2.1. *The family $(\pi_n^2, n \ge 1)$ is exponentially tight in $(\mathbf{S}^2, \rho^2)$, if*

$$\lim_j \limsup_n \frac{1}{n} \log P(L(j, \pi_n^2) < \infty) = -\infty. \tag{2.3}$$

Proof: Taking defined in (2.1) compact K_j^2 and noticing that $\{\pi_n^2 \in \mathbf{S}^2 \setminus K_j^2\} = \{L^2(j, \pi_n) < \infty\}$, we derive (1.6) from (2.3).

Lemma 2.2. *For every measurable sets $A_n, n \ge 1$ and $B_{n,i}, n \ge 1, i \ge 1$ obeying a property $\lim_{i\to\infty} \limsup_n \frac{1}{n} \log P(B_{n,i}) = -\infty$, the following equality holds*

$$\limsup_n \frac{1}{n} \log P(A_n) = \limsup_{i\to\infty} \limsup_n \frac{1}{n} \log P(A_n, \Omega \setminus B_{n,i}).$$

Proof: The required statement follows from obvious inequalities: $P(A_n) \geq P(A_n, \Omega \setminus B_{n,i})$, $P(A_n) \leq 2[P(A_n, \Omega \setminus B_{n,i}) \vee P(B_{n,i})]$.

Theorem 2.1. *Under* ($\mathbf{H}^*$) *and* ($\mathbf{H_0}$) *the family* $(\pi_n^2, n \geq 1)$ *is exponentially tight in* $(\mathbf{S}^2, \rho^2)$.

Proof: Put

$$\gamma(y) = \frac{1}{\sqrt{\inf_{|x|>y}[w(x) - w_\star]}}, \tag{2.4}$$

where $w(x)$ is the function from ($\mathbf{H}^*$). Show that with such $\gamma(y)$ (2.3) holds. In fact, by virtue of ($\mathbf{H_0}$) and Chernoff's inequality $P(u(\xi_0) > in) \leq e^{-(in)b} \int_R e^{bu(x)} \alpha_0(dx)$. Consequently, for $B_{n,i} = \{v(\xi_0) > in\}$ we get $\limsup_n \frac{1}{n} \log P(B_{n,i}) \leq -ib \to -\infty$, $i \to \infty$. Therefore by virtue of Lemmas 2.1 and 2.2, the exponential tightness of $(\pi_n^2, n \geq 1)$ takes place provided that

$$\lim_i \limsup_j \limsup_n \frac{1}{n} \log P(L(j, \pi_n^2) < \infty,\ v(\xi_0) \leq in) = -\infty. \tag{2.5}$$

To verify (2.5), use a chain of inclusions (recall that $L(j, \pi_n^2) \geq j$):

$$\begin{aligned}
\{L(j, \pi_n^2) < \infty\} \ &\subseteq\ \{\int_{(|x|>L(j,\pi_n^2)) \cup (|y|>L(j,\pi_n^2))} \pi_n^2(dx, dy) > \gamma(L(j, \pi_n^2))\} \\
&\subseteq\ \{\int_{(|x|>L(j,\pi_n^2))} \int_R \pi_n^2(dx, dy) \\
&\qquad + \int_R \int_{(|y|>L(j,\pi_n^2))} \pi_n^2(dx, dy) > \gamma(L(j, \pi_n^2))\} \\
&\subseteq\ \{2 \int_{(|x|>L(j,\pi_n^2))} \int_R \pi_n(dx) \\
&\qquad + \frac{1}{n} I(|\xi_0| > L(j, \pi_n^2)) > \gamma(L(j, \pi_n^2))\}.
\end{aligned} \tag{2.6}$$

Then (2.5) holds provided that

$$\begin{aligned}
&\limsup_j \limsup_n \frac{1}{n} \log P\Big(\int_{|x|>j} \pi_n(dx) \\
+ \ & \frac{1}{2n} I(|\xi_0| > L(j, \pi_n^2)) > \gamma(L(j, \pi_n^2)), u(\xi_0) \leq in\Big) \\
= \ & -\infty.
\end{aligned} \tag{2.7}$$

Put $Z_n = \prod_{k=0}^{n-1} \frac{e^{v(\xi_{k+1})}}{E(e^{v(\xi_{k+1})}|\xi_k)}$, where $v(x)$ is from $(\mathbf{H}^*)$. Due to Markovian property $E(e^{v(\xi_{k+1})} \mid \xi_k) = \mathbf{E}(e^{v(\xi_{k+1})} \mid \xi_0, ..., \xi_k)$, $P--$a.s. and, so $\mathbf{E}Z_n = 1$ what implies an obviously inequality

$$1 \geq EI(\int_{|x|>L(j,\pi_n^2)} \pi_n(dx) + \frac{1}{2n} I(|\xi_0| > L(j, \pi_n^2)) > \gamma(L(j, \pi_n^2), u(\xi_0) \leq in) Z_n. \quad (2.8)$$

From the definition of Z_n and $w(x)$ (see $(\mathbf{H}^*)$), it follows

$$\begin{aligned} \log Z_n &= \sum_{k=1}^{n} v(\xi_k) - \sum_{k=1}^{n} \log \int_R e^{v(y)} \pi(\xi_{k-1}, dy) \\ &= v(\xi_n) - v(\xi_0) + \sum_{k=1}^{n} w(\xi_{k-1}) \\ &= v(\xi_n) - v(\xi_0) + n w_\star + n \int_R [w(x) - w_\star] \pi_n(dx)). \end{aligned}$$

Evaluate now from below the value $\log Z_n$ on a set $\{ \int_{|x|>L(j,\pi_n^2)} \pi_n(dx) + \frac{1}{2n} I(|\xi_0| > L(j, \pi_n^2)) > \gamma(L(j, \pi_n^2)\}$. Arguing $L(j, \pi_n^2) \geq j$, we find

$$\begin{aligned} &\log Z_n \\ \geq\ & -v(\xi_0) + n w_* + n \inf_{|y|>L(j,\pi_n^2)} [w(y) - w_*] \Big(\frac{1}{2} \gamma(L(j, \pi_n^2)) - \frac{1}{2n} I(|\xi_0| > L(j, \pi_n^2) \Big) \\ \geq\ & -v(\xi_0) - \frac{1}{2} \inf_{|y|>|\xi_0|} [w(y) - w_*] + \frac{n}{2} \frac{\inf_{|y|>L(j,\pi_n^2)} [w(y) - w_*]}{\sqrt{\inf_{|y|>L(j,\pi_n^2)} [w(y) - w_*]}} + n w_* \\ =\ & -v(\xi_0) - \frac{1}{2} \inf_{|y|>|\xi_0|} [w(y) - w_*] + \frac{n}{2} \sqrt{\inf_{|y|>j} [w(y) - w_*]} + n w_* \\ \geq\ & -in - \frac{1}{2} \inf_{|y|>i} [w(y) - w_*] + \frac{n}{2} \sqrt{\inf_{|y|>j} [w(x) - w_*]} + n w_*. \end{aligned}$$

Hence, we arrive to

$$\begin{aligned} & \frac{1}{n} \log P\Big(\int_{|x|>L(j,\pi_n^2)} \pi_n(dx) \\ +\ & \frac{1}{2n} I(|\xi_0| > L(j, \pi_n^2)) > \gamma(L(j, \pi_n^2) \Big) \\ \leq\ & i + \frac{1}{2n} \inf_{|y|>i} [w(y) - w_*] - \frac{1}{2} \sqrt{\inf_{|y|>j} [w(y) - w_*]} - w_* \\ \to\ & -\infty,\ n \to \infty,\ j \to \infty, \end{aligned}$$

that is (2.7) holds.

Corollary.

$$\lim_i \limsup_n \frac{1}{n} \log P\Big(\int_{(|x|>i)\cup(|y|>i)} \pi_n^2(dx,dy) > q\Big) = -\infty, \quad q>0.$$

3. Upper bound for local LDP.

Theorem 3.1 *Assume (***F***), (***H****), and (***H$_0$***). Then for every $\lambda \in \mathbf{S}^2$*

$$\limsup_{\delta\to 0} \limsup_n \frac{1}{n} \log P(\rho^2(\pi_n^2,\lambda) \le \delta) \le - \begin{cases} \infty, & \textit{marginals are different} \\ H(\lambda|\lambda_\mu^\circ), & \textit{marginals of} \lambda (=\mu), \end{cases}$$

where $H(\lambda|\lambda_\mu^\circ)$ is the conditional entropy defined in (1.3).

Proof of this theorem is based on a sequence of auxiliary results formulated below as lemmas.

Let P and Q be probability measures on measurable space $(\Omega,\mathcal{F})$ and let $\mathcal{U}^m(Q) = \{u(\omega) : \int_\Omega u(\omega)dQ = 1\}$ be a set of non negative $\mathcal{F}$-measurable functions. Due to (1.3), the conditional entropy $H(P|Q)$ of measure P with respect to measure Q is defined as:

$$H(P|Q) = \begin{cases} \int_\Omega \log \frac{dP}{dQ}(\omega)dP, & P \ll Q \\ \infty, & \text{otherwise.} \end{cases}$$

Put $V(x) = \begin{cases} x\log x + 1 - x, & x>0 \\ 1, & x=0. \end{cases}$ It is clear, $V(x)$ is convex non negative continuous function and the following formula holds

$$H(P|Q) = \begin{cases} \int_\Omega V(\frac{dP}{dQ}(\omega))dQ, & P \ll Q \\ \infty, & \text{otherwise.} \end{cases} \tag{3.1}$$

For $u \in \mathcal{U}^m$, let $G(u) = \int_\Omega \log u(\omega)dP$. It is well known (see Donsker and Varadhan [2]) that $H(P|Q) = \sup_{u\in\mathcal{U}^m} G(u)$. In fact, for $P \ll Q$ ($h = \frac{dP}{dQ}$) and any function $u \in \mathcal{U}^m(Q)$ we have $G(u) = \int_\Omega h\log u dQ = \int_\Omega (h\log u + 1 - u)dQ$ and, since for $x \ge 0$ $\sup_{y\ge 0} x\log y + 1 - y = V(x)$, inequality holds $G(u) \le \int_\Omega V(h)dQ = G(h)$, i.e. $\sup_{u\in\mathcal{U}^m} G(u)$ is attained at point $h \in \mathcal{U}^m(Q)$. If $P \ll Q$ fails, denote by P^s the singular part of P with respect to Q. Then one can choose a set Γ such that $Q(\Gamma)=0$ and $P^s(\Omega\setminus\Gamma)=0$. Taking $u_N(\omega)=1+NI_\Gamma(\omega)$ from $\mathcal{U}^m(Q)$ we find

$$G(u_N)=\int_\Omega \log(1+NI_\Gamma(\omega))dP \geq \int_\Omega \log(1+NI_\Gamma(\omega))dP^s=\log(1+N)\to\infty,\ N\to\infty.$$

Lemma 3.1 *Let $\Omega=R^2$. Then*

$$H(P|Q)=\sup_{u\in\mathcal{U}^c(Q)} G(u),$$

where $\mathcal{U}^c(Q)$ is subset of $\mathcal{U}^m(Q)$ of continuous function.

Proof: Assume $P\ll Q$. Put $h(\omega)=\frac{dP}{dQ}(\omega)$ and

$$\left\{\begin{array}{l} f_N(\omega)\\ c_N\\ h_N(\omega)\end{array}\right. = \left\{\begin{array}{l} h(\omega)I_{]N^{-1},N]}(h(\omega))+(1-I_{]N^{-1},N]}(h(\omega)))\\ \left(\int_\Omega f_N dQ\right)^{-1}\\ c_N f_N(\omega).\end{array}\right.$$

Evidently $h_N\in\mathcal{U}^m(Q)$ and also

$$\begin{aligned} G(h_N) &= \int_\Omega \log h_N(\omega)dP\\ &= \int_\Omega h(\omega)\log h_N(\omega)dQ\\ &= \log c_N+\int_{1/N<h(\omega)\leq N} h(\omega)\log h(\omega)dQ\\ &= \log c_N+\int_{1/N<h(\omega)\leq 1} h(\omega)\log h(\omega)dQ+\int_{1<h(\omega)\leq N} h\log h(\omega)dQ. \end{aligned}$$

Since $\lim_N c_N=1$, by Lebesgue dominated theorem we have

$$\lim_N\int_{1/N<h(\omega)\leq 1} h(\omega)\log h(\omega)dQ=\int_{h(\omega)\leq 1} h(\omega)\log h(\omega)dQ$$

and by Beppo-Levy's theorem

$$\lim_N\int_{1<h(\omega)\leq N} h(\omega)\log h(\omega)dQ=\int_{1<h(\omega)} h(\omega)\log h(\omega)dQ.$$

Therefore $\lim_N G(h_N)=G(h)$.

Let N be fixed. Choose a sequence of continuous function $u_{N,n}, n\geq 1$ such that $\frac{1}{2N}\leq u_{N,n}\leq 2N$ and $Q-\lim_n u_{N,n}=h_N$ (and so $\lim_n\int_\Omega u_{N,n}dQ=1$). Put $\overline{u}_{N,n}=$

$u_{N,n}/\int_\Omega u_{N,n}dQ$ and note that $\overline{u}_{N,n} \in \mathcal{U}^c(Q)$. Due to $P \ll Q$, $P - \lim_n u_{N,n} = h_N$. Then, by Lebesgue dominated theorem $\lim_n G(u_{N,n}) = G(h_N)$. Hence, for any $\varepsilon > 0$ there exist $N(\varepsilon)$ and $n(\varepsilon)$ such that $G(h_{N(\varepsilon)}) + \varepsilon \geq G(h)$ and $G(\overline{u}_{N(\varepsilon),n(\varepsilon)}) + \varepsilon \geq G(h_{N(\varepsilon)})$.

Thus, for any $\varepsilon > 0$ there exists a function $u_\varepsilon = \overline{u}_{N(\varepsilon),n(\varepsilon)} \in \mathcal{U}^c(Q)$ such that $G(u_\varepsilon) + 2\varepsilon \geq G(h) = H(P|Q)$. The last means that under $P \ll Q$ the required result holds.

If $P \ll Q$ fails, then for fixed N choose a sequence $u_{N,n}, n \geq 1$ from $\mathcal{U}^c$ such that $1/2 \leq u_{N,n} \leq 2N, n \geq 1$ and $P^s - \lim_n u_{N,n} = u_N$, where $u_N(\omega) = 1 + NI_\Gamma(\omega)$. Then

$$
\begin{aligned}
G(u_{N,n}) &= \int_\Omega \log u_{N,n} dP = \int_\Omega \log[(u_{N,n} + 1/2) - 1/2] dP \\
&\geq \int_\Omega \log(u_{N,n} + 1/2) dP - \log 2 \geq \int_\Omega \log(u_{N,n} + 1/2) dP^s - \log 2
\end{aligned}
$$

and so $\liminf_n G(u_{N,n}) \geq \log(3/2 + N) - \log 2$ that is $\sup_{u \in \mathcal{U}^c} G(u) = \infty$.

Lemma 3.2. *Let $\lambda_{\alpha\beta}$ and $\lambda_{\mu\nu}$ be probability measures from $\mathbf{S}^2$ with marginals α, β and μ, ν respectively. Then* $\rho^2(\lambda_{\alpha\beta}, \lambda_{\mu\nu}) \geq \begin{cases} \rho(\alpha,\mu) \\ \rho(\beta,\nu) \end{cases}$ *and for any $\delta > 0$* $\{\rho^2(\lambda_{\alpha\beta}, \lambda_{\mu\nu}) \leq \delta\} \subseteq \{\rho(\mu,\nu) \leq 2\delta + \rho(\alpha,\beta)\}$.

Proof: Put $F_{\alpha\beta}$ and $F_{\mu\nu}$ distribution functions corresponding to measures $\lambda_{\alpha\beta}$ and $\lambda_{\mu\nu}$ respectively. By the definition of Levy-Prokhorov's metric

$$
\begin{aligned}
& \rho^2(\lambda_{\alpha\beta}, \lambda_{\mu\nu}) \\
= \; & \sup_{x,y} \min \left[u : F_{\alpha\beta}(x-u, y-u) - u \leq F_{\mu\nu}(x,y) \leq F_{\alpha\beta}(x+u, y+u) + u\right] \\
\geq \; & \sup_{x} \min \left[u : F_{\alpha\beta}(x-u, \infty) - u \leq F_{\mu\nu}(x,\infty) \leq F_{\alpha\beta}(x+u, \infty) + u\right] \\
= \; & \rho(\alpha,\mu)
\end{aligned}
$$

and analogously $\rho^2(\lambda_{\alpha\beta}, \lambda_{\mu\nu}) \geq \rho(\beta,\nu)$. Consequently

$$
\{\rho^2(\lambda_{\alpha\beta}, \lambda_{\mu\nu}) \leq \delta\} \subseteq \{\rho(\alpha,\mu) \leq \delta\} \cap \{\rho(\beta,\nu) \leq \delta\}.
$$

By the triangular inequality $\rho(\alpha,\beta) \leq \rho(\alpha,\mu) + \rho(\mu,\beta)$, $\rho(\mu,\nu) \leq \rho(\mu,\beta) + \rho(\nu,\beta)$ and so,

$$
\begin{aligned}
\{\rho^2(\lambda_{\alpha\beta}, \lambda_{\mu\nu}) \leq \delta\} &\subseteq \{\rho(\beta,\mu) \leq \delta + \rho(\alpha,\beta)\} \cap \{\rho(\beta,\nu) \leq \delta + \rho(\alpha,\beta)\} \\
&\subseteq \{\rho(\mu,\nu) \leq 2\delta + \rho(\alpha,\beta))\}.
\end{aligned}
$$

Let ν_1, ν_2 be probability measures from $\mathbf{S}^2$ and F_1, F_2 their distribution functions respectively. Put $a = \rho^2(\nu_1, \nu_2)$ and $x = (x_1, x_2)$.

Lemma 3.3. *For compactly supported and continuously differentiable (by $\frac{\partial^2}{\partial x_1 \partial x_2}$) function $f = f(x)$*

$$
\begin{aligned}
& |\int_{R^2} f(x)[\nu_1(dx) - \nu_2(dx)]| \\
\leq \;& a \int_{R^2} |\frac{\partial^2}{\partial x_1 \partial x_2} f(x)| dx + \int_{R^2} |\frac{\partial^2}{\partial x_1 \partial x_2} f(x)| [F(x+a) - F(x-a)] dx,
\end{aligned}
$$

where $F(x)$ is any of $F_1(x)$, $F_2(x)$.

Proof: Integrating by parts, we obtain $\int_{R^2} f(x) dF_i(x) = \int_{R^2} \frac{\partial^2}{\partial x_1 \partial x_2} f(x) F_i(x) dx$. Thereby

$$
|\int_{R^2} f(x)[\nu_1(dx) - \nu_2(dx)]| \leq \int_{R^2} |\frac{\partial^2}{\partial x_1 \partial x_2} f(x)| [F_1(x) - F_2(x)] dx.
$$

Due to the property of Levy-Prokhorov's metric, we have $F_1(x-a) - a \leq F_2(x)$ and $F_2(x) \leq F_1(x+a) + a$, that is the required result holds with $F = F_1$: $|F_1(x) - F_2(x)| \leq a + |F_1(x+a) - F_1(x-a)|$.

For $F = F_2$, the proof is similar.

Corollary. If $f = f(x)$ is compactly supported and continuous only, then, approximating it $\sup_{x \in R} |f(x) - f_\varepsilon(x)| \leq \varepsilon/2$, where for each ε function $f_\varepsilon(x)$ satisfies assumptions of the lemma, we get

$$
\begin{aligned}
& |\int_{R^2} f(x)[\nu_1(dx) - \nu_2(dx)]| \\
\leq \;& \varepsilon + a \int_{R^2} |\frac{\partial^2}{\partial x_1 \partial x_2} f_\varepsilon(x)| dx + \int_{R^2} |\frac{\partial^2}{\partial x_1 \partial x_2} f_\varepsilon(x)| [F(x+a) - F(x-a)] dx
\end{aligned}
$$

The next lemma plays substantial role in proving Theorem 3.1.

Lemma 3.4. *Assume (*$\mathbf{F}$*), (*$\mathbf{H}^*$*), and (*$\mathbf{H_0}$*). Then, for every λ from $\mathbf{S}_{\mu\mu}$,*

$$
\begin{aligned}
& \lim_{q \to 0} \limsup_{i} \limsup_{\delta \to 0} \limsup_{n} \frac{1}{n} \log P(\rho^2(\pi_n^2, \lambda) \leq \delta, \int_{(|x|>i) \cup (|y|>i)} \pi_n^2(dx, dy) \leq q) \\
\leq \;& -H(\lambda | \lambda_\mu^\circ).
\end{aligned}
$$

Proof: Put $A(n,i,q) = \{\rho^2(\pi_n^2,\lambda) \le \delta,\ \int_{(|x|>i)\cup(|y|>i)} \pi_n^2(dx,dy) \le q\}$. Taking compactly supported continuous function $v(x,y)$, put $u(x,y) = \frac{e^{v(x,y)}}{\int_R e^{v(x,z)}\pi(x,dz)}$. By virtue of (**F**) $u(x,y)$ is continuous bounded function such that $\int_{R^2} u(x,y)\lambda_\mu^\circ(dx,dy) = 1$. The last means that $u(x,y) \in \mathcal{U}^c(\lambda_\mu^\circ)$. Let $Z_n = \prod_{k=1}^n u(\xi_{k-1},\xi_k)$. By Markovian property $E(u(\xi_{k-1},\xi_k)|\xi_{k-1}) = E(u(\xi_{k-1},\xi_k)|\xi_{k-1},...,\xi_0) \quad P-$a.s. what implies $EZ_n = 1$ and in turn an obvious inequality

$$1 \ge \mathbf{E} I_{A(n,i,q)} Z_n. \tag{3.2}$$

Evaluate now from bellow, on the set $A(n,i,q)$, the value of $\log Z_n$:

$$\begin{aligned}
& \log Z_n \\
= & \sum_{k=1}^n \log u(\xi_{k-1},\xi_k) = n\int_{R^2} \log u(x,y)\pi_n^2(dx,dy) \\
\ge & \ n\int_{R^2} \log u(x,y)\lambda(dx,dy) - n|\int_{R^2} \log u(x,y)[\lambda(dx,dy) - \pi_n^2(dx,dy)]|.
\end{aligned}$$

Choosing non negative continuous function $\phi_i(x,y)$ such that $\phi_i(x,y) = 1$ on $\{|x| \le i-1\} \cap \{|y| \le i-1\}$ and $\phi_i(x,y) = 0$ on $\{|x| \ge i\} \cup \{|y| \ge i\}$, we get

$$\begin{aligned}
& \log Z_n \\
\ge & \ n\int_{R^2} \log u(x,y)\lambda(dx,dy) - n|\int_{R^2} \phi_i(x,y)\log u(x,y)[\lambda(dx,dy) - \pi_n^2(dx,dy)]| \\
& -n\sup_{x,y} |\log u(x,y)| \int_{(|x|\ge i)\cup(|y|\ge i)} \lambda(dx,dy) - n\int_{(|x|\ge i)\cup(|y|\ge i)} \pi_n^2(dx,dy) \\
= & \ n\int_{R^2} \log u(x,y)\lambda(dx,dy) - nr'(i) - nr''(i) - nr'''(i,n).
\end{aligned}$$

Denote by $F(x,y)$ the distribution function corresponding to $\lambda(dx,dy)$. By the corollary to Lemma 3.3, the following estimates (on $A(n,i,q)$) hold:

$$\begin{aligned}
& r'(i) \\
\le & \ \varepsilon + \delta\int_{R^2} |\frac{\partial^2}{\partial x_1 \partial x_2} f_\varepsilon(x,y)|dxdy \\
& + \int_{R^2} |\frac{\partial^2}{\partial x_1 \partial x_2} f_\varepsilon(x,y)|[F(x+\delta,y+\delta) - F(x-\delta,y-\delta)]dxdy \ (\le \varepsilon + q'(i,\varepsilon)),
\end{aligned}$$

where $f_\varepsilon(x,y)$ is "$\varepsilon/2$-approximation" of $\phi_i(x,y)\log u(x,y)$;

$$\begin{aligned} r''(i) &= \sup_{x,y} |\log u(x,y)| \int_{(|x|\geq i)\cup(|y|\geq i)} \lambda(dx,dy) \ (= q''(i)); \\ r'''(i,n) &\leq \sup_{x,y} |\log u(x,y)| q. \end{aligned}$$

These estimates and (3.2) imply an inequality

$$\begin{aligned} &\frac{1}{n} \log P\Big(A(n,i,q)\Big) \\ \leq & \ -\int_{R^2} \log u(x,y)\lambda(dx,dy) + \varepsilon + q'(i,\varepsilon) + q''(i) + \sup_{x,y} |\log u(x,y)|q, \end{aligned} \tag{3.3}$$

corresponding to which

$$\lim_{q\to 0} \limsup_{i} \limsup_{\delta\to 0} \limsup_{n} \frac{1}{n} \log P(A(n,i,q) \leq \varepsilon - \int_{R^2} \log u(x,y)\lambda(dx,dy).$$

Thus, by virtue of $\sup_{u\in\mathcal{U}^c} \int_{R^2} \log u(x,y)\lambda(dx,dy) = H(\lambda|\lambda^\circ_\mu)$, Lemma 3.1, and an arbitrariness of ε the required result holds.

The proof of Theorem 3.1: Assume λ from $\mathbf{S}^2$ has different marginals, say, μ and ν. Denote by π_n and π'_n marginals of π^2_n: $\pi_n(A) = \frac{1}{n}\sum_{k=1}^n I(\xi_{k-1} \in A)$, $\pi'_n(A) = \frac{1}{n}\sum_{k=1}^n I(\xi_k \in A)$. Since the total variation $||\pi_n - \pi^{'}_n|| \leq 2/n$, we get $\rho(\pi_n, \pi^{'}_n) \leq 2/n$. By Lemma 3.2 $\{\rho^2(\pi^2_n,\lambda) \leq \delta\} \subseteq \{\rho(\mu,\nu) \leq 2\delta + 2/n\}$ and so, under $2\delta + 2/n \leq \rho(\mu,\nu)$, $\log P(\rho(\pi^2_n,\lambda) \leq \delta) = -\infty$.

Assume λ from $\mathbf{S}^2$ has the same marginals. Put

$$\begin{aligned} A(n) &= \{\rho^2(\pi^2_n,\lambda) \leq \delta\} \\ B(n,i,q) &= \{ \int_{(|x|>i)\cup(|y|>i)} \pi^2_n(dx,dy) > q\}. \end{aligned}$$

The required result follows from Lemma 3.4, the corollary to Theorem 2.1, and Lemma 2.2, which, under

$$\begin{aligned} \lim_{q\to 0} \limsup_{i} \limsup_{n} \frac{1}{n} \log P(B(n,i,q)) &= -\infty \\ A(n) \cap (\Omega \setminus B(n,i,q)) &= A(n,i,q), \end{aligned}$$

can be reformulated as

$$\begin{aligned}
&\limsup_{\delta\to 0}\limsup_n \frac{1}{n}\log P(A(n)) \\
= \;& \lim_{q\to 0}\limsup_i \limsup_{\delta\to 0}\limsup_n \frac{1}{n}\log P(A(n),\ \Omega\setminus B(n,i,q)).
\end{aligned}$$

4. Lower bound for local LDP.

Theorem 4.1 *Assume (**I'**) and (**RM'**). Then for every λ from $\mathbf{S}^2$ and $\delta>0$*

$$\liminf_n \frac{1}{n}\log P(\rho^2(\pi_n^2,\lambda)\le\delta)\ge -J^2(\lambda).$$

Proof of this theorem consists in a few steps formulated below as lemmas.

Lemma 4.1. *Assume (**I'**) and (**RM'**). Then $\alpha\sim l$.*

Proof: By virtue of (**I'**) and (**RM'**) we get

$$\alpha(dy)=l(dy)\int_R p'(y|x)\alpha(dx), \tag{4.1}$$

that is $\alpha\ll l$. Also note $\frac{d\alpha}{dl}>0$ l-a.s. what implies, by Lebesgue decomposition of l with respect to α, that $l\ll\alpha$.

Lemma 4.2 *Assume (**I'**) and (**RM'**). Then $p(x|y)=\frac{d\lambda'}{d\lambda^\circ_\alpha}(x,y)p'(y|x)\quad(l\times\alpha)-a.s.$ and $p(y|x)>0$ $(l\times\alpha)$-a.s.*

Proof: By virtue of (**I'**) $\lambda'\sim\lambda^\circ_\alpha$ and, so $\pi(x,dy)\alpha(dx)=\frac{d\lambda^\circ_\alpha}{d\lambda'}(x,y)p'(y|x)l(dy)\alpha(dx)$. The first statement holds, since for every x we have $\pi(x,dy)=p(y|x)l(dy)$. To prove the second, show that $\frac{d\lambda^\circ_\alpha}{d\lambda'}>0$ $(l\times\alpha)$-a.s. To this end, use

$$\begin{aligned}
\frac{d\lambda^\circ_\alpha}{d\lambda'} &> 0\ \lambda'-\text{a.s.}\\
\lambda'(dx,dy) &= p'(y|x)l(dy)\alpha(dx)\\
p'(y|x) &>0\quad (l\times\alpha)-\text{a.s.}
\end{aligned}$$

Thereby $\lambda'(A)=0$ implies $(l\times\alpha)(A)=0$, that is $(l\times\alpha)\ll\lambda'$-a.s. and the required inequality holds.

Lemma 4.3 *Assume* (**I'**) *and* (**RM'**). *If* $H(\lambda|\lambda_\mu^\circ) < \infty$, $\lambda \in \mathbf{S}_{\mu\mu}$, *then* $\mu \ll l$.

Proof: Due to (1.3), $\lambda' \ll \lambda_\mu^\circ$ is implied by $H(\lambda|\lambda_\mu^\circ) < \infty$. In turn, (**RM'**) gives as the consequence $\lambda_\mu^\circ(dx, dy) = p(y|x)l(dy)\mu(dx)$. Let $l(A) = 0$. Then

$$\lambda_\mu^\circ(R \times A) = \int_A \int_R p(y|x)\mu(dx)l(dy) = 0.$$

Consequently $\mu(A) = \lambda(R \times A) = 0$.

Lemma 4.4. *Let* $\eta = (\eta_k)_{k\geq 0}$ *be a stationary Markov process values in* R, *having the marginal measure* $\mu(dx)$ *and transition probability* $\pi_\mu(x, dy)$. *Assume*

$$\begin{array}{rcl} \pi_\mu(x, dy) & = & p_\mu(y|x)l(dy) \ \mu - a.s. \\ p_\mu(y|x) & > & 0 \ (l \times \mu) - a.s. \end{array}$$

Then η *is* μ*-ergodic process.*

Proof: By Theorem 4 (see §1, Ch. 4 [18]), μ-ergodicity of η holds provided that an equation

$$f(x) = \int_R f(y)\pi_\mu(x, dy) \ \mu - \text{a.s.} \tag{4.2}$$

obeys unique μ-a.s. (to within a multiplicative constant) bounded solution. Let $f(x)$ be some solution of (4.2). Show that

$$\mu(x : sign \ f(x) = const) = 1. \tag{4.3}$$

Since μ is invariant measure we have

$$\mu(dy) = l(dy) \int_R p_\mu(y|x)\mu(dx). \tag{4.4}$$

(4.2) and (4.4) imply

$$\begin{array}{rcl} \int_R |f(x)|\mu(dx) & = & \int_R |\int_R f(y)p_\mu(y|x)l(dy)|\mu(dx) \\ & \leq & \int_R \int_R |f(y)|p_\mu(y|x)l(dy)\mu(dx) = \int_R |f(y)|\mu(dy) \end{array}$$

and, thus

$$\int_R \{ \int_R |f(y)| p_\mu(y|x) l(dy) - | \int_R f(y) p_\mu(y|x) l(dy)| \} \mu(dx) = 0.$$

The last means

$$\int_{f^+(y)>0} f^+(y) p_\mu(y|x) l(dy) \int_{f^-(y)>0} f^-(y) p_\mu(y|x) l(dy) = 0 \ \mu - \text{a.s.}, \tag{4.5}$$

where $a^+ = \max(0, a)$ and $a^- = -\min(a, 0)$. By Lemma 4.1 $l \sim \mu$. Thereby the violation of (4.3) contradicts to (4.5). Thus, (4.3) holds.

If $f_1(x)$ and $f_2(x)$ are solutions of (4.2), then for any constants c_1 and c_2 the function $f(x) = c_1 f_1(x) + c_2 f_2(x)$ is solution of (4.2) too. Constants c_1 and c_2 can be chosen such that to violate (4.3). Thus (4.2) obeys the desired unique solution.

Lemma 4.5. *Assume (***I'***) and (***RM'***). For every λ from $\mathbf{S}_{\mu\mu}$ with $H(\lambda|\lambda^\circ_\mu) < \infty$ there exist families λ^ε and $\mu(\varepsilon)$ ($\varepsilon \in (0,1)$) from $\mathbf{S}^2$ such that*

$$\lambda^\varepsilon \in \mathbf{S}_{\mu(\varepsilon)\mu(\varepsilon)}$$

and

$$\begin{aligned} \lambda^\varepsilon &\sim \lambda^\circ_{\mu(\varepsilon)} && (4.6)\\ H(\lambda^\varepsilon|\lambda^\circ_{\mu(\varepsilon)}) &< \infty && (4.7)\\ \lim_{\varepsilon\to 0} \rho^2(\lambda^\varepsilon, \lambda) &= 0 && (4.8)\\ \lim_{\varepsilon\to 0} H(\lambda^\varepsilon|\lambda^\circ_{\mu(\varepsilon)}) &= H(\lambda|\lambda^\circ\mu^) && (4.9) \end{aligned}$$

Proof: Let λ' be the measure, involving in (**I'**), with both marginals α. Put

$$\left\{ \begin{array}{l} \mu(\varepsilon) \\ \lambda^\varepsilon \\ \lambda^\circ_{\mu(\varepsilon)} \end{array} \right. = \left\{ \begin{array}{l} (1-\varepsilon)\mu + \varepsilon\alpha \\ (1-\varepsilon)\lambda + \varepsilon\lambda' \\ (1-\varepsilon)\lambda^\circ_\mu + \varepsilon\lambda^\circ_\alpha \end{array} \right.$$

and note that λ^ε obeys the same marginals $(1-\varepsilon)\mu + \varepsilon\alpha = \mu(\varepsilon)$.

Proof of (4.6). Let $\lambda^\circ_{\mu(\varepsilon)}(A) = 0$. Then $\lambda^\circ_\mu(A) = 0$ and $\lambda^\circ_\alpha(A) = 0$. Since $H(\lambda|\lambda^\circ_\mu) < \infty$, we get $\lambda \ll \lambda^\circ_\mu$ and so, $\lambda(A) = 0$. By virtue of (**I'**) $\lambda' \sim \lambda^\circ_\alpha$, what implies $\lambda'(A) = 0$. Consequently $\lambda^\varepsilon(A) = 0$. On the other hand, if $\lambda^\varepsilon(A) = 0$, then $\lambda(A) = 0$ and $\lambda'(A) = 0$, while the equivalence of λ' and λ°_α has as a consequence $\lambda^\circ_\alpha(A) = 0$.

Therefore $\lambda^\circ_{\mu(\varepsilon)}(A)=0$ holds provided that $\lambda^\circ_\mu(A)=0$. The last holds by virtue of $\lambda^\circ_\alpha(A)=0$ and $\lambda^\circ_\alpha(dx,dy)=p(y|x)l(dy)\alpha(dx)$ and $\lambda^\circ_\mu(dx,dy)=p(y|x)l(dy)\mu(dx)$ and also by $\mu\ll\alpha$ (Lemma 4.3)) and $\lambda^\circ_\mu\ll\lambda^\circ_\alpha$.

Thus, $\lambda^\varepsilon\sim\lambda^\circ_{\mu(\varepsilon)}$.

Proof of (4.7): Evidently $\mu\ll\mu(\varepsilon)$ and $\alpha\ll\mu(\varepsilon)$. Denote by

$$\left\{\begin{array}{l} h_\alpha(x,y)\\ h(x,y)\\ g_\mu(x)\\ g(x)\end{array}\right. = \left\{\begin{array}{l} \frac{d\lambda'}{d\lambda^\circ_\alpha}(x,y)\\ \frac{d\lambda}{d\lambda^\circ_\mu}(x,y)\\ (1-\varepsilon)\frac{d\mu}{d\mu(\varepsilon)}(x)\\ \varepsilon\frac{d\alpha}{d\mu(\varepsilon)}(x).\end{array}\right.$$

Then

$$\frac{d\lambda(\varepsilon)}{d\lambda^\circ_{\mu(\varepsilon)}}(x,y)=h(x,y)g(x)+h_\alpha(x,y)g_\mu(x)\ \ (=h^\varepsilon(x,y)).$$

Due to the definition of the conditional entropy (3.1) we obtain

$$H(\lambda^\varepsilon|\lambda^\circ_{\mu(\varepsilon)})=\int_{R^2}V(h^\varepsilon(x,y))\lambda^\circ_{\mu(\varepsilon)}(dx,dy). \tag{4.10}$$

Functions $g(x)$ and $g_\alpha(x)$ are chosen such that to satisfy $g(x)+g_\alpha(x)=1$ $\mu(\varepsilon)$-a.s. (and $\lambda^\circ_{\mu(\varepsilon)}$-a.s.). $V(x)$ is convex function and so, by Jensen's inequality

$$V(h^\varepsilon(x,y))\le g_\mu(x)V(h(x,y))+g(x)V(h_\alpha(x,y))\ \ \lambda^\circ_{\mu(\varepsilon)}-\text{a.s.} \tag{4.11}$$

The last and (3.1) imply

$$\begin{aligned} &H(\lambda^\varepsilon|\lambda^\circ_{\mu(\varepsilon)})\\ \le\ &(1-\varepsilon)\int_{R^2}V(h(x,y))\frac{d\mu}{d\mu(\varepsilon)}\lambda^\circ_{\mu(\varepsilon)}(dx,dy)+\varepsilon\int_{R^2}V(h_\alpha(x,y))\frac{d\alpha}{d\mu(\varepsilon)}\lambda^\circ_{\mu(\varepsilon)}(dx,dy)\\ =\ &(1-\varepsilon)\int_{R^2}V(h(x,y))\lambda^\circ_\mu(dx,dy)++\varepsilon\int_{R^2}V(h_\alpha(x,y))\lambda^\circ_\alpha(dx,dy)\\ =\ &(1-\varepsilon)H(\lambda|\lambda^\circ_\mu)+\varepsilon H(\lambda'|\lambda^\circ_\alpha)<\infty. \end{aligned} \tag{4.12}$$

Proof of (4.8): It is clear that the total variation $||\lambda^\varepsilon-\lambda||\le 2\varepsilon$, i.e. $\rho^2(\lambda^\varepsilon,\lambda)\to 0,\ \varepsilon\to 0$.

Proof of (4.9): Due to (4.12), $\limsup_{\varepsilon\to 0} H(\lambda^\varepsilon|\lambda^\circ_{\mu(\varepsilon)}) \leq H(\lambda|\lambda^\circ_\mu)$. The opposite inequality $\liminf_{\varepsilon\to 0} H(\lambda^\varepsilon|\lambda^\circ_{\mu(\varepsilon)}) \geq H(\lambda|\lambda^\circ_\mu)$ is derived from (4.10) by Fatou's lemma.

The proof of Theorem 4.1: It is clear that only the case $J^2(\lambda) < \infty$ has to be checked. We use the fact that in this case

$$
\begin{aligned}
&J^2(\lambda) = H(\lambda|\lambda^\circ_\mu)\\
&\lambda \in \mathbf{S}_{\mu\mu}\\
&\lambda \ll \lambda^\circ_\mu.
\end{aligned}
$$

Presuppose first that

$$\lambda \sim \lambda^\circ_\mu, \qquad \alpha_0 \sim \alpha. \tag{4.13}$$

and denote by

$$h(x,y) = \frac{d\lambda}{d\lambda^\circ_\mu}(x,y). \tag{4.14}$$

Let us define, on a mesurable space $(R^\infty, \mathcal{B}(R^\infty))$ $(R^\infty = (x_0, x_1, ...)$, $\mathcal{B}(R^\infty)$ is the Borel σ-algebra), probability measures Q and Q^μ , where Q is the distribution of the original Markov process ξ and Q^μ corresponds to a stationary Markov process with the marginal distribution μ and the transition probability

$$\pi_\mu(x,dy) = h(x,y)\pi(x,dy)(= h(x,y)p(y|x)l(dy)). \tag{4.15}$$

Due to (4.13) and Lemma 4.3, we have $\mu \ll \alpha_0$ and $p_\mu(y|x) > 0$ $(l \times \mu)$-a.s. Thereby, applying Lemma 4.4, one can conclude that Markov process $(x_k, Q^\mu)_{k\geq 0}$ is μ-ergodic. Denote by Q_n and Q^μ_n restrictions, on σ-algebra $\mathcal{B}(R^{n+1})$, of Q and Q^μ respectively. Since $\mu \ll \alpha_0$, we get $Q^\mu_n \ll Q_n, \quad n \geq 0$ herewith a process of local density $(Z_n(x^n_0) = \frac{Q^\mu_n}{Q_n}, n = 0, 1, ...$ is given by the formula:

$$Z_n(x^n_0) = \frac{d\mu}{d\alpha_0}(x_0)\exp\big(\sum_{k=1}^{n} \log h(x_{k-1}, x_k)\big). \tag{4.16}$$

Lebesgue's decomposition of Q_n with respect to Q^μ_n implies an inequality: for each Δ from $\mathcal{B}(R^{n+1})$

$$Q(\Delta) \geq \int_{\Delta} Z^{-1}(x_0^n) dQ_n^{\mu}. \tag{4.17}$$

Introduce now occupation measures $\pi_n^2(x_0^n)$: $\pi_n^2(A \times B)(x_0^n) = \frac{1}{n}\sum_{k=1}^n I(x_{k-1} \in A, x_k \in B)$ and define sets

$$\begin{aligned}
\Delta_n^1 &= \{\rho^2(\pi_n^2(x_0^n), \lambda) \leq \delta\} \\
\Delta_n^2 &= \{|\frac{1}{n}\sum_{k=1}^n h(x_{k-1}, x_k) - H(\lambda|\lambda_\mu^\circ)| \leq \beta\} \\
\Delta^c &= \{\frac{d\mu}{d\alpha_0}(x_0) > 1/c\}.
\end{aligned} \tag{4.18}$$

Put $\Delta^{n,c} = \Delta_n^1 \cap \Delta_n^2 \cap \Delta^c$. Then, taking into account (4.17), we find

$$\begin{aligned}
P(\rho_n^2(\pi_n^2, \lambda) \leq \delta) &= Q(\rho_n^2(\pi_n^2(x_0^n), \lambda) \leq \delta) \\
&\geq \int_{\Delta^{n,c}} \mathcal{Z}_n^{-1}(x_0^n) dQ_n^\mu \\
&\geq cQ_n^\mu(\Delta^{n,c}) \exp(-nH(\lambda|\lambda_\mu^\circ) - n\beta).
\end{aligned}$$

Hence

$$\liminf_n \log P(\rho^2(\pi_n^2, \lambda) \leq \delta) \geq -H(\lambda|\lambda_\mu^\circ) - \beta + \liminf_n \log Q_n^\mu(\Delta^{n,c}).$$

Assume

$$\begin{aligned}
\lim_{c\to\infty} \mu(x_0 : \frac{d\mu}{d\alpha}(x_0) > 1/c) &= 1 \\
\lim_n Q_n^\mu(\Delta_n^1) &= 1 \\
\lim_n Q_n^\mu(\Delta_n^2) &= 1.
\end{aligned} \tag{4.19}$$

Then, accordingly to an arbitrariness of β, the required lower bound holds. The first part of (4.19) takes place since, as it was mentioned above, $\mu \ll \alpha_0$. The second part is implied by the ergodicity of $(x_k, Q^\mu)_{k\geq 0}$ as long as the two-dimensional distribution of it is λ and so, due to Birkhoff-Khinchin's theorem, $\lim_n \rho^2(\pi_n^2(x_0^n), \lambda) = 0$ Q^μ-a.s. By making assumption $\infty > H(\lambda|\lambda_\mu^\circ) = \int_{R^2} \log h(x,y)\lambda(dx,dy)$ and so, due to Birkhoff-Khinchin's theorem, the last part holds too.

Thus, under (4.13), the lower bound holds.

Let the second condition ($\alpha_0 \sim \alpha$) in (4.13) be valid and let λ^ε and $\mu(\varepsilon)$ be measures from Lemma 4.5. For these measures, the first part of (4.13) is valid. Show that $\mu(\varepsilon) \ll \alpha_0$. By virtue of Lemmas 4.1 and 4.3, $\mu \ll \alpha$ what implies $\mu(\varepsilon) \ll \alpha$. The last, due to the second part in (4.13), implies the required absolute continuity.

By proved above $\liminf_n \frac{1}{n} \log P(\rho^2(\pi_n^2, \lambda^\varepsilon) \le \delta) \ge -H(\lambda^\varepsilon|\lambda^\circ_{\mu(\varepsilon)})$. Accordingly to Lemma 4.5, one can choose ε_0 such that for fixed δ and $\varepsilon \le \varepsilon_0$ an inequality takes place: $\rho^2(\lambda, \lambda^\varepsilon) \le \delta/2$. Then, taking into account Lemma 4.5, we arrive to

$$\begin{aligned}\liminf_n \frac{1}{n} \log P(\rho^2(\pi_n^2, \lambda) \le \delta) &\ge \liminf_n \frac{1}{n} \log P(\rho^2(\pi_n^2, \lambda^\varepsilon) \le \delta/2) \\ &\ge -H(\lambda^\varepsilon|\lambda^\circ_{\mu(\varepsilon)}) \\ &\to -H(\lambda|\lambda^\circ_\mu),\ \varepsilon \to 0,\end{aligned}$$

that is the required inequality takes place provided that the second part in (4.9) holds. To relinquish on it, introduce, following Donsker and Varadhan [1], a new Markov process $\xi^1 = (\xi^1_k)_{k \ge 0}$ with $\xi^1_k \equiv \xi_{k+1}$. It has the same transition probability $\pi(x, dy)$ and initial distribution $\alpha^1_0(dy) = \int_R \pi(x, dy)\alpha_0(dx) = l(dy) \int_R p(y|x)\alpha_0(dx)$. Hence and by virtue of $(\mathbf{H}_0)$, it follows that $\alpha^1_0 \sim l$ and so, by Lemma 4.1 $\alpha_0 \sim \alpha$. Therefore, by the previous proof, the lower bound holds for $\pi_n^2(\xi^1)$, where

$$\pi_n^2(A \times B(\xi^1) = \frac{1}{n} \sum_{k=1}^n I(\xi^1_{k-1} \in A, \xi^1_k \in B) = \frac{1}{n} \sum_{k=1}^n I(\xi_k \in A, \xi_{k+1} \in B),$$

that is

$$\liminf_n \frac{1}{n} \log P(\rho^2(\pi_n^2(\xi^1), \lambda) \le \delta) \ge -H(\lambda|\lambda^\circ_\mu). \tag{4.20}$$

The required result now follows from (4.20) and the fact that the total variation $||\pi_n^2 - \pi_n^2(\xi^1)|| \le 2/n$.

5. Proof of Theorem 1.2.

By Theorem 2.1 the family (π_n^2) is exponentially tight. Then by (1.8) (Puhalskii's theorem) there exists a subsequence $(\pi^2_{\tilde{n}})$, obeying the LDP with the rate function $\tilde{J}^2(\lambda)$ and so, for $\delta > 0$ the upper and lower bounds hold:

$$\begin{aligned}\liminf_{\tilde n} \frac{1}{\tilde n}\log P(\rho^2(\pi^2_{\tilde n},\lambda)\le\delta) &\ge \liminf_{\tilde n}\frac{1}{\tilde n}\log P(\rho^2(\pi^2_{\tilde n},\lambda)<\delta)\\ &\ge -\inf_{\nu:\rho^2(\nu,\lambda)<\delta}\tilde J^2(\nu)\ge -\tilde J^2(\nu). \qquad (5.1)\end{aligned}$$

and

$$\limsup_{\tilde n}\frac{1}{\tilde n}\log P(\rho^2(\pi^2_{\tilde n},\lambda)\le\delta)\le -\inf_{\nu:\rho^2(\nu,\lambda)<\delta}\tilde J^2(\nu). \qquad (5.2)$$

From the definition of inf it follows that for every $\varepsilon>0$ and $\delta>0$ there exists a measure ν^δ_ε such that $\rho^2(\nu^\delta_\varepsilon,\lambda)\le\delta$ and

$$\inf_{\nu:\rho^2(\nu,\lambda)<\delta}\tilde J^2(\nu)\ge\tilde J^2(\nu^\delta_\varepsilon)-\varepsilon. \qquad (5.3)$$

Since $\tilde J^2(\lambda)$ is rate function (see condition (0) from the definition of LDP), it is semi-continuous from below. The last implies $\liminf_{\delta\to0}\tilde J^2(\nu^\delta_\varepsilon)\ge\tilde J^2(\lambda)$ and so, accordingly to (5.2) and (5.3)

$$\begin{aligned}&\limsup_{\tilde n}\frac{1}{\tilde n}\log P(\rho^2(\pi^2_{\tilde n},\lambda)\le\delta)\\ \le\;&\limsup_{\delta\to0}[\varepsilon-\tilde J^2(\nu^\delta_\varepsilon)]\\ =\;&\varepsilon-\liminf_{\delta\to0}\tilde J^2(\nu^\delta_\varepsilon)\\ \le\;&\varepsilon-\tilde J^2(\lambda)\to-\tilde J^2(\lambda),\ \varepsilon\to0. \qquad (5.4)\end{aligned}$$

Thus, (5.1) and (5.4) have as a sequence the local LDP (see (1.9) for $(\pi^2_{\tilde n})$.

As it was mentioned in the introduction, the identity $\tilde J^2(\lambda)\equiv J^2(\lambda)$ takes place . Using this fact, one can establish the LDP upper and lower bounds for (π^2_n).

For closed set $F^2\in\mathbf{S}^2$, choose subsequences $(\pi^2_{n'})$ and (π^2_n) such that

$$\limsup_n\frac{1}{n}\log P(\pi^2_n\in F^2)=\lim_{n'}\frac{1}{n'}\log P(\pi^2_{n'}\in F^2).$$

By (1.8), one can choose a subsequence $(\pi^2_{\tilde n})$ from $(\pi^2_{n'})$ such that $(\pi^2_{\tilde n})$ obeys the LDP with rate function $\tilde J^2(\lambda)\equiv J^2(\lambda)$.

Then

$$\begin{aligned}\limsup_n \frac{1}{n}\log P(\pi_n^2 \in F^2) &= \lim_{n'} \frac{1}{n'}\log P(\pi_{n'}^2 \in F^2)\\ &= \lim_{\tilde{n}} \frac{1}{\tilde{n}}\log P(\pi_{\tilde{n}}^2 \in F^2) \le -\inf_{\lambda\in F^2} J^2(\lambda).\end{aligned}$$

The lower bound is proved in the same way.

6. Proof of Theorem 1.1.

Let λ^1 λ^2 be from $\mathbf{S}^2$. Their marginals are denoted by (μ^1,ν^1), (μ^2,ν^2). By Lemma 3.2 $\rho^2(\lambda^1,\lambda^2) \ge \begin{cases}\rho(\mu^1,\mu^2)\\ \rho(\nu^1,\nu^2).\end{cases}$ Therefore, a mapping $\lambda^1 \longrightarrow (\mu^1,\nu^1)$ is uniformly continuous in Levy-Prokhorov's metric. Hence, by the contraction principle of Varadhan [3], the family $(\pi_n, n\ge 1)$ obeys the LDP in $(\mathbf{S},\rho)$ with rate function $J'(\mu) = \inf J^2(\lambda)$, where inf is taken over all measures λ from $\mathbf{S}^2$ if only one marginal is μ. On the other hand, since $J^2(\lambda)=\infty$ for any λ with different marginals we arrive to $J'(\mu) = \inf_{\lambda\in\mathbf{S}_{\mu\mu}} H(\lambda|\lambda^\circ_\mu)$.

7. Example

Let $\xi = (\xi_k)_{k\ge 0}$ be defined by a recursion $\xi_{k+1} = f(\xi_k) + g(\xi_k)\zeta_{k+1}$, where $(\zeta_k)_{k\ge 1}$ is i.i.d. sequence of random variable, which is independent of ξ_0.

Assume

1. $f(x)$ and $g(x)$ are continuous function such that $|\frac{f(x)}{x}| \le a < 1 \qquad 1/r \le |g(x)| \le r,\ r > 0$;

2. the distribution of random variable ζ_1 obeys a density $p_\zeta(y)$ with respect to Lebesgue measure and what is more $p_\zeta(y)$ is continuous function such that $r_1\exp(-r_2|y|^q) \le p_\zeta(y) \le r_3\exp(-r_4|y|)\ \ r_i > 0, i=1,...,4,\ q\ge 1$;

3. There exists $b>0$ such that $Ee^{b|\xi_0|} < \infty$.

Show that, under above-mentioned conditions, both LDP's in $(\mathbf{S}^2,\rho^2)$ and $(\mathbf{S},\rho)$ take place.

Due to Theorems 1.1 and 1.2, the following implication has to be checked only. $\{\mathbf{1.},\mathbf{2.},\mathbf{3.}\} \Longrightarrow \{(\mathbf{F}),(\mathbf{I'}),(\mathbf{H}^\star),(\mathbf{RM'}),(\mathbf{H}_0)\}$.

By **2.**, for every x the transition probability $\pi(x,dy)$ obeys a density with respect to Lebesgue measure: $p(y|x) = p_\zeta(\frac{y-f(x)}{g(x)})$. Thereby $\{\mathbf{1.},\mathbf{2.}\} \Longrightarrow \{(\mathbf{F})\}$. As λ', one can take the two-dimensional distribution of a stationary Gaussian Markov process $(\xi'_k)_{k\geq 0}$ defined by a linear recursion $\xi'_{k+1} = a\xi'_k + \zeta'_{k+1}$, where $(\zeta'_k)_{k\geq 1}$ is i.i.d. sequence of (0,1)-Gaussian random variables, independent of ξ'_0, which is $(0, 1/(1-a^2))$-Gaussian random variable. Then $\frac{d\alpha}{dx}(x) = \sqrt{\frac{1-a^2}{2\pi}}\exp\big(-\frac{x^2}{2/(1-a^2)}\big)$ and $\lambda'(dx,dy) = \frac{1}{\sqrt{2\pi}}\exp\big(-\frac{(y-ax)^2}{2}\big)dy\alpha(dx)$. On the other hand, we get $\lambda'(dx,dy) = p_\zeta(\frac{y-f(x)}{g(x)})dy\alpha(dx)$ and so, $\lambda' \sim \lambda^\circ_\alpha$ with $\frac{d\lambda'}{d\lambda^\circ_\alpha}(x,y) = \frac{\sqrt{\frac{1-a^2}{2\pi}}\exp\left(-\frac{(y-ax)^2}{2}\right)}{p_\zeta\left(\frac{y-f(x)}{g(x)}\right)}$. Then, correspondingly to **1.**, **2.**

$$\frac{d\lambda'}{d\lambda^\circ_\alpha}(x,y) \leq \frac{1}{r_1}\exp\big(r_2|\frac{y-f(x)}{g(x)}|^q\big) \leq \frac{1}{r_1}\exp\big(\frac{r_2}{r}[|y|+a|x|]^q\big),$$

what has as a consequence $H(\lambda'|\lambda^\circ_\alpha) \leq -\log r_1 + \int_{R^2}\frac{r_2}{r}[|y|+a|x|]^q\lambda'(dx,dy) < \infty$, i.e. $\{\mathbf{1.},\mathbf{2.}\} \Longrightarrow \{(\mathbf{I'}))\}$.

To verify $(\mathbf{H}^\star)$, taking $v(x) = c|x| \;\; c \in (a\frac{r_4}{r}, \frac{r_4}{r})$, find

$$w(x) = c|x| - \log\int_R \exp(c|y|)p_\zeta(\frac{y-f(x)}{g(x)})dy.$$

Hence and due to **2.** we arrive to an inequality:

$$\begin{aligned} w(x) &\geq -\log r_3 + c|x| - \log\int_R \exp\big(c|y| - r_4|\frac{y-f(x)}{g(x)}|\big)dy \\ &\geq -\log r_3 + c|x| - \log\int_R \exp\big(c|y| - \frac{r_4}{r}|y| + a\frac{r_4}{r}|x|\big)dy \\ &= (c - a\frac{r_4}{r})|x| - \log r_3 - \log\int_R \exp\big(-[\frac{r_4}{r} - c]|y|\big)dy, \end{aligned}$$

i.e. $\{\mathbf{2.}\} \Longrightarrow \{(\mathbf{H}^\star)\}$ and, what is more, $\{\mathbf{2.}\} \Longrightarrow \{(\mathbf{RM'})\}$ and $\{\mathbf{2.},\mathbf{3.}\} \Longrightarrow \{(\mathbf{H}_0)\}$.

References

[1] M.D. Donsker and S.R.S Varadhan, "Asymptotic evaluation of certain Markov process expectations for large time, III". *Communication of Pure and Applied Mathematics*, **XXIX**, 1976, pp. 3891–462.

[2] M.D. Donsker and S.R.S Varadhan, "Asymptotic evaluation of certain Markov process expectations for large time, I". *Communication of Pure and Applied Mathematics*, **XXVIII**, 1975, pp. 1–17.

[3] S.R.S Varadhan, *Large Deviations and Applications.* SIAM, Philadelphia, 1984.

[4] M.D. Donsker and S.R.S Varadhan, "Asymptotic evaluation of certain Markov process expectations for large time, IV". *Communication of Pure and Applied Mathematics*, **XXXVI**, 1983, pp. 183–212.

[5] R.S. Ellis, "Large deviations for the empirical of a Markov chain with an application to the multivariate empirical measure." *Ann. Probab.*, **16**, 1988, pp. 1496–1508.

[6] P. Dupuis and R.S. Ellis, "A stochastic optimal control approach to the theory of large deviations". *Preprint*, 1992.

[7] J.D. Deuschel and D.W. Stroock, *Large Deviations.* Academic Press, New York, 1989.

[8] A.A Puhalskii, "On functional principle of large deviations". *New trends in Probability and Statistics., Vilnius, Lithuania, VSP/Mokslas*, 1991, pp. 198-218.

[9] M.I. Freidlin and A.D. Wentzell, *Random Perturbations of Dynamical Systems.* N.Y. Springer. 1984.

[10] R.S. Liptser and A.A. Pukhalskii, "Limit theorems on large deviations for semimartingales". *Stochastics and Stochastic Reports*, **38**, 1992, pp. 201-249.

[11] De A. Acosta, "Large deviations for empirical measures of Markov chains".*J. Theoret. Probab.*, **3**, 1990, pp. 395–431.

[12] J. Gärtener, "On large deviations for the invariant measure". *Theory of probability and its application*, **22**, N 1, 1977, pp. 27–42.

[13] S. Orey and S. Pelican, "Large deviation principles for stationary processes", *Ann. Probab.*, **16**, 1988, pp. 1481–1495.

[14] A.Yu. Veretennikov, "On large deviations for ergodic process empirical measures". *Advances in Soviet Mathematics* **12**, 1992, pp. 125–133.

[15] L. Wu, "Grandes déviation pour les processus de Markov essentiellement irréductibles. I. Temps discret". *C.R. Acad. Sci. Paris*, **312**, N 1, 1991, pp. 609–614.

[16] N.C. Jain, "Large deviation lower bounds for additive functionals of Markov processes". *Annals of Probability*, **18**, N 3, 1990, pp. 1071-1098.

[17] P. Billingsley, *Convergence of probability measures.*, John Willey and Sons, New York-London-Sydney-Toronto, 1968.

[18] A.V. Skorokhod, *Asymptotics method in the theory of stochastic differential equations.* Naukova Dumka, Kiev, 1987.

[19] P. Dupuis and R.S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations.* John Willey and Sons, to be published.