LARGE DEVIATIONS FOR ERGODIC PROCESSES IN SPLIT SPACES

ADINA OPRISAN AND ANDRZEJ KORZENIOWSKI

Department of Mathematics, University of Texas at Arlington, Box 19404, Arlington, TX 76019, USA

ABSTRACT. We study a family of stochastic additive functionals of Markov processes with locally independent increments switched by jump Markov processes in an asymptotic split phase space. Based on an average approximation, we obtain a large deviation result for this stochastic evolutionary system using a weak convergence approach. Examples, including compound Poisson processes, illustrate cases in which the rate function is calculated in an explicit form.

AMS (MOS) Subject Classification. 60F10, 60H10, 60B10.

1. Introduction

The main mathematical object of this paper is a family of coupled Markov processes $(\xi(t), x(t))$, $t \geq 0$ called the switched and switching processes, respectively. The switched process describes the evolution of the system and it is a stochastic functional of the process $\eta(t; x)$, $t \geq 0, x \in E$ with locally independent increments [6] (they are also called weakly differentiable [5] or piecewise deterministic [2] processes). In order to reduce the complexity of the phase space, the switching processes that describe the random changes in the evolution of the system, are jump Markov processes considered in a split space $E = \cup_{k=1}^{N} E_k$, $E_k \cap E_{k'} = \emptyset$, $k \neq k'$ with non-communicating components, and having the ergodic property on each class E_k. By introducing the parameter $\epsilon > 0$ one defines a jump Markov process on the split phase space with small transition probabilities between the states of the system and further merges the classes E_k, $k = 1, 2, \cdots, N$ into distinct states k, $1 \leq k \leq N$. The average limit theorem of the stochastic additive functional with fast time-scaling switching process is obtained by using the martingale characterization [11] and a solution of the singular perturbation problem for reducible-invertible operators [7]. We are interested in finding the large deviation principle for this sequence of stochastic additive functionals. Using the weak convergence approach of Dupuis and Ellis [3], a large deviation principle is derived for a sequence of random walks constructed such that they have the same distribution as the linear interpolation sequence of samples of stochastic additive functionals.

This research was supported by Mathematical Sciences Division, US Army Research Office, Grant No. W911NF-07-0283.

2. Preliminaries

Let $(E, \mathcal{E})$ be a complete, separable metric space. We will use the following notation and definitions throughout the paper.

$\mathbf{D}[0,\infty)$ the space of right continuous functions having left hand side limits. This embedded with Skorokhod metric becomes a complete, separable metric space.

$(\mathbf{C}[0,\infty), ||\cdot||)$ with $||x|| = sup_{t\geq 0}|x(t)|$, $x \in \mathbf{C}[0,\infty)$ is a complete, separable metric space.

$\mathbf{B}$ the Banach space of all bounded measurable functions with norm, $||\varphi|| = sup_{x\in E}|\varphi(x)|$, $\varphi \in \mathbf{B}$.

Stochastic space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ with the filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions: it is right-continuous and complete.

The family of cadlag Markov processes $\eta(t;x)$, $t \geq 0, x \in E$, parameterized by x, are such that $\eta(t;x(t))$ is measurable, are of locally independent increment processes determined by their infinitesimal generators

$$\mathbb{\Gamma}(x)\varphi(u) = a(u;x)\varphi'(u) + \int_{\mathbb{R}^d}[\varphi(u+v) - \varphi(u) - v\varphi'(u)]\Gamma(u, dv; x), \tag{2.1}$$

where the positive kernels $\Gamma(u, dv; x), x \in E$, are continuous and bounded on $u \in \mathbb{R}^d$, and uniformly continuous and bounded on $x \in E$, and the product $a\varphi'$ stands for the inner product $< a, \nabla\varphi >$ in $\mathbb{R}^d$.

The switching jump Markov process $x(t)$, $t \geq 0$ is defined by its infinitesimal generator Q as

$$Q\varphi(x) = q(x)\int_E P(x, dy)[\varphi(y) - \varphi(x)], \tag{2.2}$$

where the kernel $P(x, B)$ is the transition kernel of the embedded Markov chain, and $q(x), x \in E$ is the intensity of jumps function.

The Markov additive process $(\xi(t), x(t))$, $t \geq 0$ is determined by the infinitesimal generator

$$\mathbb{L}\varphi(u, x) = Q\varphi(u, x) + \mathbb{\Gamma}(x)\varphi(u, x). \tag{2.3}$$

The general scheme of phase merging is realized by the family of time-homogeneous cadlag Markov jump process $x^\epsilon(t)$, $t \geq 0, \epsilon > 0$ with the standard phase space $(E, \mathcal{E})$, on the split phase space

$$E = \bigcup_{k=1}^{N} E_k, \qquad E_k \bigcap E_{k'} = \emptyset, \qquad k \neq k'$$

given by the infinitesimal generator

$$Q^\epsilon\varphi(x) = q(x)\int_E P^\epsilon(x, dy)[\varphi(y) - \varphi(x)]. \tag{2.4}$$

The phase merging algorithm is considered under the following assumptions:

A1. The stochastic kernel in (2.4) is represented in the following form

$$P^\epsilon(x,B) = P(x,B) + \epsilon P_1(x,B)$$

where the stochastic kernel $P(x,B)$ is coordinated with the splitting as follows:

$$P(x,E_k) = 1\!\!1_k(x) := \begin{cases} 1, & x \in E_k, \\ 0, & x \notin E_k \end{cases}$$

A2. The Markov supporting process $x(t)$, $t \geq 0$, on the state space $(E, \mathcal{E})$, determined by the generator Q given in (2.2) is supposed to be uniformly ergodic in every class E_k, $1 \leq k \leq N$, with the stationary distribution $\pi_k(dx)$, $1 \leq k \leq N$, satisfying the following relations

$$\pi_k(dx)q(x) = q_k\rho_k(dx), \qquad q_k = \int_{E_k} \pi_k(dx)q(x),$$

$$\rho_k(B) = \int_{E_k} \rho_k(dx)P(x,B), \qquad \rho_k(E_k) = 1.$$

The perturbing operator $P_1(x,B)$ is a signed kernel which satisfies the conservative condition $P_1(x,E) = 0$.

A3. The average exit probabilities satisfy the following condition

$$\hat{p}_k := \int_{E_k} \rho_k(dx)P_1(x, E\backslash E_k) > 0, \qquad 1 \leq k \leq N.$$

Introduce the merging function $m(x) = k, \quad x \in E_k, 1 \leq k \leq N$, and the merged process

$$\hat{x}^\epsilon(t) := m(x^\epsilon(t/\epsilon)), \quad t \geq 0, \tag{2.5}$$

on the merged phase space $\hat{E} = \{1, ..., N\}$.

The phase merging principle establishes the weak convergence of the above process to the limit Markov process $\hat{x}(t)$.

Theorem 2.1. (*Ergodic phase merging principle*) *Under the assumptions $A1 - A3$, the following weak convergence holds*

$$\hat{x}^\epsilon(t) \Rightarrow \hat{x}(t), \qquad \epsilon \to 0.$$

The limit merged Markov process $\hat{x}(t), t \geq 0$, on the merged state space $\hat{E}$ is determined by the generator matrix

$$\hat{Q} = (\hat{q}_{kr}; 1 \leq k, r \leq N),$$

with entries

$$\hat{q}_{kr} = \hat{q}_k p_{kr}, \qquad p_{kr} = \int_{E_k} \rho_k(dx)P_1(x,E_r), \qquad 1 \leq k, r \leq N, \tag{2.6}$$

where ρ_k is the stationary distribution of the corresponding embedded Markov chain.

Theorem 2.2. (*Average approximation* [8]) *Let the stochastic evolutionary system $\xi^\epsilon(t), t \geq 0$ be represented by*

$$\xi^\epsilon(t) = \xi^\epsilon(0) + \int_0^t \eta^\epsilon\left(ds; x^\epsilon(\frac{s}{\epsilon})\right), \qquad t \geq 0, \epsilon > 0. \tag{2.7}$$

Let the process $\eta^\epsilon(t;x)$, $t \geq 0, \epsilon > 0, x \in E$ be given by the infinitesimal generators

$$\mathbb{\Gamma}^\epsilon(x)\varphi(u) = a^\epsilon(u;x)\varphi'(u) + \epsilon^{-1}\int_{\mathbb{R}^d}[\varphi(u+\epsilon v) - \varphi(u) - \epsilon v\varphi'(u)]\Gamma_\epsilon(u, dv; x). \tag{2.8}$$

Let the switching Markov process $x^\epsilon(t)$, $t \geq 0$ satisfies the phase merging condition of Theorem 2.1. Let the following conditions be valid

C1. *the drift velocity $a(u;x)$ belongs to the Banach space $\mathbf{B}^1$, with*

$$a^\epsilon(u;x) = a(u;x) + \theta^\epsilon(u;x)$$

where $\theta^\epsilon(u;x) \to 0$ as $\epsilon \to 0$ uniformly on $(u;x)$ and $\Gamma_\epsilon(u, dv; x) \equiv \Gamma(u, dv; x)$ independent of ϵ.

C2. *the operator $\gamma^\epsilon(x)\varphi(u) = \epsilon^{-1}\int_{\mathbb{R}^d}[\varphi(u+\epsilon v) - \varphi(u) - \epsilon v\varphi'(u)]\Gamma(u, dv; x)$ is negligible on $\mathbf{B}^1$:*

$$\sup_{\varphi \in C^1(\mathbb{R}^d)} ||\gamma_\epsilon(x)\varphi|| \to 0 \qquad as \qquad \epsilon \to 0$$

C3. *the convergence in probability of the initial values of $(\xi^\epsilon(t), m(x^\epsilon(\frac{t}{\epsilon}))$, $t \geq 0$ holds, that is*

$$(\xi^\epsilon(0), m(x^\epsilon(0)) \to (\xi(0), \hat{x}(0))$$

and there exists a constant $c \in \mathbb{R}_+$ such that $\sup_{\epsilon>0} \mathbb{E}|\xi^\epsilon(0)| \leq c < \infty$.

Then the stochastic evolutionary system $\xi^\epsilon(t)$, $t \geq 0$ defined by (2.7) converges weakly to the averaged stochastic system $\hat{\xi}(t)$,

$$\xi^\epsilon(t) \Rightarrow \hat{\xi}(t) \qquad as \qquad \epsilon \to 0.$$

The limit process $\hat{\xi}(t), t \geq 0$ is defined by a solution of the evolutionary equation

$$\frac{d}{dt}\hat{\xi}(t) = \hat{a}(\hat{\xi}(t); \hat{x}(t)), \qquad \hat{\xi}(0) = \xi(0), \tag{2.9}$$

where the averaged velocity is determined by

$$\hat{a}(u;k) = \int_{E_k} \pi_k(dx)a(u;x) \qquad 1 \leq k \leq N.$$

Note 2.3. The limit process $\hat{\xi}(t)$ is a random dynamical system evolving deterministically on random time intervals $[\tau_i, \tau_{i+1})$, where $\{\tau_i\}_{i=1}^{N(T)}$ are the transition times of the stationary merged process $\hat{x}(t)$ and $N(T)$ the number of transitions on $[0, T]$.

3. Large deviation principle for ergodic Markov processes

Let $x(t)$, $t \in \mathbb{R}_+$ be a time-homogeneous Markov process on a compact metric space X, $\mathcal{B}(X)$ be the Borel σ-algebra in X and $\mathcal{M}(X)$ the space of probability measures on $\mathcal{B}(X)$. Let introduce a random measure on $\mathcal{B}(X)$ by

$$\nu_t(B) = \frac{1}{t}\int_0^t 1_{\{x(s)\in B\}} ds, \quad B \in \mathcal{B}(X).$$

Theorem 3.1. *Assume that the process $x(t)$, $t \in \mathbb{R}_+$ is an ergodic Markov process. Then the following large deviation result holds*

$$-\inf_{m\in\Gamma^\circ} I(m) \le \liminf_{t\to\infty} \frac{1}{t}\log \mathbb{P}\{\nu_t \in \Gamma\} \le \limsup_{t\to\infty} \frac{1}{t}\log \mathbb{P}\{\nu_t \in \Gamma\} \le -\inf_{m\in\bar{\Gamma}} I(m)$$

where the rate function $I : \mathcal{M}(X) \to [0, +\infty]$ is defined by

$$I(m) = -\inf\{\int (\phi(x))^{-1} Q\phi(x) m(dx) : \phi \in \mathcal{D}(Q), \phi > 0\}$$

and $\Gamma \in \mathcal{B}(\mathcal{M}(X))$ be the Borel σ-algebra in $\mathcal{M}(X)$.
Typically $\Gamma = \{\nu \in \mathcal{M}(X) | d(\nu, m) > \delta, I(m) = 0\}$ where d is some metric on $\mathcal{M}(X)$.

The rate function $I(m)$ verifies the following properties

(i) $I(m) \ge 0$ for all $m \in \mathcal{M}(X)$, and $I(m) = 0$ if and only if m is the invariant measure for the ergodic Markov process,
(ii) $I(m)$ is a convex function, i.e.,

$$I(sm_1+(1-s)m_2) \le sI(m_1)+(1-s)I(m_2), \quad m_i \in \mathcal{M}(X),\ i = 1, 2,\ 0 < s < 1$$

(iii) $I(m)$ is a lower semi-continuous function, i.e.,

$$\liminf_{m_n\to m} I(m_n) \ge I(m)$$

(iv) For any $b > 0$ the set $C_b(I) = \{m : I(m) \le b\}$ is compact, and the function $I(m)$ is continuous on this compact set.

For proofs and details see [9] and [10].

We illustrate the concept of the split phase space in the following example.

Example 3.2.

Let us consider a four-state Markov process $x(t)$, $t \in \mathbb{R}_+$ on the split phase space $E = \{1, 2, 3, 4\} = E_1 \cup E_2$, $E_1 = \{1, 2\}$, $E_2 = \{3, 4\}$ generated by

$$Q = \begin{pmatrix} -\lambda_1 & \lambda_1 & 0 & 0 \\ \mu_1 & -\mu_1 & 0 & 0 \\ 0 & 0 & -\lambda_2 & \lambda_2 \\ 0 & 0 & \mu_2 & -\mu_2 \end{pmatrix}$$

One checks that the Markov process $x(t)$ is ergodic in both E_1 and E_2 with stationary distributions $\pi_1 = (\frac{\mu_1}{\lambda_1+\mu_1} \quad \frac{\lambda_1}{\lambda_1+\mu_1})$ and $\pi_2 = (\frac{\mu_2}{\lambda_2+\mu_2} \quad \frac{\lambda_2}{\lambda_2+\mu_2})$.

Now we analyze singularly perturbed Markov processes by introducing a small parameter $\epsilon > 0$ which leads to a singular perturbed system involving two-time scales, the actual time t and the stretched time $\frac{t}{\epsilon}$. Since the process $x(t)$ is ergodic on E_1, E_2, the system can be decomposed and the states of the Markov process can be aggregated.

Let $x^\epsilon(t)$ be a Markov chain on E generated by $Q + \epsilon Q_1$ with Q defined above and Q_1 given by

$$Q_1 = \begin{pmatrix} -\lambda_1 & 0 & \lambda_1 & 0 \\ 0 & -\mu_1 & 0 & \mu_1 \\ \lambda_2 & 0 & -\lambda_2 & 0 \\ 0 & \mu_2 & 0 & -\mu_2 \end{pmatrix}$$

and $x^\epsilon(\frac{t}{\epsilon})$ be a time-invariant Markov process with generator $Q^\epsilon = \frac{1}{\epsilon}Q + Q_1$.

Note that for small ϵ, the Markov process $x^\epsilon(\frac{t}{\epsilon})$ jumps more frequently within each block and less frequently from one block to another. To further understanding of the underlying process, we consider the merged process $\hat{x}^\epsilon(t) := m(x^\epsilon(\frac{t}{\epsilon}))$ obtained by aggregating the states in the k^{th} block by a single state k and study its asymptotic behavior (for many asymptotic results see [12]).

Theorem 2.1 states that the limit process is a Markov process on the merged space $\hat{E} = \{1, 2\}$ determined by generator matrix $\hat{Q} = (\hat{q}_{kr}, 1 \leq k, r \leq 2)$ with $\hat{q}_{kr}$ verifying (2.6). For this example, $q_1 = \frac{2\lambda_1\mu_1}{\lambda_1+\mu_1}$, $q_2 = \frac{2\lambda_2\mu_2}{\lambda_2+\mu_2}$, $p_{11} = -1, \quad p_{12} = 1, \quad p_{21} = 1, \quad p_{22} = -1$. Thus,

$$\hat{Q} = \begin{pmatrix} -\frac{2\lambda_1\mu_1}{\lambda_1+\mu_1} & \frac{2\lambda_1\mu_1}{\lambda_1+\mu_1} \\ \frac{2\lambda_2\mu_2}{\lambda_2+\mu_2} & -\frac{2\lambda_2\mu_2}{\lambda_2+\mu_2} \end{pmatrix} := \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}$$

Note 3.3. The merged process $\hat{x}^\epsilon(t)$, unlike its limit $\hat{x}(t)$, is not time-homogeneous.

Let us consider now the occupational time of $\hat{x}(t)$ defined by

$$\nu_t(B) = \frac{1}{t}\int_0^t \mathbb{1}_{\{\hat{x}(t)\in B\}}(s)ds$$

for any $B \in \mathcal{B}(\hat{E})$. By ergodic theorem, the measure ν_t converges to the ergodic distribution ρ as t goes to ∞. As an example, let $\mathcal{M}$ be the set of all probability measures on $\{0, 1\}$ identified with $\{(p, 1-p), 0 \leq p \leq 1\}$ and $d(x, y) = |x| + |y|$, $x, y \in \mathbb{R}^2$. Then Theorem 3.1 implies that for $\rho = (p_0, 1-p_0)$ and $\Gamma = \{(p, 1-p) \,|\, d((p, 1-p), (p_0, 1-p_0)) > \delta\}$, $\mathbb{P}(\nu_t \in \Gamma) \sim \exp(-tI(p_0 + \delta, 1 - p_0 - \delta))$ for large t with

$$I(m) = -\inf\{\int_{\hat{E}} \frac{(\hat{Q}\phi)(y)}{\phi(y)} m(dy) : \phi \in \mathcal{D}(\hat{Q}), \phi(y) > 0, \forall y \in \{0, 1\}\} =$$

$$-\inf\{\lambda p(\frac{\phi(2)}{\phi(1)}-1)+\mu(1-p)(\frac{\phi(1)}{\phi(2)}-1); \quad \phi(1),\phi(2)>0\}$$

for $m=(p,1-p)$.

The infimum is attained at $\sqrt{\frac{\mu}{\lambda}(\frac{1}{p}-1)}$ and $I(m)=\lambda p+\mu(1-p)-2\sqrt{\lambda\mu p(1-p)}$.

4. Large deviations for stochastic additive functionals

Let us consider the family of stochastic additive functionals $\xi^\epsilon(t)$, $t\geq 0$ represented by

$$\xi^\epsilon(t)=\xi^\epsilon(0)+\int_0^t \eta^\epsilon\left(ds;x^\epsilon(\frac{s}{\epsilon})\right), \qquad t\geq 0,\epsilon>0.$$

The family of coupled Markov processes $(\xi^\epsilon(t),x^\epsilon(\frac{t}{\epsilon}))$, $t\geq 0$, $\epsilon>0$ on $\mathbb{R}^d\times E$ has infinitesimal generator $\mathbb{L}^\epsilon$ given by $\mathbb{L}^\epsilon=\frac{1}{\epsilon}Q+Q_1+\mathbb{\Gamma}^\epsilon(x)$ with the domain $\mathcal{D}(\mathbb{L}^\epsilon)$ dense in $\mathbf{C}(\mathbb{R}^d\times E)$ and the limit process $(\hat{\xi}(t),\hat{x}(t))$, $t\geq 0$ is a Markov process on $\mathbb{R}^d\times\hat{E}$.

Our goal is to show the large deviation principle for this family of stochastic additive functionals with the rate function I stated as

$$-\inf_{\Gamma^\circ} I\leq \liminf_{\epsilon\to 0}\epsilon\log\mathbb{P}\{\xi^\epsilon\in\Gamma\}\leq\limsup\epsilon\log\mathbb{P}\{\xi^\epsilon\in\Gamma\}\leq -\inf_{\bar{\Gamma}} I \tag{4.1}$$

where Γ° and $\bar{\Gamma}$ represent the interior respectively the closure of the set Γ. In the particular case in which we take $\Gamma=\{\xi(t):||\xi(t)-\hat{\xi}(t)||>\delta\}$ one gets the asymptotic behavior of the $\mathbb{P}(\sup_{t\in[0,T]}||\xi^\epsilon(t)-\hat{\xi}(t)||>\delta)$.

An important consequence of the large deviation principle is due to Varadhan and it is called the Laplace principle [3] (Theorem 1.2.1).

Proposition 4.1. *If the sequence ξ^ϵ satisfies the large deviation principle on $\mathbf{D}([0,T],\mathbb{R}^d)$ with rate function $I_u(\varphi)$, then for all bounded continuous functions $h:\mathbf{D}([0,T],\mathbb{R}^d)\to\mathbb{R}$*

$$\lim_{\epsilon\to 0}\epsilon\log\mathbb{E}\{\exp[-\frac{1}{\epsilon}h(\xi^\epsilon)]\}=-inf_{\varphi\in\mathbf{D}([0,T],\mathbb{R}^d)}\{h(\varphi)+I_u(\varphi)\} \tag{4.2}$$

The Laplace principle implies the large deviation principle with the same rate function [3](Theorem 1.2.3).

Proposition 4.2. *If I_u is a rate function on $\mathbf{D}([0,T],\mathbb{R}^d)$ and the limit (4.2) is valid for all bounded continuous functions h, then the sequence ξ^ϵ satisfies the large deviation principle on $\mathbf{D}([0,T],\mathbb{R}^d)$ with rate function I.*

Lemma 4.3. *Suppose that for each fixed $k\in\hat{E}$, the family $\xi^\epsilon_t:=\xi^\epsilon_t(u;k)$, $t\geq 0$, $\epsilon>0$ satisfies the large deviation principle with the rate function $I_{u,k}(\cdot)$. If $\hat{x}_t$ is a stationary process on $\hat{E}$ then $\xi^\epsilon_t(u;\hat{x}(t))$ satisfies the large deviation principle with the rate function $I_u(\varphi)=\min\{I_{u,k}(\varphi):1\leq k\leq N\}$.*

Proof. Since for each fixed $k \in \hat{E}$, the family $\xi_t^\epsilon(u;k)$, $t \geq 0$, $\epsilon > 0$ satisfy the large deviation principle with the rate function $I_{u,k}$, we have

$$\epsilon \log \mathbb{P}(\xi^\epsilon \in \Gamma | \hat{x}^\epsilon = k) \sim -\inf_\Gamma I_{u,k}.$$

Let's denote $b_k^\epsilon := \mathbb{P}(\xi^\epsilon \in \Gamma | \hat{x}^\epsilon = k)$, $b_k := \inf_\Gamma I_{u,k}$ and $p_k = \mathbb{P}(\hat{x}^\epsilon = k)$. Thus $\epsilon \log b_k^\epsilon \sim -b_k$ and therefore $b_k^\epsilon = \exp(-\frac{1}{\epsilon} b_k + c_k^\epsilon)$ with $c_k^\epsilon = o(\frac{1}{\epsilon})$.
We want to prove that $\epsilon \log \mathbb{P}(\xi^\epsilon \in \Gamma) \sim -\min\{b_1, \cdots, b_N\}$. We may assume that $b_1 \leq b_2 \leq \cdots \leq b_N$ and $0 < p_i < 1$, $1 \leq i \leq N$ without loss of generality. Since $\mathbb{P}(\xi^\epsilon \in \Gamma) = \sum_{k=1}^N \mathbb{P}(\xi^\epsilon \in \Gamma | \hat{x}^\epsilon = k)\mathbb{P}(\hat{x}^\epsilon = k)$, it is enough to prove that $\epsilon \log(b_1^\epsilon p_1 + \cdots + b_N^\epsilon p_N) \sim -b_1$ which is equivalent to $\frac{1}{b_1^\epsilon p_1 + \cdots + b_N^\epsilon p_N} \sim \frac{1}{b_1^\epsilon p_1}$. This is true because $\frac{b_i^\epsilon}{b_1^\epsilon} = \exp(-\frac{1}{\epsilon}(b_i - b_1 + \epsilon(c_i^\epsilon - c_1^\epsilon))$ goes to 0 as ϵ goes to 0. □

Theorem 4.4. (*Main result*) *For absolutely continuous functions φ from $\mathbf{D}([0,T], \mathbb{R}^d)$, with $T > 0$ arbitrary fixed, satisfying $\varphi(0) = u$, and for each fixed $k \in \hat{E}$, define*

$$I_{u,k}(\varphi) := \int_0^T L(\varphi(t), \dot{\varphi}(t); k) dt, \tag{4.3}$$

where L is subsequently defined by (4.9). For all other functions in $\mathbf{D}([0,T], \mathbb{R}^d)$, $I_{u,k}(\varphi) := \infty$. Then the family $\xi^\epsilon(t), \epsilon > 0$ satisfies the Large deviation principle with rate function

$$I_u(\varphi) = \min\{I_{u,k}(\varphi) : 1 \leq k \leq N\} \tag{4.4}$$

Proof. This will be carried out in several steps. For the sake of clarity, it became necessary to state a number of known results, which we reformulated and adapted to our situation.

Step 1: Consider the martingale problem for the generator $\mathbb{L}^\epsilon$ and its relationship with the exponential martingale problem [11] by taking the transformation H^ϵ defined as

$$H^\epsilon f := \epsilon e^{-\frac{1}{\epsilon} f} \mathbb{L}^\epsilon e^{\frac{1}{\epsilon} f} \tag{4.5}$$

An important step is to prove the convergence of H^ϵ for an appropriate collection of sequences f^ϵ to an operator H in the sense that if f^ϵ converges to f as $\epsilon \to 0$ the $H^\epsilon f^\epsilon$ converges to Hf [4].

Let us consider the test functions $f^\epsilon(u,x) = f(u) + \epsilon \log \varphi^\epsilon(u,x)$ with $\varphi^\epsilon(u,x) = \varphi(u, m(x)) + \epsilon \varphi_1(u,x)$, where $f, \varphi^\epsilon(u,x)$ are bounded, measurable, continuous differentiable functions on $u \in \mathbb{R}^d$, with bounded first derivative, and uniformly continuous on E, convergent to the function $f(u)$ Then, $H^\epsilon f^\epsilon$ converges to Hf,

$$Hf(u;x) := a(u;x) f'(u) + \int_{\mathbf{R}^d} (e^{v f'(u)} - 1 - v f'(u)) \Gamma(u, dv; x). \tag{4.6}$$

Applying the stationary projector $\Pi : \mathbf{B}(E) \to \hat{\mathbf{E}}$, defined by $\Pi\varphi(x) := \int_E \rho(dx)\varphi(y)1\!\!1(x)$ (where $1\!\!1(x) = 1$ for all $x \in E$), we obtain

$$\hat{H}f(u;k) = \hat{a}(u;k)f'(u) + \int_{\mathbf{R}^d} (e^{vf'(u)} - 1 - vf'(u))\hat{\Gamma}(u, dv; k) \tag{4.7}$$

where

$$\hat{a}(u;k) = \int_{E_k} \pi_k(dx)a(u;x) \quad and \quad \hat{\Gamma}(u, dv; k) = \int_{E_k} \pi_k(dx)\Gamma(u, dv; k).$$

A key role is played by the function in u and p in $\mathbb{R}^d$ defined by

$$H(u, p; k) := \hat{a}(u;k)p + \int_{\mathbf{R}^d} (e^{vp} - 1 - vp)\hat{\Gamma}(u, dv; k) \tag{4.8}$$

having the following properties:
(a) for each $p \in \mathbb{R}^d$ and each $k \in \hat{E}$, $\sup_{u\in\mathbf{R}^d} H(u, p; k) < \infty$;
(b) for each $k \in \hat{E}$, $h(u, p; k)$ is a continuous function of $(u, p) \in \mathbb{R}^d \times \mathbb{R}^d$.
For u and q in $\mathbb{R}^d$ we define the Legendre-Fenchel transform

$$L(u, q; k) := \sup_{p\in\mathbf{R}^d} \{pq - H(u, p; k)\} \tag{4.9}$$

Step 2: As in Lemma 6.2.3. [3] the following properties of the Legendre-Fenchel function can be proved.

Lemma 4.5. *The functions $H(u, p; k)$ and $L(u, q; k)$ defined by (4.8) and (4.9) respectively, have the following properties*

(a) *For each $u \in \mathbb{R}^d$, $k \in \hat{E}$, $H(u, p; k)$ is a finite convex function of $p \in \mathbb{R}^d$ which is differentiable for all p. In addition, $H(u, p; k)$ is a continuous function of $(u, p) \in \mathbb{R}^d \times \mathbb{R}^d$*
(b) *For each $u \in \mathbb{R}^d, k \in \hat{E}$, $L(u, q; k)$ is a convex function of $q \in \mathbb{R}^d$. In addition, $L(u, q; k)$ is a nonnegative, lower semi-continuous function of $(u, q) \in \mathbb{R}^d \times \mathbb{R}^d$*
(c) *$L(u, q; k)$ is uniformly superlinear in the sense:*

$$\lim_{N\to\infty} \inf_{u\in\mathbf{R}^d} \inf_{q\in\mathbf{R}^d:||q||=N} \frac{1}{||q||} L(u, q; k) = \infty$$

(d) *For each $u \in \mathbb{R}^d$, $k \in \hat{E}$, the relative interior $ri(domL(u, \cdot; k)) = ri(convS_{\mu(\cdot|u,k)})$; in particular $L(u, q; k)$ equals ∞ for $u \in \mathbb{R}^d$ and $q \in (cl(convS_{\mu(\cdot|u,k)}))^c$. For any $q \in ri(domL(u, \cdot; k))$ there exists $v = v(u, q; k) \in \mathbb{R}^d$ such that $\nabla_v H(u, v(u, q; k); k) = q$. In addition,*

$$L(u, q; k) = v(u, q; k)q - H(u, v(u, q; k); k)$$

(e) *Suppose in addition that for a given $u \in \mathbb{R}^d$, $convS_{\mu(\cdot/u)}$ has nonempty interior. Then $H(u, v; k)$ is a strictly convex function of $v \in \mathbb{R}^d$, $int(domL(u, \cdot; k))$ is nonempty, for each $q \in int(domL(u, \cdot; k))$ there exists a unique value of v such that $\nabla_v H(u, v(u, q; k); k) = q$, and $L(u, \cdot; k)$ is differentiable on $int(domL(u, \cdot; k))$.*

(f) *For each u and q in $\mathbb{R}^d$, $k \in \hat{E}$,*

$$L(u,q;k) = \inf\{R(\nu(\cdot)||\mu(\cdot|u,k) : \nu \in \mathcal{P}(\mathbb{R}^d), \int_{\mathbb{R}_d} v\nu(dv) = q\}$$

and the infimum is always attained. If $L(u,q;k) < \infty$, then the infimum is attained uniquely. $R(\cdot||\cdot)$ is the relative entropy defined by $R(\nu||\theta) := \int(\log \frac{d\nu}{d\theta})d\nu$ whenever ν is absolutely continuous with respect to θ. Otherwise $R(\nu||\theta) := \infty$.

(g) *There is a stochastic kernel $\nu(dv|u,k)$ on $\mathbb{R}^d$ given $\mathbb{R}^d \times \hat{E}$ satisfying for u and q in $\mathbb{R}^d$,*

$$R(\nu(\cdot|u,k)||\mu(\cdot|u,k)) = L(u,q;k) \quad \textit{and} \quad \int_{\mathbb{R}^d} v\nu(dv|u,k) = q$$

(h) *If $\nu \in \mathcal{P}(\mathbb{R}^d)$ satisfies $R(\nu(\cdot)||\mu(\cdot|u,k)) < \infty$ for $u \in \mathbb{R}^d, k \in \hat{E}$ then $\int_{\mathbb{R}^d} ||v||\nu(dv) < \infty$ and*

$$R(\nu(\cdot|u,k)||\mu(\cdot|u,k)) \geq L(u, \int_{\mathbb{R}_d} v\nu(dv); k).$$

Step 3: To prove Laplace principle for the sequence ξ^ϵ it is sufficient to prove it for a sequence of random walks X^n constructed below.

Let h be any bounded continuous function mapping $\mathbf{D}([0,T],\mathbb{R}^d)$ into $\mathbb{R}$. We prove the Laplace limit (4.2) when $\epsilon \to 0$ along any sequence $\{\epsilon_n, n \in \mathbb{N}\}$ converging to 0. Let's fix such a sequence. By sampling the process ξ^{ϵ_n} at a sequence of times depending on ϵ_n, we define a sequence of piecewise linear processes $\{\zeta^n, n \in \mathbb{N}\}$ for which we prove Laplace principle. Then we show that the sequence is superexponentially closed to $\{\xi^{\epsilon_n}, n \in \mathbb{N}\}$.

Fix $T > 0$. For each $n \in \mathbb{N}$, let $c_n := [\frac{T}{\epsilon_n}]$ (where $[x]$ represents the integer part of x). Consider the sampled sequence $\xi^{\epsilon_n}(\frac{Tj}{c_n}), j = 0, 1, \cdots, c_n - 1$. Define $\zeta^n := \{\zeta^n(t), t \in [0,T]\}$ by

$$\zeta^n(t) = \xi^{\epsilon_n}(\frac{Tj}{c_n}) + c_n(t - \frac{Tj}{c_n}) \left(\xi^{\epsilon_n}(\frac{T(j+1)}{c_n} - \xi^{\epsilon_n}(\frac{Tj}{c_n})\right)$$

for $t \in \left[\frac{Tj}{c_n}, \frac{T(j+1)}{c_n}\right]$, which is the linear interpolation of the sampled sequence $\xi^{\epsilon_n}(\frac{Tj}{c_n}), j = 0, 1 \cdots, c_n - 1$.

For each fixed $k \in \hat{E}$, let $\{v_j^n(u;k), u \in \mathbb{R}^d, j \in \mathbb{N}_0\}$ be an i.i.d sequence of random vector fields having the common distribution

$$\mu^n(dv|u,k) := \mathbb{P}_u \left(\frac{c_n}{T}(\xi^{\epsilon_n}(\frac{T}{c_n}) - u) \in dv\right) \tag{4.10}$$

which is a stochastic kernel on $\mathbb{R}^d$ given $\mathbb{R}^d \times \hat{E}$.

We construct the random walks corresponding to the sequence of stochastic kernels $\mu^n(dv|u,k)$ as follows: for each $u \in \mathbb{R}^d, k \in \hat{E}, n \in \mathbb{N}$, consider the sequence of random variables $\{X_j^n, j = 0, 1, \cdots, c_n - 1\}$ taking values in $\mathbb{R}^d$ with

$$X_{j+1}^n := X_j^n + \frac{T}{c_n} v_j^n(X_j^n; k), \quad X_0^n = u.$$

Suppose that the the sequence of random vectors X_j^n is interpolated into a piecewise linear continuous-time process $X^n := \{X^n(t), t \in [0, T]\}$ by

$$X^n(t) = X_j^n + \left(t - \frac{Tj}{c_n}\right) v_j^n(X_j^n; k), \quad t \in \left[\frac{Tj}{c_n}, \frac{T(j+1)}{c_n}\right], j = 0, 1, \cdots, c_n - 1$$

Then the distribution of ζ^n is the same as the distribution of X^n. For each $n \in \mathbb{N}$ and $u, p \in \mathbb{R}^d, k \in \hat{E}$, define

$$H^n(u, p; k) := \log \int_{\mathbb{R}^d} e^{vp} \mu^n(dv|u, k) \tag{4.11}$$

Step 4: We will show that the function $H(u, p; k)$ defined in (4.8) can be written as the moment generating function of a stochastic kernel $\mu(dv|u, k)$.

Since the conditions of the Proposition 10.3.2 in [3] are fulfilled the next result follow.

Proposition 4.6. *For each $k \in \hat{E}$, the following conclusions hold:*

(a) *there exists a superlinear function $f : (0, \infty) \to \mathbb{R} \cup \{\infty\}$ such that for any $\epsilon > 0, \delta > 0, s \in [0, T], t \in (s, T]$*

$$\sup_{u \in \mathbb{R}^d} \mathbb{P}_{u,k}\{\sup_{s \le \sigma \le t} ||\xi^\epsilon(\sigma) - \xi^\epsilon(s)|| \ge \delta\} \le 2d \exp\left(-\frac{t-s}{\epsilon} f(\frac{\delta}{\sqrt{d}(t-s)})\right)$$

(b) *for each $p \in \mathbb{R}^d$, $\sup_{n \in \mathbb{N}} \sup_{u \in \mathbb{R}^d} H^n(u, p; k) < \infty$*

(c) *for each $p \in \mathbb{R}^d$ and each compact subset $K \subset \mathbb{R}^d$,*

$$\lim_{n \to \infty} \sup_{u \in K} |H^n(u, p; k) - H(u, p; k)| = 0 \tag{4.12}$$

(d) *for each $u \in \mathbb{R}^d$, the sequence of probability measures $\mu^n(dv|u, k), n \in \mathbb{N}$ converges weakly to a probability measure $\mu(dv|u, k)$ on $\mathbb{R}^d$ and for each $p \in \mathbb{R}^d$,*

$$H(u, p; k) = \log \int_{\mathbb{R}^d} e^{pv} \mu(dv|u, k).$$

The family $\mu(dv|u, k), u \in \mathbb{R}^d, k \in \hat{E}$ defines a stochastic kernel on $\mathbb{R}^d$ given $\mathbb{R}^d \times \hat{E}$. In addition, the function mapping $u \in \mathbb{R}^d \mapsto \mu(\cdot|u, k) \in \mathcal{P}(\mathbb{R}^d)$ is continuous in the topology of weak convergence on $\mathcal{P}(\mathbb{R}^d)$.

Step 5: In order to study the Laplace principle for the process X^n, we need to verify the asymptotic behavior of

$$W^n(u) := -\frac{1}{c_n} \log \mathbb{E}_u(\exp(-c_n h(X^n))), \tag{4.13}$$

where $\mathbb{E}_u$ denotes the expectation with respect to $\mathbb{P}_u$ and h is any bounded continuous function mapping $\mathbf{C}([0,T],\mathbb{R}^d)$ into $\mathbb{R}$. We will show that this is equal to the minimal cost of function of an associated stochastic control problem.

We now specify the stochastic control problem whose minimal cost function gives a representation for the function $W^n(u)$. The controlled process is a discrete-time process $\bar{X}^n_j$, $j = 0, 1, \cdots, c_n - 1$, and at each time t there will be a control ν^n_j giving the distributions of the controlled random variable that replaces this noise due to the increments. ν^n_j is a stochastic kernels on $(\mathbb{R}^d)^{j+1}$, denoted by $\nu^n_j(dv) = \nu^n_j(dv|\bar{X}^n_0, \cdots, \bar{X}^n_j)$. A sequence of controls $\{\nu^n_{1,j},\ j = 0, 1, \cdots, c_n - 1\}$ is called an admissible control sequence.

Then, as in [3] (Theorem 4.3.1) we get the variational representation of W^n_u as

$$W^n(u) = \inf_{\nu^n_j} \bar{\mathbb{E}}_u\{\sum_{j=0}^{c_n-1}\left[\frac{1}{c_n} R(\nu^n_j(\cdot)||\mu(\cdot|\bar{X}^n_j, k)\right] + h(\bar{X}^n\} \tag{4.14}$$

where the infimum is taken over all admissible control sequences $\{\nu^n_j\}$. For $n \in \mathbb{N}$ and $t \in [0,T]$, define the stochastic kernel

$$\nu^n(dv|t) := \begin{cases} \nu^n_j(dv), & t \in [\frac{Tj}{c_n}, \frac{T(j+1)}{c_n}), \quad j = 0, 1, \cdots, c_n - 2 \\ \nu^n_{c_n-1}(dv), & t \in [\frac{T(c_n-1)}{c_n}, T] \end{cases}$$

The following representation holds (similar as in [3] (Corollary 5.2.1))

$$W^n(u) = \inf_{\nu^n_j} \bar{\mathbb{E}}_u\{\int_0^T R(\nu^n_1(\cdot|t)||\mu(\cdot|\tilde{X}^n(t)) + h(\bar{X}^n)\} \tag{4.15}$$

where $\tilde{X}^n = \{\tilde{X}^n(t), t \in [0,T]\}$ is the piecewise constant interpolation of the controlled random variables $\{\bar{X}^n_j, j = 0, 1, \cdots, c_n - 1\}$.

Step 6: Laplace principle upper bound

Let $I_{u,k}(\varphi) := \int_0^T L(\varphi(t), \dot{\varphi}(t), k)dt$ where L is the Legendre-Fenchel transform defined in (4.9). Then $I_{u,k}$ is a rate function and

$$\limsup_{n\to\infty} \frac{1}{c_n} \log \mathbb{E}_u(\exp(-c_n h(X^n))) \le - \inf_{\varphi \in \mathbf{C}([0,T],\mathbf{R}^d)} (I_{u,k}(\varphi) + h(\varphi)) \tag{4.16}$$

Indeed, first it can be shown that $I_{u,k}$ has compact level sets in $\mathbf{C}([0,T],\mathbb{R}^d)$ by using parts (b) and (c) of the Proposition 4.6, which implies that $I_{u,k}$ is a rate function. Then using part (h) of Proposition 4.5 we will get

$$\liminf_{n\to\infty} W^n(u) \ge \inf_{\varphi \in \mathbf{C}([0,T],\mathbf{R}^d)} (I_{u,k}(\varphi) + h(\varphi)).$$

Step 7: Laplace principle lower bound.

In order to prove the Laplace principle lower bound we need to characterize the relative interior of the effective domain of $L(u,\cdot;k)$ in terms of the stochastic kernel $\mu(dv|u,k)$. This is done in part (d) of Proposition 4.5.

For A, B subsets of $\mathbb{R}^d$ define

$$A + B := \{u \in \mathbb{R}^d : u = a + b, a \in A, b \in B\}.$$

A subset C of $\mathbb{R}^d$ is called a convex cone if it has the property that for $c \in C$, $\lambda c \in C$ $\forall \lambda \in [0,\infty)$. Denote $conC$ for the convex cone of C.

We can rewrite $H(u,p;k)$ as

$$H(u,p;k) = \hat{b}(u;k)p + \int_{\mathbf{R}^d} (e^{vp} - 1)\hat{\Gamma}(u,dv;k)$$

where

$$\hat{b}(u;k) := \hat{a}(u;k) - \int_{\mathbf{R}^d} v\hat{\Gamma}(u,dv;k)$$

Let $S_{\hat{\Gamma}_{(u,k)}}$ be the support of $\hat{\Gamma}_{(u,k)}$ and define $T_{(u,k)} := \{\hat{b}(u;k)\} + conS_{\hat{\Gamma}_{(u,k)}}$.

The relative interior $ri(domL(u,\cdot;k)) = ri(T_{(u,k)})$ and the following properties hold:

(a) The sets $intT_{(u,k)}$ are independent of $(u,k) \in \mathbb{R}^d \times \hat{E}$

(b) $0 \in intT_{(u,k)}$

With similar arguments as in Theorem 6.5.1 [3] it can be proved that

$$\limsup_{n\to\infty} W^n(u,k) \leq \inf_{\varphi\in\mathbf{C}([0,T],\mathbf{R}^d)} (I_{uk}(\varphi) + h(\varphi)).$$

This gives the Laplace principle lower bound for X^n.

$$\liminf_{n\to\infty} \frac{1}{c_n} \log \mathbb{E}_u(\exp(-c_n h(X^n))) \geq - \inf_{\varphi\in\mathbf{C}([0,T],\mathbf{R}^d)} (I_{uk}(\varphi) + h(\varphi)). \tag{4.17}$$

Thus the Laplace principle is proved for the random walk X^n and therefore for the process ζ^n.

Step 8: Laplace principle holds for the sequence ξ^{ϵ_n} because $\xi^{\epsilon_n}, \zeta^n$ are superexponentially closed, i.e.

$$\limsup_{n\to\infty} \sup_{u\in\mathbf{R}^d} \epsilon_n \log \mathbb{P}_u(\rho(\xi^{\epsilon_n}, \zeta^n) > \delta) = -\infty, \tag{4.18}$$

where ρ is Skorokhod metric on $\mathbf{D}([0,T],\mathbb{R}^d)$.

Thus, by Proposition 4.2 we obtain the large deviation principle for the sequence of random variables $\xi^\epsilon_t(u;k)$ with the rate function $I_{u,k}(\varphi) = \int_0^T L(\varphi(t),\dot{\varphi}(t);k)dt$. Using Lemma 4.3 we get the large deviation principle for the sequence of stochastic additive functionals ξ^ϵ with rate function $I_u(\varphi) = \min\{I_{u,k}(\varphi) : 1 \leq k \leq N\}$.

This completes the proof of the theorem.

□

This principle has many applications, for example finding the probability of exit from a stable domain of the process. In some cases the infimum can be explicitly found by using calculus of variations. The class of absolutely continuous functions on $[0,T]$ can be identified with the Sobolev space $H^{1,1}[0,T]$, and since the Legendre-Fenchel function $L(u,q;k)$ verifies the conditions of Tonelli's existence theorem (Theorem 3.7 [1]), the existence of the minimizer will follow. If $\varphi \in AC[0,T]$ is a local minimizer of the functional $L(\varphi,\varphi')$, then φ will satisfy the Euler-Lagrange equation which will be further simplified to the Beltrami equation: $L(\varphi,\varphi') - \varphi' L_{\varphi'}(\varphi,\varphi') = C$, where C is a constant.

Example 4.7. Compound Poisson process

Consider the compound Poisson process $\xi^\epsilon(t)$, $t \geq 0$ switched by the jump Markov process $x(t)$, $t \geq 0$ defined in Example 3.2, of the form

$$\xi^\epsilon(t) = \sum_{k=1}^{\nu(t/\epsilon;x(t/\epsilon)} a_k(x(\frac{t}{\epsilon}))$$

with the infinitesimal generator given by

$$\mathbb{\Gamma}^\epsilon(x)\phi(u) = \frac{\Lambda(x)}{\epsilon}\int_{\mathbb{R}^d}[\phi(u+\epsilon v) - \phi(u)]F(dv;x).$$

Here $\nu(t;x)$, $t \geq 0$, $x \in E = \{1,2,3,4\}$ is a homogeneous Poisson process, with intensity $\Lambda(x)$ and $a_k(x), k \geq 1, x \in E$ is a sequence of i.i.d. random variables, independent of $\nu(t)$, $t \geq 0$, with common distribution $F(dv;x)$.

Using notation $\hat{a}(k) = \int_{E_k}\pi_k(dx)a(x)$, this process converges weakly,

$$\xi^\epsilon(t) \Rightarrow \int_0^t \hat{a}(\hat{x}(s))ds, \quad \text{as } \epsilon \to 0.$$

Applying the operator $H^\epsilon f^\epsilon$ as in equation (4.5) we get the limitting operator Hf as follows

$$Hf(u,x) = \Lambda(x)\int_{\mathbb{R}^d}[e^{vf'(u)} - 1]F(dv,x).$$

For tractability purposes, let's suppose that $F(dv;x)$ is independent of x. Then the projected operator $\hat{H}f$ is

$$\hat{H}f(u,k) = \hat{\Lambda}(k)\int_{\mathbb{R}^d}[e^{vf'(u)} - 1]F(dv)$$

where $\hat{\Lambda}(k) = \int_{E_k}\pi_k(dx)\Lambda(x)$. Hence, $\hat{\Lambda}(1) = \frac{2\lambda_1\mu_1}{\lambda_1+\mu_1}$ and $\hat{\Lambda}(2) = \frac{2\lambda_2\mu_2}{\lambda_2+\mu_2}$. Assume that the random variables $a_k(x)$ are distributed exponential with the parameter λ. Then the function $H(p;k), p \in \mathbb{R}, k \in \hat{E} = \{1,2\}$ defined in the relation 4.8 is

$$H(p;k) = \hat{\Lambda}(k)\frac{p}{\lambda - p}, \quad \lambda > p$$

The Legendre-Fenchel transform $L(q;k) = \sup_{p\in\mathbb{R}}\{pq - H(p;k)\}$ becomes

$$L(q;k) = \lambda q - 2\sqrt{\lambda q \hat{\Lambda}(k)} + \hat{\Lambda}(k),$$

the supremum being attained for $p = \lambda - \sqrt{\frac{\lambda\hat{\Lambda}(k)}{q}}$. Therefore, for $T > 0$ arbitrary fixed, and for absolutely continuous functions $\varphi \in \mathbf{D}([0,T],\mathbb{R})$, with $\varphi(0) = 0$, the process ξ^ϵ satisfies the large deviation principle. Its rate function is $I(\varphi) = \min_{k=1,2} I_k(\varphi)$, where $I_k(\varphi) = \int_0^T L(\varphi'(t);k)dt$ and $L(\varphi'(t)) = \lambda\varphi'(t) - 2\sqrt{\lambda\varphi'(t)\hat{\Lambda}(k)} + \hat{\Lambda}(k)$.

REFERENCES

[1] M. Buttazzo, G. Giaquinta and S. Hildebrandt. *One-dimensional Variational Problems*. Clarendon Press, Oxford, 1998.

[2] M.H.A. Davis. *Markov Models and Optimization*. Chapman & Hall, London/Glasgow/New York/Tokyo/Melbourne/Madras, 1993.

[3] P. Dupuis and R.S. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations*. John Wiley & Sons, New York/Chichester/Brisbane/Toronto/Singapore, 1997.

[4] J. Feng and T.J. Kurtz. *Large Deviations for Stochastic Processes*. American Mathematical Society, providence, Rhode Island USA, 2006.

[5] I.I Gikhman and A.V. Skorokhod. *The Theory of Stochastic Processes*. Springer-Verlag, Berlin/heidelberg/New York/London/Paris/Tokyo, 1987.

[6] V.S. Koroliuk and N. Limnios. *Stochastic Systems in Merging Phase Space*. World Scientific, New Jersey/London/Singapore/Beijing/Shanghai/Hong Kong/Taipei/Chennai, 2005.

[7] V.S. Korolyuk and V.V. Korolyuk. *Stochastic Models of Systems*. Kluwer Academic, Dordrecht, 1999.

[8] V.S. Korolyuk and N. Limnios. Average and diffusion approximation of stochastic evolutionary systems in an asymptotic split state space. *Annals Appl.Probab.*, 14:489–516, 2004.

[9] F.C. Skorokhod, A. V. Hoppensteadt and H. Salehi. *Random perturbation methods with applications in science and engineering*. Springer, New York, 2002.

[10] D. W. Stroock. *An Introduction to the Theory of Large Deviations*. Springer-Verlag, New York, 1984.

[11] D. W. Stroock and S.R.S. Varadhan. *Multidimensional Diffusion Processes*. Springer-Verlag, Berlin/Heidelberg/New York, 1979.

[12] G. G. Yin and Q. Zhang. *Continuous-Time Markov Chains and Applications*. Springer, New York/Chichester/Brisbane/Toronto/Singapore, 1998.