Probab. Theory Relat. Fields 118, 483–521 (2000)
Digital Object Identifier (DOI) 10.1007/s004400000095

A. de Acosta

A general non-convex large deviation result with applications to stochastic equations*

Received: 23 June 1999 / Revised version: 17 February 2000 /
Published online: 22 November 2000 –

Abstract. We prove an abstract large deviation result for a sequence of random elements of a vector space satisfying an "abstract exponential martingale condition". The framework naturally generates non-convex rate functions. We apply the result to solutions of Itô stochastic equations in $\mathbf{R}^d$ driven by Brownian motion and a Poisson random measure.

1. Introduction

In many large deviation problems, the object under study is a sequence $\{Y_n\}$ of random elememts of a topological vector space E and convexity considerations play an important role. Let

$$\phi_n(\xi) = \log \mathbf{E} \exp\langle Y_n, \xi\rangle, \quad \xi \in E^*,$$

where E^* is the dual space of E. It is well known that if

$$\phi(\xi) = \lim_n n^{-1}\phi_n(n\xi) \tag{1.1}$$

exists for all $\xi \in E^*$ and satisfies a suitable differentiability condition, and

$$\{\mathscr{L}(Y_n)\} \text{ is exponentially tight,} \tag{1.2}$$

then $\{\mathscr{L}(Y_n)\}$ satisfies the large deviation principle with rate function

$$\phi^*(y) = \sup_{\xi\in E^*} [\langle y, \xi\rangle - \phi(\xi)], \quad y \in E. \tag{1.3}$$

This elegant result is essentially due to Gärtner [Ga] when $E = \mathbf{R}^d$; see also Ellis [E] (in this case, (1.2) is superfluous). For arbitrary E results of this type have been obtained by Baldi [Ba], Bryc [Br], Dawson and Gärtner [Da-G], Dembo and Zeitouni [De-Z], O'Brien-Sun [OB-S] and de Acosta [deA1], [deA2]. The infinite

A. de Acosta: Department of Mathematics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106-7058, USA. e-mail: add3@po.cwru.edu

*This research was partially supported by an NSF grant.

Mathematics Subject Classification (2000): 60F10, 60H10

Key words and phrases: Abstract large deviations – Non-convex rate functions – Itô stochastic equations

dimensional result may be applied, for example, to prove a large deviation principle for Lévy processes [deA1].

Of course, the function ϕ^* defined in (1.3) is convex. It obviously follows that the result cannot be applied – at least, not directly – to situations in which the rate function is not convex. Such situations exist in abundance.

In the present paper we introduce an abstract scheme which generalizes the previously described result to a non-convex framework. The main idea is to postulate the existence of suitable functions $\Phi_n : E \times E^* \to \mathbf{R}$ such that for all $\xi \in E^*$

$$\mathbf{E}\exp[\langle Y_n, \xi\rangle - \Phi_n(Y_n, \xi)] = 1 \tag{1.4}$$

and for all $x \in E, \xi \in E^*$

$$\Phi(x, \xi) = \lim_n n^{-1}\Phi_n(x, n\xi)$$

exists. For reasons that will become clear in Section 4, (1.4) might be thought of as an "abstract exponential martingale condition". In Theorems 2.1 and 2.2 we prove that under suitable regularity conditions on Φ and a suitable form of (1.2), $\{\mathscr{L}(Y_n)\}$ satisfies the large deviation principle with rate function $\Phi^*(y, y)$, where for $y, z \in E$,

$$\Phi^*(y, z) = \sup_{\xi\in E^*}[\langle z, \xi\rangle - \Phi(y, \xi)]. \tag{1.5}$$

This rate function is generated by convex conjugation, but because of the dependence of Φ on its first variable it is in general not convex. Thus (1.5) provides a natural way of generating non-convex rate functions in a vector space context.

As an illustration of our scheme we consider the large deviation principle for diffusions. For simplicity, let $X_n^x = \{X_n^x(t), t \in [0, 1]\}$ be the solution of the stochastic equation in $\mathbf{R}$

$$X_n^x(t) = x + \int_0^t b(X_n^x(s))ds + n^{-1/2}\int_0^t \sigma(X_n^x(s))dB(s),$$

where B is standard Brownian motion, b, σ are bounded and uniformly Lipschitz with $|\sigma| \geq c > 0$ and $x \in \mathbf{R}$. Then according to the classical Freidlin-Wentzell theorem [W-F], $\{\mathscr{L}(X_n^x)\}$ satisfies the large deviation principle in $C[0, 1]$ with rate function

$$I^x(f) = \begin{cases} \frac{1}{2}\int_0^1 (f'(s) - b(f(s)))^2\sigma^{-2}(f(s))ds & \text{if } f(0) = x \text{ and} \\ & \quad f \text{ is absolutely continous} \\ \infty & \text{otherwise.} \end{cases}$$

Our abstract large deviation principle (Theorems 2.1 and 2.2) applies to this situation. Let $M[0, 1]$ be the space of finite signed measures on [0,1]. In this case condition (1.4) is

$$\mathbf{E}\exp\left[\int X_n^x\, d\lambda - \Phi_n^x(X_n^x, \lambda)\right] = 1$$

where

$$\Phi_n^x(f,\lambda) = \lambda([0,1])x + \int_0^1 G_n(f(s),\lambda([s,1]))ds$$
$$\Phi^x(f,\lambda) = \lambda([0,1])x + \int_0^1 G(f(s),\lambda([s,1]))ds, \quad f \in C[0,1], \lambda \in M[0,1]$$

and

$$G_n(y,\alpha) = b(y)\alpha + (2n)^{-1}\sigma^2(y)\alpha^2,$$
$$G(y,\alpha) = b(y)\alpha + \frac{1}{2}\sigma^2(y)\alpha^2 \quad y,\alpha \in \mathbf{R}.$$

The rate function I^x is obtained in our scheme by the variation formula

$$I^x(f) = (\Phi^x)^*(f,f) = \sup_{\lambda \in M[0,1]} \left[\int f\,d\lambda - \Phi^x(f,\lambda) \right], \quad f \in C[0,1].$$

More generally, we present in Theorem 3.1 an application of the abstract large deviation principle to a sequence of Markov processes in $\mathbf{R}^d$ which are defined as solutions of Itô stochastic equations driven by Brownian motion and a Poisson random measure (for a precise statement of the considerable breadth of the class of Markov processes defined in this way, see Cinlar-Jacod [C-J]). In the Markov context, some closely related results were obtained originally by Wentzell [W] and more recently, in an improved form, by Dupuis-Ellis [Du-E]; in the semimartingale context, a closely related result was obtained by Liptser-Pukhalskii [L-P]. Because of the different frameworks and assumptions, it is not immediately clear how Theorem 3.1 compares with these results. However, we wish to emphasize two aspects of our work which are different from those papers and, it appears to us, deserve mention: (i) the assumptions in Theorem 3.1 are explicit boundedness, Lipschitz and integrability conditions on the data of the stochastic equations and not on latent objects, such as rate functions; (ii) no non-degeneracy assumptions are made in Theorem 3.1. We also indicate in Theorem 9.1 how the methods of the present paper apply to a somewhat more general class of Markov processes, which includes those considered in [W] and [Du-E]; the lower bound in this result requires, however, a non-degeneracy assumption.

A very recent contribution to the study of large deviations for Markov processes is Feng and Kurtz [F-K], where a nonlinear semigroup and exponential martingale problem approach is developed.

The paper is organized as follows. In Section 2 we prove the abstract large deviation result. The vector space and measurability assumptions are formulated so as to apply to $D([0,1],\mathbf{R}^d)$ endowed with the uniform norm; also, for greater flexibility, the exponential tightness assumption does not involve $\{Y_n\}$ but an auxiliary sequence $\{Z_n\}$ which is superexponentially close in probability to $\{Y_n\}$.

Sections 3–9 are devoted to the application to Itô stochastic equations. In Section 3 we describe the equations and state the large deviation theorem for the solution processes, $\{X_n^x\}$. In Section 4 we establish the "abstract exponential martingale condition". In Section 5 we introduce a suitable time discretization $\{Z_n^x\}$ of $\{X_n^x\}$ and show that it is exponentially tight in $(D(T,\mathbf{R}^d), \|\cdot\|_\infty)$ and superexponentially close to $\{X_n^x\}$. In Section 6 we identify the rate function–given in the abstract framework by the variational formula (1.5) – as a classical integral expression and prove the compactness of the level sets. In Section 7 we show that under a non-degeneracy (uniform ellipticity) assumption on the diffusion coefficient of the stochastic equations, the subdifferentiability assumption in Theorem 2.2 holds. In Section 8 we show that the solutions of the stochastic equations perturbed by an independent Brownian motion with small variance–which renders the diffusion coefficient non-degenerate–are superexponentially close in probability to the solutions of the corresponding original equations. Finally, in Section 9 the items in Sections 4–8 are assembled and Theorem 3.1 is proved.

2. An abstract non-convex large deviation result

We will consider the following objects:

- E is a Hausdorff topological vector space and $\mathscr{V}$ is a fundamental system of open symmetric neighborhoods of 0.
- $\mathscr{E}$ is a σ-algebra of subsets of E such that
  - (i) $\mathscr{E}$ contains the class of compact sets.
  - (ii) $\mathscr{E}$ contains $\mathscr{V}$.
  - (iii) $(E, \mathscr{E})$ is a measurable vector space; that is, the map $(x, y) \to x + y$ is $(E \times E, \mathscr{E} \otimes \mathscr{E})/(E, \mathscr{E})$-measurable and the map $(\lambda, x) \to \lambda x$ is $(\mathbf{R} \times E, \mathscr{B}(\mathbf{R}) \times \mathscr{E})/(E, \mathscr{E})$-measurable; here $\mathscr{B}(\mathbf{R})$ is the Borel σ-algebra of $\mathbf{R}$.
- F is a subspace of the dual space E^* such that $\langle \cdot, \xi \rangle$ is $\mathscr{E}$-measurable for all $\xi \in F$.
- For a function $\Phi : E \times F \to \mathbf{R}$, we define for $x, y \in E$

$$\Phi^*(x, y) = \sup_{\xi \in F} [\langle y, \xi \rangle - \Phi(x, \xi)],$$

  note that if $\Phi(x, 0) = 0$ for all $x \in E$, then $\Phi^* \geq 0$.
- $\{a_n\}_{n \in \mathrm{N}}$ is a positive sequence with $\lim_n a_n = \infty$.

In general, $\mathscr{E}$ may be smaller than the Borel σ-algebra of E. The framework will be applied in Sections 3–9 to the following setting: $E = D([0, 1], \mathbf{R}^d)$ endowed with the uniform norm $\| \cdot \|_\infty$, $\mathscr{V}$ is the class of open balls with center at 0, $\mathscr{E}$ is the σ-algebra generated by the evaluation mappings, and $F = M([0, 1], \mathbf{R}^d)$, the space of $\mathbf{R}^d$-valued vector measures on $([0, 1], \mathscr{B}([0, 1])$. We have previously used this setting in de Acosta [deA1].

In Theorems 2.1 and 2.2 we prove large deviation upper and lower bounds, respectively, for $\{\mathscr{L}_{\mathbf{P}_n}(Y_n)\}$, where for each $n \in \mathbf{N}$, Y_n is an E-valued random vector satisfying the key assumption (4) of Theorem 2.1. For greater flexibility in the applications, it is not convenient to assume that $\{\mathscr{L}_{\mathbf{P}_n}(Y_n)\}$ is exponentially tight, which may fail to be true; we suppose instead that there exists a sequence $\{Z_n\}$ of E-valued random vectors such that $\{Y_n\}$ and $\{Z_n\}$ are superexponentially close in probability and $\{\mathscr{L}_{\mathbf{P}_n}(Z_n)\}$ is exponentially tight.

Theorem 2.1. *(upper bounds) Let $\Phi_n, \Phi : E \times F \to \mathbf{R}$ be such that*

(1) *for all $\xi \in F$, $\Phi_n(\cdot, \xi)$ is $\mathscr{E}$-measurable.*
(2) *for all $\xi \in F$, $\Phi(\cdot, \xi)$ is $\mathscr{E}$-measurable, continuous and satisfies $\Phi(x, 0) = 0$ for all $x \in E$.*
(3) *for all $\xi \in F$,*

$$b_n(\xi) \overset{\Delta}{=} \sup_{x \in E} |a_n^{-1} \Phi_n(x, a_n \xi) - \Phi(x, \xi)| \to 0 \quad \text{as } n \to \infty.$$

For each $n \in \mathbf{N}$, let Y_n, Z_n be E-valued, $\mathscr{E}$-random vectors defined on $(\Omega_n, \mathscr{A}_n, \mathbf{P}_n)$, and assume

(4) *for all $n \in \mathbf{N}, \xi \in F$,*

$$\mathbf{E}_n \exp[\langle Y_n, \xi \rangle - \Phi_n(Y_n, \xi)] = 1.$$

(5) *$\{\mathscr{L}_{\mathbf{P}_n}(Z_n)\}$ is exponentially tight.*
(6) *for every $V \in \mathscr{V}$.*

$$\lim_n a_n^{-1} \log \mathbf{P}_n\{Y_n - Z_n \in V^c\} = -\infty.$$

Assume furthermore that for all $a \geq 0$,

(7) *the level set $L_a = \{x \in E : \Phi^*(x, x) \leq a\}$ is compact.*
Then for all $A \in \mathscr{E}$,

$$\lim_n \sup a_n^{-1} \log \mathbf{P}_n\{Y_n \in A\} \leq - \inf_{x \in \bar{A}} \Phi^*(x, x).$$

Proof. Let K be a compact subset of E. We claim:

$$\lim_n \sup a_n^{-1} \log \mathbf{P}_n\{Z_n \in K\} \leq - \inf_{x \in K} \Phi^*(x, x). \tag{2.1}$$

For, assume that $\inf_{x \in K} \Phi^*(x, x) < \infty$ and let $\varepsilon > 0$. For $\xi \in F$, let

$$V(\xi) = \{x \in E : \langle x, \xi \rangle - \Phi(x, \xi) > \inf_{y \in K} \Phi^*(y, y) - \varepsilon\}.$$

Then $V(\xi)$ is open, $V(\xi) \in \mathscr{E}$ and

$$K \subset \{x \in E : \Phi^*(x, x) > \inf_{y \in K} \Phi^*(y, y) - \varepsilon\} = \bigcup_{\xi} V(\xi).$$

By compactness, there exists $\xi_1, \ldots, \xi_k \in F$ such that

$$K \subset \bigcup_{i=1}^{k} V(\xi_i).$$

Let $V \in \mathscr{V}$ be such that

$$K + V \subset \bigcup_{i=1}^{k} V(\xi_i).$$

Then

$$\{Z_n \in K\} \subset \{Y_n \in K + V\} \cup \{Y_n - Z_n \in V^c\}$$
$$\subset \{Y_n \in \bigcup_{i=1}^{k} V(\xi_i)\} \cup \{Y_n - Z_n \in V^c\}$$

and therefore

$$\mathbf{P}_n\{Z_n \in K\} \leq 2\max\{\mathbf{P}_n\{Y_n \in \bigcup_{i=1}^{k} V(\xi_i)\}, \mathbf{P}_n\{Y_n - Z_n \in V^c\}\}. \tag{2.2}$$

Now, letting $b = \inf_{x\in K} \Phi^*(x, x) - \varepsilon$, we have

$$\begin{aligned}\mathbf{P}_n\{Y_n \in \bigcup_{i=1}^{k} V(\xi_i)\} &\leq \sum_{i=1}^{k} \mathbf{P}_n\{\langle Y_n, \xi_i\rangle - \Phi(Y_n, \xi_i) > b\}\\ &\leq e^{-a_n b} \sum_{i=1}^{k} \mathbf{E}_n \exp[\langle Y_n, a_n\xi_i\rangle - a_n\Phi(Y_n, \xi_i)]\\ &\leq e^{-a_n b} \cdot k \cdot \max_i\{\mathbf{E}_n \exp[\langle Y_n, a_n\xi_i\rangle\\ &\quad -\Phi_n(Y_n, a_n\xi_i)] \cdot \exp(a_n b_n(\xi_i))\}.\end{aligned} \tag{2.3}$$

Letting $n \to \infty$, by assumptions (3), (4), and (6), and (2.2), (2.3), we obtain

$$\limsup_n a_n^{-1} \log \mathbf{P}_n\{Z_n \in K\} \leq -b = -\inf_{x\in K} \Phi^*(x, x) + \varepsilon.$$

Since ε is arbitrary, claim (2.1) follows when $\inf_{x\in K} \Phi^*(x, x) < \infty$. If $\inf_{x\in K} \Phi^*(x, x) = \infty$, the argument is similar; we omit it.

Next, let $A \in \mathscr{E}$ and let $a > 0$. By assumption (5), there exists a compact set $K_a \subset E$ such that

$$\limsup_n a_n^{-1} \log \mathbf{P}_n\{Z_n \in K_a^c\} \leq -a. \tag{2.4}$$

We have, for $V \in \mathscr{V}$,

$$\{Y_n \in A\} \subset \{Y_n \in A, Z_n \in K_a, Y_n - Z_n \in V\} \cup \{Z_n \in K_a^c\} \cup \{Y_n - Z_n \in V^c\}$$
$$\subset \{Z_n \in \overline{A + V} \cap K_a\} \cup \{Z_n \in K_a^c\} \cup \{Y_n - Z_n \in V^c\}.$$

and

$$\mathbf{P}_n\{Y_n \in A\} \leq 3\max\{\mathbf{P}_n\{Z_n \in \overline{A + V} \cap K_a\}, \mathbf{P}_n\{Z_n \in K_a^c\}, \mathbf{P}_n\{Y_n - Z_n \in V^c\}\}. \tag{2.5}$$

Letting $n \to \infty$, by assumption (6), (2.1), (2.4) and (2.5), we have

$$\begin{aligned}\limsup a_n^{-1} \log \mathbf{P}_n\{Y_n \in A\} &\leq \max\{-\inf\{\Phi^*(x, x) : x \in \overline{A + V} \cap K_a\}, -a\}\\ &\leq \max\{-\inf\{\Phi^*(x, x) : x \in \overline{A + V}\}, -a\}.\end{aligned}$$

Since a is arbitrary, we have obtained: for all $V \in \mathscr{V}$,

$$\limsup a_n^{-1} \log \mathbf{P}_n\{Y_n \in A\} \leq -\inf\{\Phi^*(x,x) : x \in \overline{A+V}\}.$$

Now $\{\overline{A+V} : V \in \mathscr{V}\}$ is a directed decreasing family of closed sets with intersection $\bar{A}$. Using assumption (7), by a well-known property of good rate functions (see e.g. [De-Z], p.119) we have

$$\sup_{V\in\mathscr{V}} \inf\{\Phi^*(x,x) : x \in \overline{A+V}\} = \inf\{\Phi^*(x,x) : x \in \bar{A}\} \qquad \square$$

We will need the following definition in Theorem 2.2. A fucntion $\phi : F \to \mathbf{R}$ is *E-Gâteaux differentiable at* $\xi \in F$ if there exists a point $\nabla\phi(\xi) \in E$ such that for all $\eta \in F$,

$$\langle \nabla\phi(\xi), \eta\rangle = \lim_{t\to 0} t^{-1}[\phi(\xi + t\eta) - \phi(\xi)].$$

We shall use the notation

$$\partial\Phi^*(x,y) \stackrel{\Delta}{=} \partial(\Phi^*(x,\cdot))(y) \subset F$$

for the subdifferential of the convex function $\Phi^*(x,\cdot)$ at $y \in E$ (for the definition of subdifferential, see e.g. [A-E], [E-T]). We emphasize that condition (10) below is a uniqueness assumption; no assertion is made about existence. This condition can often be verified easily using Gronwall's lemma. Condition (11), on the other hand, is in general more difficult to verify; in the application to stochastic equations, the verification requires proving Proposition 7.1. (It is not difficult to show, however, that in the case of the classical Friedlin-Wentzell Theorem – as in the Introduction– the verification is very simple). If Φ does not depend on x and E is a Banach space, then condition (11) follows from the Brondsted-Rockafellar theorem (see e.g. [A-E]). It would be interesting to find useful assumptions on Φ in the general abstract framework of Theorem 2.2 under which condition (11) is automatically true.

Theorem 2.2. *(lower bounds) Let E, $\mathscr{E}$, F be as described before Theorem 2.1, and assume furthermore that F separates points in E.*

Let Φ_n, Φ, $\{Y_n\}$, $\{Z_n\}$ satisfy conditions (1)–(7) of Theorem 2.1, and assume furthermore

(8) For all $x \in E$, $\Phi(x,\cdot)$ is convex and E-Gâteaux differentiable.
(9) For all $\xi \in F$, $\bar{\Phi}(\xi) \stackrel{\Delta}{=} \sup_{x\in E} |\Phi(x,\xi)| < \infty$.
(10) For all $\xi \in F$, the equation $x = \nabla\Phi(x,\xi)$ has at most one solution in E.
(11) For all $x_0 \in E$ such that $\Phi^*(x_0,x_0) < \infty$, for every neighborhood W of x_0 and for every $\varepsilon > 0$, there exists $x_1 \in W$ such that $\partial\Phi^*(x_1,x_1) \neq \phi$ and

$$\Phi^*(x_1,x_1) < \Phi^*(x_0,x_0) + \varepsilon.$$

Then for every $A \in \mathscr{E}$,

$$\liminf_n a_n^{-1} \log \mathbf{P}_n\{Y_n \in A\} \geq -\inf_{x\in A^\circ} \Phi^*(x,x).$$

Proof. Let $A \in \mathscr{E}, x_0 \in A^\circ$, and assume w.l.o.g. that $\Phi^*(x_0, x_0) < \infty$. Let $x_1 \in A^\circ$ be as in condition (11), and let $\xi \in \partial\Phi^*(x_1, x_1)$,

$$W = \{x \in E : \langle x, \xi\rangle - \Phi(x, \xi) < \Phi^*(x_0, x_0) + \varepsilon\}.$$

Then $W \in \mathscr{E}$ and W is open. Let $V = W \cap A$. Then

$$\begin{aligned}\sup_{y\in V}[\langle y, a_n\xi\rangle - \Phi_n(y, a_n\xi)] &\leq a_n \left(b_n(\xi) + \sup_{y\in V}[\langle y, \xi\rangle - \Phi(y, \xi)]\right)\\ &\leq a_n(b_n(\xi) + \Phi^*(x_0, x_0) + \varepsilon).\end{aligned}$$

It follows that

$$\begin{aligned}\mathbf{P}_n\{Y_n \in A\} &\geq \mathbf{P}_n\{Y_n \in V\}\\ &\geq \inf_{y\in V} \exp[-(\langle y, a_n\xi\rangle - \Phi_n(y, a_n\xi))] \cdot\\ &\quad \cdot \int I_V(Y_n) \exp[\langle Y_n, a_n\xi\rangle - \Phi_n(Y_n, a_n\xi)]d\mathbf{P}_n\\ &\geq \exp[-a_n(b_n(\xi) + \Phi^*(x_0, x_0) + \varepsilon)] \cdot\\ &\quad \cdot \int I_V(Y_n) \exp[\langle Y_n, a_n\xi\rangle - \Phi_n(Y_n, a_n\xi)]d\mathbf{P}_n\end{aligned}$$

and hence

$$\begin{aligned}\liminf_n a_n^{-1} \log \mathbf{P}_n\{Y_n \in A\} &\geq -(\Phi^*(x_0, x_0) + \varepsilon)\\ &\quad + \liminf_n a_n^{-1} \log \int I_V(Y_n) \exp[\langle Y_n, a_n\xi\rangle - \Phi_n(Y_n, a_n\xi)]d\mathbf{P}_n.\end{aligned}$$

Therefore in order to complete the proof it is enough to show that

$$\liminf_n \int I_V(Y_n) \exp[\langle Y_n, a_n\xi\rangle - \Phi_n(Y_n, a_n\xi)]d\mathbf{P}_n = 1,$$

or, on account of condition (4),

$$\limsup_n \mathbf{P}_{n,\xi}\{Y_n \in V^c\} = 0, \tag{2.6}$$

where $d\mathbf{P}_{n,\xi} = \exp[\langle Y_n, a_n\xi\rangle - \Phi_n(Y_n, a_n\xi)]d\mathbf{P}_n$.

For $y \in E, \eta \in F$, let

$$\begin{aligned}\Phi_{n,\xi}(y, \eta) &= \Phi_n(y, a_n\xi + \eta) - \Phi_n(y, a_n\xi)\\ \Phi_\xi(y, \eta) &= \Phi(x, \xi + \eta) - \Phi(x, \xi);\end{aligned}$$

it is not difficult to prove that for $y, z \in E$,

$$\Phi_\xi^*(y, z) = \Phi^*(y, z) - (\langle z, \xi\rangle - \Phi(y, \xi)).$$

It is easily shown that $\Phi_{n,\xi}$ and Φ_ξ satisfy (1)–(3) and

$$\mathbf{E}_{n,\xi} \exp[\langle Y_n, \eta\rangle - \Phi_{n,\xi}(Y_n, \eta)] = 1 \quad \text{for all } \eta \in F. \tag{2.7}$$

We claim now that for every compact set $K \subset E$,

$$\limsup_n a_n^{-1} \log \mathbf{P}_{n,\xi}\{Z_n \in K\} \le -\inf_{x\in K} \Phi_\xi^*(x, x). \tag{2.8}$$

In order to prove this, we show first

$$\lim_n a_n^{-1} \log \mathbf{P}_{n,\xi}\{Y_n - Z_n \in V_1^c\} = -\infty \quad \text{for all } V_1 \in \mathscr{V}, \tag{2.9}$$

$$\{\mathscr{L}_{\mathbf{P}_{n,\xi}}(Z_n)\} \text{ is exponentially tight.} \tag{2.10}$$

For,

$$\begin{aligned}\mathbf{P}_{n,\xi}\{Y_n - Z_n \in V_1^c\} &= \int I_{V_1^c}(Y_n - Z_n) \exp[\langle Y_n, a_n\xi\rangle - \Phi_n(Y_n, a_n\xi)]d\mathbf{P}_n \\ &\le (\mathbf{P}_n\{Y_n - Z_n \in V_1^c\})^{1/2} \\ &\quad \times \left(\int \exp[\langle Y_n, 2a_n\xi\rangle - 2\Phi_n(Y, a_n\xi)]d\mathbf{P}_n\right)^{1/2} \\ &\le (\mathbf{P}_n\{Y_n - Z_n \in V_1^c\})^{1/2} \\ &\quad \times \exp\left(\frac{a_n}{2}[\bar{\Phi}(2\xi) + 2\bar{\Phi}(\xi) + b_n(2\xi) + 2b_n(\xi)]\right)\end{aligned} \tag{2.11}$$

by a simple estimate, taking into account (4) and (9). Using (6), (2.11) implies (2.9). (2.10) is proved similarly using (5).

Now (2.8) follows from (2.1), (2.7), (2.9) and (2.10). We claim next that there exists $V_1 \in \mathscr{V}$ such that

$$x_1 \notin \overline{V^c + V_1}. \tag{2.12}$$

For, choosing $V_1 \in \mathscr{V}$ such that $x_1 + V_1 + V_1 \subset V$, it easily follows that $(x_1 + V_1) \cap (V^c + V_1) = \phi$, which implies (2.12). Now let K_1 be a compact subset of E such that

$$\limsup_n a_n^{-1} \log \mathbf{P}_{n,\xi}\{Z_n \in K_1^c\} \le -1.$$

We have

$$\begin{aligned}\{Y_n \in V^c\} &\subset \{Y_n \in V^c, Y_n - Z_n \in V_1, Z_n \in K_1\} \cup \{Y_n - Z_n \in V_1^c\} \cup \{Z_n \in K_1^c\} \\ &\subset \{Z_n \in K_2\} \cup \{Y_n - Z_n \in V_1^c\} \cup \{Z_n \in K_1^c\},\end{aligned}$$

where $K_2 = \overline{V^c + V_1} \cap K_1$. Therefore by (2.8),

$$\begin{aligned}\limsup_n a_n^{-1} \log \mathbf{P}_{n,\xi}\{Y_n \in V^c\} &\le \max\{\limsup_n a_n^{-1} \log \mathbf{P}_{n,\xi}\{Z_n \in K_2\}, -1\}, \\ &\le \max\{-\inf_{x\in K_2} \Phi_\xi^*(x, x), -1\}.\end{aligned}$$

In order to prove (2.6), and hence to complete the proof of the theorem, it suffices to show

$$\ell \stackrel{\Delta}{=} \inf_{x \in K_2} \Phi_\xi^*(x, x) > 0. \tag{2.13}$$

Suppose $\ell = 0$. By the compactness of K_2 and the lower semicontinuity of Φ_ξ^*, there exists $x_2 \in K_2$ such that $\Phi_\xi^*(x_2, x_2) = 0$, that is,

$$\Phi^*(x_2, x_2) - [\langle x_2, \xi \rangle - \Phi(x_2, \xi)] = 0.$$

It follows that for all $\eta \in F$, $t > 0$, by assumption (8),

$$\langle x_2, \xi + t\eta \rangle - \Phi(x_2, \xi + t\eta) \le \langle x_2, \xi \rangle - \Phi(x_2, \xi),$$

$$\langle x_2, \eta \rangle \le \lim_{t \downarrow 0} t^{-1} [\Phi(x_2, \xi + t\eta) - \Phi(x_2, \xi)] = \langle \nabla\Phi(x_2, \xi), \eta \rangle.$$

Therefore $\langle x_2, \eta \rangle = \langle \nabla\Phi(x_2, \xi), \eta \rangle$ for all $\eta \in F$, and since F separates points in E, we have $x_2 = \nabla\Phi(x_2, \xi)$. On the other hand, by (8) and well known convex analysis arguments, $\xi \in \partial\Phi^*(x_1, x_1)$ implies $x_1 \in \partial\Phi(x_1, \xi)$, hence $x_1 = \nabla\Phi(x_1, \xi)$. Now by assumption (10) we must have $x_1 = x_2$. But this is impossible on account of (2.12) and $x_2 \in \overline{V^c + V_1}$. This establishes (2.13). □

3. Statement of the application to stochastic equations

We will consider the following conditions.

$$b : \mathbf{R}^d \to \mathbf{R}^d \text{ is bounded and uniformly Lipschitz.} \tag{3.1}$$

$$\sigma : \mathbf{R}^d \to \mathbf{R}^{d \times d} \text{ is bounded and uniformly Lipschitz.} \tag{3.2}$$

Let $(U, \mathscr{U})$ be a measurable space, ν a σ-finite measure on $(U, \mathscr{U})$. We endow $\mathbf{R}^d \times U$ with σ-algebra $\mathscr{B}(\mathbf{R}^d) \otimes \mathscr{U}$. Let $g : \mathbf{R}^d \times U \to \mathbf{R}^d$ be a measurable function such that

(i) there exists a measurable function $\bar{g} : U \to \mathbf{R}^+$ such that

$$\sup_{y \in \mathbf{R}^d} \|g(y, u)\| \le \bar{g}(u), \quad u \in U.$$

(ii) for all $u \in U$, $g(\cdot, u)$ is continuous.

We assume:

There exists $C > 0$ such that for all $y, z \in \mathbf{R}^d$

$$\int_U \|g(y, u) - g(z, u)\|^2 \nu(du) \le C\|y - z\|^2. \tag{3.3}$$

$$\text{For all } a > 0, \int_U (\bar{g}(u))^2 \exp(a\bar{g}(u))\nu(du) < \infty. \tag{3.4}$$

For every $r > 0$, there exists $\gamma = \gamma(r) > 0$ such that

$$\sup\{\int_U (q(y,z,u))^2 \exp[\gamma q(y,z,u)]\nu(du) : \|y\|, \|z\| \le r, y \ne z\} < \infty, \tag{3.5}$$

where $q(y,z,u) = (\|y-z\|)^{-1}\|g(y,u) - g(z,u)\|$.

Let $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\ge 0}, \mathbf{P})$ be a probability space with a filtration satisfying the usual conditions. Let m be Lebesgue measure on $\mathbf{R}^+$. We assume that a standard $\mathbf{R}^d$-valued Brownian motion B and a sequence $\{N_n\}$ of stationary Poisson random measures on $(\mathbf{R}^+ \times U, \mathscr{B}(\mathbf{R}^d) \otimes \mathscr{U})$ with mean measures $\{m \otimes (n\nu)\}$, respectively, are defined on Ω (see [I-W]). Assume that (3.1)–(3.3) hold. For each $n \in \mathbf{N}$, $x \in \mathbf{R}^d$, let X_n^x be the strong solution of the Itô stochastic equation

$$\begin{aligned} X_n^x(t) = x &+ \int_0^t b(X_n^x(s))ds + n^{-1/2}\int_0^t \sigma(X_n^x(s))dB(s) \\ &+ n^{-1}\int_{[0,t]\times U} g(X_n^x(s-), u)\tilde{N}_n(ds \times du), \end{aligned} \tag{3.6}$$

where $\tilde{N}_n = N_n - m \otimes (n\nu)$ is the compensated random measure; the process $\{X_n^x(t) : t \ge 0\}$ exists, is unique and has sample paths in $D([0,\infty), \mathbf{R}^d)$ by [I-W], Th. 9.1, Ch. IV (see also [G]). Let $T = [0,1]$ and let $\mathscr{D}$ be the σ-algebra of subsets of $D(T, \mathbf{R}^d)$ generated by the evaluations $\pi_t(f) = f(t), t \in T, f \in D(T, \mathbf{R}^d)$.

Theorem 3.1. *Let $\mu_n^x = \mathscr{L}(\{X_n^x(t) : t \in T\})$. Assume that conditions (3.1)–(3.5) hold. Then $\{\mu_n^x\}_{n\in\mathbf{N}}$ satisfies the large deviation principle on $D(T, \mathbf{R}^d)$, endowed with the uniform norm $\|\cdot\|_\infty$ and the σ-algebra $\mathscr{D}$, with the good rate function*

$$I^x(f) = \begin{cases} \int_T G^*(f(s), f'(s))ds & if\ f(0) = x\ and\ f\ is\ absolutely\ continuous \\ \infty & otherwise, \end{cases}$$

where

$$G^*(y,z) = \sup_{\alpha\in\mathbf{R}^d}[\langle z, \alpha\rangle - G(y,\alpha)], \quad y, z \in \mathbf{R}^d$$

and

$$G(y,\alpha) = \langle b(y), \alpha\rangle + \frac{1}{2}\langle \sigma(y)(\sigma(y))'\alpha, \alpha\rangle + \int_U \varphi(\langle g(y,u), \alpha\rangle)\nu(du), \quad \alpha \in \mathbf{R}^d,$$

where $\varphi(t) = e^t - 1 - t, t \in \mathbf{R}$.

More specifically, under conditions (3.1)–(3.4) the upper bound holds:

$$for\ all\ A \in \mathscr{D}, \limsup_n n^{-1}\log\mu_n^x(A) \le -\inf_{f\in\bar{A}} I^x(f),$$

and under conditions (3.1)–(3.5) the lower bound holds:

$$for\ all\ A \in \mathscr{D}, \liminf_n n^{-1}\log\mu_n^x(A) \ge -\inf_{f\in A^0} I^x(f).$$

4. Exponential martingales

Proposition 4.1. *Let* b, σ, g, ν, X_n^x *be as in Section 3. Then for every* $f \in C^2(\mathbf{R}^d)$ *such that*

$$C_1 = \sup\{\|\nabla f(g)\| : y \in \mathbf{R}^d\} < \infty,\ C_2 = \sup\{|D_{ij} f(y)| : i, j \le d, y \in \mathbf{R}^d\} < \infty,$$

$$\exp[f(X_n^x(t)) - f(x) - \int_0^t h_n(X_n^x(s))ds], \quad t \ge 0$$

is an $\{\mathscr{F}_t\}$*-local martingale, where for* $y \in \mathbf{R}^d$,

$$\begin{aligned} h_n(y) = \langle b(y), \nabla f(y)\rangle + \frac{1}{2n} \sum_{i,j=1}^d (\sigma(y)(\sigma(y))')_{ij} (D_i f(y) D_j f(y) + D_{ij} f(y)) \\ + \int_U (\exp[f(y + \frac{1}{n} g(y,u)) - f(y)] - 1 - \langle \frac{1}{n} g(y,u), \nabla f(y)\rangle)(n\nu)(du). \end{aligned}$$

Proof. We first show that the integral in the definition of $h_n(y)$ exists and is a bounded continuous function on $\mathbf{R}^d$. We have for $y, z \in \mathbf{R}^d$

$$\begin{aligned} &\exp[f(y+z) - f(y)] - 1 - \langle z, \nabla f(y)\rangle \\ &= e^{-f(y)}[h(y+z) - h(y) - \langle z, \nabla h(y)\rangle], \quad \text{where } h(y) = \exp[f(y)], \\ &= e^{-f(y)} \frac{1}{2} \sum_{i,j=1}^d D_{ij} h(\bar{y}) z_i z_j, \quad \text{where } \bar{y} = y + \theta z, \theta \in (0,1) \\ &= e^{-f(y)} \frac{1}{2} \sum_{i,j=1}^d e^{f(\bar{y})} (D_{ij} f(\bar{y}) + D_i f(\bar{y}) D_j f(\bar{y})) z_i z_j. \end{aligned}$$

Therefore by the assumptions on f, for a suitable constant $C > 0$

$$\begin{aligned} &|\exp[f(y + \frac{1}{n} g(y,u)) - f(y)] - 1 - \langle \frac{1}{n} g(y,u), \nabla f(y)\rangle| \\ &\le \exp(C_1 \frac{1}{n} \|g(y,u)\|) \cdot C \|\frac{1}{n} g(y,u)\|^2 \end{aligned}$$

and the claim follows from (3.4). Next, by Itô's formula (see [I-W], p. 66) applied to h, we have:

$$\exp[f(X_n^x(t))] - \int_0^t \exp[f(X_n^x(s))] \cdot h_n(X_n^x(s))ds, \quad t \ge 0 \tag{4.1}$$

is an $\{\mathscr{F}_t\}$-local martingale. Since $\inf_{s \le t} \exp[f(X_n^x(s))] > 0$ for each $t \ge 0$, by [E-K], p. 66, we have from (4.1):

$$\exp[f(X_n^x(t))] \exp[-\int_0^t \exp[-f(X_n^x(s))](\exp[f(X_n^x(s))] \cdot h_n(X_n^x(s)))ds], t \ge 0$$

is an $\{\mathscr{F}_t\}$-local martingale, and the conclusion follows. □

Proposition 4.2. *For $n \in \mathbf{N}$, $y, \alpha \in \mathbf{R}^d$, let*

$$G_n(y,\alpha) = \langle b(y), \alpha\rangle + \frac{1}{2n}\langle \sigma(y)(\sigma(y))'\alpha, \alpha\rangle + \int_U \varphi(\langle \frac{1}{n} g(y,u), \alpha\rangle)(n\nu)(du),$$

where φ is as in Theorem 3.1. Then for each $n \in \mathbf{N}$, $\alpha \in \mathbf{R}^d$,

$$M_n^{(\alpha)}(t) \triangleq \exp[\langle X_n^x(t) - x, \alpha\rangle - \int_0^t G_n(X_n^x(s), \alpha)ds], \quad t \geq 0$$

is an $\{\mathscr{F}_t\}$ L^2-martingale. In particular, for all $t \geq 0$

$$\mathbf{E} M_n^{(\alpha)}(t) = 1.$$

Proof. Applying Proposition 4.1 to $f(y) = \langle y, \alpha\rangle$, we obtain: $M_n^{(\alpha)}(t)$, $t \geq 0$ is an $\{\mathscr{F}_t\}$ local martingale. Let $\tau_k \uparrow \infty$ a.s. be a localizing sequence of stopping times. Let $t \geq 0$, and let $\tau \leq t$ be a stopping time. By the optional sampling theorem.

$$\mathbf{E} M_n^{(\alpha)}(\tau \wedge \tau_k) = \mathbf{E} M_n^{(\alpha)}(0) = 1.$$

Since $M_n^{(\alpha)}(t) \geq 0$ and $M_n^{(\alpha)}(\tau \wedge \tau_k) \to M_n^{(\alpha)}(\tau)$ a.s., by Fatou's lemma

$$\mathbf{E} M_n^{(\alpha)}(\tau) \leq \liminf_k \mathbf{E} M_n^{(\alpha)}(\tau \wedge \tau_k) = 1. \tag{4.2}$$

Now

$$\mathbf{E}(M_n^{(\alpha)}(\tau))^2 = \mathbf{E}\exp[\langle X_n^x(\tau) - x, 2\alpha\rangle - \int_0^\tau G_n(X_n^x(s), 2\alpha)ds + \int_0^\tau G_n(X_n^x(s), 2\alpha) - 2\int_0^\tau G_n(X_n^x(s), \alpha)ds]. \tag{4.3}$$

Since $0 \leq \varphi(v) \leq \varphi(|v|) \leq \frac{1}{2}|v|^2 \exp(|v|)$ for $v \in \mathbf{R}$, we have by (3.1), (3.2) and (3.4): there exists a constant $C > 0$ such that for all $y, \alpha \in \mathbf{R}^d$

$$|G_n(y,\alpha)| \leq H_n(\|\alpha\|), \tag{4.4}$$

where for $a > 0$, $H_n(a) = Ca + \frac{C}{2n}a^2 + \frac{a^2}{2n}\int_U (\bar{g}(u))^2 \exp(n^{-1}a\bar{g}(u))\nu(du)$. By (4.3) and (4.4) and since $\tau \leq t$, it follows that for some constant $C' > 0$.

$$\mathbf{E}(M_n^{(\alpha)}(\tau))^2 \leq C'\mathbf{E} M_n^{(2\alpha)}(\tau) = C'.$$

Then for all $t \geq 0$,

$$\{M_n^{(\alpha)}(\tau) : \tau \leq t\} \text{ is uniformly integrable.}$$

By a standard argument (see e.g. [J-S], p. 12), $M_n^{(\alpha)}(t)$, $t \geq 0$ is an L^2-martingale. □

Lemma 4.3. *For all* $\ell > 0$

$$\mathbf{E}\exp(\ell \|X_n^x\|_\infty) < \infty,$$

where $\|X_n^x\|_\infty = \sup_{t\in T} \|X_n^x(t)\|$.

Proof. For $a > 0$, let $\tau_a = \inf\{t \geq 0 : \|X_n^x(t)\| > a\}$. Let $\{\alpha_1, \ldots, \alpha_d\}$ be a basis of $\mathbf{R}^d$. Then for some constant $c > 0$, all $y \in \mathbf{R}^d$,

$$\|y\| \leq c \sup_{j\leq d} |\langle y, \alpha_j\rangle|.$$

Therefore if $S = \{\alpha_j : j \leq d\} \cup \{-\alpha_j : j \leq d\}$, we have for $h \geq 0$

$$\mathbf{P}\{\|X_n^x(\tau_a \wedge 1)\| > h\} \leq (2d) \sup_{\alpha\in S} \mathbf{P}\{\langle X_n^x(\tau_a \wedge 1), \alpha\rangle > h/c\}. \tag{4.5}$$

As in the proof of Proposition 4.2, given $\beta \in \mathbf{R}^d$, for some constant C' independent of a, we have

$$\mathbf{E}\exp\langle X_n^x(\tau_a \wedge 1), \beta\rangle \leq C' \mathbf{E} M_n^{(\beta)}(\tau_a \wedge 1) = C'. \tag{4.6}$$

Therefore for all $\ell > 0$, $a > 0$ we have by (4.5) and (4.6) with $\beta = c\ell\alpha$

$$\begin{aligned}\mathbf{P}\{\|X_n^x\|_\infty > a\} &\leq \mathbf{P}\{\|X_n^x(\tau_a \wedge 1)\| > a\}\\ &\leq (2d)\sup_{\alpha\in S}(e^{-\ell a}\mathbf{E}\exp\langle X_n^x(\tau_a \wedge 1), c\ell\alpha\rangle)\\ &\leq (2d)C' e^{-\ell a}.\end{aligned}$$

This implies the conclusion. □

We shall denote by $M(T, \mathbf{R}^d)$ the space of vector measures defined on the Borel σ-algebra of T, with values in $\mathbf{R}^d$. For $f \in D(T, \mathbf{R}^d)$, $\lambda \in M(T, \mathbf{R}^d)$, we write

$$\langle f, \lambda\rangle \triangleq \int_T \langle f, d\lambda\rangle.$$

Let $\{X(t), t \geq 0\}$ be the canonical process and $\{\mathscr{G}_t, t \geq 0\}$ the canonical filtration on $D([0,\infty), \mathbf{R}^d)$. Let $P_n^x = \mathscr{L}_{\mathrm{P}}(\{X_n^x(t) : t \geq 0\})$. Then it is well known that

$$(D([0,\infty), \mathbf{R}^d), \{X(t), t \geq 0\}, \{\mathscr{G}_t, t \geq 0\}, \{P_n^x, x \in \mathbf{R}^d\}) \tag{4.7}$$

is a Markov process. In the next proposition and elsewhere in the paper, we will use the notation

$$X_n^x = \{X_n^x(t), t \in T\}, \quad P_n^x = \mathscr{L}_{\mathrm{P}}(X_n^x);$$

this abuse of notation should cause no confusion, since it will be clear from the context when P_n^x is a measure on $(D(T, \mathbf{R}^d), \mathscr{D})$.

Proposition 4.4. *Assume that (3.1)–(3.4) hold. For* $f \in D(T, \mathbf{R}^d)$, $\lambda \in M(T, \mathbf{R}^d)$, *let*

$$\Phi_n^x(f, \lambda) = \langle x, \lambda(T) \rangle + \int_T G_n(f(s), \lambda([s, 1]))ds.$$

Then

$$\mathbf{E} \exp[\langle X_n^x, \lambda \rangle - \Phi_n^x(X_n^x, \lambda)] = 1. \tag{4.8}$$

Proof. What we must prove is: for all $x \in \mathbf{R}^d$, $\lambda \in M(T, \mathbf{R}^d)$,

$$E_n^x \exp[\langle X, \lambda \rangle - \Phi_n^x(X, \lambda)] = 1, \tag{4.9}$$

where $X = \{X(t) : t \in T\}$.

(1) We show first that (4.9) holds for any $\lambda \in M(T, \mathbf{R}^d)$ of the form $\lambda = \sum_{j=1}^k \alpha_j \delta_{t_j}$, where $0 = t_0 < t_1 < \ldots < t_k \leq 1$ and $\alpha_j \in \mathbf{R}^d$, $j = 1, \ldots, k$. We prove the claim by induction. If $k = 1$, then $\lambda = \alpha_1 \delta_{t_1}$ and by the fact that $G_n(y, 0) = 0$ for all $y \in \mathbf{R}^d$ and Proposition 4.2, we have

$$E_n^x \exp[\langle X, \lambda \rangle - \Phi_n^x(X, \lambda)] = E_n^x \exp[\langle X(t_1) - x, \alpha_1 \rangle - \int_0^{t_1} G_n(X(s), \alpha_1)ds] = 1. \tag{4.10}$$

Let $k \in \mathbf{N}$. Assume now that (4.9) holds for all $\lambda_k \in M(T, \mathbf{R}^d)$ of the form $\lambda_k = \sum_{j=1}^k \alpha_j \delta_{t_j}$, with $\{t_j\}$ as above and let $\lambda = \sum_{j=1}^{k+1} \alpha_j \delta_{t_j}$. By the Markov property, conditioning on $\mathscr{G}_{t_k}$ in the second step,

$$\begin{aligned} E_n^x \exp[\langle X, \lambda \rangle - \Phi_n^x(X, \lambda)] &= E_n^x \exp[\sum_{j=1}^{k+1} \langle X(t_j), \alpha_j \rangle - \langle x, \sum_{j=1}^{k+1} \alpha_j \rangle \\ &\quad - \sum_{j=1}^{k+1} \int_{[t_j - 1, t_j)} G_n(X(s), \sum_{i=j}^{k+1} \alpha_i)ds] \\ &= E_n^x(\exp[\langle X, \lambda_k \rangle - \Phi_n^x(X, \lambda_k)] \cdot H(X(t_k)), \end{aligned}$$

where for $y \in \mathbf{R}^d$

$$H(y) = E_n^y \exp\left[\langle X(t_{k+1} - t_k) - y, \alpha_{k+1} \rangle - \int_0^{t_{k+1} - t_k} G_n(X(s), \alpha_{k+1})\, ds \right].$$

But $H(y) = 1$ for all $y \in \mathbf{R}^d$ by (4.10). Using now the inductive hypothesis, we conclude that (4.9) holds for λ, proving the claim.

(2) Let $\lambda \in M(T, \mathbf{R}^d)$. As in [deA1], Lemma A.2, define $\lambda_k \in M(T, \mathbf{R}^d)$ by

$$\lambda_k = \sum_{j=0}^k a_{kj} \delta_{j/k},$$

where $a_{k0} = \lambda(\{0\})$, $a_{kj} = \lambda((\frac{j-1}{k}, \frac{j}{k}))$, $1 \leq j \leq k$. Then for all $f \in D(T, \mathbf{R}^d)$,

$$\lim_k \int \langle f, d\lambda_k \rangle = \int \langle f, d\lambda \rangle. \tag{4.11}$$

Let $\|\cdot\|_v$ be the total variation norm on $M(T, \mathbf{R}^d)$. Then for all k,

$$\begin{aligned} \|\lambda_k\|_v &\le \|\lambda(\{0\})\| + \sum_{j=1}^{k} \|\lambda((\frac{j-1}{k}, \frac{j}{k}])\| \\ &\le |\lambda|(\{0\}) + \sum_{j=1}^{k} |\lambda|((\frac{j-1}{k}, \frac{j}{k}]) \\ &= |\lambda|([0,1]), \\ &= \|\lambda\|_v \end{aligned} \tag{4.12}$$

where $|\lambda|$ is the total variation measure associated to λ. Now for all $s \in T$,

$$\begin{aligned} \|\lambda_k([s,1]) - \lambda([s,1])\| &= \|\lambda([\frac{[sk]}{k}, 1]) - \lambda([s,1])\| \\ &\le |\lambda|[\frac{[sk]}{k}, s) \\ &\to 0. \end{aligned} \tag{4.13}$$

By (4.4), (4.12), (4.13) and the dominated convergence theorem, it follows that

$$\begin{aligned} \Phi_n^x(f, \lambda_k) &= \langle x, \lambda_k(T) \rangle + \int_T G_n(f(s), \lambda_k([s,1])ds \\ &\to \Phi_n^x(f, \lambda). \end{aligned} \tag{4.14}$$

By (4.4) and (4.12)

$$\exp[\langle X, \lambda_k \rangle - \Phi_n^x(X, \lambda_k)] \le \exp[\|\lambda\|_v \|X\|_\infty + H_n(\|\lambda\|_v)]. \tag{4.15}$$

Finally by Lemma 4.3, (4.11), (4.14), (4.15), part (1) of this proof and the dominated convergence theorem, we have

$$\begin{aligned} E_n^x \exp[\langle X, \lambda \rangle - \Phi_n^x(X, \lambda)] &= \lim_k E_n^x \exp[\langle X, \lambda_k \rangle - \Phi_n^x(X, \lambda_k)] \\ &= 1. \end{aligned}$$

This completes the proof. □

5. Exponential tightness, discretization and superexponential approximation

Let $\{X(t), t \ge 0\}$, $\{\mathscr{G}_t, t \ge 0\}$, $\{P_n^x, x \in \mathbf{R}^d\}$ be as in (4.7).

Lemma 5.1. *Let q be a seminorm on $\mathbf{R}^d$. Then for all $t \ge 0, s > 0, a > 0$,*

$$\sup_{x \in \mathbf{R}^d} P_n^x\{\sup_{h \le s} q(X(t+h) - X(t)) > a\} \le 2 \sup_{h \le s} \sup_{x \in \mathbf{R}^d} P_n^x\{q(X(h) - x) > \frac{a}{2}\}.$$

Proof. By the Markov property,

$$\begin{aligned} P_n^x\{\sup_{h\le s} q(X(t+h)-X(t))>a\} &= E_n^x E_n^x[I(\sup_{h\le s} q(X(t+h)-X(t))>a)|\mathscr{G}_t] \\ &= E_n^x P_n^{X(t)}(A(X(t))), \end{aligned}$$

where $A(y)=\{f\in D([0,\infty),\mathbf{R}^d):\sup_{h\le s} q(f(h)-y)>a\}$. Therefore

$$\sup_{x\in\mathbf{R}^d} \mathbf{P}_n^x\{\sup_{h\le s} q(X(t+h)-X(t))>a\} \le \sup_{y\in\mathbf{R}^d} P_n^y(A(y)). \tag{5.1}$$

Let $B_k=\{(j/2^k)s:0\le j\le 2^k\}$. For fixed $y\in\mathbf{R}^d$, let

$$\begin{aligned} \tau &= \inf\{u\in B_k: q(X(u)-y)>a\}, \\ A_k(y) &= \bigcup_{u\in B_k}\{\tau=u\}=\{\sup_{u\in B_k} q(X(u)-y)>a\}. \end{aligned}$$

For $u\in B_k$

$$\begin{aligned} P_n^y\{\tau=u\} \le\ & P_n^y(\{\tau=u\}\cap\{q(X(s)-y)>\frac{a}{2}\} \\ & +P_n^y(\{\tau=u\}\cap\{q(X(s)-X(u))>\frac{a}{2}\}. \end{aligned} \tag{5.2}$$

By the Markov property,

$$P_n^y(\{\tau=u\}\cap\{q(X(s)-X(u))>\frac{a}{2}\})=E_n^y I\{\tau=u\}P_n^{X(u)}(B(s-u,X(u)),$$

where $B(h,z)=\{f\in D([0,\infty),\mathbf{R}^d):q(f(h)-z)>\frac{a}{2}\}$. Therefore

$$P_n^y(\{\tau=u\}\cap\{q(X(s)-X(u))>\frac{a}{2}\}\le \sup_{h\le s}\sup_{z\in\mathbf{R}^d} P_n^z(B(h,z))\cdot P_n^y\{\tau=u\}. \tag{5.3}$$

By (5.2) and (5.3)

$$\begin{aligned} P_n^y(A_k(y)) &= \sum_{u\in B_k} P_n^y\{\tau=u\} \\ &\le P_n^y\{q(X(s)-y)>\frac{a}{2}\}+\sup_{h\le s}\sup_{z\in\mathbf{R}^d} P_n^z(B(h,z)) \\ &\le 2\sup_{h\le s}\sup_{z\in\mathbf{R}^d} P_n^z(B(h,z)). \end{aligned} \tag{5.4}$$

By the right-continuity of paths, $A_k(y)\uparrow A(y)$, and therefore $P_n^y(A_k(y))\uparrow P_n^y(A(y))$. The conclusion follows now from (5.1) and (5.4). □

For $y, \alpha \in \mathbf{R}^d$, let $G(y, \alpha)$ be as in Theorem 3.1, $G_n(y, \alpha)$ as in Proposition 4.2. Then for all $n \in \mathbf{N}$,

$$\frac{1}{n} G_n(y, n\alpha) = G(y, \alpha). \tag{5.5}$$

Let $\bar{G}(\alpha) = \sup_{y \in \mathbf{R}^d} |G(y, \alpha)|$. By (3.1), (3.2) and (3.4) and the elementary inequality previous to (4.4), for all $\alpha \in \mathbf{R}^d$

$$\bar{G}(\alpha) < \infty.$$

Lemma 5.2. *For all $\alpha \in \mathbf{R}^d$, $s > 0$,*

$$\mathit{lim\,sup}_n \frac{1}{n} \log \sup_{h \leq s} \sup_{x \in \mathbf{R}^d} E_n^x \exp\langle X(h) - x, n\alpha \rangle \leq s\bar{G}(\alpha).$$

Proof. By (4.10) and (5.5), for all $x \in \mathbf{R}^d$, $h \leq s$, $n \in \mathbf{N}$

$$\begin{aligned} E_n^x \exp\langle X(h) - x, n\alpha \rangle &= E_n^x (\exp[\langle X(h) - x, n\alpha \rangle - \int_0^h G_n(X(t), n\alpha) dt] \\ &\quad \cdot \exp[n \int_0^h G(X(t), \alpha) dt]) \\ &\leq \exp(ns\bar{G}(\alpha)) \end{aligned}$$

and the conclusion follows. □

Proposition 5.3. *For $x \in \mathbf{R}^d$, $t \geq 0$, let $Z_n^x(t) = X_n^x(\frac{[nt]}{n})$, and let $Z_n^x = \{Z_n^x(t) : t \in T\}$. Then for every $\varepsilon > 0$,*

$$\lim_n \frac{1}{n} \log \mathbf{P}\{||X_n^x - Z_n^x||_\infty > \varepsilon\} = -\infty.$$

Proof. Let $Z_n(t) = X(\frac{[nt]}{n})$. By Lemma 5.1,

$$\begin{aligned} \mathbf{P}\{||X_n^x - Z_n^x||_\infty > \varepsilon\} &= P_n^x\{||X - Z_n||_\infty > \varepsilon\} \\ &\leq P_n^x\{\sup_{1 \leq k \leq n} \sup_{h \leq 1/n} ||X(\frac{k-1}{n} + h) - X(\frac{k-1}{n})|| > \varepsilon\} \\ &\leq \sum_{k=1}^n P_n^x\{\sup_{h \leq 1/n} ||X(\frac{k-1}{n} + h) - X(\frac{k-1}{n})|| > \varepsilon\} \\ &\leq 2n \sup_{h \leq 1/n} \sup_{y \in \mathbf{R}^d} P_n^y\{||X(h) - y|| > \varepsilon/2\}. \end{aligned} \tag{5.6}$$

Proceeding as in the proof of Lemma 4.3,

$$P_n^y\{||X(h) - y|| > \varepsilon/2\} \leq (2d) \sup_{\alpha \in S} P_n^y\{\langle X(h) - y, \alpha \rangle > \frac{\varepsilon}{2c}\}. \tag{5.7}$$

By (5.6) and (5.7), for all $a > 0$

$$\begin{aligned}\mathbf{P}\{||X_n^x - Z_n^x||_\infty > \varepsilon\} &\le 4dn \sup_{\alpha\in S}\sup_{h\le 1/n}\sup_{y\in\mathbf{R}^d} P_n^y\{\langle X(h)-y,\alpha\rangle > \tfrac{\varepsilon}{2c}\}\\ &\le 4dn \sup_{\alpha\in S}\sup_{h\le 1/n}\sup_{y\in\mathbf{R}^d} \exp(-n(\tfrac{\varepsilon}{2c})a)E_n^y \exp\langle X(h)-y, na\alpha\rangle.\end{aligned}$$

By Lemma 5.2,

$$\limsup_n \frac{1}{n}\log \mathbf{P}\{||X_n^x - Z_n^x||_\infty > \varepsilon\} \le -(\frac{\varepsilon}{2c})a.$$

But a is arbitrary. □

Proposition 5.4. *Let $\{Z_n^x\}$ be as in Proposition 5.3. Then*

$$\{\mathscr{L}_{\mathbf{P}}(Z_n^x), n \in \mathbf{N}\} \textit{ is exponentially tight in } (D(T,\mathbf{R}^d), ||\cdot||_\infty).$$

Proof. Following [deA1], proof of Lemma 4.1, it suffices to show: for every $\varepsilon > 0$, $a > 0$, there exist $r > 0$, $m \in \mathbf{N}$ such that

$$\limsup_n \frac{1}{n}\log \mathbf{P}\{d(Z_n^x, H_m(C_r + x)) > \varepsilon\} \le -a, \tag{5.8}$$

where $C_r = \{y \in \mathbf{R}^d : q(y) \le r\}$, $q(y) = \sup_{j\le d}|\langle y, \alpha_j\rangle|$, with $\{\alpha_j\}$ as in the proof of Lemma 4.3, and for $A \subset \mathbf{R}^d$,

$$H_m(A) = \{f \in D(T,\mathbf{R}^d) : f = \sum_{j=0}^{m-1} x_j I_{[j/m, \frac{j+1}{m})} + x_m I_{\{1\}} : x_j \in A, j = 0,\dots,m\}.$$

Now, with Z_n as in the proof of Proposition 5.3,

$$\begin{aligned}\mathbf{P}\{d(Z_n^x, H_m(C_r+x)) > \varepsilon\} &= P_n^x\{d(Z_n - x, H_m(C_r)) > \varepsilon\}\\ &\le P_n^x\{Z_n - x \notin H_n(C_r)\}\\ &\quad + P_n^x\{Z_n - x \in H_n(C_r), d(Z_n - x, H_m(C_r)) > \varepsilon\}.\end{aligned} \tag{5.9}$$

Recalling that

$$Z_n(t) = X(\frac{[nt]}{n}) = \sum_{j=0}^{n-1} X(j/n) I_{[j/n, \frac{j+1}{n})}(t) + X(1)I_{\{1\}}(t),$$

we have by Lemma 5.1

$$\begin{aligned}P_n^x\{Z_n - x \notin H_n(C_r)\} &= P_n^x\{\sup_{j\le n} q(X(j/n) - x) > r\}\\ &\le 2\sup_{h\le 1}\sup_{y\in\mathbf{R}^d} P_n^y\{q(X(h)-y) > r/2\}.\end{aligned} \tag{5.10}$$

Let S be as in the proof of Lemma 4.3. Then

$$P_n^y\{q(X(h)-y)>r/2\}\le 4d\sup_{\alpha\in S}P_n^y\{\langle X(h)-y,\alpha\rangle>r/2\}. \tag{5.11}$$

Proceeding as in the proof of Proposition 5.3 and using (5.10), (5.11) and Lemma 5.2, we have

$$\limsup_n \frac{1}{n}\log P_n^x\{Z_n-x\notin H_n(C_r)\}\le -\frac{r}{2}+\sup_{\alpha\in S}\bar{G}(\alpha). \tag{5.12}$$

Next, for $n\ge m$, as in [deA1], pp. 90-91, and using Lemma 5.1,

$$\begin{aligned}
&P_n^x\{Z_n-x\in H_n(C_r), d(Z_n-x, H_m(C_r))>\varepsilon\}\\
&\le P_n^x\{\sup_{0\le i\le m-1}\ \sup_{1\le j\le \frac{n}{m}+1} q(X(\frac{[\frac{ni}{m}]+j}{n})-X(\frac{[\frac{ni}{m}]}{n}))>\varepsilon\}\\
&\le \sum_{i=0}^{m-1}P_n^x\{\sup_{1\le j\le \frac{n}{m}+1} q(X(\frac{[\frac{ni}{m}]+j}{n})-X(\frac{[\frac{ni}{m}]}{n}))>\varepsilon\}\\
&\le 2m\sup_{h\le 2/m}\ \sup_{y\in\mathbf{R}^d}P_n^y\{q(X(h)-y)>\varepsilon/2\}\\
&\le 4md\,\exp(-n(\varepsilon/2)\ell)\sup_{\alpha\in S}\ \sup_{h\le 2/m}\ \sup_{y\in\mathbf{R}^d}E_n^y\exp\langle X(h)-y, n\ell\alpha\rangle.
\end{aligned}$$

By Lemma 5.2,

$$\begin{aligned}
&\limsup_n \frac{1}{n}\log P_n^x\{Z_n-x\in H_n(C_r), d(Z_n-x, H_m(C_r))>\varepsilon\}\\
&\le -(\varepsilon/2)\ell+\frac{2}{m}\sup_{\alpha\in S}\bar{G}(\ell\alpha).
\end{aligned} \tag{5.13}$$

By (5.9), (5.12) and (5.13), we have

$$\begin{aligned}
&\limsup_n \frac{1}{n}\log \mathbf{P}\{d(Z_n^x, H_m(C_r+x))>\varepsilon\}\\
&\le \max\{-\frac{r}{2}+\sup_{\alpha\in S}\bar{G}(\alpha), -(\frac{\varepsilon}{2})\ell+\frac{2}{m}\sup_{\alpha\in S}\bar{G}(\ell\alpha)\}.
\end{aligned}$$

It is clear that for suitable choices of ℓ, r and m the right hand side of this inequality is no greater than $(-a)$. This establishes (5.8), completing the proof. □

6. Identification of the rate function and compactness of the level sets

Theorem 6.1. *Let G be as in Theorem 3.1, and for $f\in D(T,\mathbf{R}^d)$, $\lambda\in M(T,\mathbf{R}^d)$, $x\in\mathbf{R}^d$, let*

$$\Phi^x(f,\lambda)=\langle x,\lambda(T)\rangle+\int_T G(f(s),\lambda([s,1]))ds.$$

For, $f, h \in D(T, \mathbf{R}^d)$*, let*

$$(\Phi^x)^*(f,h) = \sup_{\lambda \in M(T,\mathbf{R}^d)} [\int \langle h, d\lambda\rangle - \Phi^x(f,\lambda)].$$

Then

$$(\Phi^x)^*(f,f) = \begin{cases} \int_T G^*(f(s), f'(s))ds & \text{if } f(0) = x \text{ and} \\ & f \text{ is absolutely continuous} \\ \infty & \text{otherwise,} \end{cases}$$

where

$$G^*(y,z) = \sup_{\alpha \in \mathbf{R}^d} [\langle z, \alpha\rangle - G(y,\alpha)], \qquad y, z \in \mathbf{R}^d.$$

Proof. (1) Suppose $f(0) = x$, f is absolutely continuous. Then

$$(\Phi^x)^*(f,f) \leq \int_T G^*(f(s), f'(s))ds.$$

For, let $\lambda \in M(T, \mathbf{R}^d)$. Then

$$\begin{aligned}\int \langle f, d\lambda\rangle - \Phi^x(f,\lambda) &= \int_T \langle \int_0^s f'(u)du + x, d\lambda(s)\rangle - [\langle x, \lambda(T)\rangle \\ &\quad + \int_T G(f(s), \lambda([s,1]))ds] \\ &= \int_T [\langle f'(s), \lambda([s,1])\rangle - G(f(s), \lambda([s,1]))]ds \\ &\leq \int_T G^*(f(s), f'(s))ds.\end{aligned}$$

(2) Suppose $(\Phi^x)^*(f,f) < \infty$. Then $f(0) = x$ and f is absolutely continuous. For, let $\alpha \in \mathbf{R}^d$. Taking $\lambda = \alpha\delta_0$, we have using $G(y,0) = 0$ for all $y \in \mathbf{R}^d$:

$$\begin{aligned}&\int_T \langle f, d(\alpha\delta_0)\rangle \leq \Phi^x(f, \alpha\delta_0) + (\Phi^x)^*(f,f), \\ &\langle f(0) - x, \alpha\rangle \leq (\Phi^x)^*(f,f)\end{aligned}$$

for all $\alpha \in \mathbf{R}^d$, which implies $f(0) = x$. Next, let $\rho > 0, \alpha_j \in \mathbf{R}^d$ with $||\alpha_j|| \leq 1, k \in \mathbf{N}$,

$$0 \leq s_1 < t_1 \leq s_2 < t_2 \leq \cdots \leq s_k < t_k \leq 1.$$

Taking $\lambda = \sum_{j=1}^k \rho\alpha_j(\delta_{t_j} - \delta_{s_j})$, we have

$$\rho \sum_{j=1}^k \langle f(t_j) - f(s_j), \alpha_j\rangle = \int_T \langle f, d\lambda\rangle \leq \sum_{j=1}^k \int_{s_j}^{t_j} G(f(s), \rho\alpha_j)ds + (\Phi^x)^*(f,f).$$

Let $A(\rho) = \sup\{|G(y, \alpha)| : ||y|| \le ||f||_\infty, ||\alpha|| \le \rho\}$. Then

$$\sum_{j=1}^{k} ||f(t_j) - f(s_j)|| \le \rho^{-1} A(\rho) \sum_{j=1}^{k} (t_j - s_j) + \rho^{-1} (\Phi^x)^*(f, f)$$

and the absolute continuity of f follows.

(3) Suppose $f(0) = x$ and f is absolutely continuous. Then

$$\int_T G^*(f(s), f'(s))ds \le (\Phi^x)^*(f, f). \tag{6.1}$$

For, let $\{D_k\}$ be an increasing sequence of finite subsets of $\mathbf{R}^d$ such that $0 \in D_1$ and $\cup_k D_k$ is dense. For $y, z \in \mathbf{R}^d$, let

$$L_k(y, z) = \sup_{\alpha \in D_k} [\langle z, \alpha \rangle - G(y, \alpha)].$$

Then for all $y, z \in \mathbf{R}^d$, $0 \le L_k(y, z) \uparrow G^*(y, z)$, and therefore

$$\int_T L_k(f(s), f'(s))ds \uparrow \int_T G^*(f(s), f'(s))ds. \tag{6.2}$$

Define now $F_n : T \to \mathbf{R}^2$ by

$$F_n(s) = \sum_{j=1}^{2^n} (f(\frac{j-1}{2^n}), 2^n[f(\frac{j}{2^n}) - f(\frac{j-1}{2^n})]) I_{[\frac{j-1}{2^n}, j/2^n)}(s).$$

Then arguing as in [deA3], p. 154, we have

$$F_n(s) \to (f(s), f'(s)) \quad \text{a.e. } [m],$$

where m is Lebesgue measure, and consequently

$$L_k(f(s), f'(s)) = \lim_n L_k(F_n(s)) \quad \text{a.e. } [m],$$

By Fatou's lemma, for each $k \in \mathbf{N}$

$$\int_T L_k(f(s), f'(s))ds \le \liminf_n \int_T L_k(F_n(s))ds.$$

Taking into account (6.2), it follows that in order to prove (6.1) it is enough to show: for all $k \in \mathbf{N}$,

$$\liminf_n \int_T L_k(F_n(s))ds \le (\Phi^x)^*(f, f). \tag{6.3}$$

By the definitions of L_k and F_n, for suitable choices of $\alpha_j^{(n)} \in D_k (j = 1, \ldots, 2^n)$ we have

$$\int_T L_k(F_n(s))ds = \frac{1}{2^n}\sum_{j=1}^{2^n}(\langle 2^n(f(\frac{j}{2^n}) - f(\frac{j-1}{2^n})), \alpha_j^{(n)}\rangle - G(f(j/n), \alpha_j^{(n)})).$$

Given $\varepsilon > 0$, by the uniform continuity of f and the uniform continuity of $G(\cdot, \alpha)$ on compact sets for each $\alpha \in D_k$, for all sufficiently large n we have for $s \in [\frac{j-1}{2^n}, j/2^n]$, $1 \le j \le 2^n$,

$$G(f(s), \alpha_j^{(n)}) \le G(f(j/2^n), \alpha_j^{(n)}) + \varepsilon$$

and therefore

$$\int_{[\frac{j-1}{2^n}, j/2^n]} G(f(s), \alpha_j^{(n)})ds \le \frac{1}{2^n}[G(f(j/2^n), \alpha_j^{(n)}) + \varepsilon].$$

Taking now

$$\lambda_n = \sum_{j=1}^{2^n} \alpha_j^{(n)}(\delta_{j/2^n} - \delta_{\frac{j-1}{2^n}}),$$

for all sufficiently large n we have

$$\begin{aligned}\int_T L_k(F_n(s))ds &\le \int_T \langle f, d\lambda_n\rangle - \int_T G(f(s), \lambda_n([s,1]))ds + \varepsilon\\ &\le (\Phi^x)^*(f, f) + \varepsilon,\end{aligned}$$

proving (6.3) and hence (6.2). Now (1)–(3) yield the result. □

Proposition 6.2. *For $a \ge 0$, let*

$$L_a = \{f \in D(T, \mathbf{R}^d) : (\Phi^x)^*(f, f) \le a\}.$$

Then L_a is compact for the uniform norm.

Proof. Since $\Phi^x(\cdot, \lambda)$ is continuous on $(D(T, \mathbf{R}^d), || \cdot ||_\infty)$, it follows that $(\Phi^x)^*$ is lower semicontinuous for $|| \cdot ||_\infty$, being the supremum of continuous functions. Therefore L_a is $|| \cdot ||_\infty$-closed. It remains to show that L_a is $|| \cdot ||_\infty$-relatively compact. By the Arzela'-Ascoli theorem, it is enough to show: (i) L_a is uniformly bounded, (ii) L_a is uniformly equicontinuous. We will prove (ii); the proof of (i) is similar. Let $\rho > 0$. Arguing as in the proof of Theorem 6.1, we have: for $f \in L_a$, $s, t \in T$,

$$||f(t) - f(s)|| \le \rho^{-1}A(\rho)|t - s| + \rho^{-1}a,$$

and (ii) follows. □

7. Some analytic considerations

Proposition 7.1. *Assume that b, σ, g, ν satisfy assumptions (3.1)–(3.4), and, furthermore: there exists $a > 0$ such that for all $y, \alpha \in \mathbf{R}^d$,*

$$\langle \sigma(y)(\sigma(y))'\alpha, \alpha\rangle \geq a||\alpha||^2. \tag{7.1}$$

Let Φ^x be as in Theorem 6.1, and assume $(\Phi^x)^*(f_0, f_0) < \infty$. Then for every $\varepsilon > 0$ there exists $f_1 \in D(T, \mathbf{R}^d)$ such that

(i) $||f_1 - f_0|| < \varepsilon$.
(ii) $(\Phi^x)^*(f_1, f_1) \leq (\Phi^x)^*(f_0, f_0) + \varepsilon$.
(iii) $\partial(\Phi^x)^*(f_1, f_1) \neq \phi$.

The proof of Proposition 7.1 requires several lemmas. When writing $\nabla G(y, \alpha)$ below, differentiation is taken with respect to the second variable. Throughout the section it is assumed that (7.1) and (3.1)–(3.4) hold.

Lemma 7.2. *For every $r > 0$, there exists $D(r) > 0$ such that*

$$||\nabla G(y, \alpha) - \nabla G(z, \alpha)|| \leq D(r)||y - z||$$
$$for\ y, z \in \mathbf{R}^d, ||\alpha|| \leq r.$$

Proof. We have, for $y, z, \alpha \in \mathbf{R}^d$.

$$\begin{aligned}\nabla G(y, \alpha) - \nabla G(z, \alpha) = {} & [b(y) - b(z)] + [\sigma(y)(\sigma(y))' - \sigma(z)(\sigma(z))']\alpha \\ & + \int_U (g(y, u)[\exp(\langle g(y, u), \alpha\rangle) - 1] \\ & - g(z, u)[\exp(\langle g(z, u), \alpha\rangle - 1])\nu(du).\end{aligned} \tag{7.2}$$

By assumptions (3.1) and (3.2), there exists a constant $C > 0$ such that for all $y, z, \alpha \in \mathbf{R}^d$,

$$||b(y) - b(z)|| \leq C||y - z||, \quad ||[\sigma(y)(\sigma(y))' - \sigma(z)(\sigma(z))']\alpha|| \leq C||\alpha||\,||y - z||. \tag{7.3}$$

Next, by a simple estimate

$$\begin{aligned}& ||g(y, u)[\exp(\langle g(y, u), \alpha\rangle) - 1] - g(z, u)[\exp(\langle g(z, u), \alpha\rangle) - 1]|| \\ & \leq ||g(y, u) - g(z, u)||[(e^{||\alpha||\bar{g}(u)} - 1) + ||\alpha||\bar{g}(u)e^{||\alpha||\bar{g}(u)}] \\ & \leq ||g(y, u) - g(z, u)|| \cdot 2||\alpha||\bar{g}(u)\exp(||\alpha||\bar{g}(u)),\end{aligned}$$

since $e^s - 1 \leq se^s$ for $s \geq 0$. Now by (3.3) and (3.4),

$$\begin{aligned}& \int_U ||g(y, u)[\exp(\langle g(y, u), \alpha\rangle) - 1] - g(z, u)[\exp(\langle g(z, u), \alpha\rangle - 1]||\nu(du) \\ & \leq \left(\int_U ||g(y, u) - g(z, u)||^2 \nu(du)\right)^{1/2} \left(\int_U (2||\alpha||\bar{g}(u) \exp(||\alpha||\bar{g}(u))^2 \nu(du)\right)^{1/2} \\ & \leq 2C||\alpha|| \left(\int_U (\bar{g}(u))^2 \exp(2||\alpha||\bar{g}(u))\right)^{1/2} ||y - z||.\end{aligned} \tag{7.4}$$

The conclusion follows from (7.2)–(7.4). □

Lemma 7.3. *There exists a function* $\rho : \mathbf{R}^d \times \mathbf{R}^d \to \mathbf{R}^d$ *such that*

(i) *For all* $y \in \mathbf{R}^d$, $\rho(y, \cdot) \in C^1(\mathbf{R}^d, \mathbf{R}^d)$.
(ii) *For all* $y \in \mathbf{R}^d$, $\rho(y, \cdot)$ *is a bijection of* $\mathbf{R}^d$.
(iii) *For all* $y \in \mathbf{R}^d$ *the derivative of* $\rho(y, \cdot)$ *is everywhere an injective linear map.*
(iv) *For all* $y \in \mathbf{R}^d$, $z \in \mathbf{R}^d$,

$$\nabla G(y, \rho(y, z)) = z.$$

(v) *For every* $r > 0$,

$$M(r) \stackrel{\Delta}{=} \sup\{||\rho(y, z)|| : y \in \mathbf{R}^d, ||z|| \le r\} < \infty.$$

(vi) *For every* $r > 0$, *there exists* $L(r) > 0$ *such that for all* $y, y' \in \mathbf{R}^d$, $||z||$, $||z'|| \le r$,

$$||\rho(y, z) - \rho(y', z')|| \le L(r)[||y - y'|| + ||z - z'||].$$

Proof. Since $\varphi(t) = e^t - 1 - t \ge 0$ for all $t \in \mathbf{R}$, we have

$$G(y, \alpha) = \langle b(y), \alpha\rangle + \langle \sigma(y)(\sigma(y))'\alpha, \alpha\rangle + h(y, \alpha), \quad y, \alpha \in \mathbf{R}^d, \tag{7.5}$$

where $h(y, \cdot)$ is a non-negative convex function for each $y \in \mathbf{R}^d$. It follows from (7.1) and (7.5) that $G(y, \cdot)$ is strictly convex, or, equivalently,

$$G(y, \beta) > G(y, \alpha) + \langle \beta - \alpha, \nabla G(y, \alpha)\rangle \quad \textit{for } \beta \ne \alpha. \tag{7.6}$$

For fixed $y \in \mathbf{R}^d$, $\nabla G(y, \cdot) : \mathbf{R}^d \to \mathbf{R}^d$ satisfies

(1) $\nabla G(y, \cdot) \in C^1(\mathbf{R}^d, \mathbf{R}^d)$.
(2) $\nabla G(y, \cdot)$ is a bijection of $\mathbf{R}^d$.
(3) the derivative of $\nabla G(y, \cdot)$ is everywhere an injective linear map.

For, (1) follows easily form (3.4). To prove (2), we observe that by (7.1) and (7.5), for every $z \in \mathbf{R}^d$

$$0 \le G^*(y, z) = \sup_{\alpha \in \mathbf{R}^d} [\langle z, \alpha\rangle - G(y, \alpha)] < \infty$$

and the supremum is attained, say at α_0. Then $z = \nabla G(y, \alpha_0)$. If $\nabla G(y, \alpha_0) = \nabla G(y, \alpha_1)$ with $\alpha_1 \ne \alpha_0$, then by (7.6)

$$\begin{aligned} G(y, \alpha_1) &> G(y, \alpha_0) + \langle \alpha_1 - \alpha_0, \nabla G(y, \alpha_0)\rangle \\ G(y, \alpha_0) &> G(y, \alpha_1) + \langle \alpha_0 - \alpha_1, \nabla G(y, \alpha_0)\rangle \end{aligned}$$

and adding the inequalities we obtain a contradiction. This proves (2).

A simple computation shows that for all $\alpha, \beta \in \mathbf{R}^d$,

$$\langle D^2 G(y, \alpha)\beta, \beta\rangle \ge 2a||\beta||^2, \tag{7.7}$$

establishing (3). We define

$$\rho(y,\cdot) \stackrel{\Delta}{=} (\nabla G(y,\cdot))^{-1}. \tag{7.8}$$

Then by (7.7), the properties (1)–(3) and the inverse function theorem, $\rho(y,\cdot)$ satisfies (i)–(iv). Next, by the mean value theorem (see e.g. [L], p. 103), for any $\alpha, \beta \in \mathbf{R}^d$

$$\nabla G(y,\beta) - \nabla G(y,\alpha) = \left[\int_0^1 D^2 G(y, \alpha + t(\beta - \alpha))dt\right](\beta - \alpha)$$

and therefore by (7.7)

$$\begin{aligned} &||\nabla G(y,\beta) - \nabla G(y,\alpha)||\,||\beta - \alpha|| \\ &\quad \geq \langle \nabla G(y,\beta) - \nabla G(y,\alpha), \beta - \alpha\rangle \\ &\quad \geq \int_0^1 \langle D^2 G(y, \alpha + t(\beta - \alpha))(\beta - \alpha), \beta - \alpha\rangle dt \\ &\quad \geq 2a||\beta - \alpha||^2, \end{aligned}$$

which implies: for all $\alpha, \beta, y \in \mathbf{R}^d$

$$\|\beta - \alpha\| \leq (2a)^{-1}\|\nabla G(y,\beta) - \nabla G(y,\alpha)\|. \tag{7.9}$$

Since $\nabla G(y, 0) = b(y)$, we have $\rho(y, b(y)) = 0$. Therefore for all $y, z \in \mathbf{R}^d$,

$$\begin{aligned} \|\rho(y,z)\| &= \|\rho(y,z) - \rho(y,b(y))\| \\ &\leq (2a)^{-1}\|\nabla G(y,\rho(y,z)) - \nabla G(y,\rho(y,b(y)))\| \\ &\leq (2a)^{-1}\|z - b(y)\| \end{aligned}$$

and (v) follows. Finally, for $y, y' \in \mathbf{R}^d$, $\|z\|, \|z'\| \leq r$, by (7.9), (v) and Lemma 7.2

$$\begin{aligned} \|\rho(y,z) - \rho(y',z')\| &\leq (2a)^{-1}\|\nabla G(y,\rho(y,z)) - \nabla G(y,\rho(y',z'))\| \\ &\leq (2a)^{-1}\left[\|z - z'\| + \|\nabla G(y',\rho(y',z')) - \nabla G(y,\rho(y',z'))\|\right] \\ &\leq (2a)^{-1}\left[\|z - z'\| + D(M(r))\|y - y'\|\right] \\ &\leq L(r)\left[\|z - z'\| + \|y - y'\|\right], \end{aligned}$$

where $L(r) = (2a)^{-1}\max\{1, D(M(r))\}$. □

Lemma 7.4. *For all $f \in D(T, \mathbf{R}^d)$, $\Phi^x(f,\cdot)$ is $D(T, \mathbf{R}^d)$-Gâteaux differentiable and in fact for $\lambda \in M(T, \mathbf{R}^d)$,*

$$\nabla\Phi^x(f,\lambda) = f_\lambda,$$

where

$$f_\lambda(t) = x + \int_0^t \nabla G(f(s), \lambda([s,1]))ds, \quad t \in T.$$

Proof. For $t \in \mathbf{R}$, $\lambda_1 \in M(T, \mathbf{R}^d)$

$$\Phi^x(f, \lambda + t\lambda_1) = \langle x, (\lambda + t\lambda_1)(T)\rangle + \int_T G(f(s), (\lambda + t\lambda_1)([s, 1]))ds.$$

Then

$$\frac{d}{dt}\Phi^x(f, \lambda + t\lambda_1)|_{t=0} = \langle x, \lambda_1(T)\rangle + \int_T \langle \nabla G(f(s), \lambda([s, 1])), \lambda_1([s, 1])\rangle ds.$$

By integration by parts (see, e.g. [deA1], Lemma A.4)

$$\int_T \langle \nabla G(f(s), \lambda([s, 1])), \lambda_1([s, 1])\rangle ds = \int_T \langle \int_0^t \nabla G(f(s), \lambda([s, 1]))ds, d\lambda_1(t)\rangle$$

and therefore for all $\lambda_1 \in M(T, \mathbf{R}^d)$,

$$\begin{aligned}\frac{d}{dt}\Phi^x(f, \lambda + t\lambda_1)|_{t=0} &= \int_T \langle f_\lambda(t), d\lambda_1(t)\rangle \\ &= \langle f_\lambda, \lambda_1\rangle.\end{aligned}$$

□

Lemma 7.5. *Let $h \in D(T, \mathbf{R}^d)$, and assume that h is of bounded variation. Let*

$$f(t) = x + \int_0^t h(s)ds, \quad t \in T.$$

Then there exists $\lambda \in M(T, \mathbf{R}^d)$ such that

$$\nabla G(f(s), \lambda((s, 1])) = h(s), \quad s \in [0, 1)$$

and consequently

$$f = \nabla\Phi^x(f, \lambda).$$

Proof. Let $\rho : \mathbf{R}^d \times \mathbf{R}^d \to \mathbf{R}^d$ be as in Lemma 7.3 and define

$$\psi(s) = \rho(f(s), h(s)) \quad s \in T.$$

Since h is bounded, by Lemma 7.3 there exists $C > 0$ such that

$$\|\psi(s) - \psi(s')\| \le C[\|f(s) - f(s')\| + \|h(s) - h(s')\|] \quad s, s' \in T.$$

It follows that ψ is right-continuous and of bounded variation. For $0 \le a < b \le 1$, define

$$\lambda((a, b]) = \psi_1(b) - \psi_1(a),$$

where $\psi_1(s) = -\psi(s)$ for $s \in [0, 1)$, $\psi_1(1) = 0$. Then λ extends uniquely to an element of $M(T, \mathbf{R}^d)$ (denoted in the same way) and for $s \in [0, 1)$ we have

$$\nabla G(f(s), \lambda((s, 1])) = \nabla G(f(s), \psi(s)) = h(s). \tag{7.10}$$

since $\lambda((s,1]) = \lambda(([s,1]))$ except possibly on a countable subset of T, from (7.10) and Lemma 7.5 it follows that

$$\nabla\Phi^x(f,\lambda) = f. \qquad \square$$

In the next lemma we follow the outline of Lemma 6.5.3 of [D-E], but the proof is self-contained and somewhat simpler on account of (7.1).

Lemma 7.6. *Let $f : T \to \mathbf{R}$ be absolutely continuous and assume*

$$\int_T G^*(f(t), f'(t))dt < \infty.$$

Then for every $\varepsilon > 0$, there exists $h : T \to \mathbf{R}$ such that h is absolutely continuous, $h' \in L^\infty(T,m)$, $h(0) = f(0)$ and

(i) $\|h - f\|_\infty < \varepsilon$,
(ii) $\int G^*(h(t), h'(t))dt \le \int G^*(f(t), f'(t))dt + \varepsilon$.

Proof. For $\ell > 0$, let $E_\ell = \{t \in T : \|f'(t))\| \le \ell\}$. Define

$$\varphi_\ell(t) = \int_0^t (I_{E_\ell}(s) + \frac{\|f'(s)\|}{\ell} I_{E_\ell^c}(s))ds.$$

Then $\varphi_\ell(0) = 0$, $\varphi_\ell(1) \ge 1$ and φ_ℓ is continuous and strictly increasing on T. Let $\psi_\ell \overset{\Delta}{=} \varphi_\ell^{-1} : [0, \varphi_\ell(1)] \to [0,1]$ and we define

$$h_\ell(v) = f(\psi_\ell(v)), \quad v \in T.$$

Let $M = \{v \in T : \psi_\ell$ is differentiable at v and f is differentiable at $\psi_\ell(v)\}$. Then $m(M) = 1$. In fact, it is easily verified that

$$\psi_\ell(t) = \int_0^t [I_{E_\ell}(\psi_\ell(v)) + (\frac{\|f'(\psi_\ell(v))\|}{\ell})^{-1} I_{E_\ell^c}(\psi_\ell(v))]dv, \quad t \in T,$$

and if $M_1 = \{t \in T : f$ is differentiable at $t\}$, then

$$\begin{aligned} m(\{v : f \text{ is differentiable at } \psi_\ell(v)\}) &= m(\psi_\ell^{-1}(M_1)) \\ = \int_T I_{M_1}(\psi_\ell(v))dv &= \int_{[0,\psi_\ell(1)]} I_{M_1}(\psi_\ell(\varphi_\ell(t)))\varphi_\ell'(t)dt \\ &= \int_{[0,\psi_\ell(1)]} \varphi_\ell'(t)dt = \varphi_\ell(\psi_\ell(1)) = 1. \end{aligned}$$

It follows that a.e. $[m]$,

$$\begin{aligned} h_\ell'(v) &= f'(\psi_\ell(v)) \cdot \psi_\ell'(v) \\ &= f'(\psi_\ell(v))[I_{E_\ell}(\psi_\ell(v)) + \left(\frac{\|f'(\psi_\ell(v))\|}{\ell}\right)^{-1} I_{E_\ell^c}(\psi_\ell(v))] \end{aligned}$$

which implies

$$\|h'_\ell(v)\| \le \ell \quad \text{a.e. } [m]. \tag{7.11}$$

Next,

$$\sup_{t\in T}\|h_\ell(t) - f(t)\| = \sup_{t\in T}\|f(\psi_\ell(t)) - f(t)\|.$$

But f is uniformly continuous on T and

$$\begin{aligned}\sup_{t\in T}\|\psi_\ell(t) - t\| &\le \int_0^1 \left[1 - \left(\frac{\|f'(\psi_\ell(v))\|}{\ell}\right)^{-1}\right] I_{E_\ell^c}(\psi_\ell(v))dv \\ &\le \int_{[0,\psi_\ell(1)]} I_{E_\ell^c}(t)\varphi'_\ell(t)dt \\ &\le \frac{1}{\ell}\int_{E_\ell^c}\|f'(t)\|dt \\ &\to 0 \text{ as } \ell \to \infty.\end{aligned}$$

Therefore

$$\|h_\ell - f\|_\infty < \varepsilon \text{ for sufficiently large } \ell. \tag{7.12}$$

Next,

$$\int_T G^*(h_\ell(t), h'_\ell(t))dt = \int_{[0,\psi_\ell(1)]} G^*(h_\ell(\varphi_\ell(s)), h'_\ell(\varphi_\ell(s)))\varphi'_\ell(s)ds. \tag{7.13}$$

Now

$$h_\ell(\varphi_\ell(s)) = f(s),$$

$$h'_\ell(\varphi_\ell(s)) = f'(s)/\varphi'_\ell(s). \tag{7.14}$$

By convexity, since $\varphi'_\ell(s) \ge 1$ a.e. $[m]$,

$$\begin{aligned}G^*(f(s), f'(s)(\varphi'_\ell(s))^{-1}) &\le (1 - (\varphi'_\ell(s))^{-1})G^*(f(s), 0) \\ &\quad + (\varphi'_\ell(s))^{-1}G^*(f(s), f'(s)).\end{aligned} \tag{7.15}$$

It follows easily from (7.1) and (7.5) that

$$\begin{aligned}C = \sup_{y\in\mathbf{R}^d} G^*(y, 0) &= \sup_{y\in\mathbf{R}^d}\sup_{\alpha\in\mathbf{R}^d}[-G(y,\alpha)] \\ &\le \sup_{y\in\mathbf{R}^d}[-\langle b(y), \alpha\rangle - \frac{a}{2}\|\alpha\|^2] \\ &< \infty.\end{aligned} \tag{7.16}$$

By (7.13)–(7.16),

$$\int_T G^*(h_\ell(t), h'_\ell(t))dt \leq C\int_T (\varphi'_\ell(s) - 1)ds + \int_T G^*(f(s), f'(s))ds.$$

But

$$\begin{aligned}\int_T (\varphi'_\ell(s) - 1)ds &= \int_T \left(\frac{\|f'(s)\|}{\ell} - 1\right) I_{E_\ell^c}(s)ds \\ &\leq \frac{1}{\ell}\int_T \|f'(s)\|ds \\ &\to 0 \text{ as } \ell \to \infty.\end{aligned}$$

Therefore

$$\int_T G^*(h_\ell(t), h'_\ell(t))dt \leq \int_T G^*(f(s), f'(s))ds + \varepsilon \tag{7.17}$$

for sufficiently large ℓ. Choosing $h = h_\ell$ for ℓ large enough, h has the desired properties by (7.11), (7.12) and (7.17). □

Proof of Proposition 7.1. Assume that $(\Phi^x)^*(f_0, f_0) < \infty$. By Theorem 6.1, $f_0(0) = x$, f_0 is absolutely continuous and

$$(\Phi^x)^*(f_0, f_0) = \int_T G^*(f_0(s), f'_0(s))ds.$$

By Lemma 7.6, given $\varepsilon > 0$ there exists $h \in D(T, \mathbf{R}^d)$ such that $h(0) = x$, h is absolutely continuous, $r = \|h'\|_{L^\infty(T,m)} < \infty$, $\|h - f_0\|_\infty < \varepsilon/2$ and

$$\int_T G^*(h(s), h'(s))ds \leq \int_T G^*(f_0(s), f'_0(s))ds + \varepsilon/2.$$

Let $M(r)$ be as in Lemma 7.3(v). Choose $\delta > 0$ such that

$$\sup\{|G(y,\alpha) - G(y',\alpha)| : \|y\|, \|y'\| \leq r + \|x\|, \|y - y'\| < \delta, \|\alpha\| \leq M(r)\} < \varepsilon/4.$$

Let $v \in D(T, \mathbf{R}^d)$ be a function of bounded variation such that $\|v\|_\infty \leq r$ and $\|v - h'\|_{L^1(T,m)} < \min\{\varepsilon/4M(r), \varepsilon/2, \delta\}$. Let $f_1(t) = x + \int_0^t v(s)ds$, $t \in T$. Then $\|f_1 - h\|_\infty < \delta$ and by Lemma 7.3,

$$G^*(f_1(s), v(s)) = \langle v(s), \rho(f_1(s), v(s))\rangle - G(f_1(s), \rho(f_1(s), v(s))),$$

so a.e. $[m]$

$$\begin{aligned}G^*(f_1(s), v(s)) \leq{}& \|v(s) - h'(s)\|M(r) + G^*(h(s), h'(s)) \\ &+ [G(h(s), \rho(f_1(s), v(s))) - G(f_1(s), \rho(f_1(s), v(s)))],\end{aligned}$$

which implies

$$\begin{aligned}\int_T G^*(f_1(s), f'_1(s))ds &\leq M(r)\|v - h'\|_{L^1(T,m)} + \int_T G^*(h(s), h'(s))ds + \varepsilon/4 \\ &\leq \int_T G^*(h(s), h'(s))ds + \varepsilon/2.\end{aligned}$$

It follows that $\|f_1 - f_0\|_\infty < \varepsilon$ and

$$\int_T G^*(f_1(s), f_1'(s))ds \le \int_T G^*(f_0(s), f_0'(s)) + \varepsilon.$$

By Lemma 7.5, there exists $\lambda \in M(T, \mathbf{R}^d)$ such that

$$f_1 = \nabla \Phi^x(f_1, \lambda),$$

and by elementary facts from convex analysis, this implies $\lambda \in \partial(\Phi^x)^*(f_1, f_1)$.□

8. Superexponential approximation

Let b, σ, g, ν be as in Section 3, satisfying (3.1)–(3.5). For fixed $x \in \mathbf{R}^d$, in this section we will write $X_n = X_n^x$ to simplify the notation, where X_n^x is as in Section 3. For $a > 0, n \in \mathbf{N}$, let $X_n^{(a)}$ be the strong solution of the Itô stochastic equation

$$X_n^{(a)}(t) = x + \int_0^t b(X_n^{(a)}(s))ds + n^{-1/2}\left(\int_0^t \sigma(X_n^x(s))dB(s) + aW(t)\right) + n^{-1}\int_{[0,t]\times U} g(X_n^{(a)}(s-), u)\tilde{N}_n(ds \times du), \qquad (8.1)$$

where $\{W(t) : t \ge 0\}$ is a standard $\mathbf{R}^d$-valued Brownian motion independent of $\{\{B(t), t \ge 0\}, N_n\}$.

Proposition 8.1. *For every $\delta > 0$,*

$$\lim_{a\downarrow 0} \lim_n \sup \frac{1}{n} \log \mathbf{P}\{\|X_n^{(a)} - X_n\|_\infty > \delta\} = -\infty.$$

For the proof of Proposition 8.1 we need several lemmas.

Lemma 8.2.

(1) $\lim_{r\to\infty} \limsup_n n^{-1} \log \mathbf{P}\{\|X_n\|_\infty > r\} = -\infty.$
(2) $\lim_{r\to\infty} \limsup_n n^{-1} \log \sup_{0<a\le 1} \mathbf{P}\{\|X_n^{(a)}\|_\infty > r\} = -\infty.$

Proof. Proceeding as in the proof of Lemma 4.3, let $\tau_r = \inf\{t \ge 0 : \|X_n(t)\| > r\}$. Then

$$\mathbf{P}\{\|X_n\|_\infty > r\} \le e^{-nr}(2d) \sup_{\alpha\in S} \mathbf{E}\exp[\langle X_n(\tau_r \wedge 1), n\alpha\rangle]. \qquad (8.2)$$

Now

$$\begin{aligned}\mathbf{E}\exp[\langle X_n(\tau_r \wedge 1), n\alpha\rangle] &\le \mathbf{E}\exp[\langle X_n(\tau_r \wedge 1) - x, n\alpha\rangle \\ &\quad - \int_0^{\tau_r\wedge 1} G_n(X_n(s), n\alpha)ds] \\ &\quad \cdot \exp[\sup_{y\in\mathbf{R}^d} |G_n(y, n\alpha)|] \\ &\le \exp(An), \end{aligned} \qquad (8.3)$$

where $A < \infty$, by Proposition 4.2, the inequality previous to (4.4) and assumption (3.4). The first statement follows from (8.2) and (8.3). Statement (2) is proved similarly. □

Lemma 8.3. *Let* $\ell \geq 1$, $f_\ell(y) = \ell \log(1 + \|y\|^2)$, $y \in \mathbf{R}^d$. *Then*

(1) For all $y \in \mathbf{R}^d$, $1 \leq i, j \leq d$

$$\|\nabla f_\ell(y)\| \leq \frac{2\ell\|y\|}{1 + \|y\|^2}, \quad |D_{ij} f(y)| \leq \frac{2\ell}{1 + \|y\|^2}.$$

(2) For all $y, z \in \mathbf{R}^d$

$$0 \leq \exp[f_\ell(y+\|y\|z)-f_\ell(y)]-1-\langle \|y\|z, \nabla f_\ell(y)\rangle \leq 16\ell^2 d\|z\|^2 \exp(2\ell\|z\|).$$

Proof. Simple calculations yield (1). Let $h_\ell(y) = \exp[f_\ell(y)]$, $y \in \mathbf{R}^d$. Then h is convex, and therefore

$$\begin{aligned} &\exp[f_\ell(y + \|y\|z) - f_\ell(y)] - 1 - \langle \|y\|z, \nabla f_\ell(y)\rangle \\ &\quad = \exp[-f_\ell(y)](h_\ell(y + \|y\|z) - h_\ell(y) - \langle \|y\|z, \nabla h_\ell(y)\rangle) \geq 0. \end{aligned} \tag{8.4}$$

Next, for $0 \leq \theta \leq 1$,

$$\begin{aligned} f_\ell(y + \theta\|y\|z) - f_\ell(y) &\leq f_\ell(y(1 + \theta\|z\|)) - f_\ell(y) \\ &= \langle \theta\|z\|y, \nabla f_\ell(y + \theta'\|z\|y)\rangle, \quad \theta' \in (0, \theta) \\ &\leq \|z\|\|y\|\|\nabla f_\ell(y(1 + \theta'\|z\|))\| \\ &\leq \|z\|\|y(1 + \theta'\|z\|)\|\|\nabla f_\ell(y(1 + \theta'\|z\|))\| \\ &\leq 2\ell\|z\|. \end{aligned} \tag{8.5}$$

Suppose $\|z\| \leq 1/2$. Then by (8.4) and Taylor's formula,

$$\begin{aligned} &\exp[f_\ell(y + \|y\|z) - f_\ell(y)] - 1 - \langle \|y\|z, \nabla f_\ell(y)\rangle \\ &= \exp[-f_\ell(y)]\left[\frac{1}{2}\|y\|^2 \sum_{i,j} \exp[f_\ell(\bar{y})](D_{ij} f_\ell(\bar{y}) + D_i f_\ell(\bar{y}) D_j f_\ell(\bar{y}))z_i z_j\right], \end{aligned} \tag{8.6}$$

where $\bar{y} = y + \theta\|y\|z, \theta \in (0, 1)$. By (1)

$$|D_{ij} f_\ell(\bar{y})| \leq \frac{2\ell}{1 + |\|y\| - \|y\|\|z\||^2} \leq \frac{2\ell}{1 + (\|y\|^2/4)},$$

$$|D_i f(\bar{y}) D_j f(\bar{y})| \leq \frac{4\ell^2\|\bar{y}\|^2}{(1 + \|\bar{y}\|^2)^2} \leq \frac{4\ell^2}{1 + (\|y\|^2/4)}.$$

Therefore by (8.5) and (8.6),

$$\begin{aligned}
&\exp[f_\ell(y+\|y\|z)-f_\ell(y)]-1-\langle \|y\|z,\nabla f_\ell(y)\rangle \\
&\quad\le \exp[f_\ell(y+\theta\|y\|z)-f_\ell(y)]\left(\frac{1}{2}\|y\|^2\cdot\frac{8\ell^2}{1+(\|y\|^2/4)}\cdot d\|z\|^2\right)\\
&\quad\le 16\ell^2 d\|z\|^2\exp(2\ell\|z\|).
\end{aligned}$$

On the other hand, if $\|z\|>1/2$, by (1) and (8.5)

$$\begin{aligned}
&\exp[f_\ell(y+\|y\|z)-f_\ell(y)]-1-\langle \|y\|z,\nabla f_\ell(y)\rangle \\
&\quad\le \exp(2\ell\|z\|)-1+2\ell\|z\|\\
&\quad\le 2\exp(2\ell\|z\|)\\
&\quad\le 8\|z\|^2\exp(2\ell\|z\|).
\end{aligned}$$

□

Proof of Proposition 8.1. Let

$$Z_n^{(a)}(t)=a^{-1}(X_n^{(a)}(t)-X_n(t)),\quad t\in T.$$

By (3.6) and (8.1), we have

$$\begin{aligned}
Z_n^{(a)}(t)=&\int_0^t b_n^{(a)}(s)ds+n^{-1/2}\left[\int_0^t\sigma_n^{(a)}(s)dB(s)+W(t)\right]\\
&+n^{-1}\int_{[0,t]\times U}g_n^{(a)}(s,u)\tilde N_n(ds\times du),
\end{aligned}$$

where

$$b_n^{(a)}(s)=a^{-1}[b(X_n^{(a)}(s))-b(X_n(s))],\quad \sigma_n^{(a)}(s)=a^{-1}[\sigma(X_n^{(a)}(s))-\sigma(X_n(s))],$$

$$g_n^{(a)}(s,u)=a^{-1}[g(X_n^{(a)}(s),u)-g(X_n(s),u)].$$

For $n\in\mathbf{N}$, $a>0$, $v,w\in\mathbf{R}^d$, $u\in U$, set $y=a^{-1}(v-w)$,

$$\begin{aligned}
I(n,a,v,w,u)=&\exp[f_\ell(y+\frac{1}{na}(g(v,u)-g(w,u)))-f_\ell(y)]\\
&-1-\langle\frac{1}{na}(g(v,u)-g(w,u)),\nabla f_\ell(y)\rangle,
\end{aligned}$$

$$\begin{aligned}
h_n^{(a)}(v,w)=&\langle a^{-1}(b(v)-b(w)),\nabla f_\ell(y)\rangle\\
&+\frac{1}{2n}\sum_{i,j=1}^d(a^{-2}[(\sigma(v)-\sigma(w))(\sigma(v)-\sigma(w))']_{ij}+\delta_{ij})\\
&\times(D_{ij}f_\ell(y)+D_if_\ell(y)D_jf_\ell(y))\\
&+\int_U I(n,a,v,w,u)(n\nu)du,
\end{aligned}$$

where f_ℓ is as in Lemma 8.3. Arguing as in Proposition 4.1, we have: the integral in the definition of $h_n^{(a)}(v, w)$ exists and is a bounded continuous function of (v, w). Again following the proof of Proposition 4.1, we obtain:

$$M_n^{(a)}(t) \stackrel{\Delta}{=} \exp[f_\ell(Z_n^{(a)}(t)) - \int_0^t h_n^{(a)}(X_n^{(a)}(s), X_n(s))ds], t \geq 0$$

is an $\{\mathscr{F}_t\}$-positive local martingale, hence an $\{\mathscr{F}_t\}$ -supermartingale.

Let $\delta > 0$, $\tau_a = \inf\{t \geq 0 : \|Z_n^{(a)}(t)\| > \delta/a\}$. Then for $a > 0, r > 0$,

$$\mathbf{P}\{\|X_n^{(a)} - X_n\|_\infty > \delta\} \leq \mathbf{P}\{\|X_n^{(a)}\|_\infty > r\} + \mathbf{P}\{\|X_n\|_\infty > r\} \\ +\mathbf{P}\{\|X_n^{(a)}\|_\infty \leq r, \|X_n\|_\infty \leq r, \|Z_n^{(a)}(\tau_a \wedge 1)\| > \delta/a\}. \tag{8.7}$$

Let $\varphi_\ell(p) = \ell \log(1 + p^2)$, $p \in \mathbf{R}$. Then

$$\begin{aligned} &\mathbf{P}\{\|X_n^{(a)}\|_\infty \leq r, \|X_n\|_\infty \leq r, \|Z_n^{(a)}(\tau_a \wedge 1)\| > \delta/a\} \\ &\quad \leq \exp[-\varphi_\ell(\delta/a)]\mathbf{E}(\exp[f_\ell(Z_n^{(a)}(\tau_a \wedge 1))] \cdot I(\|X_n^{(a)}\|_\infty \leq r, \|X_n\|_\infty \leq r)) \\ &\quad = \exp[-\varphi_\ell(\delta/a)]\mathbf{E}(M_n^{(a)}(\tau_a \wedge 1) \exp\left[\int_0^{\tau_a \wedge 1} h_n^{(a)}(X_n^{(a)}(s), X_n(s))ds\right] \\ &\quad\quad \times I(\|X_n^{(a)}\|_\infty \leq r, \|X_n\|_\infty \leq r)). \end{aligned} \tag{8.8}$$

Take now $\ell = n\gamma/2 \geq 1$, where $\gamma = \gamma(r) > 0$ is as in (3.5). By (3.1) and Lemma 8.3(1), for $v, w \in \mathbf{R}^d$, $a > 0$

$$\begin{aligned} |\langle a^{-1}(b(v) - b(w)), \nabla f_\ell(y)\rangle| &\leq C\|a^{-1}(v - w)\|\|\nabla f_\ell(a^{-1}(v - w))\| \\ &\leq 2\ell C = (\gamma C)n. \end{aligned} \tag{8.9}$$

By (3.2) and Lemma 8.3(1), for $v, w \in \mathbf{R}^d$, $a > 0$

$$\begin{aligned} &\frac{1}{2n}\sum_{i,j=1}^d (a^{-2}[(\sigma(v) - \sigma(w))(\sigma(v) - \sigma(w))']_{ij} + \delta_{ij}) \\ &\quad \times(D_{ij} f_\ell(y) + D_i f_\ell(y) D_j f_\ell(y)) \\ &\leq \frac{C}{2n}\sum_{i,j=1}^d \|a^{-1}(v - w)\|^2(|D_{ij} f_\ell(a^{-1}(v - w))| \\ &\quad + |D_i f_\ell(a^{-1}(v - w)) D_j f_\ell(a^{-1}(v - w))|) \\ &\leq \frac{C}{n} \cdot 8\ell^2 \cdot d = (2\gamma^2 Cd)n. \end{aligned} \tag{8.10}$$

For $v \neq w$, letting $z = (n\|v - w\|)^{-1}(g(v, u) - g(w, u))$, we have

$$I(n, a, v, w, u) = \exp[f_\ell(y + \|y\|z) - f_\ell(y)] - 1 - \langle\|y\|z, \nabla f_\ell(y)\rangle$$

and by Lemma 8.3(2), with q as in (3.5),

$$0 \leq I(n, a, v, w, u) \leq 16\ell^2 dn^{-2}(q(v, w, u))^2 \exp[2\ell n^{-1} q(v, w, u)]. \quad (8.11)$$

Then by (3.5) and (8.11),

$$C(r) = \sup\{\int_U I(n, a, v, w, u)\nu(du) : n \in \mathbf{N}, a > 0, \|v\| \leq r, \|w\| \leq r\} < \infty. \quad (8.12)$$

By (8.9), (8.10) and (8.12): for all $n \geq 2/\gamma$,

$$\sup\{|h_n^{(a)}(v, w)| : a > 0, \|v\| \leq r, \|w\| \leq r\} \leq C'(r)n, \quad (8.13)$$

where $C'(r)$ is a positive constant depending only on r. Now by (8.8) and (8.13), and taking into account that $M_n^{(a)}(0) \equiv 1$, we have

$$\mathbf{P}\{\|X_n^{(a)}\|_\infty \leq r, \|X_n\|_\infty \leq r, \|Z_n^{(a)}(\tau_a \wedge 1)\| > \delta/a\} \leq \exp[-\varphi_\ell(\delta/a)] \cdot C'(r)n. \quad (8.14)$$

By (8.7) and (8.14), for all $\delta > 0, a > 0, r > 0$,

$$\begin{aligned}
&\limsup_n \frac{1}{n} \log \mathbf{P}\{\|X_n^{(a)} - X_n\|_\infty > \delta\} \\
&\leq \max\{-\frac{\gamma(r)}{2} \log(1 + (\delta/a)^2), \limsup_n \frac{1}{n} \log \mathbf{P}\{\|X_n\|_\infty > r\}, \\
&\qquad \limsup_n \frac{1}{n} \log \sup_{0<p\leq 1} \mathbf{P}\{\|X_n^{(p)}\|_\infty > r\}\}.
\end{aligned}$$

The result follows now by first letting $a \to 0$ and then $r \to \infty$, using Lemma 8.2. □

9. Proof of Theorem 3.1

(1) **Upper bounds.** We apply Theroem 2.1 with $a_n = n$, $E = D(T, R^d)$ endowed with $\|\cdot\|_\infty$, $\mathscr{E} = \mathscr{D}$, $F = M(T, \mathbf{R}^d)$, $Y_n = X_n^x$, $Z_n = Z_n^x$ (defined in Section 5), $\Phi_n = \Phi_n^x$, $\Phi = \Phi^x$.

The properties of $\mathscr{E}$ are easily verified, taking $\mathscr{V}$ to be the set of all open balls with center at 0. The $\mathscr{E}$-measurability of $\langle \cdot, \lambda \rangle$ for $\lambda \in M(T, \mathbf{R}^d)$ is proved in [de A1], Corollary A.3.

Condition (3) of Theorem 2.1 is trivially verified, since for all $n \in \mathbf{N}$ by (5.5)

$$\frac{1}{n}\Phi_n^x(f, n\lambda) = \Phi^x(f, \lambda) \quad f \in E, \lambda \in F.$$

Condition (4) is proved in Proposition 4.4. Condition (5) is proved in Proposition 5.4 and condition (6) in Proposition 5.3. Condition (7) is proved in Proposition 6.2. Applying Theorem 2.1 and taking into account Theorem 6.1, the upper bound in Theorem 3.1 follows.

(2) **Lower bounds.** For fixed $x \in \mathbf{R}^d$, $a > 0$, let $X_n^{(a)}$ be the strong solution of equation (8.1). Then by Proposition 4.4, for all $n \in \mathbf{N}$, $\lambda \in M(T, \mathbf{R}^d)$,

$$\mathbf{E}\exp[\langle X_n^{(a)}, \lambda\rangle - \Phi_n^{(a)}(X_n^{(a)}, \lambda)] = 1,$$

where

$$\Phi_n^{(a)}(f, \lambda) = \langle x, \lambda(T)\rangle + \int_T G_n^{(a)}(f(s), \lambda([s, 1]))ds$$

and

$$G_n^{(a)}(y, \alpha) = \langle b(y), \alpha\rangle + \frac{1}{2n}\langle(\sigma(y)(\sigma(y))' + a^2 I)\alpha, \alpha\rangle + \int_U \varphi(\langle n^{-1}g(y, u), \alpha\rangle)(n\nu)(du).$$

Let

$$\Phi^{(a)}(f, \lambda) = \langle x, \lambda(T)\rangle + \int_T G^{(a)}(f(s), \lambda([s, 1]))ds,$$

where

$$G^{(a)}(y, \alpha) = \langle b(y), \alpha\rangle + \frac{1}{2}\langle(\sigma(y)(\sigma(y))' + a^2 I)\alpha, \alpha\rangle + \int_U \varphi(\langle g(y, u), \alpha\rangle)\nu(du).$$

Then the convexity of $\Phi^{(a)}(f, \cdot)$ follows from that of $G^{(a)}(y, \cdot)$ and the E-Gâteaux differentiability of $\Phi^{(a)}(f, \cdot)$ is proved in Lemma 7.4, so condition (8) holds. Condition (9) is easily verified using (3.4). Conditions (3)–(7) are verified as in the proof of the upper bound. To verify condition (10): Let $\lambda \in M(T, \mathbf{R^d}$, and assume that $f = \nabla\Phi^{(a)}(f, \lambda)$, $h = \nabla\Phi^{(a)}(h, \lambda)$. By Lemmas 7.2 and 7.4, for $t \in T$

$$\|f(t) - h(t)\| = \|\int_0^t \Big[\nabla G^{(a)}(f(s), \lambda([s, 1]) - \nabla G^{(a)}(h(s), \lambda([s, 1]))\Big] ds\| \leq D(r)\int_0^t \|f(s) - h(s)\|ds,$$

where $r = \sup\{\|f\|_\infty, \|h\|_\infty, \|\lambda\|_v\}$ Now Gronwall's lemma implies $\|f(t) - h(t)\| = 0$ for all $t \in T$, that is, $f = h$. Condition (11) is proved in Proposition (7.1). Applying Theorem 2.2, the lower bound holds for $\{\mathscr{L}_{\mathbf{P}}(X_n^{(a)})\}$ with rate function $(\Phi^{(a)})^*(f, f)$. For all $y, \alpha \in \mathbf{R}^d$,

$$G^{(a)}(y, \alpha) \downarrow G(y, \alpha)$$

and it easily follows that for all $f \in E$, $\lambda \in F$,

$$\Phi^{(a)}(f, \lambda) \downarrow \Phi^*(f, \lambda). \tag{9.1}$$

Let $A \in \mathscr{D}$, $f_0 \in A^0$, $\delta > 0$ such that $B_\delta(f_0) \subset A$. Then

$$\mathbf{P}\{X_n^{(a)} \in B_{\delta/2}(f_0)\} \leq \mathbf{P}\{X_n^x \in A\} + \mathbf{P}\{\|X_n^{(a)} - X_n^x\|_\infty \geq \delta/2\}$$

and by the previous discussion,

$$\max\left(\liminf_n \frac{1}{n} \log \mathbf{P}\{X_n^x \in A\}, \limsup_n \frac{1}{n} \log \mathbf{P}\{\|X_n^{(a)} - X_n^x\|_\infty \geq \delta/2\}\right) \geq -(\Phi^{(a)})^*(f_0, f_0). \tag{9.2}$$

By (9.1), for all $f \in E$

$$\begin{aligned} \lim_{a\downarrow 0}(\Phi^{(a)})^*(f, f) &= \sup_{a>0}(\Phi^{(a)})^*(f, f) = \sup_{a>0}\sup_{\lambda\in F}[\langle f, \lambda\rangle - \Phi^{(a)}(f, \lambda)] \\ &= \sup_{\lambda\in F}\sup_{a>0}[\langle f, \lambda\rangle - \Phi^{(a)}(f, \lambda)] \\ &= \sup_{\lambda\in F}[\langle f, \lambda\rangle - \Phi^x(f, \lambda)] \\ &= (\Phi^x)^*(f, f). \end{aligned} \tag{9.3}$$

Letting $a \downarrow 0$ in (9.2) and using Proposition 8.1 and (9.3), we obtain

$$\liminf_n \frac{1}{n} \log \mathbf{P}\{X_n^x \in A\} \geq -(\Phi^x)^*(f_0, f_0).$$

This completes the proof of the lower bound and hence of Theorem 3.1. □

In the next result we show the methods of this paper apply to a somewhat more general class of Markov processes. The upper bound is obtained under rather mild conditions; for the lower bound, we require Lipschitz and non-degeneracy assumptions. The latter assumption – condition (4) of Theorem 9.1 – is needed in order to carry through the arguments in Section 7, proving an analogue of Proposition 7.1 and thereby verifying condition (11) of Theorem 2.2. Since it is not clear to us how to appoximate superexponentially a Markov process satisfying (1)–(3), (5), (6), and (9.4) of Theorem 9.1 by Markov processes satisfying those assumptions and (4), the statement of Theorem 9.1 contains (4) as a hypothesis. In the case in which the processes are solutions of stochastic equations, we have shown in Section 8 that such a superexponential approximation is possible.

For each $n \in \mathbf{N}$, $x \in \mathbf{R}^d$, let P_n^x be a probability measure on $D([0, \infty), \mathbf{R}^d)$ endowed with the σ-algebra $\mathscr{D}([0, \infty), \mathbf{R}^d)$ generated by the evaluations. Let $\{X(t) : t \geq 0\}$, $\{\mathscr{G}_t : t \geq 0\}$ be as in Section 4. We assume that for each $n \in \mathbf{N}$,

$$(D([0, \infty), \mathbf{R}^d), \{X(t) : t \geq 0\}, \{\mathscr{G}_t : t \geq 0\}, \{P_n^x : x \in \mathbf{R}^d\})$$

is a Markov process.

Theorem 9.1. *Let $G_n, G : \mathbf{R}^d \times \mathbf{R}^d \to \mathbf{R}$. We consider the following conditions:*

(1) $G_n(\cdot, \alpha)$ is measurable for all $\alpha \in \mathbf{R}^d$.
(2) For all $r > 0$, $\lim_n \sup_{\|\alpha\|\leq r} \sup_{y\in\mathbf{R}^d} |\frac{1}{n} G_n(y, n\alpha) - G(y, \alpha)| = 0$.

(3) $G(y,0)=0$ *for all* $y\in\mathbf{R}^d$, $G(\cdot,\alpha)$ *is continuous for all* $\alpha\in\mathbf{R}^d$ *and for all* $r>0$*,*

$$\sup_{\|\alpha\|\le r}\sup_{y\in\mathbf{R}^d}|G(y,\alpha)|<\infty.$$

(4) For all $y\in\mathbf{R}^d$, $G(y,\cdot)\in C^2(\mathbf{R}^d)$ *and there exists* $c>0$ *such that for all* $\beta\in\mathbf{R}^d$

$$\inf\{\langle D^2G(y,\alpha)\beta,\beta\rangle : y\in\mathbf{R}^d,\alpha\in\mathbf{R}^d\}\ge c\|\beta\|^2.$$

(5) For all $r>0$*, there exists* $D(r)>0$ *such that for* $\|\alpha\|\le r$, $y,z\in\mathbf{R}^d$ *,*

$$\|\nabla G(y,\alpha)-\nabla G(z,\alpha)\|\le D(r)\|y-z\|.$$

(6) $\sup\{\|\nabla G(y,0)\| : y\in\mathbf{R}^d\}<\infty.$

Let $Q_n^x=\mathscr{L}_{P_n^x}(\{X(t):t\in T\})$*. Suppose that* G_n, G *satisfy assumptions (1)–(6) and, furthermore, for all* $x,\alpha\in\mathbf{R}^d$, $n\in\mathbf{N}$*,*

$$\exp[\langle X(t)-x,\alpha\rangle-\int_0^t G_n(X(s),\alpha)ds],\quad t\ge 0 \tag{9.4}$$

is $\{\mathscr{G}_t\}$*-martingale under* P_n^x*. Then* $\{Q_n^x\}$ *satisfies the large deviation principle on* $D(T,\mathbf{R}^d)$*, endowed with the uniform norm* $\|\cdot\|_\infty$ *and the* σ*-algebra* $\mathscr{D}$*, with the rate function*

$$I^x(f)=\begin{cases}\int_T G^*(f(s),f'(s))ds & \text{if } f(0)=x \text{ and}\\ & \quad f \text{ is absolutely continuous}\\ \infty & \text{otherwise.}\end{cases}$$

More specifically, the upper bound holds under assumptions (1)–(3) and (9.4), and the lower bound under assumptions (1)–(6) and (9.4).

We omit the proof; it is not difficult to carry it out by arguing as in the proof of the upper bound and the first part of the proof of the lower bound in Theorem 3.1 and retracing the relevant items in Sections 4–7.

Note added in proof. We very recently relaxed Condition (11) of Theorem 2.2 to a continuity condition not involving subdifferentials. This result will appear elsewhere.

Acknowledgements. It is a pleasure to thank Tom Kurtz for some valuable insights on Markov processes and for a very useful technical suggestion related to Section 8.

References

[deA1] de Acosta, A.: Large deviations for vector valued Lévy processes. Stoch. Proc. and Appl., **51**, 75–115 (1994)

[deA2] de Acosta, A.: Projective systems in large deviation theory II: some applications. In: Probability in Banach spaces 9, Progress in Probability, **35**, 241–250. Birkhäuser, Boston, MA, (1994)

[deA3] de Acosta, A.: Exponential tightness and projective systems in large deviation theory. In: Festschrift for Lucien Le Cam, 143–156. Springer-Verlag, New York, (1997)

[A-E] Aubin, J.P., Ekeland, I.: Applied nonlinear Analysis. Wiley, New York, (1984)

[Ba] Baldi, P.: Large deviations and stochastic homogenization. Ann. Mat. Pure Appl., **151**, 161–177 (1988)

[Br] Bryc, W.: Large deviations by the asymptotic value method. In: Diffusion Processes and Related Problems in Analysis, I, 447–472. Birkhäuser, Boston, (1990)

[C-J] Cinlar, E., Jacod, J.: Representation of semimartingale Markov processes in terms of Wiener Processes and Poisson random measures. In: Seminar on Stochastic processes 1981, 159–242. Birkhäuser, Boston (1981)

[Da-G] Dawson, D., Gärtner, J.: Large deviations from the McKean-Vlasov limit for interacting diffusions. Stochastics, **20**, 247–308 (1987)

[De-Z] Dembo, A., Zeitouni, O.: Large deviation techniques and applications. (2nd Ed.) Springer-Verlag, New York, (1998)

[Du-E] Dupuis, P., Ellis, R.S.: A weak convergence approach to the theory of large deviations. Wiley, New York, (1997)

[Ek-T] Ekeland, I., Temam, R.: Convex Analysis and variational problems. North Holland, Amsterdam, (1976)

[El] Ellis, R.S.: Large deviations for a general class of dependent random vectors. Ann. Probab., **12**, 1–12 (1984)

[Et-K] Ethier, S.N., Kurtz, T.G.: Markov processes: Characterization and convergence. Wiley, New York, (1986)

[F-K] Feng, J., Kurtz, T.G.: Exponential tightness, exponential martingale problems and nonlinear semigroup convergence in large deviations for stochastic processes. Preprint, (1999)

[Ga] Gärtner, J.: On large deviations from the invariant measure. Theory Probab. Appl., **22**, 24–39 (1977)

[Gr] Graham, C.: McKean-Vlasov Ito-Skorohod equations and non-linear diffusions with discrete jump sets. Stoch. Proc. Appl., **40**, 69–82 (1992)

[I-W] Ikeda, N., Watanabe, S.: Stochastic differential equations and diffusion processes. North-Holland, Amsterdam, (1981)

[J-S] Jacod, J., Shiryaev, A.N.: Limit theorems for stochastic processes. Springer-Verlag, New York, (1987)

[L] Lang, S.: Analysis II. Addison-Wesley, Reading, Mass., (1969)

[L-P] Liptser, R.S., Pukhalskii, A.A.: Limit theorems on large deviations for semimartingales. Stochastics and Stoch. Rep., **38**, 201–249 (1992)

[OB-S] O'Brien, G.L., Sun, J.: Large deviations in Linear Spaces. Probab. Math. Stat., **16**, 261–273 (1996)

[W] Wentzell, A.D.: Limit theorems on Large Deviations for Markov Stochastic Processes. Kluwer, Dordrecht. (1991)

[W-F] Wentzell, A.D., Freidlin, M.I.: On small random perturbations of dynamical systems. Russian Math. Surveys, **25**, 1–55 (1970)