The method of stochastic exponentials for large deviations

A. Puhalskii

Institute for Problems in Information Transmission, Ernolovoy ul., 19, Moscow 101447, Russian Federation

Received 18 September 1991; revised 10 January 1994

Abstract

We present a method for proving the large-deviation principle for processes with paths in the Skorohod space which is analogous to the method of stochastic exponentials in weak convergence. It is applied to derive new results on large deviations for semimartingales as well as for processes with independent increments.

Keywords: Large deviations; Exponential tightness; Semimartingales; Cramér condition

1. Introduction

In Puhalskii (1991) we introduced a new approach to obtaining the large deviation principle (LDP) (Varadhan, 1984). It is based on a counterpart to the Prohorov theorem on the equivalence of weak relative compactness and tightness for a family of probability measures: every subsequence of a sequence of probability measures on a Polish space contains a subsubsequence obeying the LDP whenever a tightness condition (known as exponential tightness (Deuschel and Stroock, 1989) holds (Puhalskii, 1991). This result applied to the study of large deviations of stochastic processes with paths in the Skorohod space, enabled us to work out an analogue of the method of finite-dimensional distributions in weak convergence, i.e. one can, in certain cases, prove the LDP for processes by verifying ‘finite-dimensional LDPs’ and checking exponential tightness.

In this paper, we are primarily concerned with the study of large deviations of semimartingales. Aiming at developing large-deviation theory parallel to the theory of weak convergence, we use the method of finite-dimensional distributions to obtain an analogue of the method of stochastic exponentials (Jacod and Shiryaev, 1987; Liptser and Shiryaev, 1989) which proved to be fruitful in deriving weak convergence results for semimartingales.

More specifically, our main result states that if for a sequence of \( \mathbb{R}^d \)-valued càdlàg processes \( (X^n, n \geq 1), X^n = (X^n_t)_{t \geq 0}, \) defined each on a stochastic basis \( (\Omega, \mathcal{F}, F^n = (\mathcal{F}^n_t)_{t \geq 0}, P) \), we have for each \( \lambda \in \mathbb{R}^d \) the representation

\[
e^{(\lambda_t X^n_t - X^n_0)} = \epsilon^n_t(\lambda) Y^n_t(\lambda)
\]
\((\langle \lambda, x \rangle)\) stands for the scalar product of \(\lambda, x \in \mathbb{R}^d\), where \(\mathcal{E}_n^\mu(\lambda) = (\mathcal{E}_t^\mu(\lambda))_{t \geq 0}, \mathcal{E}_0^\mu(\lambda) = 1\), is a (real-valued, càdlàg) \(\mathcal{F}^n\)-predictable positive process, and \(Y_n(\lambda) = (Y_t^n(\lambda))_{t \geq 0}, Y_0^n(\lambda) = 1\), is a (real-valued, càdlàg) \(\mathcal{F}^n\)-local martingale. Then the LDP for \((X^n)\) is implied by certain convergence of \(\mathcal{E}^n(\lambda)\) (Theorem 2.1).

As in the method of stochastic exponentials for weak convergence which requires that the limiting process be a process with independent (or conditionally independent) increments (Jacod and Shiryaev, 1987; Liptser and Shiryaev, 1989), the limiting rate function in our case is of a particular form which may be thought of as an analogue of the distribution of a process with independent increments.

When \(X^n\) are semimartingales, the above gives conditions for the LDP in terms of the Doléans-Dade exponentials of appropriate cumulants, again in analogy with weak convergence (Jacod and Shiryaev, 1987; Liptser and Shiryaev, 1989). It is to this result that the name 'method of stochastic exponentials' is due.

Another (in fact, closely related) application is to processes with independent increments. Here we are able to extend earlier results of Borovkov (1967) and (partially) Mogulskii (1976), who studied homogeneous processes with independent increments.

We state our results in Section 2 and prove them in Sections 4 and 5. In Section 2 we recall also the main points of the method of finite-dimensional distributions from Puhalskii (1991). Section 3 contains the necessary background for the proofs. The final section dwells on the similarities between weak convergence and large deviations of stochastic processes. We assume that the reader is familiar with standard definitions and facts from large-deviation theory (Varadhan, 1984; Stroock, 1984; Deuschel and Stroock, 1989) (though some of them are recalled below) and from the theory of martingales (Jacod and Shiryaev, 1987; Liptser and Shiryaev, 1989).

### 2. Main results

Let \((X^n, n \geq 1), X^n = (X^n_t)_{t \geq 0}\), be a sequence of \(\mathbb{R}^d\)-valued stochastic processes with paths in the Skorohod space \(D(\mathbb{R}^d) = D(\mathbb{R}_+, \mathbb{R}^d)\) of all \(\mathbb{R}^d\)-valued càdlàg (i.e. right-continuous with left-hand limits) functions on \(\mathbb{R}_+\). Each \(X^n\) is defined on a stochastic basis \((\Omega, \mathcal{F}, \mathcal{F}^n = (\mathcal{F}^n_t)_{t \geq 0}, P)\). We assume that \(D(\mathbb{R}^d)\) is supplied with the Skorohod–Lindvall metric (see, e.g., Jacod and Shiryaev, 1987; Liptser and Shiryaev, 1989) and is thus a Polish space. Elements of \(D(\mathbb{R}^d)\) are denoted \(X = (X_t)_{t \geq 0}\). Also \(C(\mathbb{R}^d)\) is the subspace of \(D(\mathbb{R}^d)\) consisting of continuous functions and is submitted with induced, i.e. locally uniform, topology. All the processes, which we consider, are càdlàg and we do not mention this in the sequel.

Let \(I: D(\mathbb{R}^d) \to [0, \infty]\) be a rate function on \(D(\mathbb{R}^d)\) (Varadhan, 1984; Stroock, 1984) (or a good rate function in the terminology of Deuschel and Stroock (1989)), i.e. the set \(\{X \in D(\mathbb{R}^d): I(X) \leq a\}\) is compact for every \(a \geq 0\). Introduce the set function

\[
V(A) = \sup_{x \in A} \exp(-I(X)), \quad A \subset D.
\]
Symbolically, we will write $V = \exp(-I)$, which we call *deviability* (Puhalskii, 1994) is seen to be an analogue of probability (Puhalskii, 1993, 1994), and the following definition from Puhalskii (1994) seems to be convenient.

Let $\mathcal{L}(X^n)$ denote the law of $X^n$. Say that the sequence $(\mathcal{L}(X^n), n \geq 1)$ large deviation (or LD) converges to $\exp(-I)$ as $n \to \infty$ and write $\mathcal{L}(X^n) \overset{1.d.}{\to} \exp(-I)$ ($n \to \infty$) if $\mathcal{L}(X^n)$ obeys the LDP (in $D(\mathbb{R}^d)$) with $I$ (Varadhan, 1984; Stroock, 1984) or the full LDP as in Deuschel and Stroock (1989)). Large-deviation convergence is similarly understood in other metric spaces (primarily $\mathbb{R}^n$).

We call $\exp(-I)$ a large deviation (LD) accumulation point of $(\mathcal{L}(X^n))$ if $\mathcal{L}(X^n) \overset{1.d.}{\to} \exp(-I)(n' \to \infty)$ for a subsequence $(n')$.

Say that $(\mathcal{L}(X^n))$ is $C$ exponentially tight if it is exponentially tight (i.e. for any $\varepsilon > 0$ there exists a compact $K \subset D(\mathbb{R}^d)$ such that $[P(X^n \in D(\mathbb{R}^d) \setminus K)]^{1/n} < \varepsilon$ for all $n$) and any LD accumulation point $\exp(-I)$ of $(\mathcal{L}(X^n))$ (which exists by Puhalskii (1991)) satisfies $I(X) = \infty \forall X \in D(\mathbb{R}^d) \setminus C(\mathbb{R}^d)$.

In the paper we will be using the following version of the method of finite-dimensional distributions.

**Theorem A.** Let $(\mathcal{L}(X^n), n \geq 1)$ be $C$-exponentially tight. Assume that for all $k = 1, 2, \ldots$ and $t_1 < \cdots < t_k \in U$, a dense subset of $\mathbb{R}_+$, we have (in $D(\mathbb{R}^d)^k$)

$$(\mathcal{L}(X_{t_1}^n, \ldots, X_{t_k}^n)) \overset{1.d.}{\to} \exp(-I_{t_1}, \ldots, t_k), \quad (n \to \infty),$$

where $I_{t_1}, \ldots, t_k$ are rate functions on $(\mathbb{R}^d)^k$.

Then $I(X)$, $X = (X_t)_{t \geq 0} \in D(\mathbb{R}^d)$, defined by

$$I(X) = \sup_{t_1, \ldots, t_k \in U} I_{t_1, \ldots, t_k}(X_{t_1}, \ldots, X_{t_k}),$$

is a rate function on $D(\mathbb{R}^d)$, and

$$(\mathcal{L}(X^n) \overset{1.d.}{\to} \exp(-I)) \quad (n \to \infty).$$

To prove $C$-exponential tightness, an analogue of the Aldous condition for tightness (Aldous, 1978) will be used.

Denote $T_L(F^n)$ for $L > 0$ the set of all $F^n$-stopping times $\tau \leq L$ and let $|x|$, for $x \in \mathbb{R}^d$, denote the Euclidean norm.

**Theorem B.** If

(i) $\lim_{A \to \infty} \limsup_{n \to \infty} \left[ P \left( \sup_{t \leq L} |X^n_t| > A \right) \right]^{1/n} = 0 \quad \forall L > 0,$

(ii) $\lim_{\delta \to 0} \limsup_{n \to \infty} \sup_{\tau \in T_L(F^n)} \left[ P \left( \sup_{t \geq \delta} |X^n_{\tau + t} - X^n_{\tau}| > \eta \right) \right]^{1/n} = 0, \quad \forall L > 0, \eta > 0,$

then $(\mathcal{L}(X^n))$ is $C$-exponentially tight.
Theorem A is a particular case of Theorem 4.5 in Puhalskii (1991) and Theorem B is a direct consequence of Theorems 4.4 and 4.6 there; see also Liptser and Pukhalskii (1992) (though those are proved in Puhalskii (1991) for $d = 1$, the general case does not differ).

Now we introduce the rate function on $D(\mathbb{R}^d)$ which will be the limiting one in our results. Let $G_t(\lambda), t \geq 0, \lambda \in \mathbb{R}^d$, be a real-valued function, continuous in $t$ for each $\lambda \in \mathbb{R}^d$ and $G_0(\lambda) = 0, G_t(0) = 0$. Denote $A_0$ the set of all $\mathbb{R}^d$-valued piecewise-constant functions $(\lambda(t), t \geq 0)$ of the form

$$
\lambda(t) = \lambda_0 I_{[0]}(t) + \sum_{i=1}^{k} \lambda_i I_{[t_{i-1}, t_i]}(t),
$$

where $\lambda_i \in \mathbb{R}^d, 0 \leq i \leq k, 0 = t_0 < t_1 < \cdots < t_k$, and $I_A$ is the indicator of a set $A$ (below we also use the notation $1(A)$).

For $X = (X_t), t \geq 0 \in D(\mathbb{R}^d)$ and $(\lambda(t), t \geq 0) \in A_0$, we set by definition

$$
\mathcal{I}(\lambda(t)) = \lambda_0 X_0 + \sum_{i=1}^{k} \left[ \langle \lambda_i, X_{t_i} - X_{t_{i-1}} \rangle - (G_t(\lambda_i) - G_{t_i-1}(\lambda_i)) \right]
$$

and define

$$
I(X) = \sup_{(\lambda(t)), t \geq 0 \in A_0} \mathcal{I}(\lambda(t)), X \in D(\mathbb{R}^d).
$$

**Lemma 2.1.** $I$ is a rate function on $D(\mathbb{R}^d)$.

The following particular case is encountered quite often ($\mathcal{B}(\mathbb{R}_+)$ and $\mathcal{B}(\mathbb{R}^d)$ denote the Borel $\sigma$-fields on $\mathbb{R}_+$ and $\mathbb{R}^d$, respectively).

**Lemma 2.2.** Assume that

$$
G_t(\lambda) = \int_0^t g_s(\lambda) q(ds), \quad t \geq 0, \lambda \in \mathbb{R}^d,
$$

where $q(dt)$ is a continuous locally finite measure on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ ($g(t) = 0$, $q([0, t]) < \infty, t \geq 0$), and $g_t(\lambda)$ with $g_t(0) = 0$ is real-valued, $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}^d)$-measurable and continuous in $\lambda$ for each $t \geq 0$, and for all $A > 0, t > 0$,

$$
\int_0^t \sup_{|\lambda| \leq A} |g_s(\lambda)| q(ds) < \infty.
$$

Then

$$
I(X) = \left\{ \begin{array}{ll}
\int_0^\infty \sup_{\lambda \in \mathbb{R}^d} \left( \left\langle \lambda, \frac{dX}{dq} (t) \right\rangle - g_t(\lambda) \right) q(dt), & \text{d}X \ll dq, \quad X_0 = 0, \\
\infty, & \text{otherwise,}
\end{array} \right.
$$
where \( dX \ll dq \) for \( X = (X_t)_{t \geq 0} \in D(\mathbb{R}^d) \) stands for \( X \) to be absolutely continuous with respect to \( q \) and \( dX/dq \) denotes the derivative.

Note that the last two integrals are well defined because due to the continuity of \( g_i(\lambda) \) in \( \lambda \) the supremum in the integrands may be taken over rational \( \lambda \in \mathbb{R}^d \), so that the integrands are measurable with respect to \( \mathcal{B}_q(\mathbb{R}^d) \), the \( q \)-completion of \( \mathcal{B}(\mathbb{R}^d) \).

We prove Lemmas 2.1 and 2.2 in Section 5.

For the main theorem, we need one more condition on \( G_t(\lambda) \). Denote \( H_{s,t}(x) \), \( s < t \), \( x \in \mathbb{R}^d \), the conjugate function (or the Legendre–Fenchel transform) of \( (G_t(\lambda) - G_s(\lambda)) \) (Rockafellar, 1970, Section 12):

\[
H_{s,t}(x) = \sup_{\lambda \in \mathbb{R}^d} \left( \langle \lambda, x \rangle - (G_t(\lambda) - G_s(\lambda)) \right), \quad x \in \mathbb{R}^d. \tag{2.5}
\]

\( H_{s,t}(x) \) is obviously convex in \( x \).

Let \( \text{ri(dom } H_{s,t}) \) denote the relative interior of the effective domain of \( H_{s,t} \) (Rockafellar, 1970). We require that the following hold:

(\( G \)) for all \( 0 \leq s < t \) the function \( H_{s,t}(x) \) is strictly convex on \( \text{ri(dom } H_{s,t}) \).

Condition (\( G \)) in various forms is rather common in large-deviation literature (see Gärtner, 1977; Freidlin and Wentzell, 1984; Ellis, 1984; Baldi, 1988; de Acosta, 1985, 1990, 1991), and it is well known that for it to hold it is sufficient that the closed convex hull of \( G_t(\lambda) \) be differentiable in \( \lambda \) for all \( t \geq 0 \) (Rockafellar, 1970, Theorem 26.3). It is handy to introduce the following notion of 'superexponential convergence in probability': a sequence \( (\xi_n) \) of \( \mathbb{R}^d \)-valued random variables is said to converge to an \( \mathbb{R}^d \)-valued random variable \( \xi \) superexponentially in probability if

\[
\lim_{n \to \infty} \left[ P(|\xi_n - \xi| > \varepsilon) \right]^{1/n} = 0 \quad \forall \varepsilon > 0.
\]

We denote this \( \xi_n \xrightarrow{P^{1/n}} \xi(n \to \infty) \).

We state our main result.

**Theorem 2.1.** Let for every \( \lambda \in \mathbb{R}^d \) and \( n = 1, 2, \ldots \) there exists an \( F^n \)-predictable real-valued positive process \( \mathcal{E}^n(\lambda) = (\mathcal{E}^n_t(\lambda))_{t \geq 0} \), \( \mathcal{E}^n_0(\lambda) = 1 \), such that the process \( Y^n(\lambda) = (Y^n_t(\lambda))_{t \geq 0} \) with

\[
Y^n_t(\lambda) = e^{\langle \lambda, x^n_t - x^n_0 \rangle} (\mathcal{E}^n_t(\lambda))^{-1}
\]

is an \( F^n \)-local martingale.

If \( X^n_0 \xrightarrow{P} 0 \) \( (n \to \infty) \), and for all \( T > 0 \) and \( \lambda \in \mathbb{R}^d \)

\[
\left( \sup_{t \leq T} \mathcal{E}^{1/n} \right) \left( \sup_{t \leq T} \frac{1}{n} \log \mathcal{E}^{n}_t(n\lambda) - G_t(\lambda) \right) \xrightarrow{P^{1/n}} 0 \quad (n \to \infty),
\]

then

\[
\mathcal{L}(X^n) \xrightarrow{1.d.} \exp(-1) \quad (n \to \infty).
\]
Now let $X^n$ be semimartingales. Denotes $\mu^n = \mu^n(dx,dx)$ the measure of jumps of $X^n$ (Jacod and Shiryaev, 1987; Liptser and Shiryaev, 1989):

$$\mu^n([0,t], \Gamma) = \sum_{0 < s < t} \Delta X^n_s^* 1(\Delta X^n_s^* \in \Gamma \setminus \{0\}), \quad t > 0, \quad \Gamma \in \mathcal{B}^{\infty}$$

where $\Delta X^n_s^*$ is the jump of $X^n = \{X^n_t \}_{t \geq 0}$ at $s$.

Let $v^n = v^n(ds,dx)$ be the $F^n$-compensator of $\mu^n$ (Jacod and Shiryaev, 1987; Liptser and Shiryaev, 1989). Assume that the following analogue of the Cramér condition holds:

$$\int_0^t \int_{\mathbb{R}^d} e^{\langle \lambda, x \rangle} 1(|x| > 1) v^n(ds,dx) < \infty \quad P\text{-a.s.}, \quad t > 0, \quad \lambda \in \mathbb{R}^d. \quad (2.6)$$

Then $X^n$ is a special (and even a locally square integrable) semimartingale (Jacod and Shiryaev, 1987, II.2.29). Let

$$X^n = X^n_0 + A^n + M^n$$

be its canonical decomposition, i.e. $A^n = \{A^n_t \}_{t \geq 0}$, $A^n_0 = 0$, is an $F^n$-predictable $\mathbb{R}^d$-valued process with finite variation over finite intervals, and $M^n = \{M^n_t \}_{t \geq 0}$, $M^n_0 = 0$, is an $F^n$-local (which under (2.6) is even locally square integrable) $\mathbb{R}^d$-valued martingale.

Denote $C^n = \{C^n_t \}_{t \geq 0}$, $C^n_0 = 0$, the $F^n$-predictable quadratic variation process of the continuous martingale part of $X^n$ (Jacod and Shiryaev, 1987; Liptser and Shiryaev, 1989).

We shall assume that $(A^n, C^n, v^n)$ is the ‘good’ version of the characteristics (Jacod and Shiryaev, 1987) in the sense that (2.6) holds identically and also identically $(a \wedge b = \min(a, b))$:

$$\int_0^t \int_{\mathbb{R}^d} (1 \wedge |x|^2) v^n(ds,dx) < \infty, \quad v^n(\{t\}, \mathbb{R}^d) \leq 1, \quad t > 0, \quad (2.7)$$

$$\Delta A^n_t = \int_{\mathbb{R}^d} x v^n(\{t\}, dx), \quad t > 0, \quad (2.8)$$

and for all $s < t$, $(C^n_t - C^n_s)$ is a symmetric positive-semi-definite $d \times d$-matrix.

We denote $\lambda C^n_t$ the product of $\lambda \in \mathbb{R}^d$ and $C^n_t$ and define the cumulant $G^n(\lambda) = \{G^n_t(\lambda) \}_{t \geq 0}$ by

$$G^n(\lambda) = \langle \lambda, A^n \rangle + \frac{1}{2} \langle \lambda, \lambda, C^n \rangle + \int_0^t \int_{\mathbb{R}^d} (e^{\langle \lambda, x \rangle} - 1 - \langle \lambda, x \rangle) v^n(ds,dx), \quad t > 0, \quad \lambda \in \mathbb{R}^d. \quad (2.9)$$

Let $\mathcal{E}(G^n(\lambda)) = \{\mathcal{E}(G^n(\lambda))_t \}_{t \geq 0}$ be the stochastic (or the Dolèans-Dade) exponential of $G^n(\lambda)$ Jacod and Shiryaev, 1987; Liptser and Shiryaev, 1989, i.e. an $F^n$-predictable real-valued process which is the solution to the equation

$$\mathcal{E}(G^n(\lambda))_t = 1 + \int_0^t \mathcal{E}(G^n(\lambda))_s \cdot dG^n_s(\lambda), \quad \mathcal{E}(G^n(\lambda))_0 = 1.$$
Explicitly \( \mathcal{E}(G^n(\lambda)) \) can be written as
\[
\mathcal{E}(G^n(\lambda))_t = e^{G^n(\lambda)} \prod_{s \leq t} (1 + \Delta G^n_s(\lambda)) e^{-\Delta G^n_s(\lambda)}.
\]
(2.10)

In view of (2.8) and (2.9) and the continuity of \( C^* \),
\[
\Delta G^n_s(\lambda) = \int_{\mathbb{R}^d} (e^{\langle \lambda, x \rangle} - 1) v_n(t, dx) > -1,
\]
and so (2.10) and (2.7) can be seen to yield (cf. Liptser and Shiryaev, 1989, Lemma 4.2.2)
\[
\mathcal{E}(G^n(\lambda))_t > 0, \quad t \geq 0.
\]
(2.11)

**Theorem 2.2** (cf. Jacod and Shiryaev, 1987, Theorem VIII.2.30). Let \( X^n \) be semimartingales satisfying (2.6) and \( G^n(\lambda) \) be defined by (2.9). If \( X^n \xrightarrow{P_{1/n}} 0 \ (n \to \infty) \), and for all \( T > 0, \lambda \in \mathbb{R}^d \),
\[
\left( \sup_t \mathcal{E} \sup_{t \leq T} \left| \frac{1}{n} \log \mathcal{E}(G^n(n\lambda))_t - G_t(\lambda) \right| \right)^{P_{1/n}} \to 0 \ (n \to \infty),
\]
then
\[
\mathcal{L}(X^n) \xrightarrow{1.d.} \exp(-I) \quad (n \to \infty).
\]

As our second application of Theorem 2.1, we regard \( X^n \) to be processes with independent increments (not necessarily semimartingales) satisfying the Cramér condition
\[
E e^{\langle \lambda, X^n - X^n_0 \rangle} < \infty, \quad t > 0, \lambda \in \mathbb{R}^d.
\]
(2.12)

**Theorem 2.3.** Let \( X^n \) be processes with independent increments and (2.12) hold. If \( X^n \xrightarrow{P_{1/n}} 0, \) and for all \( T > 0 \) and \( \lambda \in \mathbb{R}^d \),
\[
\left| \frac{1}{n} \log E e^{\langle \lambda, X^n - X^n_0 \rangle} - G_t(\lambda) \right| > 0 \ (n \to \infty),
\]
then
\[
\mathcal{L}(X^n) \xrightarrow{1.d.} \exp(-I) \quad (n \to \infty).
\]

Note that if \( X^n \) are semimartingales with independent increments, then the assertions of Theorems 2.2 and 2.3 coincide since \((A^n, C^n, \nu^n)\) is deterministic (Jacod and Shiryaev, 1987; Liptser and Shiryaev, 1989), conditions (2.6) and (2.12) coincide and
\[
\mathcal{E}(G^n(\lambda))_t = E e^{\langle \lambda, X^n - X^n_0 \rangle}
\]

Before proceeding with the proofs we explain the position of our results in the literature. Conditions for large deviations of (linear-space-valued) random variables in terms of the convergence of the logarithms of their exponential moments have been appearing since the work of Gärtner (1977) (see Freidlin and Wentzell, 1984; Ellis,

Large deviations for processes with independent increments under the Cramér condition were first studied by Borovkov (1967), who considered a homogeneous real-valued process (or a Lévy process); this corresponds in our notation to the case $X^n = L_n/n$, where $L = (L_t, t \geq 0)$ is a Lévy process. Then the conditions of Theorem 2.3 are trivially satisfied and the rate function is given by Lemma 2.2 with $q(dt) = dt$ and
\[
g_t(\lambda) = \langle \lambda, b \rangle + \frac{1}{2} \langle \lambda, c \lambda \rangle + \int_{\mathbb{R}^d} (e^{\langle \lambda, x \rangle} - 1 - \langle \lambda, x \rangle) K(dx),
\]
where $b \in \mathbb{R}^d$, $c$ is a symmetric positive-semidefinite $d \times d$ matrix and $K$ is a measure on $\mathbb{R}^d$ with $K(\{0\}) = 0$ for which the latter integral is finite for all $\lambda \in \mathbb{R}^d$ (cf. Jacod and Shiryaev, 1987, Corollary II.4.19).

Mogulskii (1976) obtained the LDP for multidimensional random walks, while de Acosta (1991) studied Banach-space-valued Lévy processes. In fact, these results as well as a vast majority of others studying large deviations of stochastic processes, are mostly for the uniform topology on $D([0, T], \mathbb{R}^d)$. Since the uniform topology is finer than the Skorohod one, the LDP for the former is, in general, the stronger assertion. But in the case that limiting rate function is concentrated on $C(\mathbb{R}^d)$ (in other words, if the sequence $(\mathcal{L}(X^n))$ is $C$-exponentially tight), the two are easily seen to be equivalent. Namely, the following holds.

**Theorem C.** Assume that for the Skorohod topology on $D(\mathbb{R}^d)$, $\mathcal{L}(X^n) \overset{1, d}{\longrightarrow} \exp(-1)$ ($n \to \infty$). If $I(X) = \infty$ for $X \in D(\mathbb{R}^d) \setminus C(\mathbb{R}^d)$ and $X^n$ are still random elements with respect to the locally uniform topology on $D(\mathbb{R}^d)$, then $I(X)$ is a rate function for the locally uniform topology and, in this topology, $\mathcal{L}(X^n)$ also LD converges to $\exp(-1)$.

The proof is the same as for weak convergence (cf. Billingsley, 1968, Section 18). Now to obtain the LDP in $D([0, T], \mathbb{R}^d)$ for the uniform topology, it remains to apply the contraction principle (Varadhan, 1966, 1984; Friedlin and Wentzell, 1984; Deuschel and Stroock, 1989) in the form suggested in Puhalskii (1991) (see also Puhalskii, 1994).

Thus, our results indeed extend those for homogeneous processes with independent increments with the Cramér condition in that they cover the nonhomogeneous case as well as allow the increments to be dependent. For the case that the Cramér condition (2.12) fails to hold for all $\lambda \in \mathbb{R}^d$, see Mogulskii (1976, 1993), Lynch and Sethuraman (1987), and de Acosta (1991).

### 3. Auxiliary results

This section collects some general facts required for the sequel. The following assertion is in fact trivial and has nothing to do with probability. We give it as a separate statement because it comes our way quite frequently. All the values below are real numbers, $a \lor b = \max(a, b).$
Lemma 3.1. (1) If for $a_{n, \delta} \geq 0, b_{n, \delta} \geq 0, n = 1, 2, \ldots, \delta \geq 0$, 

$$
\lim_{\delta \to 0} \lim_{n \to \infty} \sup a_{n, \delta}^{1/n} = \lim_{\delta \to 0} \lim_{n \to \infty} \inf a_{n, \delta}^{1/n} = a,
$$

$$
\lim_{\delta \to 0} \lim_{n \to \infty} \sup b_{n, \delta}^{1/n} = \lim_{\delta \to 0} \lim_{n \to \infty} \inf b_{n, \delta}^{1/n} = b,
$$

then

$$
\lim_{\delta \to 0} \lim_{n \to \infty} (a_{n, \delta} + b_{n, \delta})^{1/n} = \lim_{\delta \to 0} \lim_{n \to \infty} (a_{n, \delta} + b_{n, \delta})^{1/n} = a \vee b.
$$

(2) If for $a_n \geq 0, b_n \geq 0, n = 1, 2, \ldots$, 

$$
\lim_{n \to \infty} a_n^{1/n} = a, \quad \lim_{n \to \infty} (a_n + b_n)^{1/n} = b
$$

and $b > a$, then

$$
\lim_{n \to \infty} b_n^{1/n} = b.
$$

All this follows by the inequalities

$$
x^{1/n} \vee y^{1/n} \leq (x + y)^{1/n} \leq 2^{1/n}[x^{1/n} \vee y^{1/n}], \quad x, y \geq 0.
$$

The next lemma is a result from convex analysis on properties of conjugate functions which though it is almost obvious, we failed to find in standard manuals. For it we adopt usual definitions and notation from convex analysis (Rockafellar, 1970; see also von Tiel, 1984).

For a subset $A$ of Euclidean space, $\text{cl} \ A$ denotes its closure, $\text{ri} \ A$, the relative interior, $\text{rb} \ A = \text{cl} \ A \setminus \text{ri} \ A$, the relative boundary, $\text{conv} \ A$, the convex hull of $A$. Let $f$ be a function from $\mathbb{R}^m, m \geq 1$, into $]-\infty, \infty]$. Its conjugate (or the Legendre–Fenchel transform) $f^*$ is

$$
f^*(\lambda) = \sup_{x \in \mathbb{R}^m} (\langle \lambda, x \rangle - f(x)), \quad \lambda \in \mathbb{R}^m,
$$

and the bipolar $f^{**}$ of $f$ is the conjugate of $f^*$:

$$
f^{**}(x) = \sup_{\lambda \in \mathbb{R}^m} (\langle \lambda, x \rangle - f^*(\lambda)), \quad x \in \mathbb{R}^m.
$$

Obviously, $f^{**}$ is convex and lower semicontinuous.

By $\text{epi} f$ we denote the epigraph of $f$:

$$
\text{epi} f = \{(x, y) \in \mathbb{R}^m \times \mathbb{R} : y \geq f(x)\},
$$

and let

$$
\text{dom} f = \{x \in \mathbb{R}^m : f(x) < \infty\}
$$

be the effective domain of $f$. The convex hull $\text{conv} f$ of $f$ is defined by

$$
\text{epi(conv} f) = \text{conv(} \text{epi} f),
$$
and the lower semicontinuous hull of \( f \) by

\[
\text{epi}(\text{cl} f) = \text{cl}(\text{epi} f).
\]

If \( f \) is convex, then \( \partial f(x) \) denotes the subgradient of \( f \) at \( x \).

**Lemma 3.2.** If \( f : \mathbb{R}^m \to [-\infty, \infty] \) is a lower semicontinuous function and its bipolar \( f^{**} \) is strictly convex on \( \text{ri}(\text{dom} f^{**}) \), then \( f = f^{**} \).

**Proof.** It is obvious that \( f^{**} \leq f \). So we prove the opposite inequality. By Rockafellar (1970, Cor. 12.1.1) and the argument below (see also van Tiel, 1984, 6.15) we have

\[
f^{**} = \text{cl}(\text{conv} f).
\]  

We first prove that

\[
f^{**}(x) \geq f(x), \quad x \in \text{ri}(\text{dom} f^{**}).
\]  

Assume the contrary, i.e. that for some \( x_0 \in \text{ri}(\text{dom} f^{**}) \) and \( \gamma > 0 \) we have

\[
f(x_0) > f^{**}(x_0) + \gamma.
\]  

Since \( x_0 \in \text{ri}(\text{dom} f^{**}) \), by Rockafellar (1970, Theorem 23.4) the set \( \partial f^{**}(x_0) \) is nonempty. Let \( \lambda_0 \in \partial f^{**}(x_0) \). Then by the definition of \( f^{**} \),

\[
f^{**}(x) \geq \langle \lambda_0, x \rangle - f^*(\lambda_0)
\]  

and by Rockafellar (1970, Theorem 23.5)

\[
f^{**}(x_0) = \langle \lambda_0, x_0 \rangle - f^*(\lambda_0).
\]  

It is easy to see that the strict convexity of \( f^{**} \) implies that

\[
f^{**}(x) > \langle \lambda_0, x \rangle - f^*(\lambda_0), \quad x \neq x_0.
\]  

Indeed, if for some \( x \neq x_0 \) we had equality in (3.4), then by the convexity of \( f^{**} \) and by (3.5) for any \( z \in [x_0, x[ \)

\[
f^{**}(z) \leq \langle \lambda_0, z \rangle - f^*(\lambda_0),
\]  

which together with (3.4) would give

\[
f^{**}(z) = \langle \lambda_0, z \rangle - f^*(\lambda_0), \quad z \in [x_0, x[.
\]  

But \( [x_0, x[ \subset \text{ri}(\text{dom} f^{**}) \) (by Rockafellar 1970, Theorem 6.1) and since \( x \in \text{dom} f^{**} \) if there is equality in (3.4)). Thus \( f^{**} \) would fail to be strictly convex on \( \text{ri}(\text{dom} f^{**}) \) and (3.6) is proved.

Using the lower semicontinuity of \( f \) choose \( \epsilon > 0 \) with \( \epsilon |\lambda_0| < \gamma/3 \) such that

\[
f(x) > f(x_0) - \frac{1}{2} \gamma, \quad |x - x_0| < \epsilon.
\]  

For this \( \epsilon \), choose \( \delta > 0, \delta < \gamma/3 \), satisfying

\[
\{x: \langle \lambda_0, x \rangle - f^*(\lambda_0) + \delta \geq f^{**}(x)\} \subset \{x: |x - x_0| < \epsilon\}.
\]

\[
\{x: \langle \lambda_0, x \rangle - f^*(\lambda_0) + \delta \geq f^{**}(x)\} \subset \{x: |x - x_0| < \epsilon\}.
\]  

\[
\{x: \langle \lambda_0, x \rangle - f^*(\lambda_0) + \delta \geq f^{**}(x)\} \subset \{x: |x - x_0| < \epsilon\}.
\]
To show such a \( \delta \) exists, denote \( A_\delta \) the set on the left of (3.8). Then by (3.5) and (3.6),

\[
\bigcap_{\delta > 0} A_\delta = \{x_0\}. \quad (3.9)
\]

The hyperplane in \( \mathbb{R}^m \times \mathbb{R} \) defined by the equation

\[
y = \langle \lambda_0, x \rangle - f^*(\lambda_0) + \delta, \quad x \in \mathbb{R}^m, \quad y \in \mathbb{R},
\]

is parallel to the hyperplane \( y = \langle \lambda_0, x \rangle - f^*(\lambda_0) \). The latter has with \( \text{epi} f^{**} \) in view of (3.5) and (3.6) the only point \( x_0 \) in common. Then by Rockafellar (1970, Corollary 8.4.1) the sets \( A_\delta \) are bounded. They are closed since \( f^{**} \) is lower semicontinuous. Thus, \( A_\delta \) are compacts and (3.9) easily implies that for all \( \delta > 0 \) small enough \( A_\delta \subset \{x: |x - x_0| < \epsilon\} \) proving (3.8). For the chosen \( \delta \) and \( \epsilon \), define

\[
f_{\delta, \epsilon}(x) = \max(f^{**}(x), \langle \lambda_0, x \rangle - f^*(\lambda_0) + \delta).
\]

(3.10)

Obviously, \( f_{\delta, \epsilon} \) is convex, lower semicontinuous and \( f_{\delta, \epsilon}(x_0) > f^{**}(x_0) \) by (3.5). If we show that

\[
f_{\delta, \epsilon}(x) \leq f(x), \quad x \in \mathbb{R}^m,
\]

(3.11)

this will contradict (3.1), and (3.2) will be proved.

It is clear that (3.11) holds on \( \{x: |x - x_0| \geq \epsilon\} \) since \( f_{\delta, \epsilon}(x) = f^{**}(x) \) for these \( x \) by (3.8) and (3.10).

If \( |x - x_0| < \epsilon \) then using (3.5), (3.3) and (3.7) we have

\[
\langle \lambda_0, x \rangle - f^*(\lambda_0) + \delta = \langle \lambda_0, x - x_0 \rangle + f^{**}(x_0) + \delta
\]

\[
< \epsilon|\lambda_0| + f(x_0) - \gamma + \delta \leq \epsilon|\lambda_0| + f(x) - \frac{3}{2}\gamma + \delta < f(x)
\]

(the latter by the choice of \( \epsilon \) and \( \delta \)). Since, as we noted, \( f^{**} \leq f \), this proves (3.11) on \( \{x: |x - x_0| < \epsilon\} \). Thus (3.2) is proved.

Now if \( x \in \text{rb(dom } f^{**} \text{)} \) we have by Rockafellar (1970, Theorem 7.5) in view of the lower semicontinuity of \( f^{**} \) that for any \( z \in \text{ri(dom } f^{**} \text{)} \)

\[
f^{**}(x) = \lim_{\theta \uparrow 1} f^{**}((1 - \theta)z + \theta x).
\]

(3.12)

By Rockafellar (1970, Theorem 6.1), \( [z, x] \subset \text{ri(dom } f^{**} \text{)} \), and then by the just proved,

\[
f^{**}((1 - \theta)z + \theta x) = f(((1 - \theta)z + \theta x), \quad 0 \leq \theta < 1,
\]

whence by the lower semicontinuity of \( f \) and (3.12) we have \( f^{**}(x) > f(x) \) proving the assertion of the lemma for \( x \in \text{cl(dom } f^{**} \text{)} \).

Finally, for \( x \notin \text{cl(dom } f^{**} \text{)} \) we obviously have \( f(x) = f^{**}(x) = \infty \). The lemma is proved. \( \square \)

The following three lemmas contain large deviation properties which are rather well known but it serves our purpose better to formulate them differently.
In Lemmas 3.3 and 3.4, $X^n$ and $Y^n$ are random elements taking values in a metric space $S$ (separable in Lemma 3.3) with metric $\rho$, $I$ is a rate function on $S$. As above, for a random element $Z$, $\mathcal{L}(Z)$ denotes its law.

**Lemma 3.3.** (cf. Billingsley, 1968, Theorem 4.1). If

$$\mathcal{L}(X^n) \xrightarrow{\text{ld}} \exp(-I)(n \to \infty) \quad \text{and} \quad \rho(X^n, Y^n) \xrightarrow{\text{P}} 0 \quad (n \to \infty),$$

then

$$\mathcal{L}(Y^n) \xrightarrow{\text{ld}} \exp(-I) \quad (n \to \infty).$$

**Proof.** Immediate from the definition of LD convergence (and is the same as of Theorem 4.1 in Billingsley (1968)). Note that $\rho(X^n, Y^n)$ is a random variable due to the separability of $S$ (Billingsley, 1968, p. 25). □

The next lemma belongs to Varadhan (see Varadhan, 1966, and also Varadhan (1984) and Deuschel and Stroock (1989). We give it in the form which emphasizes similarity with a result from weak convergence.

Say that a sequence $(\xi_n)$ of real-valued random variables is uniformly exponentially integrable if

$$\lim_{A \to \infty} \limsup_{n \to \infty} \left[ E I(|\xi_n| > A) \right]^{1/n} = 0.$$

**Lemma 3.4.** (cf. Billingsley, 1968, Theorem 5.4). Let $\mathcal{L}(X^n) \xrightarrow{\text{ld}} \exp(-I)(n \to \infty)$ and $f: S \to \mathbb{R}_+$ be a continuous function. If the sequence $(f(X^n), n \geq 1)$ is uniformly exponentially integrable, then

$$\lim_{n \to \infty} \left[ E(f(X^n))^n \right]^{1/n} = \sup_{x \in S} [f(x)\exp(-I(x))].$$

In the same way as for uniform integrability (Billingsley, 1968, p. 32), the Chebyshev inequality yields that for the uniform exponential integrability of $(\xi_n)$ to hold, it is sufficient that for some $\varepsilon > 0$,

$$\sup_{n} \left[ E|\xi_n|^{n(1 + \varepsilon)} \right]^{1/n} < \infty.$$ 

This assertion one can find in Deuschel and Stroock (1989). The following lemma is a modification of Gärtner’s (1977) result in the form of Freidlin and Wentzell (1984, Chap. 5, Section 1) (see also Ellis, 1984). It will be required to prove 'finite-dimensional' LD convergence.

Let $K(\lambda), \lambda \in \mathbb{R}^m, m \geq 1$, be a real-valued function. Denote $L(x), x \in \mathbb{R}^m$, its conjugate. It is easy to see that $L$ is a rate function on $\mathbb{R}^m$.

**Lemma 3.5.** Let $L(x)$ be strictly convex on $\text{ri}(\text{dom} L)$. Assume that for a sequence $(Y^n, n \geq 1)$ of $\mathbb{R}^m$-valued random variables and for each $\lambda \in \mathbb{R}^m$ there exists a sequence
of \( \mathbb{R}^m \)-valued random variables such that

\[
\lim_{n \to \infty} \frac{1}{n} \log E[\exp(n \langle \lambda, Z^n(\lambda) \rangle)] = K(\lambda),
\]

and, as \( n \to \infty \),

\[
Y^n - Z^n(\lambda) \overset{P}{\to} 0.
\]

If for each \( \lambda \in \mathbb{R}^m \) the sequence \( \exp(\langle \lambda, Z^n(\lambda) \rangle) \), \( n \geq 1 \) is uniformly exponentially integrable, then

\[
\mathcal{L}(Y^n) \overset{1.d.}{\to} \exp(-L) \quad (n \to \infty)
\]

(in \( \mathbb{R}^m \)).

**Proof.** First, it is easy to see that \( \mathcal{L}(Y^n) \), \( n \geq 1 \) is exponentially tight. Indeed if \( Y^n = (Y^n_1, \ldots, Y^n_m) \), \( Y^n_i \in \mathbb{R} \), and \( Z^n(\lambda) = (Z^n_1(\lambda), \ldots, Z^n_m(\lambda)) \), \( Z^n_i(\lambda) \in \mathbb{R} \), and \( e_i \) denotes the \( m \)-vector with 1 in the \( i \)th place and 0 in the others, then for \( i = 1, \ldots, m; A > 1 \),

\[
[P(Y^n_i > A)]^{1/n} \leq [P(|Y^n_i - Z^n(e_i)| \geq 1)]^{1/n}
\]

\[
+ [P(Z^n(e_i) > A - 1)]^{1/n} \leq [P(|Y^n_i - Z^n(e_i)| \geq 1)]^{1/n}
\]

\[
+ e^{-(A-1)}[E \exp(nZ^n_i(e_i))]^{1/n},
\]

which implies by the assumptions that

\[
\lim_{A \to \infty} \sup_{n \to \infty} [P(Y^n_i > A)]^{1/n} = 0, \quad 1 \leq i \leq m.
\]

Analogously,

\[
\lim_{A \to \infty} \sup_{n \to \infty} [P(Y^n_i < -A)]^{1/n} = 0, \quad 1 \leq i \leq m.
\]

This proves the exponential tightness of \( \mathcal{L}(Y^n) \), \( n \geq 1 \).

Therefore, by Puhalskii (1991), there exists a subsequence \( (n') \) and a rate function \( L' \) on \( \mathbb{R}^m \) such that

\[
\mathcal{L}(Y^n) \overset{1.d.}{\to} \exp(-L') \quad (n' \to \infty),
\]

and so by Lemma 3.3 and by the assumptions,

\[
\mathcal{L}(Z^n(\lambda)) \overset{1.d.}{\to} \exp(-L')(n' \to \infty), \quad \lambda \in \mathbb{R}^m.
\]

Since the sequence \( \exp(\langle \lambda, Z^n(\lambda) \rangle) \), \( n \geq 1 \) is uniformly exponentially integrable, by Lemma 3.4,

\[
\lim_{n' \to \infty} \frac{1}{n'} \log E \exp(n' \langle \lambda, Z^n(\lambda) \rangle) = \sup_{x \in \mathbb{R}^m} (\langle \lambda, x \rangle - L'(x)).
\]

and so by the assumptions,

\[
K(\lambda) = \sup_{x \in \mathbb{R}^m} (\langle \lambda, x \rangle - L'(x)), \quad \lambda \in \mathbb{R}^m.
\]
Thus, $L = (L')^*$, and by Lemma 3.2 $L' = L$. So $L'$ is unique and equals $L$. The lemma is proved.

We end the section with a multiplicative analogue of the Lenglart-Rebolledo inequality (see, e.g., Jacod and Shiryaev, 1987; Liptser and Shiryaev, 1989). It will help us to prove the C-exponential tightness of $(\mathcal{L}(X^n))$.

**Lemma 3.6.** Let $X = (X_t)_{t \geq 0}$ and $Y = (Y_t)_{t \geq 0}$ be positive, real-valued processes on a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. If for any stopping time $\tau < \infty$,

$$E(X_{\tau}/Y_\tau) \leq 1,$$

then for any stopping time $\sigma \leq \infty$ and all $a > 0$, $b > 0$,

$$P\left(\sup_{t \leq \sigma} X_t \geq a\right) \leq \frac{b}{a} + P\left(\sup_{t \leq \sigma} Y_t \geq b\right)$$

(Here $\sup_{t \leq \infty} = \sup_{t \geq 0}$).

**Proof.** Define

$$\tau = \inf(t \geq 0: X_t \geq a) \quad (\inf \emptyset = \infty).$$

Then if $\mathbb{P}(\sigma < \infty) = 1$, we have

$$P\left(\sup_{t \leq \sigma} X_t > a\right) \leq P(X_{\sigma \wedge \tau} \geq a) \leq P(Y_{\sigma \wedge \tau} > b) + P(X_{\sigma \wedge \tau} \geq a, Y_{\sigma \wedge \tau} \leq b)$$

$$\leq P(Y_{\sigma \wedge \tau} > b) + P(X_{\sigma \wedge \tau}/Y_{\sigma \wedge \tau} \geq a/b).$$

By the Chebyshev inequality and in view of the assumptions, we then have

$$P\left(\sup_{t \leq \sigma} X_t > a\right) \leq P(Y_{\sigma \wedge \tau} > b) + b/a \leq P\left(\sup_{t \leq \sigma} Y_t > b\right) + b/a.$$

To obtain the required, note that

$$P\left(\sup_{t \leq \sigma} X_t \geq a\right) = \lim_{N \to \infty} P\left(\left.\sup_{t \leq \sigma} X_t > a - \frac{1}{N}\right)\right.$$.

Now if $\mathbb{P}(\sigma < \infty) < 1$, then by the just proved we have, for any $N > 0$,

$$P\left(\sup_{t \leq \sigma} X_t \geq a\right) \leq P\left(\sup_{t \leq \sigma} X_t > a - \frac{1}{N}\right)$$

$$= \lim_{M \to \infty} P\left(\sup_{t \leq \sigma \wedge M} X_t > a - \frac{1}{N}\right) \leq \frac{b}{a - \frac{1}{N}} + P\left(\sup_{y \leq \sigma} Y_t \geq b\right).$$

Since $N$ is arbitrary, the proof is over. \lhd
4. Proof of Theorems 2.1–2.3

We begin with Theorem 2.1. In the proof we are guided by the same ideas as in weak convergence (Jacod and Shiryaev, 1987; Liptser and Shiryaev, 1989. All the conditions and notation of Theorem 2.1 are in force.

Denote
\[ G^*_s(\lambda) = \sup_{t \leq s} |G_s(\lambda)|, \quad t \geq 0, \ \lambda \in \mathbb{R}^d. \]  
(4.1)

We introduce the stopping times \((\lambda \in \mathbb{R}^d)\)
\[ \sigma^n(\lambda) = \inf\{t \geq 0 : [\mathcal{E}_t^n(\lambda)]^{1/n} \vee [\mathcal{E}_t^n(\lambda)]^{-1/n} \geq 2e^{G^*_t(\lambda)} \} \]
or
\[ [\mathcal{E}_t^n(2n\lambda)]^{1/n} \vee [\mathcal{E}_t^n(2n\lambda)]^{-1/n} \geq 2e^{G^*_t(2\lambda)}. \]  
(4.2)

By condition \((\sup \mathcal{E}^{1/n})\) of the theorem,
\[ \lim_{n \to \infty} P(\sigma^n(\lambda) \leq t) = 0, \quad t \geq 0, \ \lambda \in \mathbb{R}^d. \]  
(4.3)

Also \(\sigma^n(\lambda)\) is an \(P\)-predictable stopping time because it is a \(d\text{ébut}\) of a predictable set (since \(\mathcal{E}^n(\lambda)\) is \(F^n\)-predictable) whose graph belongs to the set (Dellacherie, 1972, IV-T.16) (see also Jacod and Shiryaev, 1987, I.2.13). Thus, \(\sigma^n(\lambda)\) is \(P\)-a.s. announced by an increasing sequence of \(F^n\)-stopping times. Since obviously \(\sigma^n(\lambda) > 0\), we can choose an \(F^n\)-stopping time \(\tau^n(\lambda)\) such that
\[ \tau^n(\lambda) < \sigma^n(\lambda), \]  
(4.4)

\[ \left[ P\left( \tau^n(\lambda) + \frac{1}{n} \leq \sigma^n(\lambda) < \infty \right) \right]^{1/n} \vee \left[ P\left( \tau^n(\lambda) \leq n, \ \sigma^n(\lambda) = \infty \right) \right]^{1/n} \leq \frac{1}{n}. \]  
(4.5)

In view of
\[ P(\tau^n(\lambda) \leq t) \leq P(\sigma^n(\lambda) \leq t + 1) + P\left( \tau^n(\lambda) + \frac{1}{n} \leq (t + 1) \wedge \sigma^n(\lambda) \right), \]
we have from (4.3) and (4.5)
\[ \lim_{n \to \infty} [P(\tau^n(\lambda) \leq t)]^{1/n} = 0, \quad t > 0. \]  
(4.6)

Note also that by (4.2) and (4.4)
\[ [\mathcal{E}_t^n \wedge \tau^n(\lambda)(n\lambda)]^{1/n} \vee [\mathcal{E}_t^n \wedge \tau^n(\lambda)(n\lambda)]^{-1/n} < 2e^{G^*_t(\lambda)}, \]  
(4.7)

\[ [\mathcal{E}_t^n \wedge \tau^n(\lambda)(2n\lambda)]^{1/n} \vee [\mathcal{E}_t^n \wedge \tau^n(\lambda)(2n\lambda)]^{-1/n} < 2e^{G^*_t(2\lambda)}. \]  
(4.8)

**Lemma 4.1.** For all \(\lambda \in \mathbb{R}^d\), the process \(Y^n(\lambda)\) is a positive supermartingale. The process \((Y^n_t(\lambda))_{t \leq T}\), for all \(T > 0\), is an \(F^n\)-square integrable martingale, and
\[ E(Y^n_t(\lambda))^2 \leq 2^{3n}e^{G^*_t(2\lambda)} + 2G^*_t(\lambda), \quad \lambda \in \mathbb{R}^d, \ t \geq 0. \]  
(4.9)
Proof. By the assumptions of the theorem, \( Y^n(\lambda) \) is a positive \( F^n \)-local martingale. Hence it is a supermartingale. So we have to prove only (4.9). By the supermartingale property of \( Y^n(\lambda) \) and since \( Y^n_0(\lambda) = 1 \), for all stopping times \( \tau \),

\[
EY^n_\tau(\lambda) \leq 1, \quad \lambda \in \mathbb{R}^d.
\] (4.10)

In view of (4.8) and the definition of \( Y^n(\lambda) \), (4.10) with \( 2\lambda \) implies that

\[
Ee^{2n(\lambda, X^n_\tau, \tau(\lambda))} \leq 2^n e^{nG^n_\tau(2\lambda)}, \quad \lambda \in \mathbb{R}^d,
\]

which again by the definition of \( Y^n(\lambda) \) and by (4.7) yields

\[
E(Y^n_\tau \wedge \tau(\lambda))^2 \leq 2^n e^{nG^n_\tau(2\lambda)} 2^{2n} e^{2nG^n_\tau(\lambda)}
\]

proving the lemma. \( \Box \)

Lemma 4.2. Set for \( 0 = t_0 < t_1 < \cdots < t_k \),

\[
Z^n_\tau(\lambda) = X^n_{t_i \wedge \tau(\lambda)} - X^n_{t_{i-1} \wedge \tau(\lambda)}, \quad \lambda \in \mathbb{R}^d.
\]

Then, for all \( \lambda_1, \ldots, \lambda_k \in \mathbb{R}^d \), the sequence \( \exp(\sum_{i=1}^k \langle \lambda_i, Z^n(\lambda_i) \rangle) \), \( n \geq 1 \) is uniformly exponentially integrable and

\[
\lim_{n \to \infty} \frac{1}{n} \log E \left[ \exp \left( n \sum_{i=1}^k \langle \lambda_i, Z^n(\lambda_i) \rangle \right) \right] = \sum_{i=1}^k (G^n_{t_i}(\lambda_i) - G^n_{t_{i-1}}(\lambda_i)).
\]

Proof. Define

\[
\zeta^n_i = \sum_{j=1}^i \langle \lambda_j, Z^n(\lambda_j) \rangle, \quad 1 \leq i \leq k, \quad \zeta^n_0 = 0.
\] (4.11)

We first prove that

\[
E(\exp(2n\zeta^n_i)) \leq 2^{2n} \prod_{j=1}^i e^{n(G^n_{t_j}(2\lambda_j) + G^n_{t_{j-1}}(2\lambda_j))}, \quad 0 \leq i \leq k.
\] (4.12)

In view of (4.11) and the definitions of \( Z^n(\lambda) \) and \( Y^n(\lambda) \), we have by Lemma 4.1 and (4.8), for \( i \geq 1 \),

\[
E[\exp(2n\zeta^n_i) | \mathcal{F}^n_{t_{i-1}}] \leq \exp(2n\zeta^n_{i-1}) \exp(-2n \langle \lambda_i, X^n_{t_{i-1} \wedge \tau(\lambda_i)} - X^n_{t_i \wedge \tau(\lambda_i)} \rangle) \times Y^n_{t_{i-1} \wedge \tau(\lambda_i)}(2n\lambda_i) 2^n e^{nG^n_{t_i}(2\lambda_i)}
\]

Applying to the latter (4.8) again we deduce

\[
E[\exp(2n\zeta^n_i) | \mathcal{F}^n_{t_{i-1}}] \leq \exp(2n\zeta^n_{i-1}) 2^{2n} e^{n(G^n_{t_i}(2\lambda_i) + G^n_{t_{i-1}}(2\lambda_i))}.
\]

This proves (4.12).

According to the remark after Lemma 3.4, the uniform exponential integrability of \( \exp(\sum_{i=1}^k \langle \lambda_i, Z^n(\lambda_i) \rangle) \), \( n \geq 1 \) is implied by (4.12) with \( i = k \).
We prove the convergence required in the lemma by proving that for \( i = 1, \ldots, k \),
\[
\lim_{n \to \infty} \frac{1}{n} \log E \exp(n \xi_i^n) = g_i,
\]
where
\[
g_i = \sum_{j=1}^{i} (G_{t_j} - G_{t_{j-1}}), \quad 1 \leq i \leq k, \quad g_0 = 0,
\]
provided (4.13) holds for \( (i - 1) \).

For \( \delta \in ]0, \frac{1}{2}[ \) define the sets
\[
B_{\delta}^{n,1} = \{ \omega \in \Omega : \left( |(\mathcal{E}_{t_{i-1}} \wedge v(\tau_i)(n\lambda_i))^{1/n} e^{-G_{t_{i-1}}(\lambda_i)} - 1| \geq \delta \right) \},
\]
\[
B_{\delta}^{n,2} = \{ \omega \in \Omega : \left( |(\mathcal{E}_{t_{i-1}} \wedge v(\tau_i)(n\lambda_i))^{-1/n} e^{G_{t_{i-1}}(\lambda_i)} - 1| \geq \delta \right) \},
\]
\[
B_{\delta}^{n} = B_{\delta}^{n,1} \cup B_{\delta}^{n,2}, \quad A_{\delta}^{n} = \Omega \setminus B_{\delta}^{n}.
\]

By \( (\sup \mathcal{E}^{1/n}) \) and (4.6), we obviously have
\[
[P(B_{\delta}^{n,1})]^{1/n} \to 0, \quad [P(B_{\delta}^{n,2})]^{1/n} \to 0 \quad (n \to \infty)
\]
and therefore by Lemma 3.1,
\[
[P(B_{\delta}^{n})]^{1/n} \to 0, \quad [P(A_{\delta}^{n})]^{1/n} \to 1 \quad (n \to \infty).
\]

Applying the Hölder inequality we then have by (4.12) and (4.14),
\[
\lim_{n \to \infty} [E \exp(n \xi_i^n) 1(B_{\delta}^{n})]^{1/n} = 0,
\]
which in view of Lemma 3.1 again implies that (4.13) would follow from
\[
\lim_{n \to \infty} \liminf_{\delta \to 0} [E \exp(n \xi_i^n) 1(A_{\delta}^{n})]^{1/n} = \lim_{n \to \infty} \limsup_{\delta \to 0} [E \exp(n \xi_i^n) 1(A_{\delta}^{n})]^{1/n} = e^{g_i}.
\]

Denote
\[
R_{\delta}^{n} = \exp(n \xi_i^n) Y_t^n \wedge v(\tau_i)(n\lambda_i)[Y_{t_{i-1}}^n \wedge v(\tau_i)(n\lambda_i)]^{-1}.
\]

By Lemma 4.1,
\[
ER_{\delta}^{n} = E \exp(n \xi_i^n - 1)
\]
and so by our assumption
\[
\lim_{n \to \infty} (ER_{\delta}^{n})^{1/n} = e^{g_i-1}.
\]

On the other hand, (4.11), (4.16) and the definitions of \( Y^n(\lambda) \) and \( Z^n(\lambda) \) yield
\[
R_{\delta}^{n} = \exp(n \xi_i^n) \left[ \mathcal{E}_{t_{i-1}} \wedge v(\tau_i)(n\lambda_i) \right]^{-1} \mathcal{E}_{t_{i-1}} \wedge v(\tau_i)(n\lambda_i),
\]
so applying the Hölder inequality we have in view of (4.7), (4.14) and (4.12) that
\[
\lim_{n \to \infty} [ER_{\delta}^{n} 1(B_{\delta}^{n})]^{1/n} = 0,
\]
which by (4.17) and Lemma 3.1 gives
\[
\lim_{n \to \infty} \left[ E R^n_\delta(1(A^n_\delta)) \right]^{1/n} \geq e^{\delta^{-1}}. \tag{4.19}
\]
By definition, we have on \( A^n_\delta \)
\[
\mathcal{E}_{n_{i-1}}(\lambda_i)(n\lambda_i) \leq (1 + \delta)^n e^{\sigma G_{n_{i-1}}(\lambda_i)},
\]
\[
\left[ \mathcal{E}_{n_{i-1}}(\lambda_i)(n\lambda_i) \right]^{-1} \leq (1 + \delta)^n e^{-\sigma G_{n_{i-1}}(\lambda_i)}.
\]
Therefore by (4.18),
\[
R^n_\delta(1(A^n_\delta)) \leq \exp(n\zeta^n_1)(1 + \delta)^{2n} e^{-n(G_{n_{i-1}}(\lambda_i) - G_{n_{i-1}}(\lambda_i))} 1(A^n_\delta).
\]
The latter implies by (4.19) that
\[
\lim_{\delta \to 0} \lim_{n \to \infty} \liminf \left[ E \exp(n\zeta^n_1) 1(A^n_\delta) \right]^{1/n} \geq e^{\delta^{-1}} e^{\sigma G_{n_{i-1}}(\lambda_i) - G_{n_{i-1}}(\lambda_i)} = e^{\delta^{-1}}.
\]
The opposite inequality for (4.15) is proved analogously.
The lemma is proved. \( \Box \)

Now we proceed to the proof of Theorem 2.1 itself. We apply Theorem A to the sequence \( (X^n) \).

First, we check the \( C \)-exponential tightness using Theorem B. Begin with condition (i). In view of (4.10) and the definition of \( Y^n(\hat{\lambda}) \), we can apply to \( (\exp^{(n\lambda, X^n_x - X^n_0)})_{x \geq 0} \) and \( \mathcal{E}^n(n\lambda) \) Lemma 3.6 to get for all \( \lambda > 0, B > 0 \) and \( L > 0 \),
\[
P\left( \sup_{t \leq L} \exp^{(n\lambda, X^n_x - X^n_0)} \geq e^{n\lambda - A} \right) \leq e^{n(B - A)} + P\left( \sup_{t \leq L} \mathcal{E}_t^n(n\lambda) \geq e^{nB} \right), \quad \lambda \in \mathbb{R}^d. \tag{4.20}
\]
Taking \( B > G^n_\delta(\lambda) + 1 \), we have by (sup \( \mathcal{E}^{1/n} \))
\[
\lim_{n \to \infty} \left[ P \left( \sup_{t \leq L} \mathcal{E}_t^n(n\lambda) \geq e^{nB} \right) \right]^{1/n} = 0,
\]
and then (4.20) yields
\[
\lim_{n \to \infty} \left[ P \left( \sup_{t \leq L} \langle \lambda, X^n_t - X^n_0 \rangle > A \right) \right]^{1/n} \leq e^{B - A} \to 0 \quad (A \to \infty), \quad \lambda \in \mathbb{R}^d.
\]
Since \( \lambda \) is arbitrary, this obviously implies that
\[
\lim_{A \to \infty} \lim_{n \to \infty} \left[ P \left( \sup_{t \leq L} |X^n_t - X^n_0| > A \right) \right]^{1/n} = 0.
\]
As by assumption \( X^n_0 \overset{P}{\to} 0 \quad (n \to \infty) \), we obtain (i).

Turning to (ii) it is again sufficient to prove that for all \( \lambda \in \mathbb{R}^d, \lambda \neq 0, \eta > 0 \),
\[
\lim_{\delta \to 0} \lim_{n \to \infty} \sup_{t \in T(F^\delta)} \left[ P \left( \sup_{t \leq \delta} \left| \frac{\lambda}{|\lambda|} X^n_{t+\delta} - X^n_t \right| > \eta \right) \right]^{1/n} = 0. \tag{4.21}
\]

By Lemma 4.1 and Doob's stopping theorem we have for any $F^n$-stopping times $\sigma$ and $\tau$ with $\sigma \geq \tau$,

$$F(Y^n_\sigma(n\lambda) | \mathcal{F}^\sigma_t) \leq Y^n_\tau(n\lambda),$$

whence

$$E[Y^n_\sigma(n\lambda) / Y^n_\tau(n\lambda)] \leq 1.$$  \hfill (4.22)

Fixing $\tau \in T_\infty(F^n)$ define, for $t \geq 0$,

$$X_t^{n,\tau} = X_{t+\tau} - X_t^n,$$  \hfill (4.23)

$$\mathcal{F}_t^{n,\tau}(\lambda) = \mathcal{F}_t^{n,\tau}(\lambda) / \mathcal{F}_t^n(\lambda),$$  \hfill (4.24)

$$F^{n,\tau} = (\mathcal{F}_t^{n,\tau})_{t \geq 0}.$$  

Then $F^{n,\tau}$ is a filtration (i.e. an increasing right continuous family of complete $\sigma$-fields). Let $\sigma$ be an $F^{n,\tau}$-stopping time. Then $(\sigma + \tau)$ is an $F^n$-stopping time, and by (4.22) (with $\sigma = \sigma + \tau$), (4.23), (4.24) and the definition of $Y^n(\lambda)$, we have

$$E[\phi^{n,\lambda}(X_{\sigma+\tau}) / \mathcal{F}^{n,\tau}_\sigma(n\lambda)] \leq 1.$$  

As $\sigma$ is an arbitrary $F^{n,\tau}$-stopping time, by Lemma 3.6 we conclude from the latter that for any $\lambda \in \mathbb{R}^d$, $\eta > 0$, $\delta > 0$ and $\alpha > 0$,

$$P \left( \sup_{t \leq \delta} \langle \lambda, X_t^{n,\tau} \rangle > |\lambda| \eta \right) \leq \exp(\alpha \eta |\lambda|) + P \left( \sup_{t \leq \delta} \frac{1}{n} \log \mathcal{F}_t^{n,\tau}(n\lambda) \geq |\alpha| |\lambda| \right).$$  \hfill (4.25)

By (4.24),

$$\sup_{t \leq \delta} \frac{1}{n} \log \mathcal{F}_t^{n,\tau}(n\lambda) \leq \frac{1}{n} \log \mathcal{F}_\tau^n(n\lambda) - G_t(\lambda)$$

$$+ \sup_{t \leq \delta} \frac{1}{n} \log \mathcal{F}_{t+\tau}(n\lambda) - G_{t+\tau}(\lambda) + \sup_{t \leq \delta} |G_{t+\tau}(\lambda) - G_t(\lambda)|.$$  \hfill (4.26)

Since $G_t(\lambda)$ is continuous in $t$, for all sufficiently small $\delta$,

$$\sup_{|t - s| \leq \delta, \delta \leq t, s \leq L + \delta} |G_t(\lambda) - G_s(\lambda)| \leq \frac{1}{2} |\lambda|,$$

and (4.26) yields for these $\delta$ and $\tau \leq L$,

$$P \left( \sup_{t \leq \delta} \frac{1}{n} \log \mathcal{F}_t^{n,\tau}(n\lambda) \geq |\alpha| |\lambda| \right) \leq P \left( \sup_{t \leq L + \delta} \frac{1}{n} \log \mathcal{F}_t^n(n\lambda) - G_t(\lambda) \geq \frac{\alpha |\lambda|}{4} \right).$$

Substituting this into (4.25) and using $(\sup \mathcal{F}^{1/n})$ we obtain for $\lambda \neq 0$,

$$\lim_{\delta \to 0} \lim_{n \to \infty} \sup \left[ P \left( \sup_{t \leq \delta} \left\langle \frac{\lambda}{|\lambda|}, X_t^{n,\tau} \right\rangle > \eta \right) \right]^{1/n} \leq \exp(\alpha \eta |\lambda|).$$
Taking $\alpha = \eta/2$ and $|\lambda| \to \infty$, and using (4.23) we arrive at (4.21). The C-exponential tightness of $(\mathcal{L}(X^n))$ is proved.

Now we prove the 'finite-dimensional LD convergence' along $U = \mathbb{R}^+ \setminus \{0\}$. Let $0 = t_0 < t_1 < \cdots < t_k$ and $\lambda_1, \ldots, \lambda_k \in \mathbb{R}^d$. Denote $m = d \times k$, $\lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^m$, $Y^n = (X^n_{t_1} - X^n_0, \ldots, X^n_{t_k} - X^n_{t_{k-1}})$, $Z^n(\lambda) = (Z^n_1(\lambda_1), \ldots, Z^n_k(\lambda_k))$, where $Z^n_i(\lambda_i)$ are the same as in Lemma 4.2. Then by (4.6),

$$\left[ P(\|Y^n - Z^n(\lambda)\| > \epsilon) \right]^{1/n} \leq \left[ P\left( \min_{1 \leq i \leq k} \tau^n(\lambda_i) \leq t \right) \right]^{1/n} \to 0 \quad (n \to \infty).$$

Applying Lemma 4.2 we see that $(Y^n, n \geq 1)$ and $(Z^n(\lambda), n \geq 1)$, $\lambda \in \mathbb{R}^m$, satisfy the conditions of Lemma 3.5 with

$$K(\lambda) = \sum_{i=1}^k (G_{t_i}(\lambda_i) - G_{t_{i-1}}(\lambda_i)).$$

$K(\lambda)$ itself also satisfies the conditions of Lemma 3.5 since its conjugate (see (2.5))

$$L(x) = \sum_{i=1}^k H_{t_{i-1}, t_i}(x_i), \quad x = (x_1, \ldots, x_k), \quad x_i \in \mathbb{R}^d,$$

is strictly convex on $\text{ri}(\text{dom} L)$ because each of $H_{t_{i-1}, t_i}$ is strictly convex on $\text{ri}(\text{dom} H_{t_{i-1}, t_i})$ and $\text{ri}(\text{dom} L) = \bigcap_{i=1}^k \text{ri}(\text{dom} H_{t_{i-1}, t_i})$.

Therefore, by Lemma 3.4,

$$\mathcal{L}(Y^n) \xrightarrow{1.d.} \exp(-L) \quad (n \to \infty).$$

The contraction principle (Varadhan, 1966, 1984; Deuschel and Stroock, 1989) then gives by the definition of $Y^n$ that

$$\mathcal{L}((X^n_{t_1} - X^n_0, \ldots, X^n_{t_k} - X^n_0)) \xrightarrow{1.d.} \exp(-I_{t_1, \ldots, t_k}) \quad (n \to \infty),$$

where (see (4.27))

$$I_{t_1, \ldots, t_k}(x_1, \ldots, x_k) = \sum_{i=1}^k H_{t_{i-1}, t_i}(x_i - x_{i-1}), \quad x_i \in \mathbb{R}^d, \quad x_0 = 0.$$

Since $X^n_0 \xrightarrow{P^{1.n}} 0$, by Lemma 3.3,

$$\mathcal{L}(X^n_{t_1}, \ldots, X^n_{t_k}) \xrightarrow{1.d.} \exp(-I_{t_1, \ldots, t_k}) \quad (n \to \infty),$$

verifying the finite-dimensional LD convergence.

According to Theorem A, it is left to prove that for $X = (X_t)_{t \geq 0} \in D(\mathbb{R}^d)$, $I(X) = \sup_{0 < t_1 < \cdots < t_k} I_{t_1, \ldots, t_k}(X_{t_1}, \ldots, X_{t_k}),$ (4.29)

$$I(X) = \sup_{0 < t_1 < \cdots < t_k} I_{t_1, \ldots, t_k}(X_{t_1}, \ldots, X_{t_k}),$$

where $I$ is defined in (2.3).

If $X_0 = 0$, then (4.29) easily follows from (2.2) and (2.3), on the one hand, and (2.5 and (4.28), on the other hand.
Now assume that $|X_0| > 0$. Then for a $e \in \mathbb{R}^d$ with $|e| = 1$ we have $\langle e, X_0 \rangle > \varepsilon > 0$, and by the right continuity of $X = (X_t)_{t \geq 0}$, for $t$ small enough, $\langle e, X_t \rangle > \varepsilon/2$. Therefore, for these $t$ and all $A > 0$,

$$I_t(X_t) = \sup_{\lambda \in \mathbb{R}^d} (\langle \lambda, X_t \rangle - G_t(\lambda)) \geq \langle Ae, X_t \rangle - G_t(Ae) \geq \frac{1}{2} Ae - G_t(Ae).$$

Since $G_t(\lambda)$ is continuous in $\lambda$ and $G_0(\lambda) = 0$, we obtain

$$\liminf_{t \to 0} I_t(X_t) \geq \frac{1}{2} Ae,$$

which means by the arbitrariness of $A$ that

$$\sup_{t > 0} I_t(X_t) = \infty.$$

Thus, the right hand side of (4.29) is equal to $\infty$. The theorem is proved.

**Remark 1.** From Theorem 2.1 and the contraction principle it follows that in fact the finite-dimensional LD convergence holds with $U = \mathbb{R}_+^d$.

**Remark 2.** As it is seen, we do not use in the proof Lemma 2.1 because, by Theorem A, $I$ is necessarily a rate function on $D(\mathbb{R}^d)$. That is why we defer proof of Lemmas 2.1 and 2.2 till Section 5. Still we retain Lemma 2.1 in the paper because, first, it does not require condition (G); the rate function $I$ from (2.3) appears in Theorems 2.2 and 2.3 as well, and, third, the proof is very simple.

Now, the assertion of Theorem 2.2 is a direct consequence of Theorem 2.1 and the following.

**Lemma 4.3.** Under the conditions of Theorem 2.2 the process $(e^{\langle \lambda, X_t \rangle - X_{t+}G(\lambda)}^\delta \mathbb{E}^\nu(G^\nu(\lambda)))_{t \geq 0}$ is an $\mathcal{F}^\nu$-local martingale for all $\lambda \in \mathbb{R}^d$.

The proof is analogous to that of Jacod and Shiryaev (1987, Th. II.2.47(a)) (see also Liptser and Shiryaev, 1989, Theorem 4.3.1).

Theorem 2.3 also trivially follows from Theorem 2.1 since under its conditions $(e^{\langle \lambda, X_t \rangle - X_{t+}G(\lambda)}^\delta \mathbb{E}^{\lambda}(e^{\langle \lambda, X_t \rangle - X_{t+}\lambda}))_{t \geq 0}$ is a martingale.

5. Proof of Lemmas 2.1 and 2.2

**Proof of Lemma 2.1.** Denote $\Phi(a) = \{X \in D(\mathbb{R}^d): I(X) \leq a\}$, $a \geq 0$. By the definition of $I$ (see (2.2) and (2.3)) we have for $X \in \Phi(a)$ and $\lambda \in \mathbb{R}^d$, $\lambda \neq 0$,

$$\left(\frac{\lambda}{|\lambda|}, X_t - X_s \right) \leq \frac{a}{|\lambda|} + \frac{G_t(\lambda) - G_s(\lambda)}{|\lambda|}, \quad s \leq t. \quad (5.1)$$

By the continuity of $G_t(\lambda)$ in $t$ and since $\lambda$ is arbitrary, we conclude that, for all $e \in \mathbb{R}^d$ with $|e| = 1$ and all $T > 0$,

$$\lim_{\delta \to 0} \sup_{X \in \Phi(a)} \sup_{|t - s| \leq \delta, 0 \leq s, t \leq T} |\langle e, X_t - X_s \rangle| = 0.$$
which obviously implies that
\[\limsup_{\delta \to 0} \sup_{X \in \mathcal{F}(a)} |X_t - X_s| = 0. \quad (5.2)\]

From (5.1) it also easily follows that \(\sup_{X \in \mathcal{F}(a)} |X_t| < \infty\). This and (5.2) show that \(\mathcal{F}(a)\) is compact in the locally uniform topology on \(D([0,T])\) and hence in the Skorohod topology. The lemma is proved. \(\square\)

For Lemma 2.2, we need an auxiliary result.

**Lemma 5.1.** Let \(f(t, \lambda), t \geq 0, \lambda \in \mathbb{R}^d\), be a \(\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R}^+)\)-measurable real-valued function, continuous in \(\lambda\) for all \(t \geq 0\), \(f(t, 0) = 0\). Assume that \(f(t, \lambda)\) is locally majorized by an integrable function: for all \(T > 0, N > 0\), there exists a \(\mathcal{B}(\mathbb{R}^+)\)-measurable, real-valued nonnegative function \(b_{N,T}(t)\) such that

\[|f(t, \lambda)| \leq b_{N,T}(t), \quad \forall t \leq T, \quad |\lambda| \leq N,\]

and

\[\int_0^T b_{N,T}(t) q(dt) < \infty.\]

Then for all \(T > 0,\)

\[\int_0^T \sup_{\lambda \in \mathbb{R}^d} f(t, \lambda) q(dt) = \sup_{(\lambda(t)) \in A_0} \int_0^T f(t, \lambda(t)) q(dt), \quad (5.3)\]

where \(A_0\) is defined in Section 2.

**Proof.** Denote

\[F(t) = \sup_{\lambda \in \mathbb{R}^d} f(t, \lambda). \quad (5.4)\]

First of all, \(F(t)\) is \(\mathcal{B}(\mathbb{R}^+)\)-measurable and nonnegative, so that the integral on the left-hand side of (5.3) is well defined (the argument is the same as in the remark after Lemma 2.2).

Let \(A\) be the set of all \(\mathcal{B}(\mathbb{R}^+)\)-measurable \(\mathbb{R}^d\)-valued functions \((\lambda(t), t \geq 0)\) and \(A^+\), its subset of the functions \((\lambda(t), t \geq 0)\) with the property \(f(t, \lambda(t)) \geq 0\) q.a.e. We prove (5.3) by proving in succession that

\[\int_0^T F(t) q(dt) = \sup_{(\lambda(t)) \in A^+} \int_0^T f(t, \lambda(t)) q(dt), \quad (5.5)\]

and

\[\sup_{(\lambda(t)) \in A^+} \int_0^T f(t, \lambda(t)) q(dt) = \sup_{(\lambda(t)) \in A_0} \int_0^T f(t, \lambda(t)) q(dt). \quad (5.6)\]
For (5.5), take an arbitrary \( \varepsilon > 0 \) and introduce the set \( A_\varepsilon = \left\{(t, \lambda) \in [0, T] \times \mathbb{R}^d : f(t, \lambda) \geq (F(t) - \varepsilon)^+ \land \frac{1}{\varepsilon}\right\} \).

Then \( A_\varepsilon \in \mathcal{B}_d([0, T]) \otimes \mathcal{B}(\mathbb{R}^d) \) and the projection \( \pi(A_\varepsilon) \) of \( A_\varepsilon \) on \([0, T]\) defined as

\[
\pi(A_\varepsilon) = \{t \in [0, T] : (t, \lambda) \in A_\varepsilon \text{ for some } \lambda \in \mathbb{R}^d\}
\]

coincides with \([0, T]\) by the definition of \( F(t) \).

Therefore by an obvious consequence of the cross-sectional theorem (Dellacherie, 1972, I-T37), there exists an \( \mathbb{R}^d \)-valued \( \mathcal{B}_d(\mathbb{R}^d) \)-measurable function \( \lambda_\varepsilon(t) \) such that its graph \([\lambda_\varepsilon] \subset A_\varepsilon\), i.e.

\[
f(t, \lambda_\varepsilon(t)) \geq (F(t) - \varepsilon)^+ \land \frac{1}{\varepsilon}, \quad t \leq T. \tag{5.7}
\]

Now, let \( \hat{\lambda}_\varepsilon(t) \) be a function which is \( \mathcal{B}(\mathbb{R}^d) \)-measurable and coincides with \( \lambda_\varepsilon(t) \) \( q \)-almost everywhere on \([0, T]\). Obviously, \( (\hat{\lambda}_\varepsilon(t)) \in A^+ \). Then from (5.7) we have

\[
\int_0^T f(t, \hat{\lambda}_\varepsilon(t))q(dt) \geq \int_0^T (F(t) - \varepsilon)^+ \land \frac{1}{\varepsilon}q(dt). \tag{5.8}
\]

As \( \varepsilon > 0 \) is arbitrary, (5.5) is proved.

To obtain (5.6), it is sufficient to prove that for any bounded \((\lambda(t)) \in A \) and \( \varepsilon > 0 \) there exists \((\mu_e(t)) \in A_0 \) such that

\[
\left| \int_0^T f(t, \lambda(t))q(dt) - \int_0^T f(t, \mu_e(t))q(dt) \right| < \varepsilon.
\]

By the local majoration condition and the continuity of \( f(t, \lambda) \) in \( \lambda \), we have applying the Lebesgue dominated convergence theorem that the set of such \((\lambda(t)) \in A \) is closed under bounded pointwise convergence. Therefore, by the monotone class theorem (see, e.g., Dellacherie and Meyer, 1978, p. 15) it contains all the bounded functions measurable with respect to the \( \sigma \)-field generated by the functions from \( A_0 \). The lemma is proved. \( \square \)

**Proof of Lemma 2.2.** First, consider the case \( dX \ll dq, X_0 = 0 \). We apply Lemma 5.1 with

\[
f(t, \lambda) = \left\langle \lambda, \frac{dX}{dq}(t) \right\rangle - g_t(\lambda).
\]

All the conditions of Lemma 5.1 hold. In particular, we can take

\[
b_{N,T}(t) = N \left| \frac{dX}{dq}(t) \right| + \sup_{|\lambda| \leq N} |g_t(\lambda)|.
\]
Therefore, by Lemma 5.1 and in view of (2.2) and (2.3),
\[
I(X) = \int_0^\infty \sup_{\lambda \in \mathbb{R}^d} \left( \lambda \cdot \frac{dX}{dq}(t) - g_t(\lambda) \right) q(dt).
\]

If \(X_0 \neq 0\), then taking
\[
\lambda(t) = N \frac{X_0}{|X_0|} \mathbf{1}_{[0]}(t),
\]
we have by (2.3)
\[
I(X) \geq \int_0^\infty \left[ \lambda(t) dX_t - dG_t(\lambda(t)) \right] = N |X_0| \to \infty \quad (n \to \infty).
\]

Finally, assume that \(X_0 = 0\) but \(X\) is not absolutely continuous with respect to \(q\) on an interval \([0, T]\). Then we can choose \(\varepsilon > 0\) such that for every \(\delta > 0\) there exist
\[
0 \leq t_1 < \ldots < t_{2l} \leq T
\]
satisfying
\[
\sum_{i=1}^{l} q([t_{2i-1}, t_{2i}]) < \delta, \quad \sum_{i=1}^{l} |X_{t_{2i}} - X_{t_{2i-1}}| > \varepsilon.
\]

(5.9)

For \(N > 0\), take
\[
\lambda_N(t) = N \sum_{i=1}^{l} \frac{X_{t_{2i}} - X_{t_{2i-1}}}{|X_{t_{2i}} - X_{t_{2i-1}}|} \mathbf{1}_{[t_{2i-1}, t_{2i}]}(t)
\]
(of course we may assume that \(|X_{t_{2i}} - X_{t_{2i-1}}| > 0\)).

Then by (2.2), (2.3) and (5.8),
\[
I(X) \geq \int_0^\infty \left[ \lambda_N(t) dX_t - dG_t(\lambda_N(t)) \right]
\]
\[
= N \sum_{i=1}^{l} |X_{t_{2i}} - X_{t_{2i-1}}| \mathbf{1}_{[t_{2i-1}, t_{2i}]}(t) - \sum_{i=1}^{l} \int_{t_{2i-1}}^{t_{2i}} g_t(\lambda_N(t)) q(dt)
\]
\[
> N \varepsilon - \int_0^T \sup_{|\lambda| \leq N} |g_t(\lambda)||\mathbf{1}_{[t_{2i-1}, t_{2i}]}(t)| q(dt).
\]

(5.10)

By (5.8), the latter integrand goes to 0 in measure \(q\) as \(\delta \to 0\), and so by the Lebesgue-dominated convergence theorem,
\[
\lim_{\delta \to 0} \int_0^T \sup_{|\lambda| \leq N} |g_t(\lambda)||\mathbf{1}_{[t_{2i-1}, t_{2i}]}(t)| q(dt) = 0,
\]
and (5.9) gives \(I(X) > N \varepsilon\) for arbitrary \(N\). The lemma is proved. \(\square\)

6. Concluding remarks

In conclusion, we summarize some of the parallels that we found between large deviations and weak convergence of stochastic processes. These parallels, some of which were, as far as we know, first pointed out by Lynch and Sethuraman (1987), are
Table 1.

<table>
<thead>
<tr>
<th>Weak convergence</th>
<th>Large convergence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tightness</td>
<td>Exponential tightness</td>
</tr>
<tr>
<td>Convergence in probability</td>
<td>Superexponential convergence in probability</td>
</tr>
<tr>
<td>Uniform integrability</td>
<td>Uniform exponential integrability</td>
</tr>
<tr>
<td>Continuous mapping theorem</td>
<td>Contraction principle</td>
</tr>
<tr>
<td>Method of characteristic functions</td>
<td>Gärtner's theorem</td>
</tr>
<tr>
<td>Prohorov's theorem</td>
<td></td>
</tr>
<tr>
<td>Method of finite-dimensional distributions</td>
<td></td>
</tr>
<tr>
<td>Method of stochastic exponentials</td>
<td>Analogues for large deviations</td>
</tr>
</tbody>
</table>

encountered both at the level of concepts and at the level of results as presented in Table 1.

It is most remarkable that the underlying ideas are also very similar which can be seen either in this paper or in Puhalskii (1991) and also in Puhalskii (1993, 1994, to appear). As it follows from the results of O'Brien and Vervaat (1991), this analogy is deep rooted and one can develop a theory treating the large deviation principle and weak convergence theory approaches (along with other techniques available) in other large-deviation settings.

Acknowledgements

This work was carried out in part during the author's visit to the Computing Science Department, University of Newcastle-upon-Tyne. The author would like also to thank the anonymous referee for well pointed remarks.

References

C. Dellacherie, Capacités et Processus Stochastiques (Springer, Berlin, 1972)