On large deviations for SDEs with small diffusion and averaging

A.Yu. Veretennikov *

Institute for Information Transmission Problems, Bolshoy Karetnii 19, 101447, Moscow, Russia

Received 28 May 1998; received in revised form 12 January 2000; accepted 7 February 2000

Abstract

A large deviation principle is established for stochastic differential equation systems with slow and fast components and small diffusions in the slow component. © 2000 Elsevier Science B.V. All rights reserved.

MSC: 60F10; 60J60

Keywords: SDE; Averaging; Small perturbation; Large deviations

1. Introduction

We consider the SDE system of dimension $d + \ell$,

\[ \begin{align*}
  dX^i_t &= f(X^i_t, Y^i_t)\, dt + \varepsilon \sigma_1(X^i_t, Y^i_t)\, dw^1_t + \sigma_3(X^i_t, Y^i_t)\, dw^3_t, \quad X^i_0 = x, \\
  dY^i_t &= -2B(X^i_t, Y^i_t)\, dt + \varepsilon^{-1}(C_1(Y^i_t)\, dw^1_t + C_2(Y^i_t)\, dw^2_t), \quad Y^i_0 = y
\end{align*} \]

as $0 \leq t \leq 1$. Here $(w^1, w^2, w^3)$ is a 3d-dimensional Wiener process, $\varepsilon > 0$ is a small parameter, $x \in \mathbb{R}^d, y \in \mathbb{R}^\ell$, $f \in B(\mathbb{R}^{d+\ell}; \mathbb{R}^d), \sigma_1 \in B(\mathbb{R}^{d+\ell}; \mathbb{R}^{d \times d}), \sigma_3 \in B(\mathbb{R}^{d+\ell}; \mathbb{R}^{d \times d}), B \in B_{\text{loc}}(\mathbb{R}^{d+\ell}; \mathbb{R}^\ell), C_1 \in B(\mathbb{R}^\ell; \mathbb{R}^{\ell \times d}), C_2 \in B(\mathbb{R}^\ell; \mathbb{R}^{\ell \times d}); B (B_{\text{loc}})$ denotes the bounded (locally bounded) Borel functions. $\sigma_{1,3}(\cdot, y)$ and $B(\cdot, y)$ are continuous in the first variable for any $y$, linear growth and Lipschitz conditions are assumed

\[ |B(x, y)| \leq L(1 + |y|), \]

\[ |f(x, y) - f(x', y')| \leq L(|x - x'| + |y - y'|), \]

the function $CC^* = C_1 C_1^* + C_2 C_2^*$ is continuous and $C_2 C_2^*$ is uniformly nondegenerate. Moreover, either the matrix $(\sigma_1, \sigma_3)(\sigma_1, \sigma_3)^*$ is nondegenerate and continuous or all coefficients of the system are Lipschitz ones. At any rate, under such conditions system (1) has a solution unique in law.

---

* This work was supported by grants RFBR 98-01-00062, INTAS 93-1585-EXT and INTAS 94-0378.

Tel.: +7-095-299-9415; fax: +7-095-209-0579.

E-mail address: ayu@sci.lpi.ac.ru (A.Yu. Veretennikov).
The problem of large deviations for the slow component $X^\varepsilon$ in the space $C([0,1]; \mathbb{R}^d)$ as $\varepsilon \to 0$ is under consideration. It is interesting in the theory of homogenization, in KPP equation theory, etc.

In a partial case of $\sigma_1 = \sigma_3 \equiv 0$, such a problem was considered in Freidlin (1978), Freidlin and Wentzell (1984), Veretennikov (1990, 1992, 1994, 1997, 1999). The problem with a small diffusion in a slow component and with independent Wiener processes $w^1, w^2$ (i.e. with $C_1 = \sigma_3 \equiv 0$) which is also a partial case of (1) was studied recently in Liptser (1996) and Veretennikov (1997). Different but close problems with another time scales were studied in Makhno (1995) for periodic coefficients. Finally, the problem with identical Wiener processes (i.e. only $w^1$ in both components), unit matrices $C_1$ and $\sigma_1$ and under a periodicity assumption, that is, if a fast component takes its values in the torus, i.e., a compact manifold, the large deviation principle was announced in Kleptsyna et al. (1996). In this paper no periodic assumption is made and the noncompact state space is, indeed, one of the main difficulties. We assume the following stability assumption: for any $a > 0$,

$$\lim_{|x| \to \infty} \sup_{|\beta| < a} \sup_{x} ||CC^*(y)||^{-1}B^\beta(x,y) y|/y| = -\infty,$$

$$B^\beta(x, y) = B(x, y) + C_1(y)\sigma_1^*(x, y)\beta.$$

(2)

Here $B^\beta y$ is a scalar product of two vectors. If the matrix $CC^*$ is constant then assumption (2) is equivalent to the same equality with $\beta = 0$. Close assumptions were used earlier in Donsker and Varadhan (1983), Veretennikov (1992) and other papers.

We apply the approach based on Frobenius-type theorems for positive operators, cf. Veretennikov (1994,1997). For the compact state space the use of the Frobenius-type theorem was proposed by Freidlin. However, in noncompact spaces this idea does not work directly. In the previous papers by the author the noncompact state space case was mainly investigated for less general SDE systems. For this aim a compactification method via stopping times was proposed.

Note that in the periodic case there is no diffculty in applying a “standard” Frobenius-type theorem for compact operators. Our goal in this paper is again to establish the large deviation principle (LDP) for this general SDE system and in a noncompact state space which is, in fact, the main diffculty for one cannot apply the Frobenius-type theorem. The approach based on similar compactifications via stopping times is used.

We do not touch in this paper more general “full dependence” systems with a coefficient $C(x, y)$, that is, dependent on $x$. That problem was solved in the simplest “compact” case – i.e. $Y^\varepsilon_t$ on a compact manifold – with $\sigma_1 = \sigma_3 \equiv 0$, see Veretennikov (1999). It is therefore probable that in our problem the case $C(x, y)$ could be also done, though, one should expect a considerable change of all calculus. On the other hand, one can expect quite the same effect due to small diffusion terms $\varepsilon \sigma_1$ and $\varepsilon \sigma_3$. So we retain the present exposition with $C(y)$.

In the recent paper Freidlin and Sowers (1999) a solution of a similar problem is established for the “compact” periodic case and with $B = B(y)$. More precisely, the authors consider one equation with two small parameters of the form

$$dX_t = \varepsilon^{1/2} A(X_t/\delta) dW_t + B(X_t/\delta) dt, \quad X_0 = x.$$
Our setting would correspond to “regime 2”, that is, \( \varepsilon^2 = c\delta \) had we considered a periodic case and denoted \( Y_t = \delta^{-1} X_t \); this is exactly the case studied in Kleptsyna et al. (1996). One can also find in Freidlin and Sowers (1999) the investigation of other regimes and applications to the KPP equation theory. The results of the present paper can be also applied to a wavefront propagation.

2. Main result

Let \( \tilde{w}_t^i = \varepsilon^{-1} w_{t/\varepsilon^2}^i, \ i = 1, 2 \) and \( y_t^{\psi, \beta} \) denote a solution of an SDE

\[
dy_t = B^{(t)}(x, y_t) dt + (C_1(y_t) d\tilde{w}_t^1 + C_2(y_t) d\tilde{w}_t^2), \quad y_0 = y.
\]

Assumption (2) provides, in particular, the existence of an invariant probability measure \( x \) of the Markov process \( y_t, t \geq 0 \). This invariant measure is unique due to the assumptions on matrix \( C \). We will also use the notation \( y_t^{\psi, \beta} \) for the solution of the equation

\[
dy_t = (B^{(t)}(\psi_t, y_t) dt + (C_1(y_t) d\tilde{w}_t^1 + C_2(y_t) d\tilde{w}_t^2), \quad y_0 = y
\]
given nonrandom \( \psi_t, 0 \leq t < \infty \).

The family of processes \( X^\varepsilon \) satisfies a large deviation principle (LDP) in the space \( C([0, 1]; \mathbb{R}^d) \) with a normalizing coefficient \( \varepsilon^{-2} \) and a rate function \( S(\phi) \) if three conditions are satisfied:

\[
\limsup_{\varepsilon \to 0} \limsup_{\delta \to 0} \varepsilon^2 \log P\left( X^\varepsilon \in F \right) \leq - \inf_{\phi \in F} S(\phi), \quad \forall F \text{ closed}, \tag{3}
\]

\[
\liminf_{\varepsilon \to 0} \liminf_{\delta \to 0} \varepsilon^2 \log P\left( X^\varepsilon \in G \right) \geq - \inf_{\phi \in G} S(\phi), \quad \forall G \text{ open}, \tag{4}
\]

and \( S \) is a “good” rate function, that is, for any \( s \geq 0 \) the set

\[
\Phi(s) := \{ \phi \in C([0, 1]; \mathbb{R}^d) : S(\phi) \leq s, \ \phi(0) = x \}
\]
is compact in \( C([0, 1]; \mathbb{R}^d) \) for any \( s \geq 0 \).

Suppose \( S \) is a “good” rate function. Then it is known that estimates (3) and (4) are equivalent to the following two bounds, for any \( s \geq 0, \ \phi \in C([0, 1]; \mathbb{R}^d) \):

\[
\limsup_{\varepsilon \to 0} \limsup_{\delta \to 0} \varepsilon^2 \log P(\rho(X^\varepsilon, \Phi(s)) \geq \delta) \leq -s
\]

(\( \rho \) is the distance in \( C([0, 1]; \mathbb{R}^d) \)) and

\[
\liminf_{\varepsilon \to 0} \liminf_{\delta \to 0} \varepsilon^2 \log P(\rho(X^\varepsilon, \phi) \leq \delta) \geq -S(\phi),
\]

see Freidlin and Wentzell (1984, Theorem 3.3.2).

**Theorem 1.** Let assumption (2) be satisfied. Then the family of processes \( X^\varepsilon \) satisfies the LDP in \( C([0, 1]; \mathbb{R}^d) \) with a normalizing coefficient \( \varepsilon^{-2} \) and a rate function

\[
S(\phi) = \int_0^1 L(\phi_t, \dot{\phi}_t) dt,
\]
where
\[ L(x, \alpha) = \inf_{\beta} (x\beta - H(x, \beta)), \]
\[ H(x, \beta) = \lim_{t \to \infty} t^{-1} \log \mathbb{E} \exp \left( \int_0^t f^\beta(x, y^\alpha_s) \, ds \right), \]
\[ f^\beta(x, y) = \beta f(x, y) + (|\sigma_1^\beta(x, y)|^2 + |\sigma_3^\beta(x, y)|^2)/2. \]

The limit \( H(x, \beta) \) exists, the functions \( H \) and \( L \) are convex in \( \beta \) and \( \alpha \) variables correspondingly, \( L \) is nonnegative and \( H \) is differentiable in the origin.

Note the different types of dependence of the answer on different diffusion coefficients. Functions \( \sigma_1 \) and \( \sigma_3 \) are included similar to the expression for \( f^\beta \) while \( B^\beta \) only depends on \( \sigma_1 \).

Remark 2. The differentiability of \( H \) in the point \( \beta = 0 \) is not used in the proof of the LDP. So we only mention that this property can be shown by arguments from Ellis (1985), cf. Gulinsky and Veretennikov (1993, Remark 8.1).

Remark 3. \( S(\phi) = 0 \) iff \( \phi \) is a solution of the equation \( \phi_t = x + \int_0^t \tilde{f}(\phi_s) \, ds \), where \( \tilde{f}(x) = f(x, y) \mu(d'y) \). One can show easily that \( \tilde{f} \) is bounded and continuous, hence, there exists a solution of the latter equation. In general, the rate function is equal to zero on any solution and the LDP describes deviations from the set of all solutions.

3. Auxiliary results

Let us denote \( \tau_0 = 0, \quad \tau_1 = \inf(k = 1, 2, \ldots: |y^\alpha_{k-1}| \leq R, |y^\alpha_k| \leq R), \quad \tau_{n+1} = \inf(k = \tau_n + 1, \tau_n + 2, \ldots: |y^\alpha_k| \leq R, |y^\alpha_{k+1}| \leq R), \quad n(t) = \sup(n = 0, 1, 2, \ldots: \tau_n \leq t). \) The following lemmas are standard, see Veretennikov (1992, 1994). The news is that \( y^\alpha_k \) now depends on \( \beta \) and the bounds should be uniform in \( |\beta| < a \) for any \( a \).

Lemma 1. Let assumption (2) be satisfied. Then for any \( a, c, v > 0 \) and \( C > 1 \) there exists such \( R_0 \) that for any \( R \geq R_0 \), any \( |\beta| < a \) and any \( x, y \)
\[ E_y \exp(c\tau_1) \leq C \exp(v(|y| - R)_+). \]

Lemma 2. Let assumption (2) be satisfied. Then for any \( a, \delta, C > 0 \) there exist such \( R_0, t_0 \) that for any \( R \geq R_0 , t \geq t_0, \ |\beta| < a \) the following estimate holds:
\[ P(|t^{-1}n(t) - v_R| > \delta) \leq \exp(-Ct), \]
where
\[ \inf_{|\beta| < a} \lim_{R \to \infty} v_R = 1. \]

For any \( a, \delta, C, v > 0 \) for \( R \) large enough, any \( |\beta| < a \) and any \( t \geq t(\delta), y, x \)
\[ P(|t^{-1}n(t) - v_R| > \delta) \leq \exp(-Ct) \exp(v(|y| - R + 1)_+). \]
Denote $B_R = \{ y : |y| \leq R \}$, $t^\beta_t = t_{[\nu_R]}$, $C(B_R) = \{ \text{the space of all continuous functions on } B_R \}$, $C^+(B_R)$ is the cone of nonnegative functions from $C(B_R)$. For any $R$ let us consider the semigroup of operators $T_t^\beta$, $t > 0$ on $C(B_R)$ defined by the formula

$$T_t^\beta g(y) = E_y g(y_{t^\beta t}^\beta) \exp \left( \int_0^{t^\beta t} f^\beta(x, y_s^\beta) \, ds \right).$$

Operators $T_t^\beta$ are well-defined on $C(B_R)$ for any $a > 0$ and any $|\beta| < a$ if $R$ is large enough by virtue of Lemma 1. A positive operator $T$ is called 1-bounded if $f > 0$, $f \neq 0$ implies $C(f)^{-1} \leq Tf \leq C(f)$, cf. Krasnosel’skii et al. (1989).

**Lemma 3.** Let assumption (2) be satisfied. Then operator $T_1^\beta$ is 1-bounded if $|\beta| < a$ and $R \geq R_a$.

One can add that $C(f)$ may be chosen uniform w.r.t. $|\beta| < a$.

**Lemma 4.** Let assumption (2) be satisfied. Then the spectral radius $r(T_1^\beta)$ is a simple eigenvalue of $T_1^\beta$ separated from the rest of the spectrum and its eigenfunction $e_R^\beta$ belongs to cone $C^+(B_R)$. Moreover, function $r(T_1^\beta)$ is differentiable and convex in $\beta$ ($|\beta| < a$, $R \geq R_a$) and function $e_R^\beta$ is strictly positive (that is, separated from zero) uniformly in $|\beta| < a$.

**Lemma 5.** Let $a > 0$, $\beta \in E', |\beta| < a$, and let assumption (2) be satisfied. Then there exists such $R_a > 0$ that for any $R \geq R_a$ there exists a limit

$$\mu_R(x', x, \beta) = \lim_{t \to 0} t^{-1} \log E_y \exp \left( \int_0^{t^\beta_t} f^\beta(x', y_s^\beta) \, ds \right).$$

Function $\mu_R(x', x, \beta)$ is differentiable in $\beta$ for $|\beta| < a$, $R \geq R_a$, and convex in $\beta$. There exists such a $C(R, a)$ that for any $y$, $|\beta| < a$ and $t \geq t(y)$

$$t^{-1} \log E_y \exp \left( \int_0^{t^\beta_t} f^\beta(x', y_s^\beta) \, ds \right) - \mu_R(x', x, \beta) \leq C(R) t^{-1}.$$

**Lemma 6.** Let assumption (2) be satisfied. Then for any $a > 0$ for $R \geq R_a$ function $\mu_R$ is uniformly continuous in $(x', x, \beta)$, $|\beta| < a$.

4. **Proof of Theorem 1**

(A) Let us prove the existence of limit (5). Fix $\beta \in E'$, $|\beta| < a$. For $R$ large enough we have

$$t^{-1} \log E \exp \left( \int_0^{t^\beta_t} f^\beta(x, y_s^\beta) \, ds \right)$$

$$= \nu_R(v_R t)^{-1} \log E_y \exp \left[ \left( \int_0^{t^\beta_t} + \int_0^{t^\beta_t} \right) f^\beta(x, y_s^\beta) \, ds \right].$$
By virtue of Lemmas 2, 5 and using Hölder’s inequality, one obtains similar to Veretennikov (1994), the proof of Theorem 1, that for any $a > 0$, $\delta > 0, x, y, \lambda > 0$ there exist such constants $\lambda_0, R, C(R), \kappa > 0$ that for any $|x| \leq a$, $y, t \geq t_0$,

$$-\delta - C(R)t^{-1} + t^{-1} \log \{\exp(-Kt); 1 - \exp(-\lambda t)h_R(y)\} \leq t^{-1} \log E_\gamma \left( \int_0^t f^\beta(x, y_s^\gamma) \, ds \right) - \nu_Rh_R(x, \beta)$$

$$\leq \delta + C(R)t^{-1} + t^{-1} \log \min\{\exp(-Kt); 1 + \exp(-\lambda t)h_R(y)\},$$

where $K = ||f||_b$, $h_R(y) = \exp(\kappa(|y| - R)_+)$. Since $\nu_Rh_R$ does not depend on $t$ while the other terms do not depend on $R$, one obtains the existence of the limit $H$ in (5) and its continuity. It is easy to show that $H$ is convex. Notice that $\nu_R$ depends also on $\beta$, however, $\lim_{R \to \infty} \nu_R = 1$ uniformly in $|\beta| \leq a$.

(B) Let us choose a small $\Delta > 0$ s.t. $T/\Delta = m$ is an integer. Let $\psi_\gamma$ be a right-continuous stepwise function with values in $\mathbb{R}^d$, $\phi_\gamma$ be a smooth function with values in $\mathbb{R}^d$ such that $\rho_{0Y}(\phi, \phi^\gamma) < \lambda$ and $\rho_{0T}(\phi, \psi) < \lambda$.

In the sequel we omit the index $\tilde{\varepsilon}$ in $X, Y$. Let

$$\tilde{X}_t = X_0 + \int_0^t f(\psi_s, Y_s) \, ds + \sigma_1(\psi_s, Y_s) \, dw_s + \sigma_2(\psi_s, Y_s) \, dw_s^1.$$ 

To compare $X^\varepsilon$ and $\tilde{X}^\gamma$ we use the bound for stochastic exponents with $\eta_t = (\sigma_1(X_t, Y_t) - \sigma_1(\psi_t, Y_t))$ and $\eta_t = (\sigma_2(X_t, Y_t) - \sigma_2(\psi_t, Y_t))$, $w = w^{1,3}$

$$P \left( \sup_{t \leq 1} \left| \int_0^t \eta_s \, dw_s \right| > \delta ; \sup_{t \leq 1} |\eta_s| \leq \delta \right) \leq 4 \exp(-\delta^2 \delta^{-2} \varepsilon^{-2}/4).$$

If $\delta$ is small enough w.r.t. $\delta'$ then the last expression can be made less than $\exp(-C\varepsilon^{-2})$ with any $C > 0$. So, since $f$ is Lipschitz, we get due to Gronwall’s inequality,

$$P((\omega; \rho_{0Y}(X, \phi) < \delta) \setminus (\omega; \rho_{0T}(\tilde{X}^\gamma, \phi^\gamma) < \delta')) \leq \exp(-C\varepsilon^{-2})$$

with any $C > 0$ if $\lambda$, $\delta$ and $\delta'$ are small enough w.r.t. $\delta'$.

(C) We introduce stopping times $\{\tau_{nR}^\gamma\}$ so that each random value $A_{\tau_R}^\gamma$ is close (cf. Lemma 2) to the nonrandom one $n\Delta$ while a certain auxiliary process which is close to $Y_t$ belongs to the ball $(|y| \leq R)$ for any $t = n\Delta, t = 1, 2, \ldots$. This is a sort of regularization after which one can use the results about positive operators described above in Lemmas 3 and 4. Let $A_{\tau}^0 = 0$, $A_{\tau}^1 = \varepsilon^2 t_{\Delta^2}^1$. Here $\tau^1_{nR}$ is a sequence of stopping times which is similar to the one defined in the beginning of the previous section. However, it is not completely the same and it is convenient to define it as follows:

$$\tau^0_{nR} = 0,$$

$$y^{(0)}_t = y + \int_0^t B^\beta(x_0, y^{(0)}_s) \, ds + \int_0^t C_1(y^{(0)}_s) \, dw_s + \int_0^t C_2(y^{(0)}_s) \, dw_s^1,$$

$$\tau^1_{nR} = \inf(k = t, t^1 + 1, t^1 + 2, \ldots; |y^{(0)}_k| \leq R, |y^{(0)}_k| \leq R), \quad n \geq 0.$$ 

Further, let

$$\tilde{X}^\gamma_t = x_0 + \int_0^t \left[ f(\psi_s, y^{(0)}_{s\varepsilon^2}) \, ds + \varepsilon(\sigma_1(\psi_s, y^{(0)}_{s\varepsilon^2}) \, dw_s + \sigma_2(\psi_s, y^{(0)}_{s\varepsilon^2}) \, dw_s^1) \right].$$
0 \leq t \leq A^1_R$. Denote $A^2_R = \mathbb{E}_t^2 \{ |A^2_t|^2 \}$, where
\[ \tau_n = A^1_R / k^2, \quad \tau^2 = \inf(k = 0, 1, 2, \ldots : |y^{(1)}_k| \leq R), \]
\[ y^{(1)}_t = y^{(0)}_t + \int_{A^1_R / k^2}^t B^0(x(1), y^{(1)}_s) \, ds \]
\[ + \int_{A^1_R / k^2}^t (C_1(y^{(1)}_s) \, dw^2_s + C_2(y^{(1)}_s) \, dw^3_s), \]
\[ \hat{X}^\psi_t = \hat{X}^\psi_{A^1_R} + \int_{A^1_R / k^2}^t f(x, y^{(1)}_s) \, ds + \sigma_1(x, y^{(1)}_s) \, dw^1_s + \sigma_2(x, y^{(1)}_s) \, dw^3_s \]
as $t \geq A^2_R$,
\[ \tau^2_{n+1} = \inf(k = \tau^2_n + 1, \tau^2_n + 2, \ldots : |y^{(1)}_k| \leq R, |y^{(1)}_k| \leq R), \quad n \geq 0, \ldots . \]

Notice that $\psi$ may not be constant between $A^k_R$ and $A^2_R$ and this was the reason to define the sequences $\{ \tau^2_n \}$ here instead of the use of the stopping times defined in Section 3.

(D) Let $\delta'' > 0$. Since $f$ is bounded and due to the same arguments as in step (B), we obtain for any $C > 0$,
\[ P((\omega: \rho_{0T}(\hat{X}^\psi, \varphi) < \delta') \setminus (\omega: \rho_{0T}(\hat{X}^\psi, \varphi^A) < \delta') \]
\[ \setminus (\omega: \cup_k |A^k_R - T_k| > \delta'' A)) < \exp(-Ce^{-2}) \]
if $\delta', \delta''$ and $A$ are small enough, where
\[ \hat{X}^{\psi, A} := (\hat{X}^\psi_{A^1_R}, \hat{X}^\psi_{A^2_R}, \ldots, \hat{X}^\psi_{A^m_R}), \quad \varphi^A = (\varphi_A, \ldots, \varphi_{mA}). \]

Moreover, by virtue of Lemma 2 for any $C > 0$, $P(\cup_k |A^k_R - T_k| > \delta'' A) \leq \exp(-Ce^{-2})$ if $R$ is large enough. Hence, it suffices to estimate from below the probability
\[ P(\omega: \rho_{0T}(\hat{X}^{\psi, A}, \varphi^A) < \delta'). \]

(E) Let $p > 1$, $p^{-1} + q^{-1} = 1$ and let $A, \delta_{m-1}'$ and $\varepsilon$ be small enough. Let us show the bound uniform in $|\varphi^{m-1}| \leq R, |\varphi^{m-1} - \psi_{m-1}| \leq \delta_{m-1}'$,
\[ P(|\hat{X}^\psi_{A^k_R} - \varphi^\xi_{T_k}| \hat{\tau}^\psi_{A^k_R} - \varphi^\xi_{T_k} = \varphi^{m-1}) \]
\[ \geq (\exp(-o(1)/\varepsilon^2))^{\psi/p} P(\hat{X}^\psi_{A^k_R} - \varphi^\xi_{T_k} < \delta' |\hat{X}^\psi_{A^k_R} - \varphi^{m-1} = \varphi^{m-1})^{1/p}, \]
where $o(1) \rightarrow 0$ as $A + \delta_{m-1}' \rightarrow 0$ ($o(1)$ may depend on $p$). Consider the exponent
\[ g = \exp \left( -e^{-1} \int_{A^k_R}^{A^k_{R-1}} \mathbb{E}^{-1}(\hat{Y}_s)(B^0(X_s, \hat{Y}_s) - B^0(\psi, \hat{Y}_s)) \, dw_s \right. \]
\[ - \frac{1}{2} e^{-1} \int_{A^k_{R-1}}^{A^k_R} \left. \mathbb{E}^{-1}(\hat{Y}_s)(B^0(X_s, \hat{Y}_s) - B^0(\psi, \hat{Y}_s)) \right| dw_s \right) \]
and a probability measure \( d\nu^0 = g \, d\nu \). We have,

\[
P(\|\bar{X}_t^\phi - \varphi_t^\phi \|_{L_2}^2 = \varphi_m - 1) = E^q \, g^{-1}(\|\bar{X}_t^\phi - \varphi_t^\phi \|_{L_2}^2 = \varphi_m - 1)
\]

\((E^q \) is the expectation w.r.t. \( P^q \)). Due to Hölder’s inequality

\[
E \zeta \eta \ge (E^q \zeta^1 \nu)^p (E^q \eta^{-q})^{-p/q}
\]

for \( \zeta, \eta \ge 0 \), we have for conditional expectations

\[
E^q(g^{-1}(\|\bar{X}_t^\phi - \varphi_t^\phi \|_{L_2}^2 < \delta'))(\bar{X}_t^\phi - \varphi_t^\phi \|_{L_2}^2 = \varphi_m - 1)
\]

\[
\ge (E^q(g^{\rho})|\bar{X}_t^\phi - \varphi_t^\phi \|_{L_2}^2 = \varphi_m - 1)^{-\rho/q}
\]

\[\times (E^q(I(\|\bar{X}_t^\phi - \varphi_t^\phi \|_{L_2}^2 < \delta'))|\bar{X}_t^\phi - \varphi_t^\phi \|_{L_2}^2 = \varphi_m - 1)^p.
\]

A similar upper bound also holds true. An easy calculus shows that

\[
(E^q(g^{\rho})|\bar{X}_t^\phi - \varphi_t^\phi \|_{L_2}^2 = \varphi_m - 1)) = \exp(o(1)/c^2),
\]

where \( o(1) \to 0 \) as \( A, \delta_{m-1}^\prime \to 0 \). So we get (6). The rest of the proof follows considerations in Veretennikov (1994, Proof of Theorem 2). However, a lot of important details should be changed. So we expose it, except a few standard points.

(F) Let \( \beta \in R^d \), \( |\beta| < a \), and let \( \delta_{m-1}^\prime \leq \delta \) be small enough. Denote

\[
h_{\epsilon, R}(\beta; \varphi_m) = \epsilon^2 \log E(\exp(\epsilon^{-2}\beta(\bar{X}_t^\phi - \bar{X}_t^\phi_{m-1})))|F_{\bar{X}_t^\phi_{m-1}}), \quad \varphi_m = \varphi_{m-1}^\phi.
\]

Due to Girsanov’s theorem,

\[
h_{\epsilon, R}(\beta; \varphi_m) = \epsilon^2 \log E \left[ \exp \left( \int_{\bar{X}_t^\phi_{m-1} \varphi_m}^{\varphi_m} \beta(\psi_s^\phi, y_s^\psi, \beta) \, ds \right) \right]
\]

\[
+ \epsilon \beta \int_{\bar{X}_t^\phi_{m-1} \varphi_m}^{\varphi_m} \sigma(\psi_s^\phi, y_s^\psi, \beta) \, ds
\]

\[
+ \epsilon \sigma(\psi_s^\phi, y_s^\psi, \beta) \int_{\bar{X}_t^\phi_{m-1} \varphi_m}^{\varphi_m} \beta(\psi_s^\phi, y_s^\psi, \beta) \, ds
\]

\[
\geq \epsilon^2 \int_{\bar{X}_t^\phi_{m-1} \varphi_m}^{\varphi_m} \sigma(\psi_s^\phi, y_s^\psi, \beta)^2 \, ds.
\]

for \( |\beta| < a \) and \( h_{\epsilon, R}(\beta; \varphi_m) = +\infty \) for any other \( \beta \). Let \( \psi_s = \psi_T, s \geq T \). Note that \( h_{\epsilon, R}(\beta; \varphi_m) < \infty \) for \( R \geq R(a) \) due to Lemma 3. By virtue of Lemma 5, for \( |\beta| < a \), \( R \geq R(a) \) there exist limits

\[
h_R(\beta) = \lim_{\epsilon \to 0} h_{\epsilon, R}(\beta; \varphi_m) = \int_{\bar{X}_t^\phi_{m-1} \varphi_m}^{\varphi_m} H_R(\psi_s, \beta) \, ds.
\]
uniformly in $|y^{m-1}| \leq R$ and

$$
H_R(\psi, \beta) = \lim_{\epsilon \to 0} \epsilon^2 \Delta^{-1} \log E \left( \exp \left( \int_{A_{m-1}^{\epsilon}/\epsilon} f^\beta(\psi, y^\beta, \beta) \, dx \right) \right)_{y_{A_{m-1}^{\epsilon}/\epsilon}^\beta = y^{m-1}}.
$$

(G) Consider a new probability measure on $\sigma$-field $F_{x_R}$,

$$
dP^\beta_R = \exp(-\epsilon^2 (\beta(\hat{X}_{x_R}^\psi - \vartheta^{m-1}) - \epsilon^2 \Delta H^\epsilon_{m-1}(\vartheta^{m-1}, y, \beta))) \, dP,
$$

where

$$
- \epsilon^2 \Delta H^\epsilon_{m-1}(\vartheta^{m-1}, y, \beta) = \log E(\exp(-\epsilon^2 \beta(\hat{X}_{x_R}^\psi - \vartheta^{m-1}))|F_{x_R}^\beta; \hat{X}_{x_R}^\psi = \vartheta^{m-1}, y_{x_R}^0 = y).
$$

With the obvious notation $E^\beta_R$, we get

$$
P(\hat{X}_{x_R}^\psi - \varphi^\beta_T < \delta'(\hat{X}_{x_R}^{m-1} = \vartheta^{m-1}, y_{x_R}^0 = y))
$$

$$
= E^\beta_R \{ I(\hat{X}_{x_R}^\psi - \varphi^\beta_T < \delta') \exp(-\epsilon^2 \beta(\hat{X}_{x_R}^\psi - \vartheta^{m-1})

- \epsilon^2 \Delta H^\epsilon_{m-1}(\vartheta^{m-1}, y, \beta))|\hat{X}_{x_R}^{m-1} = \vartheta^{m-1}, y_{x_R}^0 = y) \}.
$$

(H) Let us choose the value $\beta$. Denote by $L_R(\vartheta, \alpha)$ the Fenchel–Legendre transformation of $H_R(\vartheta, \beta)$. The function $L_R$ is differentiable in $\beta$ for any $|\beta| < a$ for $R \geq R_a$. Hence, after some arbitrary small perturbation of $\varphi$, we may choose $\beta_0 = \beta_0(R)$ such that $|\beta_0| < a$ and

$$
L_R(\psi_{T_{m-1}}, \alpha) > L_R(\psi_{T_{m-1}}, (\varphi^\beta_{T_{m-1}} - \varphi^\beta_{T_{m-1}})/\Delta) + \beta_0(\varphi^\beta_{T_{m-1}} - \varphi^\beta_{T_{m-1}})/\Delta,
$$

$$
\alpha \neq (\varphi^\beta_{T_{m-1}} - \varphi^\beta_{T_{m-1}})/\Delta
$$

(cf. Freidlin and Wentzell, 1984, Proof of Theorem 7.4.1). Then

$$
H_R(\beta_0) = \sup(\beta_0 \alpha - L_R(\psi_{T_{m-1}}, \alpha))
$$

$$
= \beta_0(\varphi^\beta_{T_{m-1}} - \varphi^\beta_{T_{m-1}})/\Delta - L_R^{m-1}(\psi_{T_{m-1}}, (\varphi^\beta_{T_{m-1}} - \varphi^\beta_{T_{m-1}})/\Delta).
$$

Denote $\hat{F}_{m-1} = F_{x_R}^{m-1}$. Then we have for $\beta = \beta_0$,

$$
E^\beta_R \{ I(\hat{X}_{x_R}^\psi - \varphi^\beta_T < \delta'')
$$

$$
\times \exp(-\epsilon^2 \beta(\hat{X}_{x_R}^\psi - \varphi^\beta_{T_{m-1}}) + \epsilon^2 \Delta H^\epsilon_{m-1}(\psi_{T_{m-1}}, \beta))|\hat{F}_{m-1}\}
$$

$$
\geq \exp(-\epsilon^2 \Delta L_R(\psi_{T_{m-1}}, (\varphi^\beta_{T_{m-1}} - \varphi^\beta_{T_{m-1}})/\Delta) + \nu))
$$

$$
\times E^\beta_R \{ I(\hat{X}_{x_R}^\psi - \varphi^\beta_T < \delta'')|\hat{F}_{m-1}\}.
$$

(1) We have,

$$
E^\beta_R \{ I(\hat{X}_{x_R}^\psi - \varphi^\beta_T \geq \delta'')|\hat{F}_{m-1}\} \leq \exp(-\epsilon^2 \Delta s_m)
$$


for some $s_m > 0$. This estimate follows from the upper bound for the family of measures $P^\beta_R$ in the same way as in the proof of Theorem 5.1.2 from Freidlin and Wentzell (1984). This implies
\[
P[I(|X^\psi_{n,m} - \varphi^i_T| < \delta'')|F_{m-1}]
\geq \exp(-\varepsilon^{-2}A(L_R(\psi_{m-1}, (\varphi^i_{m-1} - \varphi^i_{m-1})/A) + \nu))
[1 - \exp(-\varepsilon^{-2}s_mA)].
\]
For $\varepsilon > 0$ small enough, $1 - \exp(-\varepsilon^{-2}s_mA) \geq \frac{1}{2}$. Then
\[
P[I(|\tilde{X}^\psi_{n,m} - \varphi^i_T| < \delta'')|F_{m-1}] \geq \exp(-\varepsilon^{-2}A(L_R(\psi_{m-1}, (\varphi^i_{m-1} - \varphi^i_{m-1})/A) + \nu)).
\]
Similarly by induction we have for $R > |\nu|$, 
\[
P([\tilde{X}^\psi_{n,m-1} - \varphi^i_T| < \delta', \ldots, |X^\psi_{n,m} - \varphi^i_T| < \delta')
\geq \exp\left(-\varepsilon^{-2}A \sum_{i=1}^{m} (L_R(\psi_{m-1}, (\varphi^i_{m-1} - \varphi^i_{m-1})/A) + \nu)\right)
= \exp\left(-\varepsilon^{-2}A \sum_{i=0}^{m-1} (L_R(\psi_T, (\varphi^i_T - \varphi^i_T)/A) + \nu)\right).
\]
This implies the estimate
\[
P(\rho(\tilde{X}^\psi, \varphi^i) < \delta'') \geq \exp(-\varepsilon^{-2}(S^R_{0T}(\varphi^i) + \nu)),
\]
where
\[
S^R_{0T}(\varphi) = \int_0^T L_R(\psi_s, \varphi_s) \, ds
\]
(cf. to Lemma 7.5.1 from Freidlin and Wentzell, 1984).

(J) By virtue of Lemma 7.5.2 from Freidlin and Wentzell (1984) and differentiability of $H_R$ in $\beta$ for $|\beta| < a$ we deduce that for any sequence of step functions $\psi^u$ which converges to $\varphi$ uniformly there exists such sequence $\phi^u \in C[0, T; E']$ which also converges uniformly to $\varphi$ that
\[
\limsup_{n \to \infty} \int_0^T L_R(\psi^u_s, \phi^u_s) \, ds \leq S^R_{0T}(\varphi) = \int_0^T L_R(\psi_s, \phi_s) \, ds
\]
if $|\dot{\phi}_s| < a$ for any $s$. If $|\dot{\phi}_s| \geq a$ for some $s$ and $S^R_{0T}(\varphi) < \infty$, we may just take $\phi^u_s = \int_0^s \dot{\phi}_s I(|\dot{\phi}_s| < a) \, ds$. For a large enough we obtain $\rho_{0T}(\varphi, \phi^u) < \delta/2$ and $S^R_{0T}(\phi^u) \leq S^R_{0T}(\varphi) + \delta/2$ since the integral $\int_0^T L_R(\psi_s, \phi_s) \, ds$ is finite. Hence, we obtain
\[
P(\rho_{0T}(X^\xi, \varphi) < \delta) \geq \exp(-\varepsilon^{-2}(\rho S^R_{0T}(\varphi) + \nu)), \quad \varepsilon \to 0
\]
with any $p > 1$.

Moreover, it follows from the cited Lemma 7.5.2 (cf. the proof of Lemma 2.1 from Veretennikov, 1992) that, without loss of generality, one can assume $\varphi$ to be such that the vector $\phi_s$ belongs to the affine interior of the set $\{x: L(\psi_s, x) < \infty\}$ for any $s$ where $L(\psi_s, \phi_s) < \infty$. Then we get from the pointwise convergence $H_R(\psi_s, \cdot) \to H(\psi_s, \cdot)$, $R \to \infty$.
that $L_{\varepsilon}(\varphi, \Phi_{s}) \to L(\varphi, \Phi_{s})$, $R \to \infty$ for any such $s$. Hence, due to Fatou’s lemma we obtain
\[ S_{0T}^{R}(\varphi) \leq S_{0T}(\varphi) + v \quad (v > 0) \]
as $R \to \infty$. So,
\[ P(\rho_{0T}(X^{\varepsilon}, \varphi) \leq \varepsilon) \geq \exp(-\varepsilon^{-2}(S_{0T}(\varphi) + v)), \quad \varepsilon \to 0. \]

This bound is uniform in $x \in \mathbb{E}^{d}$, $|y| \leq r$ and $\varphi \in \Phi_{r}(s)$ for any $r, s > 0$.

(K) The first inequality being proved, the second one follows from standard considerations, see the corresponding part of Theorem 7.4.1 from Freidlin and Wentzell (1984), and we omit it. Theorem 1 is proved. □

Acknowledgements

The author is grateful to S. Makhno and to the anonymous referee for useful comments and remarks.

References