The Annals of Probability
2003, Vol. 31, No. 1, 413–443

AVERAGING PRINCIPLE OF SDE WITH SMALL DIFFUSION: MODERATE DEVIATIONS

BY A. GUILLIN

Université Blaise Pascal

Consider the following stochastic differential equation in $\mathbb{R}^d$:

$$dX_t^\varepsilon = b(X_t^\varepsilon, \xi_{t/\varepsilon})\,dt + \sqrt{\varepsilon}a(X_t^\varepsilon, \xi_{t/\varepsilon})\,dW_t,$$
$$X_0^\varepsilon = x_0,$$

where the random environment (ξ_t) is an exponentially ergodic Markov process, independent of the Wiener process (W_t), with invariant probability measure π, and ε is some small parameter. In this paper we prove the moderate deviations for the averaging principle of X^ε, that is, deviations of (X_t^ε) around its limit averaging system $(\bar{x}_t)$ given by $d\bar{x}_t = \bar{b}(\bar{x}_t)\,dt$ and $\bar{x}_0 = x_0$ where $\bar{b}(x) = \mathbb{E}_\pi(b(x,\cdot))$. More precisely we obtain the large deviation estimation about

$$\left(\eta_t^\varepsilon = \frac{X_t^\varepsilon - \bar{x}_t}{\sqrt{\varepsilon}h(\varepsilon)}\right)_{t\in[0,1]}$$

in the space of continuous trajectories, as ε decreases to 0, where $h(\varepsilon)$ is some deviation scale satisfying $1 \ll h(\varepsilon) \ll \varepsilon^{-1/2}$. Our strategy will be first to show the exponential tightness and then the local moderate deviation principle, which relies on some new method involving a conditional Schilder's theorem and a moderate deviation principle for inhomogeneous integral functionals of Markov processes, previously established by the author in Guillin (2001).

1. Introduction. The main subject of this paper is the formulation of moderate deviations for the averaging principle of a stochastic differential equation (SDE) with small diffusion, initiated by Freidlin (1978), in the case where the fast random environment does not depend on the Wiener process driving the SDE. Let us first present more precisely the framework of this study.

Let $\xi = (\xi_t)_{t\geq 0}$ be some ergodic Markov process with values in a general state space E. Consider now the diffusion process $X^\varepsilon = (X_t^\varepsilon)_{t\in[0,1]}$, jointly defined with ξ on some stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, given by the following stochastic differential equation in $\mathbb{R}^d$:

$$\begin{aligned} dX_t^\varepsilon &= b(X_t^\varepsilon, \xi_{t/\varepsilon})\,dt + \sqrt{\varepsilon}a(X_t^\varepsilon, \xi_{t/\varepsilon})\,dW_t,\\ X_0^\varepsilon &= x_0, \end{aligned} \tag{1.1}$$

Received January 2001; revised March 2002.
AMS 2000 subject classifications. 60F10, 60J60, 60J25.
Key words and phrases. Moderate deviations, averaging principle, SDE with small diffusion, Markov process, exponential ergodicity.

where (W_t) is a standard Wiener process independent of ξ, and ε is a small parameter. In such a model, X^ε is considered as the "slow process," while ξ is the fast random environment (its time scale being of order $1/\varepsilon$). As usual, we will call b (resp. a) the drift (resp. diffusion) term.

Now suppose the following ergodic property: there exists a vector field $\bar{b}$ such that

$$\lim_{T\to\infty} \frac{1}{T}\int_{t_0}^{t_0+T} \big(b(x,\xi_s)\big)\,ds = \bar{b}(x), \qquad \mathbb{P}\text{-a.s.}$$

uniformly in $x \in \mathbb{R}^d$ and $t_0 \geq 0$. Now let us denote for $t \in [0,1]$, the solution $\bar{x} = (\bar{x}_t)_{t\in[0,1]}$ of the averaged deterministic system in $\mathbb{R}^d$,

$$\begin{aligned} \dot{\bar{x}}_t &= \bar{b}(\bar{x}_t), \\ \bar{x}_0 &= x_0. \end{aligned} \tag{1.2}$$

The motivation linked to the study of these models can be found, for example, in stochastic mechanics [see Freidlin (1978)], where a polar change (or an appropriate change linked to the considered Hamiltonian) may give an amplitude evolving slowly whereas the phase is on an accelerated time scale, or in climate models [Kiefer (2000)], where climate–weather interactions may be studied within the averaging framework, climate being the slow motion and weather the fast one.

The *averaging principle* asserts that, under quite minimal assumptions and the previous ergodic property, we have

$$\lim_{\varepsilon\to 0} \mathbb{P}\Big(\sup_{t\in[0,1]} |X_t^\varepsilon - \bar{x}_t|\Big) = 0.$$

The study of such a convergence has been extensively developed in both the deterministic ($a \equiv 0$) and the stochastic context: see, for example, Bogolubov and Mitropolskii (1961) and Sanders and Verhulst (1985) for the deterministic case, Liptser (1994) for the stochastic one and more particularly Khasminskii (1980), Freidlin (1978) and Freidlin and Wentzell (1998) for when (ξ_t) is a diffusion process and Liptser (1994) for its proof in an even more general case. We finally refer to the recent works of Pardoux and Veretennikov (2000, 2001) for diffusion approximation (averaging of a singularly perturbed SDE).

Here we will study deviations of X^ε from the averaged solutions $\bar{x}$, as ε decreases to 0, that is, the asymptotic behavior of the trajectory for $t \in [0,1]$,

$$\eta_t^\varepsilon = \frac{X_t^\varepsilon - \bar{x}_t}{\sqrt{\varepsilon}h(\varepsilon)}, \tag{1.3}$$

where $h(\varepsilon)$ is some deviation scale which strongly influences the asymptotic behavior of η^ε. If $h(\varepsilon)$ is identically equal to 1, we are in the domain of the central limit theorem (CLT), which was first established by Freidlin and Wentzell (1998) with $a \equiv 0$ under some mixing conditions on the fast process; see also Liptser and

Stoyanov (1990) for the general semimartingale case, and see Rachad (1999) and Bernard and Rachad (2000) for when the fast process is a diffusion depending on the slow process.

The case $h(\varepsilon) = 1/\sqrt{\varepsilon}$ provides some large deviations estimations which have been extensively studied in the past 20 years. The first work seems to go back to Freidlin and Wentzell (1998), where $a \equiv 0$ and the fast process is a diffusion whose drift depends on the slow process. There has been much generalization in many directions in recent years; here we mention only three of them: Liptser (1996) presented, for large deviations for two scaled diffusions (the fast process is independent of the slow one), a combination of Freidlin and Wentzell's and Donsker and Varadhan's results (i.e, large deviations of the slow process coupled with the empirical measure of the fast process); Veretennikov (1999a) established the large deviation principle (LDP) when the fast process depends only on the slow process through the drift term, but allowed the same Wiener process for the diffusion term of each component; finally Veretennikov (1999b, c) proves the LDP when $a \equiv 0$ but the fast process has a "full" dependence (in drift and diffusion coefficient) on the slow process.

To fill the gap between the CLT scale $[h(\varepsilon) = 1]$ and the large deviations scale $[h(\varepsilon) = \varepsilon^{-1/2}]$, we are naturally led to study moderate deviations, that is, when the deviation scale satisfies

$$h(\varepsilon) \stackrel{\varepsilon \to 0}{\longrightarrow} +\infty, \qquad \sqrt{\varepsilon} h(\varepsilon) \stackrel{\varepsilon \to 0}{\longrightarrow} 0. \tag{1.4}$$

The moderate deviations enable us to refine the central limit theorem in the following sense: the rate function is a quadratic one which is the rate of large deviations of Gaussian processes. Moreover, if the central limit theorem gives asymptotic estimations for $\mathbb{P}(\|X^\varepsilon - \bar{x}\| \sim x\sqrt{\varepsilon})$ and the large deviation principle for $\mathbb{P}(\|X^\varepsilon - \bar{x}\| \sim x)$, the moderate deviations furnish estimations for $\mathbb{P}(\|X^\varepsilon - \bar{x}\| \sim x\sqrt{\varepsilon} h(\varepsilon))$. Finally, the conditions imposed on ξ for the moderate deviations will be less restrictive than the conditions for the LDP. When $a \equiv 0$ and $\bar{b}(x_0) = 0$ (i.e., the average system stays in the position x_0), Baier and Freidlin (1977) or Freidlin and Wentzell (1998) have imposed abstract conditions to prove the moderate deviation principle (MDP). We may cite also, in a slightly different direction, the work of Liptser and Spokoiny (1999), who obtained an upper bound of moderate deviations for integral functionals of the slow process or Klebaner and Liptser (1999) for randomly perturbed discrete dynamical systems. Very recently, still in the setting $a \equiv 0$, the author Guillin (2001) obtained the full MDP under the sole condition of the exponential ergodicity of the fast Markov process [and reasonable conditions on b, without the assumption $\bar{b}(x_0) = 0$].

The main goal of this paper will be to extend the approach of Guillin (2001) to the case where a small diffusion term is present, with the additional assumption that the fast component is independent of the Wiener process driving the diffusion. The main difficulty arises in the combination of the effect of ξ and the usual

Freidlin–Wentzell theory for LDP of small diffusion. It can be seen as the analog of the large deviations results of Liptser (1996) in the moderate deviations context (but with more general fast process).

Our method is based on two ideas: the use of explicit criteria [Liptser and Puhalskii (1992) or Feng and Kurtz (2000)] for exponential tightness in the space of continuous mappings; the exponential ergodicity of the fast Markov process and a theorem on moderate deviations of inhomogeneous integral functionals of a Markov process [Guillin (2001), Theorem 2] combined with a conditional Schilder theorem, stated in Lemma 2, Section 4.1.2.

The paper is organized as follows. In Section 2, we present the moderate deviations of η^ε. Section 3 contains some preliminary definitions and results which are essential for the proof of the theorem. The last section is devoted to the proof of our theorem.

2. Main result.

Let us start with some definitions.

2.1. *Formulation of an LDP.* Let $C^0([0,1],\mathbb{R}^d)$ be the space of continuous trajectories from $[0, 1]$ to $\mathbb{R}^d$ starting from 0, equipped with the supremum norm topology. Recall the definition of an LDP [see Deuschel and Stroock (1989) or Dembo and Zeitouni (1998) for instance]. The family $Z^\varepsilon = (Z^\varepsilon_t)_{t\in[0,1]}$ obeys the LDP in $C^0([0,1],\mathbb{R}^d)$ with speed $v(\varepsilon) \to 0$ and rate function I with respect to (w.r.t.) the supremum norm if the following hold:

1. There exists $I : C^0([0,1],\mathbb{R}^d) \to [0,\infty]$ such that I is inf-compact; that is, the level sets ($[I \le L]$ for $L \ge 0$) are compact.
2. For any open set O in $C^0([0,1],\mathbb{R}^d)$,
$$\liminf_{\varepsilon\to 0} v(\varepsilon)\log \mathbb{P}(Z^\varepsilon \in O) \ge -\inf_{z\in O} I(z).$$
3. For any closed subset F in $C^0([0,1],\mathbb{R}^d)$,
$$\limsup_{\varepsilon\to 0} v(\varepsilon)\log \mathbb{P}(Z^\varepsilon \in F) \le -\inf_{z\in F} I(z).$$

The function I is often called a *good rate function* associated with the LDP of (Z^ε).

2.2. *Assumptions on the Markov process* $(\xi_t)_{t\ge 0}$. We then set the basic definitions for the fast Markov process: $\xi = (\xi_t)_{t\ge 0}$ will denote a continuous time Markov process, defined on the probability space $(\Omega^\xi, \mathcal{F}^\xi, (\mathcal{F}^\xi_t), \mathbb{P}_\xi)$, with values in a general Polish space E, with transition semigroup $(P_t)_{t\in\mathbb{R}^+}$. The operator P_t acts on bounded measurable functions f and probability measures μ on E respectively via the relations $P_t f(x) = \int_E P_t(x,dy) f(y)$ and $\mu P_t(A) = \int_E \mu(dx) P_t(x,A)$. Throughout this paper, we will moreover suppose that the probability measure π is invariant, that is, $\pi = \pi P_t$ for all $t \ge 0$, and that the process ξ is π-irreducible and aperiodic [see Nummelin (1984) or Down, Meyn and Tweedie (1995) for more detailed explanations].

The Markov process (ξ_t) is called *exponentially ergodic* if

$$\|P_t(x,\cdot)-\pi\|_{\mathrm{var}} \le M(x)\rho^t, \qquad t\ge 0, \tag{2.1}$$

for some constant $\rho<1$ and some finite $M(x)$ which is moreover in $L^1(\pi)$, where $\|\cdot\|_{\mathrm{var}}$ is the total variation norm. This definition is a natural extension to the continuous time case of the geometrical ergodicity of Markov chains [see Nummelin (1984)]. See the application after the statement of Theorem 1 for an example of processes satisfying such a condition. We will suppose throughout this paper that the initial measure μ (the law of ξ_0) satisfies

$$\int_E M(x)\mu(dx)<\infty. \tag{2.2}$$

This condition means intuitively that ξ starting with the initial probability measure μ conserves the exponential behavior of (2.1). Note that every Dirac measure satisfies this condition as $M(x)$ is finite.

2.3. *Assumptions on the diffusion* (1.1). Now let us fix the main assumptions on the coefficients of the diffusion (1.1):

(A) The diffusion term $a(x,z):\mathbb{R}^d\times E\to\mathbb{R}^{d\times n}$ is a bounded continuous mapping with bounded (uniformly in z) first order derivative in x.
(B) The drift $b(x,z):\mathbb{R}^d\times E\to\mathbb{R}^d$ is a bounded continuous mapping with bounded (uniformly in z) first and second order derivatives in x.
(C) W is a Wiener process in $\mathbb{R}^n$, independent of ξ, defined on the stochastic basis $(\Omega^W,\mathcal{F}^W,(\mathcal{F}_t^W),\mathbb{P}_W)$.

The probability space $(\Omega,\mathcal{F},(\mathcal{F}_t),\mathbb{P})$ is then the product probability space of those of ξ and W. Recall that X^ε is the diffusion given in (1.1) and that under the ergodicity assumption on ξ the averaged drift

$$\bar b(x)=\int_E b(x,z)\pi(dz)$$

is well defined for all $x\in\mathbb{R}^d$. The averaged solution $\bar x$ is then given by (1.2). Moreover, we introduce the notation

$$\forall\, 1\le i,j\le d,\qquad B^i_j(x,y)=\frac{\partial b^i}{\partial x_j}(x,y),\qquad B(x,y)=\big(B^i_j(x,y)\big)_{1\le i,j\le d},$$

and set

$$\overline{B}(x)=\int_E B(x,z)\pi(dz).$$

We also define the *averaged diffusion* coefficient

$$\overline{a^2_s}=\int_E a(\bar x_s,z)a^t(\bar x_s,z)\pi(dz), \tag{2.3}$$

where a^t denotes the transposed matrix. Finally set

$$f(t,\cdot) = b(\bar{x}_t,\cdot) - \bar{b}(\bar{x}_t). \tag{2.4}$$

2.4. *The rate functions.* Let us introduce the following good rate function, which is the Schilder rate function with diffusion term $(\overline{a_s^2})^{1/2}$,

$$I_W(\gamma) = \begin{cases} \displaystyle\int_0^1 \sup_{\lambda\in\mathbb{R}^d} \big(\lambda^t \dot{\gamma}(s) - \tfrac{1}{2}\lambda^t \overline{a_s^2}\lambda\big)\,ds, \\ \qquad\qquad \text{if } d\gamma(s) = \dot{\gamma}(s)\,ds,\ \gamma(0) = 0, \\ +\infty, \qquad \text{otherwise.} \end{cases} \tag{2.5}$$

Note that $I_W(\gamma) = \infty$ if γ does not belong to the Cameron–Martin space $\mathcal{H}$ defined by

$$\mathcal{H} = \left\{\phi;\, d\phi(t) = \dot{\phi}(t)\,dt \text{ and } \int_0^1 \|\dot{\phi}(t)\|^2\,dt < \infty\right\}.$$

Now define the variance associated with the Markov process ξ and functional f which maps $[0,1]\times E$ to $\mathbb{R}^d$ with $f = (f_1,\dots,f_d)$,

$$\begin{aligned} \sigma_\xi^2\big(f(t,\cdot)\big)_{ij} &= \int_E f_i(t,\cdot)\int_0^\infty P_s f_j(t,\cdot)\,ds\,d\pi \\ &\quad + \int_E f_j(t,\cdot)\int_0^\infty P_s f_i(t,\cdot)\,ds\,d\pi. \end{aligned} \tag{2.6}$$

To get more explicit expression of the rate function, we will always assume

(D) $\qquad \sigma_\xi^2\big(f(t,\cdot)\big) \ge \omega I \qquad$ for some constant $\omega > 0$, for all t,

where I is the usual identity matrix. Then the following good rate function I_ξ^f is well defined:

$$I_\xi^f(\gamma) = \begin{cases} \displaystyle\frac{1}{2}\int_0^1 \langle\dot{\gamma}(s), \sigma_\xi^2\big(f(s,\cdot)\big)^{-1}\dot{\gamma}(s)\rangle\,ds, \\ \qquad\qquad \text{if } d\gamma(s) = \dot{\gamma}(s)\,ds,\ \gamma(0) = 0, \\ +\infty, \qquad \text{otherwise.} \end{cases} \tag{2.7}$$

Finally recall the definition

$$\eta_t^\varepsilon = \frac{X_t^\varepsilon - \bar{x}_t}{\sqrt{\varepsilon}h(\varepsilon)}$$

and that the deviation scale $h(\varepsilon)$ satisfies $1 \ll h(\varepsilon) \ll \varepsilon^{-1/2}$.

We can now state our result.

THEOREM 1. *Assume that ξ is an exponentially ergodic Markov process and the distribution of ξ_0 satisfies* (2.2). *Suppose moreover conditions* (A)–(D). *Then*

η^ε satisfies an LDP in $C^0([0,1],\mathbb{R}^d)$ with speed $h^{-2}(\varepsilon)$ and good rate function S given for $\gamma \in C^0([0,1],\mathbb{R}^d)$ by

$$S(\gamma) = I\left(\gamma - \int_0^\cdot \bar{B}(\bar{x}_s)\gamma(s)\,ds\right), \tag{2.8}$$

where I is defined by

$$I(\gamma) = \inf\{I_\xi^f(\gamma - \psi) + I_W(\psi); \psi \in C^0([0,1],\mathbb{R}^d)\} \tag{2.9}$$

$$= \begin{cases} \frac{1}{2}\int_0^1 \left\|\left(\sigma_\xi^2(f(s,\cdot)) + \overline{a_s^2}\right)^{-1/2}\dot{\gamma}(s)\right\|^2 ds, \\ \qquad\qquad \textit{if } d\gamma(s) = \dot{\gamma}(s)\,ds,\ \gamma(0) = 0, \\ +\infty, \qquad \textit{otherwise}. \end{cases} \tag{2.10}$$

Here, I_W is given by (2.5) *and I_ξ^f by* (2.7).

It is well known that, by (2.10), I is a good rate function [see Deuschel and Stroock (1989), Lemma 1.3.8]. From the expression of the rate function (2.8), it seems that the LDP is inherited from the addition of two independent LDPs, one coming from the drift part and the other from the diffusion term. In fact, the two driving LDPs are not independent but behave as if they were. This curious phenomenon is due to the assumption of independence between the fast process ξ and the Wiener process W and the fast convergence of the quadratic variation of the diffusion part to a deterministic one, allowing us to separate drift and diffusion LDPs. This result can be seen as a version of the "averaged Schilder's theorem."

Note also that assumption (D) of the uniform invertibility of $\sigma_\xi^2(f(t,\cdot))$ is not necessary for the LDP in this result but only for the explicit expression (2.8), (2.10) of the rate functions (see comments after Theorem 3 in the next section).

REMARK. To the author's knowledge, it is the first time that such an MDP is stated for an SDE with small diffusion and a fast process which is not necessarily given by a diffusion.

REMARK. Note that there are no nondegeneracy assumptions for the diffusion coefficient of (1.1). In fact when $a \equiv 0$, Theorem 1 gives back the result of Theorem 5 in Guillin (2001).

REMARK. The results presented here can be extended to the space $\mathcal{C}$ of continuous functions on $[0,\infty)$ equipped with the local supremum topology defined by the metric r:

$$\forall X', X'' \in \mathcal{C}, \qquad r(X', X'') = \sum_{n\geq 1} \frac{1}{2^n}\left(1 \wedge \sup_{t\leq n} |X'_t - X''_t|\right).$$

Then the LDP stated in Theorem 1 still holds in $(\mathcal{C}, r)$, by applying the Dawson–Gärtner theorem [see Dembo and Zeitouni (1998), Theorem 4.6.1], which states that it is sufficient to check the LDP in $C^0([0,T],\mathbb{R}^d)$ for all T in the uniform metric (i.e., Theorem 1).

EXAMPLE (ξ is the Ornstein–Uhlenbeck process). Let us consider the following system of SDEs in dimension 1:

$$\begin{aligned} dX_t^\varepsilon &= b(X_t^\varepsilon, \xi_{t/\varepsilon})\,dt + \sqrt{\varepsilon}a(X_t^\varepsilon, \xi_{t/\varepsilon})\,dW_t, \qquad X_0^\varepsilon = x_0, \\ d\xi_t &= -\tfrac{1}{2}\xi_t\,dt + dV_t, \end{aligned} \tag{2.11}$$

where V and W are two independent Wiener processes. Assume also that conditions (A) and (B) are satisfied. It is well known that the Ornstein–Uhlenbeck process is a positive recurrent Markov process with invariant distribution $\pi(dx) = (2\pi)^{-1/2}\exp(-x^2/2)\,dx$. Moreover, using Down, Meyn and Tweedie [(1995), Theorems 5.3 and 5.2c] or criteria (L) in Guillin [(2001), Theorem 3], we easily obtain that the Ornstein–Uhlenbeck process is exponentially ergodic. Suppose for simplicity that $\int_{\mathbb{R}} b(x_0,z)\pi(dz) = 0$ so that $\bar{x}_t = x_0$ for all t, and that the initial distribution of ξ_0 is a Dirac measure. Consequently, by Theorem 1, η^ε satisfies an LDP in $C^0([0,1],\mathbb{R})$ with speed $h^{-2}(\varepsilon)$ and good rate function S_{OU} given, for absolutely continuous γ starting from 0, by

$$S_{\mathrm{OU}}(\gamma) = \frac{1}{2(\sigma^2 + \overline{a^2})}\int_0^1 \big(\dot{\gamma}(s) - \bar{B}(x_0)\gamma(s)\big)^2\,ds \tag{2.12}$$

with

$$\begin{aligned} \sigma^2 &= 4\int_{\mathbb{R}}\left(\int_{-\infty}^z b(x_0,u)e^{-u^2/2}\,du\right)^2 e^{z^2}\,dz, \\ \overline{a^2} &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} a^2(x_0,z)e^{-z^2/2}\,dz, \end{aligned}$$

which are well defined under assumptions (A) and (B).

We next present some definitions useful for proving an LDP, and some preliminary results linked to the MDP of exponentially ergodic Markov processes proved in Guillin (2001) which will be a recurrent tool in our proof.

3. Some preliminary definitions and results. We first recall the notions of exponential tightness and local LDP which provide sufficient conditions to prove an LDP. Let us denote by $Z^\varepsilon = (Z_t^\varepsilon)_{t\in[0,1]}$ a $C^0([0,1],\mathbb{R}^d)$-valued family.

DEFINITION 1. The family (Z^ε) is said to be exponentially tight [with speed $v(\varepsilon) \to 0$] in the space $C^0([0,1],\mathbb{R}^d)$, if there exists an increasing sequence of compact sets $(K_j)_{j\geq 1}$ of $C^0([0,1],\mathbb{R}^d)$ such that

$$\lim_j \limsup_{\varepsilon\to 0} v(\varepsilon)\log\mathbb{P}(Z^\varepsilon \notin K_j) = -\infty \tag{3.1}$$

[Deuschel and Stroock (1989), Dembo and Zeitouni (1998)].

DEFINITION 2. The family (Z^ε) is said to satisfy the local LDP [with speed $v(\varepsilon)\to 0$] in $C^0([0,1],\mathbb{R}^d)$ with rate function $\hat{I}$ if, for any $z\in C^0([0,1],\mathbb{R}^d)$,

$$\begin{aligned}\lim_{\delta\to 0}\limsup_{\varepsilon\to 0} v(\varepsilon)\log\mathbb{P}\big(Z^\varepsilon\in B(z,\delta)\big) &= \lim_{\delta\to 0}\liminf_{\varepsilon\to 0} v(\varepsilon)\log\mathbb{P}\big(Z^\varepsilon\in B(z,\delta)\big)\\ &= -\hat{I}(z),\end{aligned}$$

where $B(z,\delta)$ is the ball of radius δ centered in z [Freidlin and Wentzell (1998)].

The following theorem is well known; see Deuschel and Stroock (1989), Dembo and Zeitouni (1998) or Liptser and Puhalskii (1992).

THEOREM 2. *The exponential tightness and the local LDP for the family (Z^ε) in $C^0([0,1],\mathbb{R}^d)$ with local rate function $\hat{I}$ imply the full LDP in $C^0([0,1],\mathbb{R}^d)$ for this family with rate function $I(z)\equiv\hat{I}(z)$ which is inf-compact.*

Let $g:[0,1]\times E\to\mathbb{R}^d$ be a measurable mapping and set

$$l_t^\varepsilon = \frac{1}{\sqrt{\varepsilon}h(\varepsilon)}\int_0^t g(s,\xi_{s/\varepsilon})\,ds$$

for $t\in[0,1]$. The following result is crucial for the proof of Theorem 1.

THEOREM 3 [Guillin (2001), Theorem 2]. *Assume that g satisfies the following conditions*:

(G_1) *g is a bounded measurable mapping*;
(G_2) $\int_E g(t,x)\pi(dx)=0$ *for all t*;
(G_3) $\omega_g(\delta)=\sup_{|s-t|\leq\delta,x\in E}|g(s,x)-g(t,x)|$, *the modulus of continuity*, *satisfies*

$$\lim_{\delta\to 0}\frac{\omega_g(\delta)}{\sqrt{\delta}} = 0.$$

Suppose that ξ is exponentially ergodic and the distribution of ξ_0 satisfies (2.2). *Assume moreover that $\sigma^2(g(t,\cdot))$ is invertible uniformly in time*, *that is*, *satisfies* (D). *Then $l^\varepsilon=(l_t^\varepsilon)_{t\in[0,1]}$ satisfies an LDP in $C^0([0,1],\mathbb{R}^d)$ with speed $h^{-2}(\varepsilon)$ and good rate function I_ξ^g given by* (2.7).

We also present a direct consequence of this theorem.

COROLLARY 1. *Under the assumptions on g and ξ of Theorem* 3, $\sup_{t\in[0,1]}\|l_t^\varepsilon\|$ *satisfies an LDP in $\mathbb{R}$ with speed $h^{-2}(\varepsilon)$ and good rate function J.*

In particular,

$$\lim_{x\to\infty} J(x) = +\infty, \tag{3.2}$$

$$\lim_{j\to\infty}\lim_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\Big(\sup_{t\in[0,1]}\|l_t^\varepsilon\|\ge j\Big) = -\infty. \tag{3.3}$$

PROOF. In fact, as the LDP of l^ε holds w.r.t. the uniform topology and $\gamma\to\sup_{t\in[0,1]}\gamma(t)$ is continuous, the contraction principle [see Dembo and Zeitouni (1998), Theorem 4.2.1] implies that $\sup_{t\in[0,1]}\|l_t^\varepsilon\|$ satisfies an LDP in $\mathbb{R}$ with speed $h^{-2}(\varepsilon)$ and good rate function J given by, for nonnegative x,

$$J(x) = \frac{1}{2}\inf_{\gamma\,:\,\sup_s\|\gamma(s)\|=x}\left\{\int_0^1\|\sigma_\xi^2(g(s,\cdot))^{-1}\dot\gamma(s)\|^2\,ds\right\}.$$

As J is inf-compact (i.e., for each positive L, $[J\le L]$ is compact), limit (3.2) is obvious, and (3.3) follows from the upper bound of the LDP of $\sup_t|l_t^\varepsilon|$ and (3.2). $\square$

REMARK. In fact, the LDP of Theorem 3 and the conclusion of this corollary hold without the assumption that $\sigma^2(g(t,\cdot))$ is invertible uniformly in time, as we have a good rate function, but in that case the rate function is expressed as the limit of the finite dimensional rate function of a homogenized functional and then has no explicit formulation [see Guillin (2001), formulas (2.17) and (2.35)].

The next result, which seems to be new, gives the MDP in the special simple case when the drift term is identically equal to 0 and the diffusion term does not depend on the slow variable. Set, for some function $c:[0,1]\times E\to\mathbb{R}^{d\times n}$,

$$r_t^\varepsilon = \frac{1}{h(\varepsilon)}\int_0^t c(s,\xi_{s/\varepsilon})\,dW_s$$

for $t\in[0,1]$.

PROPOSITION 1. *Suppose that ξ is an exponentially ergodic Markov process independent of W and suppose that the distribution of ξ_0 satisfies* (2.2). *Assume moreover that c is a bounded mapping satisfying condition* (G3) *of Theorem* 3. *Then (r^ε) satisfies an LDP in $C^0([0,1],\mathbb{R}^d)$ with speed $h^{-2}(\varepsilon)$, with rate function $\tilde I_W$ given by*

$$\tilde I_W(\gamma) = \begin{cases}\displaystyle\int_0^1\sup_{\lambda\in\mathbb{R}^d}\big(\lambda^t\dot\gamma(s)-\tfrac12\lambda^t\overline{c_s^2}\lambda\big)\,ds, & \\ \qquad\qquad\textit{if } d\gamma(t)=\dot\gamma(t)\,dt,\ \gamma(0)=0, & \\ +\infty, \quad \textit{otherwise}, & \end{cases} \tag{3.4}$$

where

$$\overline{c_s^2} = \int_E c(s,z)c^t(s,z)\pi(dz). \tag{3.5}$$

PROOF. This proposition is in fact a direct consequence of Theorem 3 and Theorem 2.1 of Puhalskii (1990) for large deviations of semimartingales where we only have to check the condition called (sup C) in Puhalskii (1990), because, in our case, there are no finite variation or jump parts so that conditions (A), (a), (sup B), (K+L) and (L) in Puhalskii (1990) are trivially satisfied. In our context, it is equivalent to check that there exists a continuous function C_t such that, for all positive δ,

$$\lim_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\Big(\sup_{t\in[0,1]} \big|h^2(\varepsilon)\langle r^\varepsilon\rangle_t - C_t\big| > \delta \Big) = -\infty, \tag{sup C}$$

where

$$\langle r_t^\varepsilon\rangle = \frac{1}{h^2(\varepsilon)} \int_0^t c(s,\xi_{s/\varepsilon})c^t(s,\xi_{s/\varepsilon})\,ds$$

is the predictable quadratic variation process of the continuous martingale (r_t^ε). Let us prove that (sup C) holds with $C_t = \int_0^t \overline{c_s^2}\,ds$, where $\overline{c_s^2}$ is defined in (3.5). Since, under our hypothesis, $c(s,z)c^t(s,z)$ and $\overline{c_s^2}$ both satisfy condition (G3), so does $c(s,z)c^t(s,z) - \overline{c_s^2}$, which is also bounded. By definition of $\overline{c_s^2}$, (G2) holds, and we may then apply Theorem 3, so that

$$\frac{1}{\sqrt{\varepsilon}h(\varepsilon)} \int_0^t \big(c(s,\xi_{s/\varepsilon})c^t(s,\xi_{s/\varepsilon}) - \overline{c_s^2}\big)\,ds$$

satisfies an LDP in $C^0([0,1],\mathbb{R}^d)$. Consequently Corollary 1, relation (3.3), yields, $\forall\,\delta > 0$,

$$\begin{aligned}\lim_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\bigg(\sup_{t\in[0,1]} \bigg\| \frac{1}{\sqrt{\varepsilon}h(\varepsilon)} \int_0^t \big(c(s,\xi_{s/\varepsilon})c^t(s,\xi_{s/\varepsilon}) - \overline{c_s^2}\big)\,ds \bigg\| > \frac{\delta}{\sqrt{\varepsilon}h(\varepsilon)} \bigg) \\ = -\infty,\end{aligned}$$

which is the condition (sup C) we needed to prove. □

4. Proof of Theorem 1. To simplify notation, we shall give the proof only in the case $d = n = 1$. Throughout the proof, K will denote a generic constant which may change from line to line, independent of time and ε. Let us first recall the definition of η^ε, for $t \in [0,1]$,

$$\begin{aligned}\eta_t^\varepsilon &= \frac{X_t^\varepsilon - \bar{x}_t}{\sqrt{\varepsilon}h(\varepsilon)} \\ &= \frac{1}{\sqrt{\varepsilon}h(\varepsilon)} \int_0^t \big(b(X_s^\varepsilon,\xi_{s/\varepsilon}) - \bar{b}(\bar{x}_s)\big)\,ds + \frac{1}{h(\varepsilon)} \int_0^t a(X_s^\varepsilon,\xi_{s/\varepsilon})\,dW_s.\end{aligned}$$

Now introduce $\hat{\eta}_t^\varepsilon$, defined for $t \in [0, 1]$,

$$\hat{\eta}_t^\varepsilon = \frac{1}{\sqrt{\varepsilon}h(\varepsilon)} \int_0^t \big(b(\bar{x}_s, \xi_{s/\varepsilon}) - \bar{b}(\bar{x}_s)\big)\, ds + \frac{1}{h(\varepsilon)} \int_0^t a(\bar{x}_s, \xi_{s/\varepsilon})\, dW_s + \int_0^t \bar{B}(\bar{x}_s)\hat{\eta}_s^\varepsilon\, ds$$

and

$$\lambda_t^\varepsilon = \frac{1}{\sqrt{\varepsilon}h(\varepsilon)} \int_0^t \big(b(\bar{x}_s, \xi_{s/\varepsilon}) - \bar{b}(\bar{x}_s)\big)\, ds, \tag{4.1}$$

$$\widehat{M}_t^\varepsilon = \frac{1}{h(\varepsilon)} \int_0^t a(\bar{x}_s, \xi_{s/\varepsilon})\, dW_s \tag{4.2}$$

so that

$$\hat{\eta}_t^\varepsilon = \lambda_t^\varepsilon + \widehat{M}_t^\varepsilon + \int_0^t \bar{B}(\bar{x}_s)\hat{\eta}_s^\varepsilon\, ds. \tag{4.3}$$

In the first step of the proof (Section 4.1), we will establish that $\hat{\eta}^\varepsilon$ satisfies the LDP of Theorem 1 and then in Section 4.2, we prove that η^ε and $\hat{\eta}^\varepsilon$ are exponentially equivalent w.r.t. the LDP; that is, for all positive δ,

$$\lim_{\varepsilon \to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\Big(\sup_{t \in [0,1]} |\eta_t^\varepsilon - \hat{\eta}_t^\varepsilon| > \delta \Big) = -\infty. \tag{4.4}$$

Theorem 1 will then follow from those two claims thanks to Theorem 4.2.13 of Dembo and Zeitouni (1998). Let us begin with the LDP of $\hat{\eta}^\varepsilon$.

4.1. *The LDP of* $\hat{\eta}^\varepsilon$. First write

$$R^\varepsilon = \lambda^\varepsilon + \widehat{M}^\varepsilon.$$

PROPOSITION 2. *If R^ε satisfies the LDP in $C^0([0, 1], \mathbb{R}^d)$ of speed $h^{-2}(\varepsilon)$ and good rate function I, then $\hat{\eta}^\varepsilon$ satisfies the LDP in $C^0([0, 1], \mathbb{R}^d)$ of speed $h^{-2}(\varepsilon)$ and good rate function S given by* (2.8).

PROOF. By definition of $\hat{\eta}^\varepsilon$, we have

$$\hat{\eta}_t^\varepsilon = R_t^\varepsilon + \int_0^t \bar{B}(\bar{x}_s)\hat{\eta}_s^\varepsilon\, ds.$$

We want to prove that the application mapping $\hat{\eta}^\varepsilon$ to R^ε is continuous. In other words, if $T : C^0([0, 1], \mathbb{R}^d) \to C^0([0, 1], \mathbb{R}^d)$ is defined by $T(u) = v$, where $\forall t \in [0, 1]$, $v(t) = u(t) + \int_0^t \bar{B}(\bar{x}_s)v(s)\, ds$, we want to prove that T is continuous.

Let $u, u' \in C^0([0,1], \mathbb{R}^d)$ and let $T(u) = v$ and $T(u') = v'$. Since $\bar{B}$ is bounded,

$$\begin{aligned} |v(t) - v'(t)| &= \left| u(t) - u'(t) + \int_0^t \bar{B}(\bar{x}_s)\big(v(s) - v'(s)\big)\, ds \right| \\ &\le |u(t) - u'(t)| + K \int_0^t |v(s) - v'(s)|\, ds \end{aligned}$$

and, by Gronwall's lemma,

$$|v(t) - v'(t)| \le e^K \sup_{0} |u(t) - u'(t)| \qquad \forall t \in [0,1],$$

which proves the continuity of T. Therefore, by the contraction principle, the LDP of $\hat{\eta}^\varepsilon$ follows from the one of R^ε with good rate function given by

$$\begin{aligned} S(\gamma) &= \inf_{\phi} \left\{ I(\phi);\ \gamma(t) = \phi(t) + \int_0^t \bar{B}(\bar{x}_s)\gamma(s)\, ds\ \forall t \in [0,1] \right\} \\ &= I\left(\gamma(\cdot) - \int_0^{\cdot} \bar{B}(\bar{x}_s)\gamma(s)\, ds \right), \end{aligned}$$

which is exactly (2.8). $\square$

We will often use the following lemma.

LEMMA 1. *Under the assumptions of Theorem* 1, λ^ε (*resp.* $\widehat{M}^\varepsilon$) *satisfies an LDP in* $C^0([0,1], \mathbb{R}^d)$ *with speed* $h^{-2}(\varepsilon)$ *and good rate function* I_ξ^f *given by* (2.7), [*resp.* I_W *given by* (2.5)].

PROOF. Recall that [see (4.3)]

$$\lambda_t^\varepsilon = \frac{1}{\sqrt{\varepsilon} h(\varepsilon)} \int_0^t f(s, \xi_{s/\varepsilon})\, ds \tag{4.5}$$

with $f(s,z) = b(\bar{x}_s, z) - \bar{b}(\bar{x}_s)$. Now note that, under (B), f is in fact bounded and Lipschitz continuous and is centered w.r.t. π by definition, so that condition (G) of Theorem 3 is fulfilled and then the LDP of rate I_f^ξ follows.

Under (A), a is bounded and Lipschitz continuous, so the conditions of Proposition 5 are satisfied with $c(s,z) = a(\bar{x}_s, z)$, hence $r^\varepsilon = \widehat{M}^\varepsilon$, and $\hat{I}_W$ is in this case the rate I_W. Consequently, $\widehat{M}^\varepsilon$ satisfies the desired LDP of rate I_W. $\square$

Thanks to Proposition 6, we have reduced the search for the LDP of $\hat{\eta}^\varepsilon$ to that of R^ε. This last LDP will be established in three steps: the exponential tightness, the local LDP and the explicit expression of the rate function.

4.1.1. *Exponential tightness of* R^ε. We have to find here a sequence of compacts K_j such that R^ε satisfies (3.1). This condition is not so easy to check by a direct approach. Sufficient conditions for exponential tightness in the space of continuous trajectories were given by Liptser and Puhalskii [(1992), Theorem 3.1],

$$\lim_{j\to\infty}\limsup_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\Big(\sup_{t\in[0,1]}|R_t^\varepsilon|>j\Big)=-\infty, \tag{4.6}$$

$$\lim_{\delta\to 0}\limsup_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\sup_{\tau\le 1-\delta}\mathbb{P}\Big(\sup_{t\le\delta}|R_{\tau+t}^\varepsilon-R_\tau^\varepsilon|>r\Big)=-\infty \qquad \forall\, r>0, \tag{4.7}$$

where τ is a stopping time w.r.t. $(\mathcal{F}_t)_{t\ge 0}$.

We will in fact prove a weaker form of (4.7) sufficient for the exponential tightness as remarked by Feng and Kurtz [(2000), Section 3.1, Remark 3.2],

$$\lim_{\delta\to 0}\sup_{s\in[0,1]}\limsup_{\varepsilon\to 0}\sup_{s\le t\le s+\delta}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}(|R_t^\varepsilon-R_s^\varepsilon|>r)=-\infty \qquad \forall\, r>0, \tag{4.8}$$

with the convention $R_t^\varepsilon=R_1^\varepsilon$ for $t\ge 1$.

PROOF OF (4.6). By definition of R^ε, we have, for all $t\in[0,1]$,

$$|R_t^\varepsilon|\le|\lambda_t^\varepsilon|+|\widehat{M}_t^\varepsilon| \tag{4.9}$$

so that, for (4.6), it is sufficient to establish

$$\lim_{j\to\infty}\limsup_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\Big(\sup_{t\in[0,1]}|\lambda_t^\varepsilon|>j\Big)=-\infty, \tag{4.10}$$

$$\lim_{j\to\infty}\limsup_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\Big(\sup_{t\in[0,1]}|\widehat{M}_t^\varepsilon|>j\Big)=-\infty. \tag{4.11}$$

By Lemma 1 and the proof of Corollary 1, (4.10) follows immediately.

Note that, by assumption (A), the diffusion coefficient a is uniformly bounded. Therefore, by definition of the predictable quadratic process $\langle M_t^\varepsilon\rangle$, we have

$$\langle M_1^\varepsilon\rangle=\frac{1}{h^2(\varepsilon)}\int_0^1 a^2(\bar{x}_s,\xi_{s/\varepsilon})\,ds,$$

and Bernstein's inequality [see Revuz and Yor (1994), pages 153–154] yields

$$\begin{aligned}\mathbb{P}\Big(\sup_{t\in[0,1]}|M_t^\varepsilon|>j\Big)&=\mathbb{P}\Big(\sup_{t\in[0,1]}|M_t^\varepsilon|>j;\ \langle M^\varepsilon\rangle_1\le\frac{K}{h^2(\varepsilon)}\Big)\\&\le\exp\Big(-\frac{j^2h^2(\varepsilon)}{2K}\Big),\end{aligned}$$

which obviously implies (4.11). The estimation (4.6) is proved. $\square$

PROOF OF (4.8). We essentially use the same trick. First, we obviously have

$$|R_t^\varepsilon - R_s^\varepsilon| \le |\lambda_t^\varepsilon - \lambda_s^\varepsilon| + \big|\widehat{M}_t^\varepsilon - \widehat{M}_s^\varepsilon\big|.$$

Hence, for (4.8), it is enough to prove

$$\lim_{\delta\to 0} \sup_{s\in[0,1]} \limsup_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\Big(\sup_{s\le t\le s+\delta} |\lambda_t^\varepsilon - \lambda_s^\varepsilon| > r\Big) = -\infty \qquad \forall r > 0,$$

$$\lim_{\delta\to 0} \sup_{s\in[0,1]} \limsup_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\Big(\sup_{s\le t\le s+\delta} |M_t^\varepsilon - M_s^\varepsilon| > r\Big) = -\infty \qquad \forall r > 0.$$

The first limit is once again a consequence of the MDP of λ^ε obtained in Lemma 1 and the exponential tightness deduced from Corollary 1. The second limit can be obtained by Bernstein's inequality. Estimation (4.8) is proved. □

4.1.2. *The local MDP of R^ε.* We will first prove the lower bound of the local MDP, which requires less effort, before establishing the more delicate local upper bound.

The key remark is the following variant of Schilder's theorem, which can be proved in the same way as Theorem 1.3.27 of Deuschel and Stroock [(1989), Lemmas 1.3.8, 1.3.14, 1.3.21].

LEMMA 2. *Let $(\mathcal{M}^\varepsilon)$ be a sequence of continuous Gaussian martingales, and let C_t be a deterministic continuously increasing functional. Let ψ be in the Cameron–Martin space and let $\eta > 0$. There exists $\delta = \delta(\psi, \eta) > 0$ such that, if for ε small enough*

$$\sup_{t\in[0,1]} |h^2(\varepsilon)\langle \mathcal{M}^\varepsilon\rangle_t - C_t| < \delta, \tag{4.12}$$

we have for all $a > 0$ small enough, for sufficiently small ε (depending on a, ψ, η),

$$\begin{aligned} \exp\big(-h^2(\varepsilon)\big(S_C(\psi) + \eta\big)\big) &\le \mathbb{P}\big(\mathcal{M}^\varepsilon \in B(\psi, a)\big) \\ &\le \exp\big(-h^2(\varepsilon)\big(S_C(\psi) - \eta\big)\big), \end{aligned} \tag{4.13}$$

where

$$S_C(\psi) = \begin{cases} \displaystyle\int_0^1 \sup_{\lambda\in\mathbb{R}^d} \big(\lambda^t \dot\psi(s) - \tfrac{1}{2}\lambda^t \dot C_s \lambda\big)\, ds, & \textit{if } d\psi(s) = \dot\psi(s)\, ds,\ \psi(0) = 0, \\ +\infty, & \textit{otherwise}. \end{cases}$$

Note that, when $C_t = \int_0^t \overline{a_s^2}\, ds$, we have $S_C = I_W$ of Theorem 1.

Before proving the lower and upper bounds of the local MDP, let us present the following trivial but crucial fact. By the independence of ξ and the Wiener process W, $\widehat{M^\varepsilon}$ is a $\mathbb{P}_\xi$-Gaussian continuous martingale w.r.t. the filtration generated by the Wiener process, with predictable quadratic process

$$\langle \widehat{M^\varepsilon} \rangle_t = \frac{1}{h^2(\varepsilon)} \int_0^t a^2(\bar{x}_s, \xi_{s/\varepsilon})\, ds \qquad \text{w.r.t. } \mathbb{P}_\xi.$$

It will be essential for the use of Lemma 2 in our context.

LOWER BOUND OF THE LOCAL LDP OF R^ε. Given an arbitrary path γ in $C^0([0,1],\mathbb{R})$, and positive numbers a, η, we suppose here that $I(\gamma) < \infty$ (trivial otherwise). By definition of $I(\gamma)$, choose some ψ such that

$$I_\xi(\gamma - \psi) + I_W(\psi) < I(\gamma) + \eta. \tag{4.14}$$

First, note that

$$\mathbb{P}(R^\varepsilon \in B(\gamma, a)) \geq \mathbb{E}\big(\mathbb{P}(R^\varepsilon \in B(\gamma, a),\ B^\xi_{\varepsilon,\delta} | \xi)\big),$$

where the event $B^\xi_{\varepsilon,\delta}$ is defined by

$$B^\xi_{\varepsilon,\delta} = \left\{ \sup_{t\in[0,1]} \left| \int_0^t \big(a^2(\bar{x}_s, \xi_{s/\varepsilon}) - \overline{a^2_s}\big)\, ds \right| \leq \delta \right\},$$

and δ will be chosen later. Note that, under our assumptions (as a is Lipshitz continuous), functional $a^2(\bar{x}_s, z) - \overline{a^2_s}$ satisfies condition (G) of Theorem 3; then, applying Corollary 1 (see details in the proof of Proposition 1), we have, for all positive δ,

$$\lim_{\varepsilon \to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\big((B^\xi_{\varepsilon,\delta})^c\big) = -\infty. \tag{4.15}$$

Now, by the preceding remark on $(\widehat{M^\varepsilon})$ and Lemma 2, there exist $\tilde{a}(\psi) \in (0, a/2)$ and $\delta > 0$ such that for sufficiently small ε, on the event $B^\xi_{\varepsilon,\delta}$,

$$\mathbb{P}\big(\widehat{M^\varepsilon} \in B(\psi, \tilde{a}(\psi)) | \xi\big) \geq \exp\big(-h^2(\varepsilon)(I_W(\psi) + \eta)\big) \qquad \text{a.s.} \tag{4.16}$$

Since the events $B^\xi_{\varepsilon,\delta}$ and $\{\lambda^\varepsilon \in B(\gamma - \psi, a/2)\}$ are $\sigma(\xi)$-measurable, we have

$$\begin{aligned}
\mathbb{P}(R^\varepsilon \in B(\gamma, a)) &\geq \mathbb{E}\Big(\mathbb{P}\big(R^\varepsilon \in B(\gamma, a),\ B^\xi_{\varepsilon,\delta}, \widehat{M^\varepsilon} \in B(\psi, \tilde{a}(\psi)) | \xi\big)\Big) \\
&\geq \mathbb{E}\Big(\mathbb{P}\big(\lambda^\varepsilon \in B(\gamma - \psi, a/2),\ B^\xi_{\varepsilon,\delta}, \widehat{M^\varepsilon} \in B(\psi, \tilde{a}(\psi)) | \xi\big)\Big) \\
&\geq \mathbb{E}\Big(\mathbb{1}_{\lambda^\varepsilon \in B(\gamma-\psi, a/2)} \mathbf{1}_{B^\xi_{\varepsilon,\delta}} \mathbb{P}\big(\widehat{M^\varepsilon} \in B(\psi, \tilde{a}(\psi)) | \xi\big)\Big) \\
&\geq \mathbb{P}\big(\lambda^\varepsilon \in B(\gamma - \psi, a/2),\ B^\xi_{\varepsilon,\delta}\big) \exp\big(-h^2(\varepsilon)(I_W(\psi) + \eta)\big),
\end{aligned} \tag{4.17}$$

where we have used estimation (4.16) in the last step. Note now

$$\liminf_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\Big(\lambda^\varepsilon\in B\Big(\gamma-\psi,\frac{a}{2}\Big)\Big)$$

$$(4.18)\qquad \le\max\left\{\begin{array}{l}\liminf_{\varepsilon\to 0}\dfrac{1}{h^2(\varepsilon)}\log\mathbb{P}\Big(\lambda^\varepsilon\in B\Big(\gamma-\psi,\dfrac{a}{2}\Big),B^\xi_{\varepsilon,\delta}\Big)\\ \limsup_{\varepsilon\to 0}\dfrac{1}{h^2(\varepsilon)}\log\mathbb{P}\big((B^\xi_{\varepsilon,\delta})^c\big)\end{array}\right\}.$$

On the other hand, λ^ε satisfies, by Lemma 1, an LDP with good rate function I^f_ξ, so that

$$\liminf_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\Big(\lambda^\varepsilon\in B\Big(\gamma-\psi,\frac{a}{2}\Big)\Big)\ge -I^f_\xi(\gamma-\psi).$$

Consequently, this together with (4.15) and (4.18) yields

$$\liminf_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\Big(\lambda^\varepsilon\in B\Big(\gamma-\psi,\frac{a}{2}\Big),\ B^\xi_{\varepsilon,\delta}\Big)\ge -I^f_\xi(\gamma-\psi).$$

Combining this last inequality with (4.17), we get, for every given $\eta>0$ and $a>0$,

$$\begin{aligned}\liminf_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\big(R^\varepsilon\in B(\gamma,a)\big)&\ge -I^f_\xi(\gamma-\psi)-I_W(\psi)-\eta\\ &\ge -I(\gamma)-2\eta,\end{aligned}$$

which is the desired lower bound of the local LDP.

We now turn to the more delicate proof of the upper bound of the local LDP.

UPPER BOUND OF THE LOCAL LDP OF R^ε. Fix an arbitrary γ in $C_0([0,1],\mathbb{R})$. For a positive number a, we have

$$\begin{aligned}\mathbb{P}\big(R^\varepsilon\in B(\gamma,a)\big)&=\mathbb{E}\big(\mathbb{E}(\mathbb{1}_{R^\varepsilon\in B(\gamma,a)}|\xi)\big)\\ &=\mathbb{E}\Big(\mathbb{E}\big(\mathbb{1}_{R^\varepsilon\in B(\gamma,a)}\mathbf{1}_{B^\xi_{\varepsilon,\delta}}|\xi\big)\Big)+\mathbb{E}\Big(\mathbb{E}\big(\mathbb{1}_{R^\varepsilon\in B(\gamma,a)}\mathbf{1}_{(B^\xi_{\varepsilon,\delta})^c}|\xi\big)\Big).\end{aligned}$$

Therefore,

$$(4.20)\qquad \mathbb{P}\big(R^\varepsilon\in B(\gamma,a)\big)\le\mathbb{E}\Big(\mathbb{E}\big(\mathbf{1}_{B^\xi_{\varepsilon,\delta}}\mathbb{P}(R^\varepsilon\in B(\gamma,a))|\xi\big)\Big)+\mathbb{P}\big((B^\xi_{\varepsilon,\delta})^c\big).$$

We have already noted in (4.15) that $\mathbb{P}((B^\xi_{\varepsilon,\delta})^c)$ is negligible with respect to the LDP.

For L positive, the level set $K_L=[\psi;\ I_W(\psi)\le L]$ is compact, since I_W is a good rate function. Set $\eta\in(0,1/2)$ independently of L. Recall that, due to the independence of ξ and W, $(\widehat{M^\varepsilon})$ is a $\mathbb{P}_\xi$-Gaussian continuous martingale. Then, by Lemma 2, for any $\psi \in K_L$ there exists $\tilde{a} = \tilde{a}(\psi) > 0$ such that, for sufficiently small ε,

$$\mathbf{1}_{B^{\xi}_{\varepsilon,\delta}} \mathbb{P}\big(\widehat{M}^{\varepsilon} \in B(\psi, \tilde{a})|\xi\big) \le \exp\big(-h^2(\varepsilon)\big(I_W(\psi) - \eta\big)\big) \qquad \text{a.s.} \tag{4.21}$$

On the other hand, by the LDP satisfied by λ^{ε}_t, there exists $\hat{a} = \hat{a}(\psi) > 0$ such that, for sufficiently small ε,

$$\mathbb{P}\big(\lambda^{\varepsilon} \in B(\gamma - \psi, \hat{a})\big) \le \exp\big(-h^2(\varepsilon) D\big(I_{\xi}(\gamma - \psi), \eta\big)\big), \tag{4.22}$$

where

$$D(x, \eta) = (x - \eta) \wedge \frac{1}{\eta}$$

is introduced to avoid the problem of $I_{\xi}(\gamma - \psi)$ being possibly infinite. Since K_L is compact, we may choose a finite collection of mappings $\psi_1, \dots, \psi_N \in K_L$ such that $K_L \subset G^L_N = \bigcup_{l=1}^N B(\psi_l, a_l)$, where

$$a_l = \tilde{a}(\psi_l) \wedge \left(\frac{\hat{a}(\psi_l)}{2}\right).$$

Hence, on the event $B^{\xi}_{\varepsilon,\delta}$, we have almost surely

$$\begin{aligned}
\mathbb{P}\big(R^{\varepsilon} \in B(\gamma, a)|\xi\big) &\le \mathbb{P}\big(R^{\varepsilon} \in B(\gamma, a),\ \widehat{M}^{\varepsilon} \in G^L_N|\xi\big) + \mathbb{P}\big(\widehat{M}^{\varepsilon} \notin K_L|\xi\big) \\
&\le \sum_{i=1}^N \mathbb{P}\big(\lambda^{\varepsilon} \in B(\gamma - \psi_i, a + a_i),\ \widehat{M}^{\varepsilon} \in B(\psi_i, a_i)|\xi\big) \\
&\quad + \mathbb{P}\big(\widehat{M}^{\varepsilon} \notin K_L|\xi\big).
\end{aligned}$$

Now, note that the event $[\lambda^{\varepsilon} \in B(\gamma - \psi_i, a + a_i)]$ is $\sigma(\xi)$-measurable, and we have almost surely

$$\begin{aligned}
\mathbf{1}_{B^{\xi}_{\varepsilon,\delta}} \mathbb{P}\big(R^{\varepsilon} \in B(\gamma, a)|\xi\big) &\le \sum_{i=1}^N \mathbb{1}_{\lambda^{\varepsilon} \in B(\gamma - \psi_i, a + a_i)} \mathbf{1}_{B^{\xi}_{\varepsilon,\delta}} \mathbb{P}\big(\widehat{M}^{\varepsilon} \in B(\psi_i, a_i)|\xi\big) \\
&\quad + \mathbb{P}\big(\widehat{M}^{\varepsilon} \notin K_L|\xi\big).
\end{aligned}$$

Substituting estimations (4.21) and (4.22) in the previous upper bound and using (4.20), we get for $a < \min_i a_i$ and for all ε small enough

$$\begin{aligned}
\mathbb{P}\big(R^{\varepsilon} \in B(\gamma, a)\big) &\le \sum_{i=1}^N \exp\Big(-h^2(\varepsilon)\big(D\big(I_{\xi}(\gamma - \psi_i), \eta\big) + I_W(\psi_i) - \eta\big)\Big) \\
&\quad + \mathbb{P}\big(\widehat{M}^{\varepsilon} \notin K_L\big) + \mathbb{P}\big((B^{\xi}_{\varepsilon,\delta})^c\big).
\end{aligned} \tag{4.23}$$

Remark also that, by the upper bound of the LDP of Proposition 1,

$$\limsup_{\varepsilon \to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\big(\widehat{M}^{\varepsilon} \notin K_L\big) \le -L.$$

From (4.23), we get for $0 < a < \min_i a_i$ and for any given L (i.e., with N constant)

$$\begin{aligned}
\limsup_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\big(R^\varepsilon \in B(\gamma, a)\big)
&\le \max_{1\le i\le N} \max\big\{-L, -\big(D\big(I_\xi(\gamma-\psi_i), \eta\big) + I_W(\psi_i) - \eta\big)\big\} \\
&\le \max\Big\{-L, -\min_{1\le i\le N} D\big(I_\xi(\gamma-\psi_i) + I_W(\psi_i), 2\eta\big)\Big\} \\
&\le \max\Big\{-L, -\inf_\psi D\big(I_\xi(\gamma-\psi) + I_W(\psi), 2\eta\big)\Big\},
\end{aligned}$$

where $\eta < 1/2$ ensures the second inequality. Consequently

$$\begin{aligned}
\limsup_{a\to 0} \limsup_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\big(R^\varepsilon \in B(\gamma, a)\big) \\
\le \max\Big\{-L, -\inf_\psi D\big(I_\xi(\gamma-\psi) + I_W(\psi), 2\eta\big)\Big\};
\end{aligned}$$

letting L go to infinity, we get

$$\limsup_{a\to 0} \limsup_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\big(R^\varepsilon \in B(\gamma, a)\big) \le -\inf_\psi D\big(I_\xi(\gamma-\psi) + I_W(\psi), 2\eta\big).$$

As η is arbitrarily small, we obtain the desired local upper bound.

4.1.3. *Identification of the rate function I.* As problems may arise of noncommutativity of matrices in higher dimensions, let us return in this subsection to dimension d. Recall first the definition of I, with the notation $\sigma_s^2 = \sigma_\xi^2(f(s,\cdot))$,

$$\begin{aligned}
I(\gamma) &= \inf\{I_\xi^f(\gamma-\psi) + I_W(\psi);\ \psi \in C^0([0,1], \mathbb{R}^d)\} \\
&= \inf_\psi \Big\{ \tfrac{1}{2} \int_0^1 \langle \dot\gamma(s) - \dot\psi(s), (\sigma_s^2)^{-1}\big(\dot\gamma(s) - \dot\psi(s)\big)\rangle\, ds \\
&\qquad + \int_0^1 \sup_{\beta\in\mathbb{R}^d} \{\beta^t \dot\psi(s) - \tfrac{1}{2}\beta^t \overline{a_s^2}\beta\}\, ds \Big\}.
\end{aligned}$$

Note that

$$I_W(\psi) = \begin{cases} \tfrac{1}{2} \int_0^1 \langle \dot\psi(s), \big(\overline{a_s^2}\big)_g^{-1} \dot\psi(s)\rangle\, ds, & \text{if } \dot\psi \in \operatorname{range}(\overline{a_\cdot^2}), \\ +\infty, & \text{otherwise,} \end{cases}$$

where $(\overline{a_s^2})_g^{-1}$ denotes the usual generalized inverse of $\overline{a_s^2}$.

In fact, the main difficulty in explicitly identifying this rate function is the noninvertibility of $\overline{a_s^2}$. To circumvent this problem, we will in fact approximate $\overline{a_s^2}$ by

$$A_s^\eta = \overline{a_s^2} + \eta I_d,$$

where I_d is the identity matrix and η a positive number. Now introduce

$$I^\eta(\gamma) = \inf_\psi\{I_\xi^f(\gamma - \psi) + I_W^\eta(\psi)\},$$

where

$$\begin{aligned} I_W^\eta(\psi) &= \int_0^1 \sup_{\beta\in\mathbb{R}^d}\{\beta^t\dot\psi(s) - \tfrac12\beta^t A_s^\eta\beta\}\,ds \\ &= \tfrac12\int_0^1 \|(A_s^\eta)^{-1/2}\dot\psi(s)\|^2\,ds. \end{aligned}$$

We will divide the proof into two steps: the explicit expression of I^η and the convergence of I^η to I as η decreases to 0.

EXPRESSION OF I^η. By the invertibility of A_s^η uniformly in s, we have

$$\begin{aligned} &I_\xi^f(\gamma - \psi) + I_W^\eta(\psi) \\ &\quad = \tfrac12\int_0^1 \|(\sigma_s^2)^{-1/2}(\dot\gamma(s) - \dot\psi(s))\|^2 + \|(A_s^\eta)^{-1/2}\dot\psi(s)\|^2\,ds \\ &\quad = \tfrac12\int_0^1 \|(\sigma_s^2)^{-1/2}\dot\gamma(s)\|^2 - \|((\sigma_s^2)^{-1} + (A_s^\eta)^{-1})^{-1/2}(\sigma_s^2)^{-1}\dot\gamma(s)\|^2\,ds \\ &\quad\quad + \tfrac12\int_0^1 \|((\sigma_s^2)^{-1} + (A_s^\eta)^{-1})^{1/2}\dot\psi(s) \\ &\quad\quad\quad\quad - ((\sigma_s^2)^{-1} + (A_s^\eta)^{-1})^{-1/2}(\sigma_s^2)^{-1}\dot\gamma(s)\|^2\,ds. \end{aligned}$$

Now minimizing in ψ, so that the second integral in the last equality vanishes,

$$\begin{aligned} I^\eta(\gamma) &= \tfrac12\int_0^1 \|(\sigma_s^2)^{-1/2}\dot\gamma(s)\|^2 - \|((\sigma_s^2)^{-1} + (A_s^\eta)^{-1})^{-1/2}(\sigma_s^2)^{-1}\dot\gamma(s)\|^2\,ds \\ &= \tfrac12\int_0^1 \langle\dot\gamma(s), ((\sigma_s^2)^{-1} - ((\sigma_s^2)^{-1}((\sigma_s^2)^{-1} + (A_s^\eta)^{-1})^{-1}(\sigma_s^2)^{-1}))\dot\gamma(s)\rangle\,ds \\ &= \tfrac12\int_0^1 \langle\dot\gamma(s), [(\sigma_s^2)^{-1}((\sigma_s^2)^{-1} + (A_s^\eta)^{-1})^{-1}(A_s^\eta)^{-1}]\dot\gamma(s)\rangle\,ds \\ &= \tfrac12\int_0^1 \langle\dot\gamma(s), (\sigma_s^2 + A_s^\eta)^{-1}\dot\gamma(s)\rangle\,ds. \end{aligned}$$

Consequently,

$$\lim_{\eta\to0} I^\eta(\gamma) = \tfrac12\int_0^1 \langle\dot\gamma(s), (\sigma_\xi^2(f(s,\cdot)) + \overline{a_s^2})^{-1}\dot\gamma(s)\rangle\,ds, \tag{4.24}$$

which is exactly the desired expression (2.10).

CONVERGENCE OF I^{η} TO I. We first want to prove that I_W^{η} converges to I_W as η tends to 0. Note that

$$
\begin{aligned}
I_W(\psi) &= \int_0^1 \sup_{\beta\in\mathbb{R}^d}\{\beta^t\psi - \tfrac{1}{2}\beta^t\overline{a_s^2}\beta\}\,ds \\
&= \int_0^1 \sup_{\beta\in\mathbb{R}^d}\{\beta^t\psi - \tfrac{1}{2}\beta^t(\overline{a_s^2}+\eta I_d)\beta + \tfrac{1}{2}\eta\beta^t\beta\}\,ds \\
&\geq \int_0^1 \sup_{\beta\in\mathbb{R}^d}\{\beta^t\psi - \tfrac{1}{2}\beta^t(\overline{a_s^2}+\eta I_d)\beta\}\,ds \\
&= I_W^{\eta}(\psi).
\end{aligned}
\tag{4.25}
$$

On the other hand, we have

$$
\lim_{\eta\to 0}\langle\dot\psi(s), (\overline{a_s^2}+\eta I_d)^{-1}\dot\psi(s)\rangle = \begin{cases} \langle\dot\psi(s), (\overline{a_s^2})_g^{-1}\dot\psi(s)\rangle, & \text{if } \dot\psi(s)\in\operatorname{range}(\overline{a_s^2}), \\ +\infty, & \text{otherwise.}\end{cases}
$$

Hence Fatou's lemma implies

$$
\begin{aligned}
\liminf_{\eta\to 0} I_W^{\eta}(\psi) &= \liminf_{\eta\to 0}\tfrac{1}{2}\int_0^1 \langle\dot\psi(s), (\overline{a_s^2}+\eta I_d)^{-1}\dot\psi(s)\rangle\,ds \\
&\geq \tfrac{1}{2}\int_0^1 \liminf_{\eta\to 0}\langle\dot\psi(s), (\overline{a_s^2}+\eta I_d)^{-1}\dot\psi(s)\rangle\,ds \\
&= I_W(\psi).
\end{aligned}
$$

Combining this bound with inequality (4.25), we deduce

$$
\lim_{\eta\to 0} I_W^{\eta}(\psi) = I_W(\psi). \tag{4.26}
$$

The function $\eta \to I_W^{\eta}$ is decreasing and relation (4.25) yields that $\eta \to I^{\eta}$ is decreasing and

$$
I^{\eta}(\gamma) \leq I(\gamma).
$$

Now, since all the rate functions here have compact level sets, (4.26) implies

$$
\begin{aligned}
\lim_{\eta\to 0} I^{\eta}(\gamma) &= \sup_{\eta>0}\inf_{\psi}\{I_{\xi}^{f}(\gamma-\psi) + I_W^{\eta}(\psi)\} \\
&= \inf_{\psi}\sup_{\eta>0}\{I_{\xi}^{f}(\gamma-\psi) + I_W^{\eta}(\psi)\} \\
&= I(\gamma).
\end{aligned}
$$

This ends the proof since the limit of I^{η} has been previously shown to be (4.24).

4.2. *Exponential equivalence of* η^{ε} *and* $\hat\eta^{\varepsilon}$. We divide this section into two steps, giving successive equivalent forms of η^{ε} w.r.t. the LDP, the last one being $\hat\eta^{\varepsilon}$.

Step 1. Consider the following diffusion: for $t \in [0,1]$,

$$(4.27)\qquad \begin{aligned} d\tilde{X}_t^\varepsilon &= b(\tilde{X}_t^\varepsilon, \xi_{t/\varepsilon})\,dt + \sqrt{\varepsilon}\,a(\bar{x}_t, \xi_{t/\varepsilon})\,dW_t,\\ \tilde{X}_0^\varepsilon &= x_0 \end{aligned}$$

and, for $t \in [0,1]$,

$$(4.28)\qquad \tilde{\eta}^\varepsilon = \frac{\tilde{X}_t^\varepsilon - \bar{x}_t}{\sqrt{\varepsilon}h(\varepsilon)}.$$

We will prove in this step that η^ε and $\tilde{\eta}^\varepsilon$ are exponentially equivalent w.r.t. the LDP, that is, for all positive r,

$$(4.29)\qquad \lim_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log \mathbb{P}\Big(\sup_{t\in[0,1]}|\eta_t^\varepsilon - \tilde{\eta}_t^\varepsilon| > r\Big) = -\infty.$$

To this end, observe that, using the Lipschitz property of b,

$$\begin{aligned} |\eta_t^\varepsilon - \tilde{\eta}_t^\varepsilon| &= \left|\frac{X_t^\varepsilon - \tilde{X}_t^\varepsilon}{\sqrt{\varepsilon}h(\varepsilon)}\right|\\ &= \bigg|\int_0^t \frac{b(X_s^\varepsilon, \xi_{s/\varepsilon}) - b(\tilde{X}_s^\varepsilon, \xi_{s/\varepsilon})}{\sqrt{\varepsilon}h(\varepsilon)}\,ds\\ &\qquad + \frac{1}{h(\varepsilon)}\int_0^t \big(a(X_s^\varepsilon, \xi_{s/\varepsilon}) - a(\bar{x}_s, \xi_{s/\varepsilon})\big)\,dW_s\bigg|\\ &\le K\int_0^t \left|\frac{X_s^\varepsilon - \tilde{X}_s^\varepsilon}{\sqrt{\varepsilon}h(\varepsilon)}\right|ds + \left|\frac{1}{h(\varepsilon)}\int_0^t \big(a(X_s^\varepsilon, \xi_{s/\varepsilon}) - a(\bar{x}_s, \xi_{s/\varepsilon})\big)\,dW_s\right|\\ &= K\int_0^t |\eta_s^\varepsilon - \tilde{\eta}_s^\varepsilon|\,ds + \left|\frac{1}{h(\varepsilon)}\int_0^t \big(a(X_s^\varepsilon, \xi_{s/\varepsilon}) - a(\bar{x}_s, \xi_{s/\varepsilon})\big)\,dW_s\right|. \end{aligned}$$

Then, writing

$$\tilde{M}_t^\varepsilon = \frac{1}{h(\varepsilon)}\int_0^t \big(a(X_s^\varepsilon, \xi_{s/\varepsilon}) - a(\bar{x}_s, \xi_{s/\varepsilon})\big)\,dW_s,$$

Gronwall's lemma implies

$$|\eta_t^\varepsilon - \tilde{\eta}_t^\varepsilon| \le e^K \sup_{s\in[0,1]}|\tilde{M}_s^\varepsilon| \qquad \forall\, t \in [0,1].$$

Therefore, (4.29) will follow if we can prove that, for each positive r,

$$\lim_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log \mathbb{P}\Big(\sup_{t\in[0,1]}|\tilde{M}_t^\varepsilon| > r\Big) = -\infty.$$

For any fixed arbitrary positive L, observe

$$\mathbb{P}\Big(\sup_{t\in[0,1]}|\tilde{M}_t^\varepsilon| > r\Big) \le \mathbb{P}\Big(\sup_{t\in[0,1]}|\tilde{M}_t^\varepsilon| > r;\ \langle\tilde{M}_1^\varepsilon\rangle \le L\varepsilon\Big) + \mathbb{P}\big(\langle\tilde{M}_1^\varepsilon\rangle > L\varepsilon\big).$$

Note now that, by Bernstein's inequality,

$$\frac{1}{h^2(\varepsilon)} \log \mathbb{P}\Big(\sup_{t\in[0,1]} |\tilde{M}_t^\varepsilon| > r;\ \langle \tilde{M}^\varepsilon \rangle_1 \le L\varepsilon \Big) \le -\frac{r^2}{2L\varepsilon\, h^2(\varepsilon)}$$

and since $\varepsilon h^2(\varepsilon) \to 0$ when ε decreases to 0, we have, for all arbitrary positive L and r,

$$\lim_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\Big(\sup_{t\in[0,1]} |\tilde{M}_t^\varepsilon| > r;\ \langle \tilde{M}_1^\varepsilon \rangle \le L\varepsilon \Big) = -\infty.$$

To conclude we only have to establish that

$$\lim_{L\to\infty} \lim_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\Big(\frac{1}{\varepsilon} \langle \tilde{M}_1^\varepsilon \rangle > L \Big) = -\infty. \tag{4.30a}$$

Note first that, since b is Lipschitz continuous,

$$\begin{aligned}
|\eta_t^\varepsilon| &= \left| \frac{1}{\sqrt{\varepsilon} h(\varepsilon)} \int_0^t \big(b(X_s^\varepsilon, \xi_{s/\varepsilon}) - \bar{b}(\bar{x}_s) \big)\, ds + \frac{1}{h(\varepsilon)} \int_0^t a(X_s^\varepsilon, \xi_{s/\varepsilon})\, dW_s \right| \\
&= \left| \frac{1}{\sqrt{\varepsilon} h(\varepsilon)} \int_0^t \big(b(X_s^\varepsilon, \xi_{s/\varepsilon}) - b(\bar{x}_s, \xi_{s/\varepsilon}) \big)\, ds + \lambda_t^\varepsilon + \frac{1}{h(\varepsilon)} \int_0^t a(X_s^\varepsilon, \xi_{s/\varepsilon})\, dW_s \right| \\
&\le K \int_0^t |\eta_s^\varepsilon|\, ds + \left| \lambda_t^\varepsilon + \frac{1}{h(\varepsilon)} \int_0^t a(X_s^\varepsilon, \xi_{s/\varepsilon})\, dW_s \right|
\end{aligned}$$

and then by Gronwall's lemma,

$$\sup_{t\in[0,1]} |\eta_t^\varepsilon| \le \sup_{t\in[0,1]} \left| \lambda_t^\varepsilon + \frac{1}{h(\varepsilon)} \int_0^t a(X_s^\varepsilon, \xi_{s/\varepsilon})\, dW_s \right|.$$

Therefore, since a is Lipschitz continuous,

$$\begin{aligned}
\frac{1}{\varepsilon} \langle \tilde{M}_1^\varepsilon \rangle &= \int_0^1 \left(\frac{a(X_s^\varepsilon, \xi_{s/\varepsilon}) - a(\bar{x}_s, \xi_{s/\varepsilon})}{\sqrt{\varepsilon} h(\varepsilon)} \right)^2 ds \\
&\le K \int_0^1 \left(\frac{X_s^\varepsilon - \bar{x}_s}{\sqrt{\varepsilon} h(\varepsilon)} \right)^2 ds \\
&= K \int_0^1 |\eta_s^\varepsilon|^2\, ds \\
&\le K \Big(\sup_{t\in[0,1]} |\eta_t^\varepsilon| \Big)^2 \\
&\le K \left(\sup_{t\in[0,1]} |\lambda_t^\varepsilon| + \sup_{t\in[0,1]} \left| \frac{1}{h(\varepsilon)} \int_0^t a(X_s^\varepsilon, \xi_{s/\varepsilon})\, dW_s \right| \right)^2.
\end{aligned}$$

Consequently, (4.30a) is obtained if

$$\lim_{L\to\infty}\lim_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\Big(\sup_{t\in[0,1]}|\lambda_t^\varepsilon|>L\Big)=-\infty, \tag{4.30b}$$

$$\lim_{L\to\infty}\lim_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\bigg(\sup_{t\in[0,1]}\bigg|\frac{1}{h(\varepsilon)}\int_0^t a(X_s^\varepsilon,\xi_{s/\varepsilon})\,dW_s\bigg|>L\bigg)=-\infty. \tag{4.30c}$$

Note that the first statement is exactly (4.10) and the second follows directly from Bernstein's inequality and the boundedness of a, as for (4.11). The desired negligibility is thus proved.

Step 2. Let us introduce the process $\hat{\eta}^\varepsilon$ defined for $t\in[0,1]$ by

$$\hat{\eta}_t^\varepsilon=\lambda_t^\varepsilon+\widehat{M}_t^\varepsilon+\int_0^t\bar{B}(\bar{x}_s)\hat{\eta}_s^\varepsilon\,ds,$$

where λ^ε is defined by (4.1) and $\widehat{M}^\varepsilon$ by (4.2). We want to establish in this step that $\tilde{\eta}^\varepsilon$ and $\hat{\eta}^\varepsilon$ are exponentially equivalent w.r.t. the LDP, that is, for each positive δ,

$$\lim_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\Big(\sup_{t\in[0,1]}|\hat{\eta}_t^\varepsilon-\tilde{\eta}_t^\varepsilon|>\delta\Big)=-\infty. \tag{4.31}$$

To simplify the notation, put

$$\Phi(s,\xi_{s/\varepsilon})=B(\bar{x}_s,\xi_{s/\varepsilon})-\bar{B}(\bar{x}_s), \tag{4.32}$$

$$\begin{aligned}\Psi\big(s,\tilde{X}_s^\varepsilon,\xi_{s/\varepsilon}\big)=\frac{1}{\sqrt{\varepsilon}h(\varepsilon)}\big[&b\big(\tilde{X}_s^\varepsilon,\xi_{s/\varepsilon}\big)-b(\bar{x}_s,\xi_{s/\varepsilon})\\&-\sqrt{\varepsilon}h(\varepsilon)B(\bar{x}_s,\xi_{s/\varepsilon})\tilde{\eta}_s^\varepsilon\big].\end{aligned} \tag{4.33}$$

First note that, by definition of $\tilde{\eta}_t^\varepsilon$ [see (4.27) and (4.28)], we have the following crucial decomposition:

$$\begin{aligned}\tilde{\eta}_t^\varepsilon&=\lambda_t^\varepsilon+\frac{1}{h(\varepsilon)}\int_0^t a(\bar{x}_s,\xi_{s/\varepsilon})\,dW_s+\int_0^t\bar{B}(\bar{x}_s)\tilde{\eta}_s^\varepsilon\,ds\\&\quad+\int_0^t\big(B(\bar{x}_s,\xi_{s/\varepsilon})-\bar{B}(\bar{x}_s)\big)\tilde{\eta}_s^\varepsilon\,ds\\&\quad+\frac{1}{\sqrt{\varepsilon}h(\varepsilon)}\int_0^t\big[b\big(\tilde{X}_s^\varepsilon,\xi_{s/\varepsilon}\big)-b(\bar{x}_s,\xi_{s/\varepsilon})-\sqrt{\varepsilon}h(\varepsilon)B(\bar{x}_s,\xi_{s/\varepsilon})\tilde{\eta}_s^\varepsilon\big]\,ds\\&=\lambda_t^\varepsilon+\widehat{M}_t^\varepsilon+\int_0^t\bar{B}(\bar{x}_s)\tilde{\eta}_s^\varepsilon\,ds+\int_0^t\Phi(s,\xi_{s/\varepsilon})\eta_s^\varepsilon\,ds+\int_0^t\Psi\big(s,\tilde{X}_s^\varepsilon,\xi_{s/\varepsilon}\big)\,ds.\end{aligned} \tag{4.34}$$

Therefore,

$$|\tilde{\eta}_t^\varepsilon-\hat{\eta}_t^\varepsilon|\le\int_0^1|\bar{B}(\bar{x}_s)||\tilde{\eta}_s^\varepsilon-\hat{\eta}_s^\varepsilon|\,ds+\bigg|\int_0^t\Phi(s,\xi_{s/\varepsilon})\eta_s^\varepsilon\,ds+\int_0^t\Psi\big(s,\tilde{X}_s^\varepsilon,\xi_{s/\varepsilon}\big)\,ds\bigg|.$$

By this decomposition, Gronwall's lemma entails

$$|\tilde{\eta}_t^\varepsilon - \hat{\eta}_t^\varepsilon| \le e^{\|\bar{B}\|}\biggl(\sup_{t\in[0,1]}\biggl|\int_0^t \Phi(s,\xi_{s/\varepsilon})\tilde{\eta}_s^\varepsilon\,ds\biggr| + \sup_{t\in[0,1]}\biggl|\int_0^t \Psi(s,\tilde{X}_s^\varepsilon,\xi_{s/\varepsilon})\,ds\biggr|\biggr).$$

Thus, we have to prove the negligibility of the last two terms of this inequality, that is, for all positive δ,

$$\lim_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\biggl(\sup_{t\in[0,1]}\biggl|\int_0^t \Psi(s,\tilde{X}_s^\varepsilon,\xi_{s/\varepsilon})\,ds\biggr| > \delta\biggr) = -\infty, \tag{4.35}$$

$$\lim_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\biggl(\sup_{t\in[0,1]}\biggl|\int_0^t \Phi(s,\xi_{s/\varepsilon})\tilde{\eta}_s^\varepsilon\,ds\biggr| > \delta\biggr) = -\infty. \tag{4.36}$$

To this end, note the following:

$$|\Phi(s,\xi_{s/\varepsilon})| \le \|B\| + \|\bar{B}\|, \tag{4.37}$$

$$|\Psi(s,\tilde{X}_s^\varepsilon,\xi_{s/\varepsilon})| \le K|\tilde{\eta}_s^\varepsilon|, \tag{4.38}$$

$$|\Psi(s,\tilde{X}_s^\varepsilon,\xi_{s/\varepsilon})| \le \sqrt{\varepsilon}h(\varepsilon)K|\tilde{\eta}_s^\varepsilon|^2, \tag{4.39}$$

and, by Gronwall's lemma, decomposition (4.34), inequalities (4.37), (4.38) and the boundedness of $\bar{B}$ imply

$$|\tilde{\eta}_t^\varepsilon| \le K\Bigl(\sup_{t\in[0,1]}|\lambda_t^\varepsilon| + \sup_{t\in[0,1]}|\widehat{M}_t^\varepsilon|\Bigr). \tag{4.40}$$

We will use here the techniques introduced in Guillin [(2001), Section 3] to prove (4.35) and (4.36).

Begin with (4.35). Combining inequalities (4.39) and (4.40), we have

$$\begin{aligned}
&\mathbb{P}\biggl(\sup_{t\in[0,1]}\biggl|\int_0^t \Psi(s,\tilde{X}_s^\varepsilon,\xi_{s/\varepsilon})\,ds\biggr| > \delta\biggr)\\
&\quad\le \mathbb{P}\biggl(\int_0^1 |\tilde{\eta}_s^\varepsilon|^2\,ds > \frac{\delta}{K\sqrt{\varepsilon}h(\varepsilon)}\biggr)\\
&\quad\le \mathbb{P}\biggl(\sup_{t\in[0,1]}|\lambda_t^\varepsilon|^2 > \frac{\delta}{4K\sqrt{\varepsilon}h(\varepsilon)}\biggr) + \mathbb{P}\biggl(\sup_{t\in[0,1]}|\widehat{M}_t^\varepsilon|^2 > \frac{\delta}{4K\sqrt{\varepsilon}h(\varepsilon)}\biggr).
\end{aligned}$$

Note that the negligibility of these last two terms is easily deduced from (4.10) and (4.11).

Let us deal now with (4.36), which is much more difficult. First, using (4.34), we get

$$\begin{aligned}
\int_0^t \Phi(s,\xi_{s/\varepsilon})\tilde{\eta}_s^\varepsilon\,ds &= \int_0^t \Phi(s,\xi_{s/\varepsilon})\bigl(\lambda_s^\varepsilon + \widehat{M}_s^\varepsilon\bigr)\,ds\\
&\quad + \int_0^t \Phi(s,\xi_{s/\varepsilon})\int_0^s \bar{B}(\bar{x}_u)\tilde{\eta}_u^\varepsilon\,du\,ds
\end{aligned}$$

$$+\int_0^t \Phi(s,\xi_{s/\varepsilon})\int_0^s \Phi(u,\xi_{u/\varepsilon})\tilde{\eta}_u^\varepsilon\,du\,ds$$
$$+\int_0^t \Phi(s,\xi_{s/\varepsilon})\int_0^s \Psi\big(u,\tilde{X}_u^\varepsilon,\xi_{u/\varepsilon}\big)\,du\,ds.$$

Hence by Gronwall's lemma, applied to $\int_0^t \Phi(s,\xi_{s/\varepsilon})\tilde{\eta}_s^\varepsilon\,ds$,

$$\begin{aligned}
&\left|\int_0^t \Phi(s,\xi_{s/\varepsilon})\tilde{\eta}_s^\varepsilon\,ds\right| \\
&\quad\le e^{(\|B\|+\|\bar{B}\|)}\sup_{t\in[0,1]}\biggl(\left|\int_0^t \Phi(s,\xi_{s/\varepsilon})\lambda_s^\varepsilon\,ds\right| \\
&\qquad\qquad + \left|\int_0^t \Phi(s,\xi_{s/\varepsilon})\widehat{M}_s^\varepsilon\,ds\right| \\
&\qquad\qquad + \left|\int_0^t \Phi(s,\xi_{s/\varepsilon})\int_0^s \bar{B}(\bar{x}_u)\tilde{\eta}_u^\varepsilon\,du\,ds\right| \\
&\qquad\qquad + \left|\int_0^t \Phi(s,\xi_{s/\varepsilon})\int_0^s \Psi\big(u,\tilde{X}_u^\varepsilon,\xi_{u/\varepsilon}\big)\,du\,ds\right|\biggr) \\
&\quad= e^{(\|B\|+\|\bar{B}\|)}\sup_{t\in[0,1]}\big(|\mathrm{I}_t^\varepsilon|+|\mathrm{II}_t^\varepsilon|+|\mathrm{III}_t^\varepsilon|+|\mathrm{IV}_t^\varepsilon|\big).
\end{aligned}\tag{4.41}$$

Now we have to prove negligibility with respect to the moderate deviations [i.e., with speed $h^{-2}(\varepsilon)$] of the four terms on the right-hand side of this last inequality. For the term $\mathrm{IV}_t^\varepsilon$, note that

$$\sup_{t\in[0,1]}|\mathrm{IV}_t^\varepsilon|\le(\|B\|+\|\bar{B}\|)\sup_{t\in[0,1]}\left|\int_0^t \Psi\big(s,\tilde{X}_s^\varepsilon,\xi_{s/\varepsilon}\big)\,ds\right|$$

and (4.35) thus provides the needed negligibility of $\sup_{t\in[0,1]}|\mathrm{IV}_t^\varepsilon|$. The third term needs more effort.

(a) *Negligibility of* $\mathrm{III}_t^\varepsilon$. First, by integrating by parts,

$$\begin{aligned}
\mathrm{III}_t^\varepsilon &= \int_0^t \Phi(s,\xi_{s/\varepsilon})\int_0^s \bar{B}(\bar{x}_u)\tilde{\eta}_u^\varepsilon\,du\,ds \\
&= \int_0^t \Phi(s,\xi_{s/\varepsilon})\,ds\int_0^t \bar{B}(\bar{x}_u)\tilde{\eta}_u^\varepsilon\,du \\
&\quad - \int_0^t \bar{B}(\bar{x}_u)\tilde{\eta}_u^\varepsilon\int_0^u \Phi(s,\xi_{s/\varepsilon})\,ds\,du.
\end{aligned}\tag{4.42}$$

For the first term on the right-hand side, for all $\delta > 0$, $L > 0$, we have, using (4.40),

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{t\in[0,1]}\Big|\int_0^t \Phi(s,\xi_{s/\varepsilon})\,ds \int_0^t \bar{B}(\bar{x}_u)\tilde{\eta}_u^\varepsilon\,du\Big| > \delta\Big)\\
&\qquad\le \mathbb{P}\Big(\sup_{t\in[0,1]}\Big|\int_0^t \Phi(s,\xi_{s/\varepsilon})\,ds\Big| > \frac{\delta}{L}\Big)\\
&\qquad\quad + \mathbb{P}\Big(\sup_{t\in[0,1]}\Big|\int_0^t \bar{B}(\bar{x}_u)\tilde{\eta}_u^\varepsilon\,du\Big| > \delta L\Big)\\
&\qquad\le \mathbb{P}\Big(\sup_{t\in[0,1]}\Big|\int_0^t \Phi(s,\xi_{s/\varepsilon})\,ds\Big| > \frac{\delta}{L}\Big)\\
&\qquad\quad + \mathbb{P}\Big(\sup_{s\in[0,1]}|\lambda_s^\varepsilon| > \frac{L\delta}{K}\Big) + \mathbb{P}\Big(\sup_{s\in[0,1]}|\widehat{M}_s^\varepsilon| > \frac{L\delta}{K}\Big).
\end{aligned}\tag{4.43}$$

The negligibility of the last two terms has been established previously in (4.10) and (4.11) after a suitable asymptotic of L; and for the first note that

$$\begin{aligned}
&\limsup_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\Big(\sup_{t\in[0,1]}\Big|\int_0^t \Phi(s,\xi_{s/\varepsilon})\,ds\Big| > \frac{\delta}{L}\Big)\\
&\quad = \limsup_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\Big(\sup_{t\in[0,1]}\Big|\frac{1}{\sqrt{\varepsilon}h(\varepsilon)}\int_0^t \Phi(s,\xi_{s/\varepsilon})\,ds\Big| > \frac{\delta}{L\sqrt{\varepsilon}h(\varepsilon)}\Big).
\end{aligned}$$

Under our assumptions, it is easily seen, as B is Lipschitz continuous, that Φ satisfies condition (G) of Theorem 3; then the desired negligibility follows from Corollary 1 and the fact that $\delta/(L\sqrt{\varepsilon}h(\varepsilon)) \to \infty$ for given δ and L.

Now we deal with the second term of the integration by parts (4.42). First observe that, by the boundedness of $\bar{B}$ and (4.40),

$$\begin{aligned}
\Big|\int_0^t \bar{B}(\bar{x}_u)\tilde{\eta}_u^\varepsilon \int_0^u \Phi(s,\xi_{s/\varepsilon})\,ds\,du\Big| &\le K\sup_{u\in[0,1]}|\tilde{\eta}_u^\varepsilon|\sup_{u\in[0,1]}\Big|\int_0^u \Phi(s,\xi_{s/\varepsilon})\,ds\Big|\\
&\le K\sup_{u\in[0,1]}|\lambda_u^\varepsilon|\sup_{u\in[0,1]}\Big|\int_0^u \Phi(s,\xi_{s/\varepsilon})\,ds\Big|\\
&\quad + K\sup_{u\in[0,1]}|\widehat{M}_u^\varepsilon|\sup_{u\in[0,1]}\Big|\int_0^u \Phi(s,\xi_{s/\varepsilon})\,ds\Big|.
\end{aligned}$$

Now, by the technique used for the first term of the integration by parts (4.42), we also obtain, for all $\delta > 0$,

$$\lim_{\varepsilon\to 0}\frac{1}{h^2(\varepsilon)}\log\mathbb{P}\Big(\sup_{t\in[0,1]}\Big|\int_0^t \bar{B}(\bar{x}_u)\tilde{\eta}_u^\varepsilon\int_0^u \Phi(s,\xi_{s/\varepsilon})\,ds\,du\Big| > \delta\Big) = -\infty.$$

We have thus proved the negligibility of each term of (4.43), which yields the desired negligibility of III^ε, that is, for all $\delta > 0$,

$$\lim_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\Big(\sup_{t\in[0,1]} |\mathrm{III}_t^\varepsilon| > \delta\Big) = -\infty.$$

(b) *Negligibility of* I_t^ε *and* $\mathrm{II}_t^\varepsilon$. In fact, we will only establish the negligibility of $\mathrm{II}_t^\varepsilon$, as the proof for I_t^ε is absolutely similar. Consequently, we only prove here that, for all positive δ,

$$\limsup_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\Big(\sup_{t\in[0,1]} |\mathrm{II}_t^\varepsilon| > \delta\Big) = -\infty. \tag{4.44}$$

Recall that $\mathrm{II}_t^\varepsilon = \int_0^t \Phi(s, \xi_{s/\varepsilon})\widehat{M}_s^\varepsilon\, ds$. First, by Lemma 1, $(\widehat{M}^\varepsilon)$ satisfies an LDP with good rate function I_W, and $I_W(\gamma) = +\infty$ if γ does not belong to the Cameron–Martin space. Let $L > 0$ and $\delta > 0$. As the level set $K_L = [\gamma; I_W(\gamma) \le L]$ is compact, we may choose, for any $\eta \in (0, \delta/(2(\|B\| + \|\bar{B}\|))$, a finite collection of mappings $\gamma_1, \dots, \gamma_N \in K_L$ (with N depending on δ and L) such that $K_L \subset \bigcup_{l=1}^N B(\gamma_l, \eta/2)$.

A crucial fact is that each γ_l is in the Cameron–Martin space and satisfies, by the Cauchy–Schwarz inequality,

$$|\gamma_l(t) - \gamma_l(s)| = \left|\int_s^t \dot{\gamma}_l(u)\, du\right| \le \sqrt{|t-s|}\sqrt{\int_s^t \dot{\gamma}_l^2(u)\, du},$$

which implies, by the absolute continuity of γ_l,

$$\lim_{\varepsilon\to 0} \sup_{|s-t|\le\varepsilon} \frac{|\gamma_l(t) - \gamma_l(s)|}{\sqrt{|t-s|}} = 0.$$

Note that it is exactly the last statement in condition (G) of Theorem 3. Then, for all positive δ,

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{t\in[0,1]} |\mathrm{II}_t^\varepsilon| > \delta\Big)\\
&\quad = \mathbb{P}\Big(\sup_{t\in[0,1]} |\mathrm{II}_t^\varepsilon| > \delta;\, \widehat{M}^\varepsilon \in B(K_l, \eta)\Big)\\
&\qquad + \mathbb{P}\Big(\sup_{t\in[0,1]} |\mathrm{II}_t^\varepsilon| > \delta;\, \widehat{M}^\varepsilon \notin B(K_L, \eta)\Big)\\
&\quad \le \sum_{l=1}^N \mathbb{P}\Big(\sup_{t\in[0,1]} |\mathrm{II}_t^\varepsilon| > \delta;\, \widehat{M}^\varepsilon \in B(\gamma_l, \eta)\Big) + \mathbb{P}\big(\widehat{M}^\varepsilon \notin B(K_l, \eta)\big)\\
&\quad \le \sum_{l=1}^N \mathbb{P}\Big(\sup_{t\in[0,1]} \Big|\mathrm{II}_t^\varepsilon - \int_0^t \Phi(s, \xi_{s/\varepsilon})\gamma_l(s)\, ds\Big| > \frac{\delta}{2};\, \widehat{M}^\varepsilon \in B(\gamma_l, \eta)\Big)\\
&\qquad + \sum_{l=1}^N \mathbb{P}\Big(\sup_{t\in[0,1]} \Big|\int_0^t \Phi(s, \xi_{s/\varepsilon})\gamma_l(s)\, ds\Big| > \frac{\delta}{2}\Big) + \mathbb{P}\big(\widehat{M}^\varepsilon \notin B(K_l, \eta)\big).
\end{aligned} \tag{4.45}$$

First observe that for all l, on the event $[\widehat{M}^\varepsilon \in B(\gamma_l, \eta)]$,

$$\sup_{t\in[0,1]}\left|\mathrm{II}_t^\varepsilon - \int_0^t \Phi(s, \xi_{s/\varepsilon})\gamma_l(s)\,ds\right| \le \eta \int_0^1 |\Phi(s, \xi_{s/\varepsilon})|\,ds$$
$$\le \eta(\|B\| + \|\bar{B}\|),$$

and, for $\eta < \delta/(2(\|B\| + \|\bar{B}\|))$, we have

$$\sum_{l=1}^N \mathbb{P}\left(\sup_{t\in[0,1]}\left|\mathrm{II}_t^\varepsilon - \int_0^t \Phi(s, \xi_{s/\varepsilon})\gamma_l(s)\,ds\right| > \frac{\delta}{2};\ \widehat{M}^\varepsilon \in B(\gamma_l, \eta)\right) = 0. \tag{4.46}$$

For the second term of (4.45), note that the mapping $\tilde{f}$, defined by

$$\tilde{f}(s, \xi_{s/\varepsilon}) = \big(B(\bar{x}_s, \xi_{s/\varepsilon}) - \bar{B}(\bar{x}_s)\big)\gamma(s),$$

satisfies condition (G) of Theorem 3, since Φ is Lipschitz continuous and γ satisfies (G3). Hence, by Theorem 3 and Corollary 1, for each $\gamma \in K_L$, we have

$$\begin{aligned}
&\limsup_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\left(\sup_{t\in[0,1]}\left|\int_0^t \Phi(s, \xi_{s/\varepsilon})\gamma(s)\,ds\right| > \frac{\delta}{2}\right) \\
&\quad = \limsup_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\left(\sup_{t\in[0,1]}\left|\frac{1}{\sqrt{\varepsilon}h(\varepsilon)}\int_0^t \tilde{f}(s, \xi_{s/\varepsilon})\,ds\right| > \frac{\delta}{2\sqrt{\varepsilon}h(\varepsilon)}\right) \\
&\quad = -\infty.
\end{aligned} \tag{4.47}$$

By Proposition 1, $\widehat{M}^\varepsilon$ satisfies an LDP upper bound. Hence, if ε is sufficiently small,

$$\begin{aligned}
\limsup_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\big(\widehat{M}^\varepsilon \notin B(K_L, \eta)\big) &\le \inf_{\gamma\notin K_L} I_W(\gamma) \\
&\le -L.
\end{aligned} \tag{4.48}$$

Then, as all the summations considered in (4.45) are finite sums for given L, taking the limit in (4.45), combining relations (4.46)–(4.48), we get that, for all positive δ,

$$\limsup_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\left(\sup_{t\in[0,1]} |\mathrm{II}_t^\varepsilon| > \delta\right) \le -L,$$

and as L is arbitrary, letting L tend to infinity, we obtain the desired relation (4.44) and with it relation (4.36).

Finally, $\tilde{\eta}^\varepsilon$ and $\hat{\eta}^\varepsilon$ share the same LDP, and by Step 1, we conclude that η^ε and $\hat{\eta}^\varepsilon$ are exponentially equivalent w.r.t. the LDP.

Acknowledgments. The author thanks Liming Wu for a number of fruitful discussions and advice, and gratefully acknowledges the careful review by the anonymous referees, which led to significant improvement in the presentation of the paper and to identifying the rate function of Theorem 1.

REFERENCES

BAIER, U. and FREIDLIN, M. I. (1977). Theorems on large deviations and stability under random perturbations. *Dokl. Akad. Nauk USSR* **235** 253–256.

BERNARD, P. and RACHAD, A. (2000). Moyennisation d'un oscillateur stochastique quasi-conservatif. *C. R. Acad. Sci. Paris* **331** 1029–1032.

BOGOLUBOV, N. and MITROPOLSKII, A. (1961). *Asymptotics Methods in Non-linear Mechanics*. Gordon and Breach, New York.

DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviations Techniques and Their Applications*, 2nd ed. Jones and Bartlett, Boston.

DEUSCHEL, J. D. and STROOCK, D. W. (1989). *Large Deviations*. Academic Press, New York.

DOWN, D., MEYN, P. and TWEEDIE, R. (1995). Exponential and uniform ergodicity of Markov processes. *Ann. Probab.* **23** 1671–1691.

FENG, J. and KURTZ, T. (2000). Large deviations for stochastic processes. Preprint. Available at http://www.math.wisc.edu/kurtz/feng.

FREIDLIN, M. I. (1978). The averaging principle and theorem on large deviations. *Uspekhi Mat. Nauk* **33** 107–160.

FREIDLIN, M. I. and WENTZELL, W. D. (1998). *Random Perturbations of Dynamical Systems*, 2nd ed. Springer, New York.

GUILLIN, A. (2001). Moderate deviations of inhomogeneous functionals of Markov processes and application to averaging. *Stochastic Process. Appl.* **92** 287–313.

KHASMINSKII, R. Z. (1968). On the principle of averaging for the Itô stochastic differential equations. *Kybernetika* (*Czechoslovakia*) **4** 260–279.

KHASMINSKII, R. Z. (1980). *Stochastic Stability of Differential Equations*. Sijthoff and Noordhoff, Rockville, MD.

KIEFER, Y. (2000). Averaging and climate models. In *Stochastic Climate Models*. Birkhäuser, Boston.

KLEBANER, F. C. and LIPTSER, R. S. (1999). Moderate deviations for randomly perturbed dynamical systems. *Stochastic Process. Appl.* **80** 157–176.

LIPTSER, R. S. (1994). The Bogolubov averaging principle for semimartingales. *Proc. Steklov Inst. Math.* **4** 000–000.

LIPTSER, R. S. (1996). Large deviations for two scaled diffusions. *Probab. Theory Related Fields* **106** 71–104.

LIPTSER, R. S. and PUHALSKII, A. A. (1992). Limit theorems on large deviations for semimartingales. *Stochastics Stochastics Rep.* **38** 201–249.

LIPTSER, R. S. and SPOKOINY, V. (1999). Moderate deviations type evaluation for integral functional of diffusion processes. *Electron. J. Probab.* **4** 1–25.

LIPTSER, R. S. and STOYANOV, J. (1990). Stochastic version of the averaging principle for diffusion type processes. *Stochastics Stochastics Rep.* **32** 145–163.

NUMMELIN, E. (1984). *General Irreducible Markov Chains and Non-negative Operators*. Cambridge Univ. Press.

PARDOUX, E. and VERETENNIKOV, A. Y. (2000). On Poisson equation and diffusion approximation, 2. Preprint.

PARDOUX, E. and VERETENNIKOV, A. Y. (2001). On Poisson equation and diffusion approximation, 1. *Ann. Probab.* **29** 1061–1085.

PUHALSKII, A. A. (1990). Large deviations of semimartingales via convergence of the predictable characteristics. *Stochastics Stochastics Rep.* **49** 27–85.

RACHAD, A. (1999). Principes de moyennisation pour des oscillateurs stochastiques non linéaires. Thèse, Univ. Blaise Pascal.

REVUZ, D. and YOR, M. (1994). *Continuous Martingales and Brownian Motion*, 2nd ed. Springer, Berlin.

SANDERS, J. A. and VERHULST, F. (1985). *Averaging Method in Non-linear Dynamical Systems*. Springer, Berlin.

VERETENNIKOV, A. Y. (1999a). On large deviations for stochastic differential equations with a small diffusion and averaging. *Theory Probab. Appl.* **43** 335–337.

VERETENNIKOV, A. Y. (1999b). On large deviations in the averaging principle for stochastic differential equations with complete dependence. *Theory Probab. Appl.* **43** 664–666.

VERETENNIKOV, A. Y. (1999c). On large deviations in the averaging principle for stochastic differential equations with complete dependence. *Ann. Probab.* **27** 284–296.

WU, L. (2001). Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stochastic Process. Appl.* **91** 205–238.

CEREMADE
UNIVERSITÉ PARIS DAUPHINE
PLACE MARÉCHAL DE LATTRE DE TASSIGNY
75776 PARIS CEDEX 16
FRANCE
E-MAIL: guillin@ceremade.dauphine.fr