MODERATE DEVIATION PRINCIPLE FOR DYNAMICAL SYSTEMS WITH SMALL RANDOM PERTURBATION

YUTAO MA, RAN WANG, AND LIMING WU

Abstract. Consider the stochastic differential equation in $\mathbb{R}^d$

$$\begin{cases} dX_t^\varepsilon = b(X_t^\varepsilon)dt + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)dB_t; \\ X_0^\varepsilon = x_0, \quad x_0 \in \mathbb{R}^d \end{cases}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ is C^1 such that $\langle x, b(x)\rangle \le C(1+|x|^2)$, $\sigma : \mathbb{R}^d \to \mathcal{M}(d\times n)$ is locally Lipschitzian with linear growth, and B_t is a standard Brownian motion taking values in $\mathbb{R}^n$. Freidlin-Wentzell's theorem gives the large deviation principle for X^ε for small ε. In this paper we establish its moderate deviation principle.

MSC 2010: 60F10; 60H10.

Keywords: Freidlin-Wentzell's theorem; large deviations; moderate deviations, Talagrand's transportation inequality.

1. Introduction and main result

Consider the stochastic differential equation

$$\begin{cases} dX_t^\varepsilon = b(X_t^\varepsilon)dt + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)dB_t; \\ X_0^\varepsilon = x_0, \end{cases} \tag{1.1}$$

where $x_0 \in \mathbb{R}^d$ is the starting point, $b : \mathbb{R}^d \to \mathbb{R}^d$ is the macroscopic vector field, $\sigma : \mathbb{R}^d \to \mathcal{M}(d\times n)$ (the space of all real $d\times n$ matrices) and (B_t) is the standard Brownian motion taking values in $\mathbb{R}^n$ defined on some well filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. We always assume

(**H**): *σ is locally Lipschitzian, b is C^1 and there exists some positive constant C such that*

$$\max\left\{ \mathrm{tr}(\sigma\sigma^*(x)),\ \langle x, b(x)\rangle \right\} \le C(1+|x|^2),\ \forall x \in \mathbb{R}^d.$$

Here $\langle x, y\rangle = x\cdot y$ is the Euclidean inner product and $|x| := \sqrt{\langle x, x\rangle}$.

The stochastic differential equation (1.1) has a unique non-explosive solution denoted by X_t^ε (see [10]). Given $T > 0$, when ε goes to 0, $\sup_{t\in[0,T]} |X_t^\varepsilon - X_t^0| \to 0$ in probability, where X^0 is the solution of ordinary differential equation (the non-perturbed dynamical system)

$$\begin{cases} dX_t^0 = b(X_t^0)dt; \\ X_0^0 = x_0. \end{cases} \tag{1.2}$$

Let $C([0,T], \mathbb{R}^d)$ be the Banach space of continuous functions $\gamma : [0,T] \to \mathbb{R}^d$ equipped with the sup-norm $\|\gamma\| := \sup_{t\in[0,T]} |\gamma(t)|$. For a precise asymptotic estimate of deviations of X^ε from the dynamical system X^0, Freidlin-Wentzell's theorem (see [6, 7, 9]) tells us that $\{X_t^\varepsilon\}$ satisfies the large deviation principle (LDP in short) on $C([0,T],\mathbb{R}^d)$ with the good rate function given by

$$I_{LD}(\gamma) = \frac{1}{2}\int_0^T \langle (\sigma\sigma^*(\gamma_t))^{-1}(\dot\gamma_t - b(\gamma_t)), \dot\gamma_t - b(\gamma_t)\rangle dt, \tag{1.3}$$

if γ is absolutely continuous with $\gamma(0) = x_0$, and $I(\gamma) = +\infty$ otherwise. In the expression above, $\langle A^{-1}x, x\rangle$ for a symmetric non-negative definite (not necessarily positive definite) matrix $A \in \mathcal{M}(d\times d)$ is defined by

$$\langle A^{-1}x, x\rangle := \sum_{k=1}^d \frac{1}{\lambda_k}\langle x, e_k\rangle^2 \tag{1.4}$$

with the convention $c/0 = +\infty 1_{c>0}$ for $c \geq 0$, where (e_k) is the orthonormal base of eigenvectors associated to the eigenvalues (λ_k) of A. Freidlin-Wentzell's theorem is a significant generalization of Schilder's theorem for Brownian motion ([13]), and it was usually proved under the global Lipschitz condition on σ, b. It still holds under the weaker condition **(H)** above by Lemma 3.1 below (this is more or less known).

Extensions of the Freidlin-Wentzell theory attract many recent studies. We mention here only

1) Boué-Dupuis [2] provided a new variational approach for the large deviations of Brownian functionals. Their beautiful approach turns out to be simple and efficient, even in the infinite dimension case ([11]).

2) The vector field b has some jumps. Chiang and Sheu [5] obtained its LDP with a completely different rate function related to the large deviations of local times.

3) Extensions to infinite dimension diffusions such as stochastic partial differential equations (SPDE's in short). See [3, 4, 11, 12] and references therein.

Quite surprisingly the problem of moderate deviation principle (MDP in short) for $\{X_t^\varepsilon\}_{t\in[0,T]}$ was left open (up to our knowledge). That is the subject of this paper. More precisely, we shall study the asymptotic behavior of

$$\eta_t^\varepsilon := \frac{Y_t^\varepsilon}{h(\varepsilon)}, \quad \text{with} \quad Y_t^\varepsilon := \frac{X_t^\varepsilon - X_t^0}{\sqrt{\varepsilon}}, \tag{1.5}$$

where

$$h(\varepsilon)\to +\infty \text{ and } \sqrt{\varepsilon}h(\varepsilon) \to 0, \quad \text{as } \varepsilon \to 0. \tag{1.6}$$

Through this paper we always assume that $h(\varepsilon)$ satisfies (1.6).

Theorem 1.1. *Assume the condition* **(H)**. *Then as* $\varepsilon \to 0$,

(1) **(CLT)** $Y^\varepsilon = (Y_t^\varepsilon)_{t\in[0,T]}$ *converges in probability on* $C([0,T],\mathbb{R}^d)$ *to the Gaussian Ornstein-Uhlenbeck process* Y^0, *determined by*

$$\begin{cases} dY_t^0 = Db(X_t^0)Y_t^0 dt + \sigma(X_t^0)dB_t; \\ Y_0^0 = 0 \end{cases} \tag{1.7}$$

where $Db = (\frac{\partial}{\partial x_j} b^i)_{1\leq i,j\leq d}$ *is the Jacobian matrix of* b.

(2) **(MDP)** $\eta^\varepsilon = \left(\frac{X_t^\varepsilon - X_t^0}{\sqrt{\varepsilon}h(\varepsilon)}\right)_{t\in[0,T]}$ *obeys the LDP on the space* $C([0,T],\mathbb{R}^d)$ *with speed* $h^2(\varepsilon)$ *and with rate function*

$$I(\gamma) = \frac{1}{2}\int_0^T \langle(\sigma\sigma^*(X_t^0))^{-1}(\dot\gamma_t - Db(X_t^0)\gamma_t), \dot\gamma_t - Db(X_t^0)\gamma_t\rangle dt, \tag{1.8}$$

if γ *is absolutely continuous with* $\gamma_0 = 0$, *and* $I(\gamma) := +\infty$ *otherwise, where the definition (1.4) is used in the expression above. More precisely, for any Borel measurable subset* A *of* $C([0,T],\mathbb{R}^d)$,

$$-\inf_{\gamma\in A^o} I(\gamma) \le \liminf_{\varepsilon\to 0} h^{-2}(\varepsilon)\log\mathbb{P}\left(\eta^\varepsilon \in A\right)$$
$$\le \limsup_{\varepsilon\to 0} h^{-2}(\varepsilon)\log\mathbb{P}\left(\eta^\varepsilon \in A\right) \le -\inf_{\gamma\in\bar{A}} I(\gamma),$$

where A^o *and* $\bar{A}$ *denote the interior and the closure of* A, *respectively.*

The paper is organized as follows. An outline of the proof of Theorem 1.1 is presented in the next section. In section 3 we first prove that under the assumption (**H**), X^ε is bounded in the sense of Freidlin-Wentzell's LDP, so we reduce our study to the case where σ and b are globally Lipschitzian. The details of the proof are given in Section 4, with the help of Talagrand's transportation inequality on path space established in [8] and measure concentration [1, 8, 14].

2. An outline for the proof of Theorem 1.1

Obviously

$$Y_t^\varepsilon = \frac{X_t^\varepsilon - X_t^0}{\sqrt{\varepsilon}}, \quad t \ge 0,$$

satisfies

$$dY_t^\varepsilon = \frac{1}{\sqrt{\varepsilon}}(b(X_t^\varepsilon) - b(X_t^0))dt + \sigma(X_t^\varepsilon)dB_t.$$

As X^ε is close to X^0, we have intuitively

$$dY_t^\varepsilon \approx Db(X_t^0)Y_t^\varepsilon dt + \sigma(X_t^0)dB_t,$$

in other words Y^ε should be close to Y^0 determined by (1.7). As

$$d\left(\frac{Y_t^0}{h(\varepsilon)}\right) = Db(X_t^0)\left(\frac{Y_t^0}{h(\varepsilon)}\right)dt + \frac{\sigma(X_t^0)}{h(\varepsilon)}dB_t,$$

then by Schilder's theorem (together with the contraction principle) or Freidlin-Wentzell's theorem, $Y^0/h(\varepsilon)$ obeys the LDP with the speed $h^2(\varepsilon)$ and the rate function $I(\gamma)$ given by (1.8). Hence by [6, Theorem 4.2.13], for the MDP in Theorem 1.1, it is enough to show that $\eta^\varepsilon = Y^\varepsilon/h(\varepsilon)$ is $h^2(\varepsilon)$-exponentially equivalent to $Y^0/h(\varepsilon)$, i.e.,

$$\limsup_{\varepsilon\to 0} h^{-2}(\varepsilon)\log\mathbb{P}\left(\|\eta^\varepsilon - \frac{Y^0}{h(\varepsilon)}\| > \delta\right) = -\infty, \ \forall\delta > 0. \tag{2.1}$$

This turns out to be quite difficult and we shall prove it under the global Lipschitz condition, by means of Talagrand's T_2-transportation inequality on path space established by Djellout-Guillin-Wu [8] and the corresponding concentration inequality (Bobkov-Götze's criterion [1]), in the last section.

3. Reduction to the global Lipschitzian case

At first, we shall prove that under the assumption (**H**), X^ε is bounded in the sense of LDP. For any $R \geq 0$, let

$$\tau_R^\varepsilon := \inf\{t; |X_t^\varepsilon| \geq R\}.$$

Lemma 3.1. *Under the assumption* (**H**),

$$\lim_{R\to+\infty} \limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\left(\tau_R^\varepsilon \leq T\right) = -\infty.$$

Proof. Let $f(x) = \log(1+|x|^2)$. By Itô's formula,

$$\begin{aligned} df(X_t^\varepsilon) &= \langle \sqrt{\varepsilon}\nabla f(X_t^\varepsilon), \sigma(X_t^\varepsilon)dB_t\rangle + \langle b(X_t^\varepsilon), \nabla f(X_t^\varepsilon)\rangle dt + \frac{\varepsilon}{2}\sum_{i,j}(\sigma\sigma^*)_{ij}(X_t^\varepsilon)\frac{\partial^2}{\partial x_i \partial x_j} f(X_t^\varepsilon)dt \\ &= \langle \sqrt{\varepsilon}\nabla f(X_t^\varepsilon), \sigma(X_t^\varepsilon)dB_t\rangle + \mathcal{L}^\varepsilon f(X_t^\varepsilon)dt, \end{aligned}$$

where $\mathcal{L}^\varepsilon$ is the generator of X^ε.

Consider the local martingale

$$M_t^\varepsilon = \sqrt{\varepsilon}\int_0^t \langle \nabla f(X_s^\varepsilon), \sigma(X_s^\varepsilon)dB_s\rangle.$$

By the linear growth of σ in **(H)**, its quadratic variation process $\langle M^\varepsilon\rangle$ satisfies

$$\langle M^\varepsilon\rangle_t = \varepsilon\int_0^t |\sigma(X_s^\varepsilon)^*\nabla f(X_s^\varepsilon)|^2 ds \leq 4\varepsilon C t.$$

Notice that for all $\varepsilon \in (0,1]$,

$$\mathcal{L}^\varepsilon f(x) = \frac{\varepsilon \operatorname{tr}(\sigma\sigma^*(x)) + 2\langle b(x), x\rangle}{1+|x|^2} - \frac{2|x|^2}{(1+|x|^2)^2} \leq 3C,$$

where C is the constant in **(H)**. Consequently for all $t \geq 0$ and $\varepsilon \in (0,1]$,

$$f(X_t^\varepsilon) \leq f(x_0) + M_t^\varepsilon + 3Ct.$$

For any $R > 0$ large enough so that $c(R,T) := \log(1+R^2) - [\log(1+|x_0|^2) + 3CT] > 0$, we have by the Bernstein inequality for continuous local martingale,

$$\begin{aligned} \mathbb{P}(\tau_R^\varepsilon \leq T) &= \mathbb{P}\left(\sup_{t\in[0,T]} |X_t^\varepsilon| \geq R\right) = \mathbb{P}\left(\sup_{t\in[0,T]} f(X_t^\varepsilon) \geq \log(1+R^2)\right) \\ &\leq \mathbb{P}\left(\sup_{t\in[0,T]} |M_t^\varepsilon| \geq c(R,T)\right) \leq 2\exp\left\{-\frac{c(R,T)^2}{8\varepsilon CT}\right\}, \end{aligned}$$

where the desired result follows. □

Now for any $R > 0$ large enough so that $c(R,T) > 0$, let $\sigma^{(R)}(x) = \sigma(x)$ and $b^{(R)}(x) = b(x)$ for $x \in \mathbb{R}^d$ with $|x| \leq R$, such that $\sigma^{(R)}$ is globally Lipschitzian and bounded, and $b^{(R)}$ is C^1 with $Db^{(R)}$ uniformly continuous and bounded. Consider the solution $X_t^{\varepsilon,R}$ of the corresponding stochastic differential equation (1.1) with (σ, b) replaced by $(\sigma^{(R)}, b^{(R)})$. We have by Lemma 3.1 and the proof above,

$$\begin{aligned}&\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\left(X_t^\varepsilon \neq X_t^{\varepsilon,R} \text{ for some } t \in [0,T]\right) \\ &\leq \limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\left(\tau_R^\varepsilon \leq T\right) \leq -\frac{c(R,T)^2}{8CT}.\end{aligned}$$

Hence by the approximation lemma ([6]), Freidlin-Wentzell's theorem still holds under the weaker condition **(H)** (as claimed in the Introduction); and $(X^\varepsilon - X^0)/h(\varepsilon)$ and $(X^{\varepsilon,R} - X^0)/h(\varepsilon)$ obey the same LDP by [6, Theorem 4.2.13].

In other words considering $(\sigma^{(R)}, b^{(R)})$ if necessary, we may and will suppose that **(L)** *b is C^1 with Db uniformly continuous and bounded on $\mathbb{R}^d$, and*

$$\frac{1}{2} tr[(\sigma(x) - \sigma(y))(\sigma(x) - \sigma(y))^*] + \langle x - y, b(x) - b(y)\rangle \leq L|x - y|^2, \tag{3.1}$$

for some constant $L \in \mathbb{R}$ and for all $x, y \in \mathbb{R}^d$, and $tr(\sigma\sigma^*) \leq M$ for some positive constant M.

When $L < 0$, condition (3.1) means that the diffusion is dissipative and it is equivalent to Bakry-Emery's Γ_2 condition if $\sigma = \sqrt{2} I_d$ (I_d being the identity matrix in $\mathcal{M}(d \times d)$).

4. Talagrand's inequality and Proof of Theorem 1.1

4.1. An L^2-estimate for Y^ε.

Lemma 4.1. *Under the assumption* (**L**), $Y^\varepsilon = (X^\varepsilon - X^0)/\sqrt{\varepsilon}$ *satisfies*

$$\mathbb{E}|Y_t^\varepsilon|^2 \leq M(e^{2Lt} - 1)/2L, \quad t \geq 0. \tag{4.1}$$

Proof. Under the assumption (**L**), by Itô's formula,

$$\begin{aligned} d|X_t^\varepsilon - X_t^0|^2 &= 2\langle X_t^\varepsilon - X_t^0, b(X_t^\varepsilon) - b(X_t^0)\rangle dt \\ &\quad + 2\langle X_t^\varepsilon - X_t^0, \sqrt{\varepsilon}\sigma(X_t^\varepsilon) dB_t\rangle + \varepsilon\, \mathrm{tr}(\sigma\sigma^*)(X_t^\varepsilon) dt \\ &\leq 2L|X_t^\varepsilon - X_t^0|^2 dt + \varepsilon M dt + 2\langle X_t^\varepsilon - X_t^0, \sqrt{\varepsilon}\sigma(X_t^\varepsilon) dB_t\rangle. \end{aligned}$$

Hence

$$\mathbb{E}|X_t^\varepsilon - X_t^0|^2 \leq 2L \int_0^t \mathbb{E}|X_s^\varepsilon - X_s^0|^2 ds + \varepsilon M t.$$

By Gronwall's inequality, we have

$$\mathbb{E}|X_t^\varepsilon - X_t^0|^2 \leq \int_0^t e^{2L(t-s)} \varepsilon M ds = \varepsilon M(e^{2Lt} - 1)/2L,$$

which implies the desired result by the very definition of Y^ε. □

4.2. **Talagrand's T_2-transportation inequality.** First we shall introduce some useful notions and notations (see Villani [14] for an extensive study of such quantities). Given a metric space (E,d) equipped with its Borel σ field, and $1 \le p < +\infty$, the L^p-Wasserstein distance between two probability measures μ and ν on E is defined as

$$W_p(\mu,\nu) := \inf \left(\iint d(x,y)^p d\pi(x,y) \right)^{1/p},$$

where the infimum runs over all couplings π of (μ,ν).

A probability measure μ is said to satisfy Talagrand's T_2-transportation inequality $T_2(C_T)$ on (E,d), where $C_T > 0$ is some constant (here the index T is referred to Talagrand), if for all probability measure ν

$$W_2(\nu,\mu)^2 \le 2C_T H(\nu|\mu),$$

where $H(\nu|\mu)$ is the Kullback-Leibler information or relative entropy of ν with respect to μ:

$$H(\nu|\mu) = \begin{cases} \int \log(\frac{d\nu}{d\mu})d\nu, & \text{if } \nu \ll \mu; \\ +\infty, & \text{otherwise.} \end{cases}$$

Now let $\mu = P^\varepsilon$, the law of $X^\varepsilon = (X_t^\varepsilon)_{t\in[0,T]}$, which is also a probability measure on the Hilbert space

$$E = L^2([0,T];\mathbb{R}^d) = \left\{ \varphi : [0,T] \to \mathbb{R}^d \text{ measurable } ; \|\varphi\|_2^2 = \int_0^T |\varphi(t)|^2 dt < +\infty \right\}$$

(up to dt-equivalence), equipped with the metric $d_2(\varphi_1,\varphi_2) := \|\varphi_1 - \varphi_2\|_2$. By Djellout-Guillin-Wu [8, (5.5) and Remark 5.9], we have

Lemma 4.2. *Assume that σ, b are locally Lipschitzian, σ is bounded and (3.1) holds. Then for any $\varepsilon \in (0,1]$, the law P^ε of $X^\varepsilon = (X_t^\varepsilon)_{t\in[0,T]}$ satisfies on $(L^2([0,T];\mathbb{R}^d), d_2)$ Talagrand's T_2-inequality $T_2(\varepsilon C_T)$, where*

$$C_T = \frac{\|\sigma\|_\infty^2 (e^{(\delta+2L)T} - 1)}{\delta(\delta+2L)} \tag{4.2}$$

with $\delta > 0$ arbitrary, $\frac{e^{(\delta+2L)T}-1)}{\delta+2L} := T$ if $\delta + 2L = 0$ (this type convention will be used later too) and $\|\sigma\|_\infty = \sup_{x\in\mathbb{R}^d, z\in\mathbb{R}^n, |z|=1} |\sigma(x)z|$.

Note that $\|\sigma\|_\infty \le \sqrt{\sup_{x\in\mathbb{R}^d} \mathrm{tr}(\sigma\sigma^*(x))} \le \sqrt{M}$.

Next we use this key lemma to get the following crucial measure concentration inequality.

Lemma 4.3. *Assume that σ, b are locally Lipschitzian, σ is bounded and (3.1) holds. Then for any $\varepsilon \in (0,1]$ and $r > 0$*

$$\mathbb{P}\left(\left(\int_0^T |Y_t^\varepsilon|^2 dt \right)^{\frac{1}{2}} - \mathbb{E}\left(\int_0^T |Y_t^\varepsilon|^2 dt \right)^{\frac{1}{2}} \ge r \right) \le \exp\{-\frac{r^2}{2C_T}\}, \tag{4.3}$$

where the constant C_T is given in (4.2).

Proof. Since $\|Y^\varepsilon\|_2 = \Phi(X^\varepsilon)$ where $\Phi(\varphi) = \|(\varphi - X^0)/\sqrt{\varepsilon}\|_2$ is Lipschitzian on $(L^2([0,T];\mathbb{R}^d), d_2)$ with the Lipschitzian coefficient $1/\sqrt{\varepsilon}$, the concentration inequality (4.3) follows from Lemma 4.2, by Bobkov-Götze's criterion [1, Theorem 3.1]. □

Remark 4.1. The concentration inequality (4.3) is of important independent interest. If the diffusion is dissipative, i.e. $L < 0$ in condition (3.1), by setting $\delta = -L$ in (4.2), we see that Talagrand's constant

$$C_T \le \frac{\|\sigma\|_\infty^2}{L^2}$$

which is independent of T. Notice also for the bias of X^ε from X^0 in L^2, we have $\mathbb{E}\int_0^T |X_t^\varepsilon - X_t^0|^2 dt = \varepsilon \int_0^T |Y_t^\varepsilon|^2 dt \le \varepsilon MT/(2|L|)$ (by (4.1)). For other concentration inequalities which can be derived from Talagrand's T_2-inequalities, see [8, 14].

4.3. **Proof of the MDP in Theorem 1.1.** As explained in §3, we may and will assume the condition **(L)**. For the MDP in part (2), by what is said in the outline of proof in §2, it is sufficient to show that $Y^\varepsilon/h(\varepsilon)$ and $Y^0/h(\varepsilon)$ are $h^2(\varepsilon)$-exponentially equivalent, i.e. (2.1).

We start by observing

$$\begin{aligned} d(Y_t^\varepsilon - Y_t^0) &= Db(X_t^0)(Y_t^\varepsilon - Y_t^0)dt + \big(\sigma(X_t^\varepsilon) - \sigma(X_t^0)\big)\, dB_t \\ &\quad + \frac{1}{\sqrt{\varepsilon}}\big(b(X_t^\varepsilon) - b(X_t^0) - Db(X_t^0)(X_t^\varepsilon - X_t^0)\big)\, dt \\ &=: Db(X_t^0)(Y_t^\varepsilon - Y_t^0)dt + dZ_t^\varepsilon \end{aligned} \tag{4.4}$$

where

$$Z_t^\varepsilon = \int_0^t \big(\sigma(X_s^\varepsilon) - \sigma(X_s^0)\big)\, dB_s + \frac{1}{\sqrt{\varepsilon}}\int_0^t \big(b(X_s^\varepsilon) - b(X_s^0) - Db(X_s^0)(X_s^\varepsilon - X_s^0)\big)\, ds.$$

The solution of (4.4) is given by

$$Y_t^\varepsilon - Y_t^0 = Z_t^\varepsilon + \int_0^t Db(X_s^0)J(s,t)Z_s^\varepsilon ds,$$

where $J(s,t)$ satisfies the matrix differential equation

$$J(s,s) = I_d, \quad \frac{d}{dt}J(s,t) = Db(X_t^0)J(s,t), \quad 0 \le s \le t.$$

Since $\langle Db(x)y, y\rangle \le L|y|^2$ $(y \in \mathbb{R}^d)$ by condition (3.1), we have

$$|J(s,t)y| \le e^{L(t-s)}|y|, \ \forall y \in \mathbb{R}^d.$$

Setting $\|Db\|_\infty := \sup_{(x,z)\in(\mathbb{R}^d)^2, |z|=1} |Db(x)z|$, we get for all $t \in [0,T]$,

$$|Y_t^\varepsilon - Y_t^0| \le |Z_t^\varepsilon| + \int_0^t \|Db\|_\infty e^{L(t-s)}|Z_s^\varepsilon| ds \le \left(1 + \|Db\|_\infty \frac{(e^{LT}-1)}{L}\right)\|Z^\varepsilon\|. \tag{4.5}$$

From (4.5) we see that the desired MDP-equivalence (2.1) between $Y^\varepsilon/h(\varepsilon)$ and $Y^0/h(\varepsilon)$ follows from

Proposition 4.4. *For any $r > 0$,*

$$\limsup_{\varepsilon\to 0} h^{-2}(\varepsilon)\log \mathbb{P}\left(\frac{\|Z^\varepsilon\|}{h(\varepsilon)} > r\right) = -\infty.$$

Proof. Notice that

$$\begin{aligned}|Z_t^\varepsilon| &\le |\int_0^t \left(\sigma(X_s^\varepsilon)-\sigma(X_s^0)\right)dB_s| + \int_0^t |\frac{1}{\sqrt{\varepsilon}}\left(b(X_s^\varepsilon)-b(X_s^0)-Db(X_s^0)(X_s^\varepsilon-X_s^0)\right)|ds \\ &=: |M_t^\varepsilon| + \int_0^t \frac{1}{\sqrt{\varepsilon}}|\left(b(X_s^\varepsilon)-b(X_s^0)-Db(X_s^0)(X_s^\varepsilon-X_s^0)\right)|ds.\end{aligned} \tag{4.6}$$

Next we estimate the last two terms.

(a) For the continuous martingale M_t^ε, let $\langle M^\varepsilon\rangle_t$ be its quadratic variation process. For any $\eta>0$,

$$\begin{aligned}\mathbb{P}\left(\sup_{0\le t\le T}|M_t^\varepsilon|\ge\frac{rh(\varepsilon)}{2}\right) &\le \mathbb{P}\left(\sup_{0\le t\le T}|M_t^\varepsilon|\ge\frac{rh(\varepsilon)}{2},\ \langle M^\varepsilon\rangle_T\le\eta\right)+\mathbb{P}\Big(\langle M^\varepsilon\rangle_T\ge\eta\Big) \\ &\le 2\exp\left\{-\frac{r^2h^2(\varepsilon)}{8\eta}\right\}+\mathbb{P}\Big(\langle M^\varepsilon\rangle_T\ge\eta\Big).\end{aligned} \tag{4.7}$$

Since

$$\begin{aligned}\langle M^\varepsilon\rangle_T &= \int_0^T \mathrm{tr}(\sigma(X_t^\varepsilon)-\sigma(X_t^0))(\sigma(X_t^\varepsilon)-\sigma(X_t^0))^* dt \\ &\le L_1\int_0^T |X_t^\varepsilon-X_t^0|^2dt = \varepsilon L_1\int_0^T |Y_t^\varepsilon|^2dt\end{aligned} \tag{4.8}$$

for some constant L_1. When ε is small enough, by Lemma 4.1

$$\mathbb{E}\left(\int_0^T |Y_t^\varepsilon|^2dt\right)^{\frac12}\le\frac12\left(\frac{\eta}{\varepsilon L_1}\right)^{\frac12}.$$

Then it follows from Lemma 4.3

$$\begin{aligned}\mathbb{P}\left(\langle M^\varepsilon\rangle_T\ge\eta\right) &\le \mathbb{P}\left(\int_0^T|Y_t^\varepsilon|^2dt\ge\frac{\eta}{\varepsilon L_1}\right) \\ &\le \mathbb{P}\left(\left(\int_0^T|Y_t^\varepsilon|^2dt\right)^{\frac12}-\mathbb{E}\left(\int_0^T|Y_t^\varepsilon|^2dt\right)^{\frac12}\ge\frac12\left(\frac{\eta}{\varepsilon L_1}\right)^{\frac12}\right) \\ &\le \exp\left\{-\frac{\eta}{8\varepsilon C_T L_1}\right\}.\end{aligned}$$

Noting that $\varepsilon h^2(\varepsilon)\to0$ as $\varepsilon\to0$ and $\eta>0$ is arbitrary, we obtain by (4.7)

$$\limsup_{\varepsilon\to0} h^{-2}(\varepsilon)\log\mathbb{P}\left(\sup_{0\le t\le1}|M_t^\varepsilon|\ge\frac{h(\varepsilon)r}{2}\right)=-\infty. \tag{4.9}$$

(b) Because b is C^1 and Db is uniformly continuous, for any $\eta>0$, there exists some constant $\delta>0$, such that

$$|b(x)-b(y)-Db(y)(x-y)|\le\eta|x-y|,\quad \text{if } |x-y|\le\delta.$$

When $|X_t^\varepsilon-X_t^0|\le\delta$,

$$\frac{1}{\sqrt{\varepsilon}}|b(X_t^\varepsilon)-b(X_t^0)-Db(X_t^0)(X_t^\varepsilon-X_t^0)|\le\frac{\eta}{\sqrt{\varepsilon}}|X_t^\varepsilon-X_t^0|=\eta|Y_t^\varepsilon|.$$

Thus

$$\begin{aligned}&\mathbb{P}\left(\int_0^T \frac{1}{\sqrt{\varepsilon}}\left|\left(b(X_t^\varepsilon)-b(X_t^0)-Db(X_t^0)(X_t^\varepsilon-X_t^0)\right)\right|dt\geq \frac{h(\varepsilon)r}{2}\right)\\&\leq \mathbb{P}\left(\|X^\varepsilon-X^0\|\geq\delta\right)+\mathbb{P}\left(\int_0^T|Y_t^\varepsilon|dt\geq\frac{h(\varepsilon)r}{2\eta}\right).\end{aligned}\tag{4.10}$$

By Freidlin-Wentzell's theorem,

$$\limsup_{\varepsilon\to 0}\varepsilon\log\mathbb{P}\left(\|X^\varepsilon-X^0\|\geq\delta\right)<0,$$

so

$$\limsup_{\varepsilon\to 0}h^{-2}(\varepsilon)\log\mathbb{P}\left(\|X^\varepsilon-X^0\|\geq\delta\right)=-\infty.\tag{4.11}$$

When ε is small enough, $\mathbb{E}\left(\int_0^T|Y_t^\varepsilon|^2dt\right)^{\frac{1}{2}}\leq\frac{h(\varepsilon)r}{4\eta\sqrt{T}}$ by Lemma 4.1. Therefore by Cauchy-Schwarz inequality and Lemma 4.3, we have

$$\begin{aligned}&\mathbb{P}\left(\int_0^T|Y_t^\varepsilon|dt\geq\frac{h(\varepsilon)r}{2\eta}\right)\leq\mathbb{P}\left(\left(\int_0^T|Y_t^\varepsilon|^2dt\right)^{\frac{1}{2}}\geq\frac{h(\varepsilon)r}{2\eta\sqrt{T}}\right)\\&\leq\mathbb{P}\left(\left(\int_0^T|Y_t^\varepsilon|^2dt\right)^{\frac{1}{2}}-\mathbb{E}\left(\int_0^T|Y_t^\varepsilon|^2dt\right)^{\frac{1}{2}}\geq\frac{h(\varepsilon)r}{4\eta\sqrt{T}}\right)\\&\leq\exp\left\{-\frac{h^2(\varepsilon)r^2}{32\eta^2TC_T}\right\}.\end{aligned}\tag{4.12}$$

Due to the arbitrariness of η, plugging (4.11) and (4.12) into (4.10) we obtain

$$\begin{aligned}\limsup_{\varepsilon\to 0}h^{-2}(\varepsilon)\log\mathbb{P}&\left(\int_0^T\frac{1}{\sqrt{\varepsilon}}\left|\left(b(X_t^\varepsilon)-b(X_t^0)-Db(X_t^0)(X_t^\varepsilon-X_t^0)\right)\right|dt\geq\frac{h(\varepsilon)r}{2}\right)\\&=-\infty.\end{aligned}\tag{4.13}$$

The desired result follows from (4.6), (4.9) and (4.13). □

4.4. **Proof of the CLT in Theorem 1.1.** This is much simpler. We may assume **(L)** by Lemma 3.1. For the CLT in part (1) of Theorem 1.1, by (4.5) it suffices to show

$$\lim_{\varepsilon\to 0}\mathbb{P}(\|Z^\varepsilon\|>r)=0,\ \forall r>0.\tag{4.14}$$

Using (4.6) and observing

$$\mathbb{E}\|M^\varepsilon\|^2\leq 4\mathbb{E}\langle M^\varepsilon\rangle_T\leq\varepsilon L_1\mathbb{E}\int_0^T|Y_t^\varepsilon|^2dt\to 0,$$

by Doob's maximal inequality, (4.8) and Lemma 4.1, we are led to show

$$\mathbb{P}\left(\int_0^T\frac{1}{\sqrt{\varepsilon}}\left|\left(b(X_t^\varepsilon)-b(X_t^0)-Db(X_t^0)(X_t^\varepsilon-X_t^0)\right)\right|dt>r\right)\to 0,\ \text{ as }\varepsilon\to 0.$$

This is obvious from (4.10) by taking there $h(\varepsilon)=2$ and the boundedness of $\mathbb{E}\int_0^T|Y_t^\varepsilon|^2dt,\varepsilon\in(0,1]$ in Lemma 4.1. □

Remark 4.2. It would be important and very interesting to generalize the MDP of this paper to infinite dimensional diffusions such as SPDE's.

References

[1] S. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163: 1-28, 1999.

[2] M. Boué and P. Dupuis. A variational representation for certain functionals of Brownian motion. *Ann. Probab.* 26(4): 1641-1659, 1998.

[3] C. Cardon-Weber. Large deviations for a Burger's-type SPDE, *Stochastic Process. Appl.,* 84: 53-70, 1999.

[4] S. Cerrai and M. Röckner. Large deviations for stochastic reaction-diffusion systems with multiplicative noise and non-Lipschitz reaction term, *Ann. Probab.,* 32(1B): 1100-1139, 2004.

[5] T.S. Chiang and S. J. Sheu. Large deviation of diffusion processes with discontinuous drift and their occupation times. *Ann. Probab.,* 28: 140-165, 2000.

[6] A. Dembo and O. Zeitouni. Large deviations techniques and applications. Second Edition. Applications of Mathematics 38. Springer-Verlag, 1998.

[7] J.D. Deuschel and D.W. Stroock. Large deviations. Pure and Applied Mathematicas 137. Academic Press, 1989.

[8] H. Djellout, A. Guillin and L. Wu. Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *Ann. Probab.,* 32: 2702-2732, 2004.

[9] M.I. Freidlin and A.D. Wentzell. Random perturbation of Dynamical systems. Translated by Szuc, J. Springer. Berlin, 1984.

[10] I. Ikeda and S. Watanabe. Stochastic differential equations and Diffusion processes. North-Holland, Amsterdam, 1981.

[11] W. Liu. Large deviations for stochastic evolution equations with small multiplicative noise. *Appl. Math. Optim.,* 61(1): 27-56, 2010.

[12] M. Röckner, F.Y. Wang and L. Wu. Large deviations for stochastic generalized porous media equations. *Stochastic Process. Appl.,* 116(12): 1677-1689, 2006.

[13] M. Schilder. Some asymptotic formulas for Wiener integrals. Trans. Amer. Math. Soc. 125: 63-85, 1966.

[14] C. Villani. Optimal transport. Old and new. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] 338. Springer-Verlag, Berlin, 2009.

Yutao MA, School of Mathematical Sciences & Lab. Math. Com. Sys., Beijing Normal University, 100875, Beijing China. Partially supported by the Fundamental Research Funds for the Central Universities.

E-mail address: mayt@bnu.edu.cn

Ran Wang, School of Mathematics and Statistics, Wuhan University, 430072 Hubei, China. Partially supported by NSFC 11001208 and the Fundamental Research Funds for the Central Universities.

E-mail address: whwangran@yahoo.com.cn

Liming Wu, Institute of Applied Math., Chinese Academy of Sciences, 100190 Beijing, China and Laboratoire de Math. CNRS-UMR 6620, Université Blaise Pascal, 63177 Aubière, France. Partially supported by Thousand Talents Program of the Chinese Academy and le projet ANR EVOL.

E-mail address: Li-Ming.Wu@math.univ-bpclermont.fr