ESAIM: PS 27 (2023) 558–575
https://doi.org/10.1051/ps/2023009

ESAIM: Probability and Statistics
www.esaim-ps.org

A MODERATE DEVIATION PRINCIPLE FOR STOCHASTIC HAMILTONIAN SYSTEMS

Jie Xu[1,*], Jiayin Gong[1] and Jie Ren[2]

Abstract. We prove a moderate deviation principle for stochastic differential equations (SDEs) with non-Lipschitz conditions. As an application of our result, we also study the stochastic Hamiltonian systems.

Mathematics Subject Classification. 60H10, 60F10, 65C20, 34F05.

Received July 12, 2022. Accepted April 2, 2023.

1. Introduction

Large deviation principle (LDP) for SDEs with small noise has been extensively studied since the original work of Freidlin and Wentzell [10]. The Freidlin-Wentzell type LDP describes the asymptotic behavior of the trajectory of stochastic dynamical system with small noise. A general approach for studying LDP is based on the discretization method and contraction principle (see, *e.g.* [7, 11, 20]). Later, Budhiraja, Dupuis and Maroulas in [2, 3] proposed the weak convergence method to establish the LDP that is based on a variational representation. This method amounts to establishing the weak convergence of perturbations of SDEs with small noise in the random directions of the Cameron-Martin space of the driving Bronwnian motion. It is completely different from the proof of Freidlin and Wentzell, it can avoid some complicated discretization approximations and exponential probability estimates, and only needs to verify moment estimations and moment convergences. Recently, the method has been widely applied in many papers (see, *e.g.* [1, 4, 15, 17, 24, 28]).

Like LDP, the moderate deviation principle (MDP) has rightly received considerable attention. MDP gives estimates on probability of deviations of a smaller order than LDP and with a rate function that is a quadratic form. There are many results on the MDP in various frameworks, for example, [22] for 2D stochastic Navier-Stokes equations, [23] for stochastic reaction- diffusion equations with multiplicative noise, [21] for stochastic differential delay equations with polynomial growth and [14] for dynamical systems with small random perturbation. Moreover, the MDP has been extensively investigated for stochastic dynamical system with irregular coefficients (cf. [6, 18, 19, 30, 31], etc.).

Specially, the Donsker-Varadhan type long time LDP and MDP for stochastic damping Hamiltonian systems were established by Wu [25]. The Freidlin–Wentzell type LDP for stochastic Hamiltonian systems was studied by the third author of this paper in [16]. It is natural to ask whether the MDP continues to hold for stochastic

Keywords and phrases: Moderate deviation principle, Non-Lipschitz condition, Stochastic Hamiltonian systems.

[1] College of Mathematics and Information Science, Henan Normal University, Xinxiang, Henan 453007, PR China.
[2] College of Mathematics and Information Science, Henan University Of Economics and Law, Zhengzhou, Henan 450002, PR China.

[*] Corresponding author: xujiescu@163.com

Hamiltonian systems. In this paper, we will give a positive answer. Our proof is based on the weak convergence approach. The core idea is to characterize the tightness by Arzelà-Ascoli theorem. The main difficulty of this paper is to deal with the non-Lipschitz property of the coefficients of SDEs, which requires the help of Lyapnov function technique and exponential martingale technique and stochastic Gronwall's inequality.

Consider the following stochastic nonlinear oscillator equation:

$$\begin{cases} \mathrm{d}Z_t = U_t \mathrm{d}t, & Z_0 = z, \\ \mathrm{d}U_t = [C_0 U_t - \nabla V(Z_t)]\mathrm{d}t + \sqrt{\epsilon}\Theta(Z_t)\mathrm{d}W_t, & U_0 = u, \end{cases} \tag{1.1}$$

where $C_0 \in \mathbb{R}$, $V \in C^2(\mathbb{R})$ and $\nabla V(\cdot)$ satisfies (A1) in Section 4, $\Theta \in C^2(\mathbb{R})$ has a bounded first order derivative, and W_t is a one-dimensional Brownian white noise.

Intuitively, we can see that (1.1) is equivalent to the following SDEs:

$$\mathrm{d}X_t = f_t(X_t)\mathrm{d}t + \sqrt{\epsilon} g_t(X_t)\mathrm{d}W_t, \quad X_0 = x,$$

where $x = (z, u)$, $f_t(x) = f_t(z, u) := (u, C_0 u - \nabla V(u))$, $g_t(x) = g(z, u) := (0, \Theta(z))$.

Consider the following general SDEs with small perturbation:

$$\mathrm{d}X_t^\epsilon = b_t(X_t^\epsilon)\mathrm{d}t + \sqrt{\epsilon}\sigma_t(X_t^\epsilon)\mathrm{d}W_t, \quad X_0^\epsilon = x, \tag{1.2}$$

where $b : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^m$ are Borel measurable, and $\{W_t\}_{t \geqslant 0}$ is an m-dimensional standard Brownian motion on the classical Wiener space $(\Omega, \mathscr{F}, P)$. Moreover, the coefficients are non-global Lipschitz and super-linear growth. Under assumptions (H1)–(H3) (see Sect. 2), by [29], there is a unique strong solution for equation (1.2).

As $\epsilon \downarrow 0$, the solution X_t^ε of equation (1.2) will tend to the solution of the following deterministic equation

$$\mathrm{d}X_t^0 = b_t(X_t^0)\mathrm{d}t, \quad X_0^0 = x. \tag{1.3}$$

We shall investigate the deviations of X^ϵ from the deterministic solution X^0, as $\epsilon \downarrow 0$, that is, the asymptotic behavior of the trajectory,

$$Y_t^\epsilon = \frac{1}{\sqrt{\epsilon}h(\epsilon)}(X_t^\epsilon - X_t^0), \quad t \in [0, T],$$

in which $h(\epsilon)$ is some deviation scale. In particular
(1) For $h(\epsilon) = 1/\sqrt{\epsilon}$, it is corresponding LDP, which has been established by Ren in [16].
(2) For $h(\epsilon) = 1$, it is associated with the central limit theorem (CLT), which is our future work.
(3)When the deviation scale satisfies

$$h(\epsilon) \to +\infty, \quad \sqrt{\epsilon}h(\epsilon) \to 0, \quad as \quad \epsilon \to 0, \tag{1.4}$$

it is concerned with MDP, which is our main interest in this article. Throughout this paper, we always assume that $h(\epsilon)$ satisfies equation (1.4).

The organization of the paper as follows. In Section 2, we give some preliminaries and recall some useful results. Section 3 is devoted to the proof of the MDP. In Section 4, we apply our result to stochastic Hamiltonian systems.

Throughout the paper, C with or without indexes will denote different constants, whose values may change in different places. If there is no special declaration, all expectations $\mathbb{E}$ are taken with respect to P.

2. Preliminaries

In this section, we recall some basic notations, notions, assumptions and give some important theorems and lemmas that will be used later. Let us first introduce some notations and notions.

• For the simplicity of notation, we restrict our discussion to time interval $[0,1]$. We shall use $|\cdot|$ and $\langle\cdot,\cdot\rangle$ to denote the Euclidean norm and inner product. By σ^* and $\nabla^j\sigma$ we denote the transpose and the j-order gradient of matrix σ.

• We set $\mathscr{C}$ as the space of $\mathbb{R}^d$-valued continuous functions on $[0,1]$ with the norm:

$$\|f\|_{\mathscr{C}} := \sup_{s\in[0,1]} |f(s)|.$$

• Let $\mathbb{H}$ be the Cameron-Martin space of functions $\hbar$ which are absolutely continuous and whose derivative is square integrable, *i.e.*,

$$\mathbb{H} := \left\{ \hbar : [0,1] \to \mathbb{R}^m; \|\hbar\|_{\mathbb{H}}^2 = \int_0^1 |\dot{\hbar}(s)|^2 \mathrm{d}s < \infty \right\}.$$

Moreover, $\mathbb{H}$ is a Hilbert space with inner product

$$\langle \hbar_1, \hbar_2 \rangle := \int_0^1 \langle \dot{\hbar}_1(s), \dot{\hbar}_2(s)\rangle \mathrm{d}s.$$

• Let $\mathscr{A}$ denote the class of $\mathbb{R}^m$-valued $\{\mathscr{F}_t\}$-predictable processes $\hbar$ belonging to $\mathbb{H}$ a.s.

• For $N > 0$, let

$$B_N := \{\hbar \in \mathbb{H} : \|\hbar\|_{\mathbb{H}} \leqslant N\}.$$

then, by Kolmogorov and Fomin [13], it is easy to find that B_N is a compact Polish space under the weak topology, B_N is endowed with the weak topology in this paper.

• Define

$$\mathscr{A}_N := \{\hbar \in \mathscr{A} : \hbar \in B_N, a.s.\}.$$

Next, we recall the definition of large deviations from [9, 10]. Let $\mathscr{E}$ be a Polish space with the Borel σ-field $\mathscr{B}(\mathscr{E})$.

Definition 2.1. A family $\{Y^\epsilon\}_{\epsilon>0}$ of $\mathscr{E}$-valued random elements is said to satisfy the large deviation principle on $\mathscr{E}$ with rate function I and with the speed function $h^2(\epsilon)$ which is a sequence of positive numbers tending to $+\infty$ as $\epsilon \to 0$, if the following conditions hold:

(1) $I : \mathscr{E} \to [0, +\infty]$, for each $a < \infty$, the level set $\{f \in \mathscr{E} : I(f) \leqslant a\}$ is compact.
(2) for $A \in \mathscr{B}(E)$, we define $I(A) := \inf_{f\in A} I(f)$,

$$-I(A^o) \leqslant \liminf_{\epsilon\to 0} \frac{1}{h^2(\epsilon)} \log P(Y^\epsilon \in A) \leqslant \limsup_{\epsilon\to 0} \frac{1}{h^2(\epsilon)} \log P(Y^\epsilon \in A) \leqslant -I(\bar{A}),$$

where A^o and $\bar{A}$ are the interior and closure of the set A.

We use the method of weak convergence to prove the MDP. By [5], it proves the equivalence of the Laplace principle and LDP, so it is enough to prove the following theorem.

Theorem 2.2. *Suppose* $\{\Gamma^\epsilon\}_{\epsilon>0}$ *satisfies the following assumptions*: *there exists a measurable map* $\Gamma^0 : \mathbb{H} \to \mathscr{E}$ *such that*

(i) *Consider* $N < \infty$ *and a family* $\{\varphi^\epsilon\} \subset \mathscr{A}_N$ *such that* φ^ϵ *converges in distribution to* $\varphi \in \mathscr{A}_N$ *as* $\epsilon \to 0$. *Then* $\Gamma^\epsilon(W_\cdot + (\sqrt{\epsilon})^{-1} \int_0^\cdot \dot{\varphi}^\epsilon(s)\mathrm{d}s)$ *converges in distribution to* $\Gamma^0(\int_0^\cdot \dot{\varphi}(s)\mathrm{d}s)$.

(ii) *For every* $N < \infty$ *the set* $\{\Gamma^0(\int_0^\cdot \dot{\varphi}(s)\mathrm{d}s); \varphi \in B_N\}$ *is a compact subset of* $\mathscr{E}$.

Then the family $\{\Gamma^\epsilon\}_{\epsilon>0}$ *satisfies a large deviation principle in* $\mathscr{E}$ *with rate function* I *given by*

$$I(g) := \inf_{\{\varphi \in \mathbb{H}; g = \Gamma^0(\int_0^\cdot \dot{\varphi}(s)\mathrm{d}s)\}} \Big\{\frac{1}{2}\int_0^T |\dot{\varphi}(s)|^2 \mathrm{d}s\Big\}, \quad g \in \mathscr{E},$$

with the convention $\inf \emptyset = \infty$.

The following stochastic Gronwall's inequality can be found in [26], Lemma 3.8.

Lemma 2.3. *Let* $\xi(t)$ *and* $\eta(t)$ *be two nonnegative càdlàg* $\mathscr{F}_t$*-adapted processes,* A_t *a continuous nondecreasing* $\mathscr{F}_t$*-adapted process with* $A_0 = 0$, M_t *a local martingale with* $M_0 = 0$. *Suppose that*

$$\xi(t) \leqslant \eta(t) + \int_0^t \xi(s)\mathrm{d}A_s + M_t, \quad \forall t > 0.$$

Then for any $0 < q < \kappa < 1$ *and stopping time* τ , *we have*

$$\Big(\mathbb{E}(\xi(t)^*)^q\Big)^{\frac{1}{q}} \leqslant \Big(\frac{\kappa}{1-\kappa}\Big)^{1/q}\Big(\mathbb{E}e^{\kappa A_\tau/(1-\kappa)}\Big)^{(1-\kappa)/\kappa}\mathbb{E}(\eta(\tau)^*), \tag{2.1}$$

where $\xi(t)^* := \sup_{s\in[0,t]} \xi(s)$.

We call $\mathscr{V} \in C^2(\mathbb{R}^d)$ a Lyapunov function if

$$\mathscr{V}(x) \geqslant 1 \text{ and } \lim_{|x|\to\infty} \mathscr{V}(x) = +\infty.$$

We give the following assumptions throughout this paper:
(H1) b and σ are continuous in x and $b(x)$ is differentiable with respect to x. For a Lyapunov function $\mathscr{V} \in C^2(\mathbb{R}^d)$ and $\alpha \in [0,1]$, there exist constants $C_1, C_2, C_3, C_4 > 0$ such that for all $t \in [0,1]$, $x, y \in \mathbb{R}^d$,

$$|\sigma_t^*(x) \cdot \nabla\mathscr{V}(x)|^2 \leqslant C_1 \cdot \mathscr{V}^{2-\alpha}(x), \tag{2.2}$$

$$|\nabla b_t(x) - \nabla b_t(y)| \leqslant C_2 \cdot (\mathscr{V}^\alpha(x) + \mathscr{V}^\alpha(y)) \cdot |x-y|, \tag{2.3}$$

$$|\sigma_t(x) - \sigma_t(y)|^2 \leqslant C_3 \cdot (\mathscr{V}^\alpha(x) + \mathscr{V}^\alpha(y)) \cdot |x-y|^2, \tag{2.4}$$

$$\langle x-y, b_t(x) - b_t(y)\rangle \leqslant C_4 \cdot (\mathscr{V}^\alpha(x) + \mathscr{V}^\alpha(y)) \cdot |x-y|^2. \tag{2.5}$$

where ∇ is the gradient along the x direction, *i.e.*, $\nabla b = \big(\frac{\partial}{\partial x_j} b^i\big)_{1\leqslant i,j\leqslant d}$ is the Jacobain matrix of b, and

$$L_t := \frac{1}{2}\sum_{i,j=1}^d \sum_{k=1}^m (\sigma_t^{ik}\sigma_t^{jk})(x)\partial_i\partial_j + \sum_{i=1}^d b_t^i(x)\partial_i.$$

(H2) There exist constants $C_5, C_6 \in \mathbb{R}$ such that for all $t \in [0,1]$,

$$\sum_{i=1}^{d} b_t^i(x)\partial_i \mathscr{V}(x) \leqslant C_5 \cdot \mathscr{V}(x), \tag{2.6}$$

$$\sum_{i,j=1}^{d}\sum_{k=1}^{m}(\sigma_t^{ik}\sigma_t^{jk})(x)\partial_i\partial_j \mathscr{V}(x) \leqslant C_6 \cdot \mathscr{V}(x). \tag{2.7}$$

(H3) There exist constants $C_7, C_8 > 0$ such that for all $t \in [0,1]$,

$$|b_t(x)| + |\nabla b_t(x)| \leqslant C_7 \cdot (1 + \mathscr{V}^{\alpha}(x)|x|), \quad |\sigma_t(0)| \leqslant C_8. \tag{2.8}$$

Remark 2.4. As shown below, conditions (2.2), (2.6) and (2.7) guarantee the exponential integrability; conditions (2.4) and (2.5) guarantee the pathwise uniqueness; (2.3) and (2.8) are the fundamental conditions for proving the MDP.

3. Moderate deviation principle

In this section, we shall prove the MDP. First of all, we give the main result of this section.

Theorem 3.1. *Assume that* (H1)-(H3) *hold. Then Y^{ϵ} satisfies a LDP with the speed $h^2(\epsilon)$ and with the rate function I, which is defined by*

$$I(g) := \frac{1}{2}\inf_{\{\varphi\in\mathbb{H};\ g=\Gamma^0(\int_0^{\cdot}\dot{\varphi}(s)\mathrm{d}s)\}}\Big\{\int_0^T |\dot{\varphi}(s)|^2\mathrm{d}s\Big\},$$

where $\Gamma^0(\int_0^{\cdot}\dot{\varphi}(s)\mathrm{d}s) := Y^{\varphi}(\cdot)$ satisfies the following equation:

$$Y_t^{\varphi} = \int_0^t \nabla b_s(X_s^0)Y_s^{\varphi}\mathrm{d}s + \int_0^t \sigma_s(X_s^0)\dot{\varphi}(s)\mathrm{d}s. \tag{3.1}$$

Before giving the proof of Theorem 3.1, we need to give some necessary lemmas below.

Lemma 3.2. *Assume that* (H1)–(H3) *hold, then equation (3.1) has a unique solution for any $\varphi \in B_N$. Moreover, there exists a constant $C > 0$ such that*

$$\sup_{\varphi\in B_N}\Big\{\sup_{t\in[0,1]}|Y_t^{\varphi}|\Big\} \leqslant C.$$

Proof. By [16], Theorem 3.2, equation (1.3) has a unique solution and satisfies

$$\sup_{t\in[0,1]}|X_t^0| + \sup_{t\in[0,1]}\mathscr{V}(X_t^0) \leqslant C. \tag{3.2}$$

We can obtain Y_t^n by introducing the following truncated equation:

$$Y_t^n = \int_0^t \nabla b_s^n(X_s^0)Y_s^n\mathrm{d}s + \int_0^t \sigma_s^n(X_s^0)\dot{\varphi}(s)\mathrm{d}s, \tag{3.3}$$

where

$$\nabla b_t^n(x) := \nabla b_t(x)\chi_n(x), \quad \sigma_t^n(x) := \sigma_t(x)\chi_n(x)$$

and $\chi_n : \mathbb{R}^d \to \mathbb{R}$ is a family of non-negative smooth cutoff functions satisfying

$$\sup_{x\in\mathbb{R}^d} |\chi_n'(x)| \leqslant 1, \quad \chi_n(x) = \begin{cases} 1, & |x| \leqslant n, \\ 0, & |x| \geqslant n+1. \end{cases}$$

By equation (2.8), we have

$$|\nabla b_t^n(x)| \leqslant C_n, \quad |\sigma_t^n(x)| \leqslant C_n.$$

For any $N > 0$ and $|x|, |y| \leqslant N$, by equations (2.3), (2.4) and (2.8), we can get

$$|\sigma_t^n(x) - \sigma_t^n(y)| + |\nabla b_t^n(x) - \nabla b_t^n(y)| \leqslant C_n \cdot |x - y|.$$

Since the classical Carathéodory's existence theorem for ODEs in [8], (3.3) has a unique solution Y_t^n. For $\varphi \in B_N$, by equations (2.4), (2.8), (3.2) and Hölder's inequality, we have

$$\begin{aligned} |Y_t^n| &\leqslant \int_0^t C_8 \cdot (1 + \mathscr{V}^\alpha(X_s^0)|X_s^0|)|Y_s^n|\mathrm{d}s + \int_0^t \big(|\sigma_s^n(X_s^0) - \sigma_s^n(0)| + |\sigma_s^n(0)|\big) \cdot |\dot\varphi(s)|\mathrm{d}s \\ &\leqslant C\int_0^t |Y_s^n|\mathrm{d}s + \int_0^t \big[C_3 \cdot (\mathscr{V}^\alpha(X_s^0) + \mathscr{V}^\alpha(0)) \cdot |X_s^0| + C_8\big] \cdot |\dot\varphi(s)|\mathrm{d}s \\ &\leqslant C\int_0^t |Y_s^n|\mathrm{d}s + C\Big(\int_0^t |\dot\varphi(s)|^2\mathrm{d}s\Big)^{\frac{1}{2}} \\ &\leqslant C + C\int_0^t |Y_s^n|\mathrm{d}s. \end{aligned}$$

Applying Gronwall's inequality, we have

$$\sup_{t\in[0,1]} |Y_t^n| \leqslant C, \tag{3.4}$$

where C is a constant independent of n. By a method similar to Lemma 3.5 below (or see [19], Lem. 3.6), we can get $\{Y^n\}_{n\in N}$ is compatible using the stopping time and Gronwall's inequality. Let

$$Y_t^\varphi := \lim_{n\to\infty} Y_t^n,$$

it is easy to verify that Y_t^φ is a non-explosive solution of equation (3.1). By equation (3.4), it is easily seen that

$$\sup_{\varphi\in B_N} \Big\{ \sup_{t\in[0,1]} |Y_t^\varphi| \Big\} \leqslant C.$$

Next, we prove the uniqueness. Suppose that Y^1 and Y^2 are two different solutions of equation (3.1). By equations (2.8) and (3.2), we have

$$|Y_t^1 - Y_t^2| \leqslant \int_0^t |\nabla b_s(X_s^0)Y_s^1 - \nabla b_s(X_s^0)Y_s^2|\mathrm{d}s$$

$$\begin{aligned}&\leqslant\int_0^t C_7(1+\mathscr{V}^\alpha(X_s^0)|X_s^0|)|Y_s^1-Y_s^2|\mathrm{d}s\\&\leqslant C\int_0^t |Y_s^1-Y_s^2|\mathrm{d}s.\end{aligned}$$

Gronwall's inequality yields the uniqueness and we complete the proof. □

Lemma 3.3. *Assume that* (H1)–(H3) *hold, for any* $0<N<\infty$*, the family*

$$K_N:=\Big\{\Gamma^0\Big(\int_0^\cdot \dot\varphi(s)\mathrm{d}s\Big);\varphi\in B_N\Big\}$$

is compact in $\mathscr{C}$.

Proof. According to the compactness of B_N, it is enough to prove that Γ^0 is continuous. Let $\varphi_n\to\varphi$ weakly in B_N, by equation (3.1), we can write

$$\begin{aligned}Y_t^{\varphi_n}-Y_t^{\varphi}=&\int_0^t\nabla b_s(X_s^0)(Y_s^{\varphi_n}-Y_s^{\varphi})\mathrm{d}s+\int_0^t\sigma_s(X_s^0)(\dot\varphi_n(s)-\dot\varphi(s))\mathrm{d}s\\:=&I_1^n(t)+I_2^n(t).\end{aligned}$$

By equations (2.4), (2.8) and (3.2), we have

$$\begin{aligned}\int_0^1|\sigma_s(X_s^0)|^2\mathrm{d}s&\leqslant\int_0^1\big(2|\sigma_s(X_s^0)-\sigma_s(0)|^2+2|\sigma_s(0)|^2\big)\mathrm{d}s\\&\leqslant\int_0^1 2\big(C_3\cdot(\mathscr{V}^\alpha(X_s^0)+\mathscr{V}^\alpha(0))\cdot|X_s^0|^2+C_8^2\big)\mathrm{d}s\leqslant C.\end{aligned}\tag{3.5}$$

Since $\varphi_n\to\varphi$ weakly in B_N and using equation (3.5), we can get from the definition of weak convergence in B_N

$$\int_0^t\sigma_s(X_s^0)(\dot\varphi_n(s)-\dot\varphi(s))\mathrm{d}s\to 0.$$

Thus

$$I_2^n(t)\to 0.$$

For $\varphi,\varphi_n\in B_N$, by equation (3.5) and Hölder's inequality, we have

$$\begin{aligned}\sup_{0\leqslant t\leqslant 1}|I_2^n(t)|&\leqslant\int_0^1|\sigma_s(X_s^0)(\dot\varphi_n(s)-\dot\varphi(s))|\mathrm{d}s\\&\leqslant\Big(\int_0^1|\sigma_s(X_s^0)|^2\mathrm{d}s\Big)^{\frac12}\Big(\int_0^1|\dot\varphi_n(s)-\dot\varphi(s)|^2\mathrm{d}s\Big)^{\frac12}\\&\leqslant C\sqrt{2N}.\end{aligned}\tag{3.6}$$

For $0 \leqslant s \leqslant t \leqslant 1$, $\varphi, \varphi_n \in B_N$, by equations (2.4), (2.8), (3.2) and Hölder's inequality, we get

$$\begin{aligned}|I_2^n(t) - I_2^n(s)| \leqslant & \int_s^t \big|\sigma_r(X_r^0)(\dot{\varphi}_n(r) - \dot{\varphi}(r))\big|\mathrm{d}u \\ \leqslant & \Big(\int_s^t |\sigma_r(X_r^0)|^2 \mathrm{d}r\Big)^{\frac{1}{2}} \Big(\int_s^t |\dot{\varphi}_n(r) - \dot{\varphi}(r)|^2 \mathrm{d}r\Big)^{\frac{1}{2}} \\ \leqslant & C\sqrt{2N}\Big(\int_s^t \big(C_3 \cdot (\mathscr{V}^\alpha(X_r^0) + \mathscr{V}^\alpha(0)) \cdot |X_r^0| + C_8\big)\mathrm{d}r\Big)^{\frac{1}{2}} \\ \leqslant & C\sqrt{2N}|t-s|^{\frac{1}{2}}.\end{aligned} \tag{3.7}$$

By Ascoli-Arzelà's lemma, equations (3.6) and (3.7) can imply that $\{I_2^n\}_{n\geqslant 1}$ are compact in $\mathscr{C}$, then

$$\lim_{n\to\infty} \|I_2^n(t)\|_{\mathscr{C}} = 0. \tag{3.8}$$

By equations (2.8) and (3.2), we have for any $u \in [0,1]$

$$\begin{aligned}\sup_{0\leqslant t\leqslant u} |Y_t^{\varphi_n} - Y_t^{\varphi}| \leqslant & \int_0^u C_7(1 + \mathscr{V}^\alpha(X_s^0)|X_s^0|)|Y_s^{\varphi_n} - Y_s^{\varphi}|\mathrm{d}s + \sup_{0\leqslant t\leqslant 1} |I_2^n(t)| \\ \leqslant & C\int_0^u (\sup_{0\leqslant r\leqslant s} |Y_r^{\varphi_n} - Y_r^{\varphi}|)\mathrm{d}s + \|I_2^n(t)\|_{\mathscr{C}}.\end{aligned} \tag{3.9}$$

Using equations (3.8), (3.9) and Gronwall's inequality, we find that

$$\sup_{0\leqslant t\leqslant 1} |Y_t^{\varphi_n} - Y_t^{\varphi}| \leqslant e^C \cdot \|I_2^n(t)\|_{\mathscr{C}} \to 0.$$

The proof is complete. □

Let $Y^\epsilon := (X^\epsilon - X^0)/\sqrt{\epsilon}h(\epsilon)$, then Y^ϵ satisfies the the following equation:

$$Y_t^\epsilon = \frac{1}{\sqrt{\epsilon}h(\epsilon)}\int_0^t \big[b_s(X_s^0 + \sqrt{\epsilon}h(\epsilon)Y_s^\epsilon) - b_s(X_s^0)\big]\mathrm{d}s + \frac{1}{h(\epsilon)}\int_0^t \sigma_s(X_s^0 + \sqrt{\epsilon}h(\epsilon)Y_s^\epsilon)\mathrm{d}W_s.$$

By the Yamada-Watanabe theorem (see [27]), there exists a measurable map $\Gamma^\epsilon : \mathscr{C} \to \mathscr{C}$ such that $Y_\cdot^\epsilon = \Gamma^\epsilon(W_\cdot)$. For any $\varphi^\epsilon \in \mathscr{A}_N$, we can get

$$\mathbb{E}\Big(\exp\Big\{\frac{h^2(\epsilon)}{2}\int_0^1 |\dot{\varphi}^\epsilon(s)|^2 \mathrm{d}s\Big\}\Big) < \infty,$$

that is, the Novikov's condition holds. Then we define a probability measure P_ϵ by

$$\mathrm{d}P_\epsilon := R_\epsilon \mathrm{d}P = \exp\Big\{-h(\epsilon)\int_0^1 \dot{\varphi}^\epsilon(s)\mathrm{d}W_s - \frac{h^2(\epsilon)}{2}\int_0^1 |\dot{\varphi}^\epsilon(s)|^2\mathrm{d}s\Big\}\mathrm{d}P.$$

By Girsanov theorem, we have

$$\widetilde{W}_t = W_t + h(\epsilon)\int_0^t \dot{\varphi}^\epsilon(s)\mathrm{d}s$$

is a Brownian motion under the probability measure P_ϵ. Moreover, we can obtain that $Y_\cdot^{\epsilon,\varphi^\epsilon} = \Gamma^\epsilon\big(W_\cdot + h(\epsilon)\int_0^\cdot \dot\varphi^\epsilon(s)\mathrm{d}s\big)$ satisfies the following equation:

$$
\begin{aligned}
Y_t^{\epsilon,\varphi^\epsilon} &= \frac{1}{\sqrt{\epsilon}h(\epsilon)}\int_0^t \big[b_s(X_s^0+\sqrt{\epsilon}h(\epsilon)Y_s^{\epsilon,\varphi^\epsilon}) - b_s(X_s^0)\big]\mathrm{d}s \\
&\quad + \frac{1}{h(\epsilon)}\int_0^t \sigma_s(X_s^0+\sqrt{\epsilon}h(\epsilon)Y_s^{\epsilon,\varphi^\epsilon})\mathrm{d}W_s + \int_0^t \sigma_s(X_s^0+\sqrt{\epsilon}h(\epsilon)Y_s^{\epsilon,\varphi^\epsilon})\dot\varphi^\epsilon(s)\mathrm{d}s.
\end{aligned}
\tag{3.10}
$$

By equations (1.3) and (3.10), let $X_t^{\epsilon,\varphi^\epsilon} = X_t^0 + \sqrt{\epsilon}h(\epsilon)Y_t^{\epsilon,\varphi^\epsilon}$, we can write

$$
X_t^{\epsilon,\varphi^\epsilon} = x + \int_0^t b_s(X_s^{\epsilon,\varphi^\epsilon})\mathrm{d}s + \sqrt{\epsilon}\int_0^t \sigma_s(X_s^{\epsilon,\varphi^\epsilon})\mathrm{d}W_s + \sqrt{\epsilon}h(\epsilon)\int_0^t \sigma_s(X_s^{\epsilon,\varphi^\epsilon})\dot\varphi^\epsilon(s)\mathrm{d}s. \tag{3.11}
$$

The following exponential integrability is critical.

Lemma 3.4. *For any $\beta > 0$ and $\alpha \in (0,1]$, we have*

$$
\mathbb{E}\Bigg(\exp\Bigg[\sup_{t\in[0,1]}\big(\beta(e^{-\lambda(t)}\mathscr{V}^\alpha(X_t^0+\sqrt{\epsilon}h(\epsilon)Y_t^{\epsilon,\varphi^\epsilon})\big)\Bigg]\Bigg) \leqslant 2(C_1(\alpha\beta)^2+1)\cdot\exp\big[\beta\cdot\mathscr{V}^\alpha(x)\big]
$$

where $\lambda(t) = t(|C_5| + |C_6|/2 + C_1\alpha\beta/4 + 2/(\alpha\beta)) + \sqrt{C_1}\int_0^t|\dot\varphi^\epsilon(s)|\mathrm{d}s$.

Proof. The proof is similar to [16], Lemma 3.7, so we omit the details. □

Lemma 3.5. *Assume that* (H1)–(H3) *hold. Then equation (3.10) has a unique solution.*

Proof. We denote $Y_t = Y_t^{\epsilon,\varphi^\epsilon}$ for the sake of simplicity. We use the truncation method in Lemma 3.2 to obtain the following truncated equation :

$$
\begin{aligned}
Y_t^n =& \frac{1}{\sqrt{\epsilon}h(\epsilon)}\int_0^t \big[b_s^n(X_s^0+\sqrt{\epsilon}h(\epsilon)Y_s^n) - b_s^n(X_s^0)\big]\mathrm{d}s \\
&+ \frac{1}{h(\epsilon)}\int_0^t \sigma_s^n(X_s^0+\sqrt{\epsilon}h(\epsilon)Y_s^n)\mathrm{d}W_s + \int_0^t \sigma_s^n(X_s^0+\sqrt{\epsilon}h(\epsilon)Y_s^n)\dot\varphi^\epsilon(s)\mathrm{d}s.
\end{aligned}
\tag{3.12}
$$

Then equation (3.12) has a unique strong solution Y_t^n. For $m < n$, let

$$
Z_t^{nm} := Y_t^n - Y_t^m
$$

and

$$
\tau_n := \inf\{t\in[0,1] : |X_t^0 + \sqrt{\epsilon}h(\epsilon)Y_t^n| \geqslant n\}(=1 \text{ if } \{\cdots\} = \emptyset).
$$

By Itô's formula, we get

$$
\begin{aligned}
|Z_t^{nm}|^2 &= \frac{2}{\sqrt{\epsilon}h(\epsilon)}\int_0^t \langle Z_s^{nm}, b_s^n(X_s^0+\sqrt{\epsilon}h(\epsilon)Y_s^n) - b_s^m(X_s^0+\sqrt{\epsilon}h(\epsilon)Y_s^m)\rangle\mathrm{d}s \\
&\quad + 2\int_0^t \langle Z_s^{nm}, (\sigma_s^n(X_s^0+\sqrt{\epsilon}h(\epsilon)Y_s^n) - \sigma_s^m(X_s^0+\sqrt{\epsilon}h(\epsilon)Y_s^m))\dot\varphi^\epsilon(s)\rangle\mathrm{d}s
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{2}{h(\epsilon)}\int_0^t\langle Z_s^{nm},(\sigma_s^n(X_s^0+\sqrt{\epsilon}h(\epsilon)Y_s^n)-\sigma_s^m(X_s^0+\sqrt{\epsilon}h(\epsilon)Y_s^m))\mathrm{d}W_s\rangle\\
&+\frac{1}{h^2(\epsilon)}\int_0^t|\sigma_s^n(X_s^0+\sqrt{\epsilon}h(\epsilon)Y_s^n)-\sigma_s^m(X_s^0+\sqrt{\epsilon}h(\epsilon)Y_s^m)|^2\mathrm{d}s\\
=:&\ J_1(t)+J_2(t)+J_3(t)+J_4(t).
\end{aligned}
$$

Noticing that

$$
\begin{aligned}
b^n_{s\wedge\tau_m}(X^0_{s\wedge\tau_m}+\sqrt{\epsilon}h(\epsilon)Y^n_{s\wedge\tau_m})&=b^m_{s\wedge\tau_m}(X^0_{s\wedge\tau_m}+\sqrt{\epsilon}h(\epsilon)Y^n_{s\wedge\tau_m}),\\
\sigma^n_{s\wedge\tau_m}(X^0_{s\wedge\tau_m}+\sqrt{\epsilon}h(\epsilon)Y^n_{s\wedge\tau_m})&=\sigma^m_{s\wedge\tau_m}(X^0_{s\wedge\tau_m}+\sqrt{\epsilon}h(\epsilon)Y^n_{s\wedge\tau_m}).
\end{aligned}
$$

For J_1, by equation (2.5), we derive

$$
\begin{aligned}
&J_1(t\wedge\tau_m)\\
&=\frac{2}{\epsilon h^2(\epsilon)}\int_0^t\langle\sqrt{\epsilon}h(\epsilon)Z^{nm}_{s\wedge\tau_m},b^m_{s\wedge\tau_m}(X^0_{s\wedge\tau_m}+\sqrt{\epsilon}h(\epsilon)Y^n_{s\wedge\tau_m})-b^m_{s\wedge\tau_m}(X^0_{s\wedge\tau_m}+\sqrt{\epsilon}h(\epsilon)Y^m_{s\wedge\tau_m})\rangle 1_{s\leqslant\tau_m}\mathrm{d}s\\
&\leqslant 2\int_0^t C_4|Z^{nm}_{s\wedge\tau_m}|^2(\mathscr{V}^\alpha(X^0_{s\wedge\tau_m}+\sqrt{\epsilon}h(\epsilon)Y^n_{s\wedge\tau_m})+\mathscr{V}^\alpha(X^0_{s\wedge\tau_m}+\sqrt{\epsilon}h(\epsilon)Y^m_{s\wedge\tau_m}))\mathrm{d}s\\
&\leqslant C_m\int_0^t|Z^{nm}_{s\wedge\tau_m}|^2\mathrm{d}s.
\end{aligned}
$$

For J_2, by equation (2.4), we have

$$
\begin{aligned}
&J_2(t\wedge\tau_m)\\
&=2\int_0^t\langle Z^{nm}_{s\wedge\tau_m},\big(\sigma^m_{s\wedge\tau_m}(X^0_{s\wedge\tau_m}+\sqrt{\epsilon}h(\epsilon)Y^n_{s\wedge\tau_m})-\sigma^m_{s\wedge\tau_m}(X^0_{s\wedge\tau_m}+\sqrt{\epsilon}h(\epsilon)Y^m_{s\wedge\tau_m})\big)\dot{\varphi}^\epsilon(s\wedge\tau_m)\rangle 1_{s\leqslant\tau_m}\mathrm{d}s\\
&\leqslant 2\int_0^t C_3|Z^{nm}_{s\wedge\tau_m}|^2(\mathscr{V}^\alpha(X^0_{s\wedge\tau_m}+\sqrt{\epsilon}h(\epsilon)Y^n_{s\wedge\tau_m})+\mathscr{V}^\alpha(X^0_{s\wedge\tau_m}+\sqrt{\epsilon}h(\epsilon)Y^m_{s\wedge\tau_m}))|\dot{\varphi}^\epsilon(s\wedge\tau_m)|\mathrm{d}s\\
&\leqslant C_m\int_0^t|Z^{nm}_{s\wedge\tau_m}|^2\cdot|\dot{\varphi}^\epsilon(s\wedge\tau_m)|\mathrm{d}s.
\end{aligned}
$$

Similarly, we deal with J_4. For $\varphi^\epsilon\in\mathscr{A}_N$, by Hölder's inequality, we have

$$
\begin{aligned}
|Z^{nm}_{t\wedge\tau_m}|^2&\leqslant C_m\int_0^t|Z^{nm}_{s\wedge\tau_m}|^2(1+|\dot{\varphi}^\epsilon_{s\wedge\tau_m}|)\mathrm{d}s+|J_3(t\wedge\tau_m)|\\
&\leqslant C_m(1+N)\Big(\int_0^t|Z^{nm}_{s\wedge\tau_m}|^4\mathrm{d}s\Big)^{1/2}+|J_3(t\wedge\tau_m)|.
\end{aligned}
\tag{3.13}
$$

By equations (2.4), (3.13) and Burkholder-Davis-Gundy inequality, it is easy to derive that

$$
\begin{aligned}
\mathbb{E}\Big(\sup_{0\leqslant u\leqslant t}|Z^{nm}_{u\wedge\tau_m}|^4\Big)&\leqslant C_m\mathbb{E}\int_0^t|Z^{nm}_{s\wedge\tau_m}|^4\mathrm{d}s+C\mathbb{E}\Big(\sup_{0\leqslant u\leqslant t}|J_3(u\wedge\tau_m)|^2\Big)\\
&\leqslant C_m\mathbb{E}\int_0^t|Z^{nm}_{s\wedge\tau_m}|^4\mathrm{d}s.
\end{aligned}
$$

Combining this with Gronwall's inequality, we have

$$\mathbb{E}\sup_{0\leqslant u\leqslant t}|Z^{nm}_{u\wedge\tau_m}|^4=0,$$

that is, for almost all ω,

$$Y^n_{t\wedge\tau_m}=Y^m_{t\wedge\tau_m},$$

thus $\tau_m\leqslant\tau_n$, $\{Y^n\}_{n\in N}$ is compatible.

Let

$$\tau_\infty:=\lim_{n\to\infty}\tau_n. \tag{3.14}$$

For $t<\tau_n, Y_t:=Y^n_t$. Observe that $X^{\epsilon,\varphi^\epsilon}_t:=X^0_t+\sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_t$ is non-explosive by Lemma 3.4. Thus $\tau_\infty=1$ and $Y_t=\lim_{n\to\infty}Y^n_t$. Moreover, we show X^0_t is non-explosive from [16], Theorem 3.2. Then we can claim Y_t is a non-explosive solution.

It remains to prove the uniqueness. Suppose that Y' and Y'' are two different solutions to equation (3.10). Let

$$Z_t=Y'_t-Y''_t.$$

By Itô's formula we have

$$|Z_t|^2=|Z_0|^2+\int_0^t|Z_s|^2(\mathrm{d}M_s+N_s\mathrm{d}s), \tag{3.15}$$

where

$$\begin{aligned}N_t:=&\frac{2}{\sqrt{\epsilon}h(\epsilon)}|Z_t|^{-2}\langle Z_t,b_t(X^0_t+\sqrt{\epsilon}h(\epsilon)Y'_t)-b_t(X^0_t+\sqrt{\epsilon}h(\epsilon)Y''_t)\rangle\\&+2|Z_t|^{-2}\langle Z_t,(\sigma_t(X^0_t+\sqrt{\epsilon}h(\epsilon)Y'_t)-\sigma_t(X^0_t+\sqrt{\epsilon}h(\epsilon)Y''_t))\dot{\varphi}^\epsilon(t)\rangle\\&+\frac{1}{h^2(\epsilon)}|Z_t|^{-2}\mathrm{tr}[(\sigma_t(X^0_t+\sqrt{\epsilon}h(\epsilon)Y'_t)-\sigma_t(X^0_t+\sqrt{\epsilon}h(\epsilon)Y''_t))\\&\qquad\cdot(\sigma^*_t(X^0_t+\sqrt{\epsilon}h(\epsilon)Y'_t)-\sigma^*_t(X^0_t+\sqrt{\epsilon}h(\epsilon)Y''_t))],\\M_t:=&\frac{2}{h(\epsilon)}\int_0^t|Z_s|^{-2}\langle Z_s,(\sigma_s(X^0_s+\sqrt{\epsilon}h(\epsilon)Y'_s)-\sigma_s(X^0_s+\sqrt{\epsilon}h(\epsilon)Y''_s))\mathrm{d}W_s\rangle.\end{aligned}$$

By equations (2.4) and (2.5), we can get

$$\begin{aligned}\langle M\rangle_t&=\frac{4}{h^2(\epsilon)}\int_0^t|Z_s|^{-4}\cdot|(\sigma^*_s(X^0_s+\sqrt{\epsilon}h(\epsilon)Y'_s)-\sigma^*_s(X^0_s+\sqrt{\epsilon}h(\epsilon)Y''_s))Z_s|^2\mathrm{d}s\\&\leqslant 4C_3\int_0^t(\mathscr{V}^\alpha(X^0_s+\sqrt{\epsilon}h(\epsilon)Y'_s)+\mathscr{V}^\alpha(X^0_s+\sqrt{\epsilon}h(\epsilon)Y''_s))\mathrm{d}s\end{aligned} \tag{3.16}$$

and

$$N_t\leqslant(2C_4+2\sqrt{C_3}+C_3)(\mathscr{V}^\alpha(X^0_t+\sqrt{\epsilon}h(\epsilon)Y'_t)+\mathscr{V}^\alpha(X^0_t+\sqrt{\epsilon}h(\epsilon)Y''_t))\cdot(1+|\dot{\varphi}^\epsilon(t)|). \tag{3.17}$$

Using Lemma 3.4, we can obtain that $M_t + \int_0^t N_s \mathrm{d}s$ is a continuous semimartingale. Thus, the unique solution of equation (3.15) is

$$|Z_t|^2 = |Z_0|^2 \cdot \exp\left\{ M_t - \frac{1}{2}\langle M\rangle_t + \int_0^t N_s \mathrm{d}s \right\}.$$

For any $\beta > 0$ and $p \geqslant 1$, by equations (3.16), (3.17) and Lemma 3.4, there exists a time $T_0 > 0$ such that

$$\mathbb{E}\left(\exp\left[8p^2\langle M\rangle_{T_0}\right]\right) + \mathbb{E}\left(\exp\left[2p\int_0^{T_0} N_s \mathrm{d}s\right]\right) \leqslant C < +\infty, \tag{3.18}$$

where $T_0 = T_0(\alpha, p, N, C_1, C_2, C_3, C_4) > 0$.

Hence, by Novikov's criterion, for any $q \in [1, 4p]$,

$$t \mapsto \mathscr{E}(qM)_t := \exp\left\{ qM_t - \frac{q^2}{2}\langle M\rangle_t \right\}$$

is a continuous exponential $(\mathscr{F}_t)$-martingale on $[0, T_0]$. By similar calculations with [29], Lemma 2.3, we have

$$\mathbb{E}\left(\sup_{t\in[0,T_0]} |Z_t|^{2p} \right) \leqslant C \cdot |Z_0|^{2p}.$$

Combining this with $Z_0 = 0$ gives the desired result. The proof is complete. □

Lemma 3.6. *Assume that* (H1)-(H3) *hold. For any given $N < \infty$, let $\varphi^\epsilon, \varphi \in \mathscr{A}_N$ be such that φ^ϵ converges in distribution to φ, as $\epsilon \to 0$. Then $\Gamma^\epsilon(W_\cdot + h(\epsilon)\int_0^\cdot \dot\varphi^\epsilon(s)\mathrm{d}s)$ converges in distribution to $\Gamma^0(\int_0^\cdot \dot\varphi(s)\mathrm{d}s)$ in $\mathscr{C}$.*

Proof. At first, we show that $\{Y^{\epsilon,\varphi^\epsilon}\}_{\epsilon\in(0,1)}$ is tight in $\mathscr{C}$. By virtue of the Arzelà-Ascoli theorem (see [12], Thm. 4.11), it is sufficient to verify that

(a) $\sup_{\epsilon\in(0,1)} \mathbb{E}|Y_t^{\epsilon,\varphi^\epsilon}|^\gamma < \infty$;

(b) $\sup_{\epsilon\in(0,1)} \mathbb{E}|Y_t^{\epsilon,\varphi^\epsilon} - Y_s^{\epsilon,\varphi^\epsilon}|^\alpha \leqslant C|t-s|^{1+\beta}, \quad 0 \leqslant s \leqslant t \leqslant 1,$

for some positive constants α, β, γ and C.

Let

$$\tau_n := \inf\{t \in [0,1] : |Y_t^{\epsilon,\varphi^\epsilon}| \geqslant n\} (= 1 \text{ if } \{\cdots\} = \emptyset).$$

By Itô's formula, we have for any $p \geqslant 2$,

$$\begin{aligned}
|Y_{t\wedge\tau_n}^{\epsilon,\varphi^\epsilon}|^p = {} & \frac{p}{\sqrt{\epsilon}h(\epsilon)} \int_0^t |Y_{s\wedge\tau_n}^{\epsilon,\varphi^\epsilon}|^{p-2} \cdot \langle Y_{s\wedge\tau_n}^{\epsilon,\varphi^\epsilon}, b_{s\wedge\tau_n}(X_{s\wedge\tau_n}^0 + \sqrt{\epsilon}h(\epsilon)Y_{s\wedge\tau_n}^{\epsilon,\varphi^\epsilon}) - b_{s\wedge\tau_n}(X_{s\wedge\tau_n}^0)\rangle \mathrm{d}s \\
& + p\int_0^t |Y_{s\wedge\tau_n}^{\epsilon,\varphi^\epsilon}|^{p-2} \cdot \langle Y_{s\wedge\tau_n}^{\epsilon,\varphi^\epsilon}, \sigma_{s\wedge\tau_n}(X_{s\wedge\tau_n}^0 + \sqrt{\epsilon}h(\epsilon)Y_{s\wedge\tau_n}^{\epsilon,\varphi^\epsilon})\dot\varphi^\epsilon(s)\rangle \mathrm{d}s \\
& + \frac{p}{h(\epsilon)} \int_0^t |Y_{s\wedge\tau_n}^{\epsilon,\varphi^\epsilon}|^{p-2} \cdot \langle Y_{s\wedge\tau_n}^{\epsilon,\varphi^\epsilon}, \sigma_{s\wedge\tau_n}(X_{s\wedge\tau_n}^0 + \sqrt{\epsilon}h(\epsilon)Y_{s\wedge\tau_n}^{\epsilon,\varphi^\epsilon})\mathrm{d}W_s\rangle \\
& + \frac{p(p-1)}{2h^2(\epsilon)} \int_0^t |Y_{s\wedge\tau_n}^{\epsilon,\varphi^\epsilon}|^{p-2} \cdot |\sigma_{s\wedge\tau_n}(X_{s\wedge\tau_n}^0 + \sqrt{\epsilon}h(\epsilon)Y_{s\wedge\tau_n}^{\epsilon,\varphi^\epsilon})|^2 \mathrm{d}s \\
=: {} & L_1(t) + L_2(t) + M_1(t) + L_3(t).
\end{aligned}$$

For L_1, by equation (2.5), it is easy to derive that

$$L_1(t) \leqslant pC_4 \int_0^t (\mathscr{V}^\alpha(X^0_{s\wedge\tau_n} + \sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}) + \mathscr{V}^\alpha(X^0_{s\wedge\tau_n}))|Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}|^p\mathrm{d}s.$$

For L_2, by equations (2.4), (2.8), (3.2) and Young's inequality, we get

$$\begin{aligned}
L_2(t) &\leqslant p\int_0^t |Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}|^{p-2}\cdot|\sigma_{s\wedge\tau_n}(X^0_{s\wedge\tau_n} + \sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n})|^2\mathrm{d}s + p\int_0^t |Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}|^p\cdot|\dot{\varphi}^\epsilon(s)|^2\mathrm{d}s\\
&\leqslant C\int_0^t |Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}|^{p-2}\cdot|\sigma_{s\wedge\tau_n}(X^0_{s\wedge\tau_n} + \sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}) - \sigma_{s\wedge\tau_n}(X^0_{s\wedge\tau_n})|^2\mathrm{d}s\\
&\quad + C\int_0^t |Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}|^{p-2}\cdot\big(|\sigma_{s\wedge\tau_n}(X^0_{s\wedge\tau_n}) - \sigma_{s\wedge\tau_n}(0)|^2 + |\sigma_{s\wedge\tau}(0)|^2\big)\mathrm{d}s + p\int_0^t |Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}|^p\cdot|\dot{\varphi}^\epsilon(s)|^2\mathrm{d}s\\
&\leqslant C\int_0^t \sqrt{C_3}|Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}|^p\cdot(\mathscr{V}^\alpha(X^0_{s\wedge\tau_n} + \sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}) + \mathscr{V}^\alpha(X^0_{s\wedge\tau_n}))\mathrm{d}s\\
&\quad + C\int_0^t |Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}|^{p-2}\cdot\big(\sqrt{C_3}(\mathscr{V}^\alpha(X^0_{s\wedge\tau_n}) + \mathscr{V}^\alpha(0))|X^0_{s\wedge\tau_n}| + C_8\big)\mathrm{d}s + p\int_0^t |Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}|^p\cdot|\dot{\varphi}^\epsilon(s)|^2\mathrm{d}s\\
&\leqslant C\int_0^t |Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}|^p\cdot(\mathscr{V}^\alpha(X^0_{s\wedge\tau_n} + \sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}) + \mathscr{V}^\alpha(X^0_{s\wedge\tau_n}) + 1 + |\dot{\varphi}^\epsilon(s)|^2)\mathrm{d}s + C.
\end{aligned}$$

Similarly, we deal with L_3. Putting all these estimates of $L_1(t) - L_3(t)$ together, for $\varphi^\epsilon \in \mathscr{A}_N$, we can obtain that

$$|Y^{\epsilon,\varphi^\epsilon}_{t\wedge\tau_n}|^p \leqslant C\int_0^t |Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}|^p\cdot(\mathscr{V}^\alpha(X^0_{s\wedge\tau_n} + \sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}) + \mathscr{V}^\alpha(X^0_{s\wedge\tau_n}) + 1 + |\dot{\varphi}^\epsilon(s)|^2)\mathrm{d}s + M_1(t) + C.$$

By stochastic Gronwall's inequality (2.1), for any $0 < q < \kappa < 1$ and stopping time τ', we have

$$\mathbb{E}\Big(\sup_{t\in[0,1\wedge\tau']} |Y^{\epsilon,\varphi^\epsilon}_{t\wedge\tau_n}|^{pq}\Big) \leqslant C\Big(\mathbb{E}\exp\{\frac{\kappa}{1-\kappa}A_{1\wedge\tau'}\}\Big)^{1-1/\kappa}, \tag{3.19}$$

where

$$A_t = \int_0^t (\mathscr{V}^\alpha(X^0_{s\wedge\tau_n} + \sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}) + \mathscr{V}^\alpha(X^0_{s\wedge\tau_n}) + 1 + |\dot{\varphi}^\epsilon(s)|^2)\mathrm{d}s.$$

Similar to the proof of equation (3.18), there exists a time T_0 such that

$$\mathbb{E}\left(\exp\left\{\frac{\kappa}{1-\kappa}\int_0^{T_0}(\mathscr{V}^\alpha(X^0_{s\wedge\tau_n} + \sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n})\mathrm{d}s\right\}\right) \leqslant C < +\infty. \tag{3.20}$$

Since A_t is a continuous adapted process and we define

$$\tau'_N := \inf\{t \in [0,1] : A_t > N\}(= 1 \text{ if } \{\cdots\} = \emptyset),$$

then

$$P\Big(\lim_{n\to\infty}\tau'_N = 1\Big) = 1. \tag{3.21}$$

In equation (3.19), we replace τ' by τ'_N and insert equation (3.20) into it, we have

$$\mathbb{E}\Big(\sup_{t\in[0,T_0\wedge\tau'_N]}|Y^{\epsilon,\varphi^\epsilon}_{t\wedge\tau_n}|^{pq}\Big)\leqslant C.$$

Letting $N,n\to\infty$, by equation (3.21) and Fatou's lemma

$$\mathbb{E}\Big(\sup_{t\in[0,T_0]}|Y^{\epsilon,\varphi^\epsilon}_{t}|^{pq}\Big)\leqslant C. \tag{3.22}$$

For $i\in\mathbb{N}$, according to the proof of [29], Theorem 2.4, by using the uniqueness of solution of equation (3.10) and shifting time technique, we can obtain that

$$Y^{\epsilon,\varphi^\epsilon}_t(x,\omega)=Y^{\epsilon,\varphi^\epsilon}_{t-iT_0}(Y^{\epsilon,\varphi^\epsilon}_{iT_0}(x,\omega),\theta_{iT_0}(\omega)),\quad \forall t\in[iT_0,(i+1)T_0].$$

Combining this with equation (3.22) gives that

$$\mathbb{E}\Big(\sup_{t\in[0,1]}|Y^{\epsilon,\varphi^\epsilon}_{t}|^{pq}\Big)\leqslant C. \tag{3.23}$$

In the sequel, by equations (2.4), (2.5), (2.8), (3.2), (3.23), Hölder's inequality and Burkholder-Davis-Gundy inequality, we have

$$\begin{aligned}
&\mathbb{E}|Y^{\epsilon,\varphi^\epsilon}_{t\wedge\tau_n}-Y^{\epsilon,\varphi^\epsilon}_{s\wedge\tau_n}|^p\\
\leqslant&\frac{C}{\epsilon^{\frac{p}{2}}h^p(\epsilon)}\mathbb{E}\Big(\int_s^t|b_{u\wedge\tau_n}(X^0_{u\wedge\tau_n}+\sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{u\wedge\tau_n})-b_{u\wedge\tau_n}(X^0_{u\wedge\tau_n})|\mathrm{d}u\Big)^p\\
&+C\mathbb{E}\Big(\int_s^t|\sigma_{u\wedge\tau_n}(X^0_{u\wedge\tau_n}+\sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{u\wedge\tau_n})\dot{\varphi}^\epsilon(u)|\mathrm{d}u\Big)^p\\
&+C\mathbb{E}\Big|\int_s^t\sigma_{u\wedge\tau_n}(X^0_{u\wedge\tau_n}+\sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{u\wedge\tau_n})\mathrm{d}W_u\Big|^p\\
\leqslant&C|t-s|^{p-1}\mathbb{E}\int_s^t\big(\mathscr{V}^\alpha(X^0_{u\wedge\tau_n}+\sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{u\wedge\tau_n})+\mathscr{V}^\alpha(X^0_{u\wedge\tau_n})\big)^p|Y^{\epsilon,\varphi^\epsilon}_{u\wedge\tau_n}|^p\mathrm{d}u\\
&+C\mathbb{E}\Big(\int_s^t|\sigma_{u\wedge\tau_n}(X^0_{u\wedge\tau_n}+\sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{u\wedge\tau_n})|^2\mathrm{d}u\Big)^{\frac{p}{2}}\Big[\Big(\int_s^t|\dot{\varphi}^\epsilon(u)|^2\mathrm{d}u\Big)^{\frac{p}{2}}+1\Big]\\
\leqslant&C|t-s|^{p-1}\int_s^t\big\{\mathbb{E}\big(\mathscr{V}^\alpha(X^0_{u\wedge\tau_n}+\sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{u\wedge\tau_n})+\mathscr{V}^\alpha(X^0_{u\wedge\tau_n})\big)^{2p}\big\}^{\frac{1}{2}}\big\{\mathbb{E}|Y^{\epsilon,\varphi^\epsilon}_{u\wedge\tau_n}|^{2p}\big\}^{\frac{1}{2}}\mathrm{d}u\\
&+C|t-s|^{\frac{p}{2}-1}\mathbb{E}\int_s^t|\sigma_{u\wedge\tau_n}(X^0_{u\wedge\tau_n}+\sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{u\wedge\tau_n})-\sigma_{u\wedge\tau_n}(X^0_{u\wedge\tau_n})+\sigma_{u\wedge\tau_n}(X^0_{u\wedge\tau_n})|^p\mathrm{d}u\\
\leqslant&C|t-s|^{\frac{p}{2}-1}\mathbb{E}\int_s^t\big(\mathscr{V}^\alpha(X^0_{u\wedge\tau_n}+\sqrt{\epsilon}h(\epsilon)Y^{\epsilon,\varphi^\epsilon}_{u\wedge\tau_n})+\mathscr{V}^\alpha(X^0_{u\wedge\tau_n})\big)^p|Y^{\epsilon,\varphi^\epsilon}_{u\wedge\tau_n}|^p\mathrm{d}u\\
&+C|t-s|^{\frac{p}{2}-1}\mathbb{E}\int_s^t\big[\big(\mathscr{V}^\alpha(X^0_{u\wedge\tau_n})+\mathscr{V}^\alpha(0)\big)^p|X^0_{u\wedge\tau_n}|^p+C_8\big]\mathrm{d}u+C|t-s|^p\\
\leqslant&C|t-s|^{\frac{p}{2}}.
\end{aligned}$$

Therefore, we can choose $p > 2$ such that (a) and (b) hold. Thus $\{Y^{\epsilon,\varphi^\epsilon}\}_{\epsilon\in(0,1)}$ is tight in $\mathscr{C}$. Consequently, it suffices to show that Y^φ is the unique limit point of $\{Y^{\epsilon,\varphi^\epsilon}\}_{\epsilon\in(0,1)}$. Let

$$M_t^\epsilon = \frac{1}{h(\epsilon)}\int_0^t \sigma_s(X_s^0 + \sqrt{\epsilon}h(\epsilon)Y_s^{\epsilon,\varphi^\epsilon})\mathrm{d}W_s.$$

Since $\{Y^{\epsilon,\varphi^\epsilon}\}_{\epsilon\in(0,1)}$ is tight, we can choose a subsequence of $(Y^{\epsilon,\varphi^\epsilon}, \varphi^\epsilon, M^\epsilon)$ convergent weakly to $(\overline{Y}, \varphi, 0)$ as $\epsilon \to 0$. By Skorokhod representation theorem there exist a probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$, and on this basis, a Brownian motion $\tilde{W}$ and also a family of $\tilde{\mathscr{F}}$-predictable process $\{\tilde{\varphi}\}$, φ taking values on $\mathscr{A}_N$, such that $\{(\varphi^\epsilon, \varphi, W)\}$ has the same law as $\{(\tilde{\varphi}^\epsilon, \tilde{\varphi}, \tilde{W})\}$ for each ϵ and

$$\lim_{\epsilon\to 0}\langle \tilde{\varphi}^\epsilon - \tilde{\varphi}, g\rangle = 0, \quad g \in \mathbb{H}, \;\; \tilde{P} - a.s.$$

For the sake of simplicity, we drop off the $\tilde{\cdot}$ in the notation. Thus, we may assume

$$(Y^{\epsilon,\varphi^\epsilon}, \varphi^\epsilon, M^\epsilon) \to (\overline{Y}, \varphi, 0), \quad \tilde{P} - a.s.$$

By virtue of the uniqueness of strong solution of equation (3.10) and taking $\epsilon \to 0$ on both sides of equation (3.10), we infer that $\overline{Y}$ also satisfies equation (3.1). Thus the desired assertion follows from the uniqueness. ☐

Proof of Theorem 3.1 By Theorem 2.2, we only need to prove that condition (i) and (ii) hold. Condition (i) has been established in Lemma 3.6. Condition (ii) is fulfilled in Lemma 3.3.

4. Application to stochastic Hamiltonian systems

In this section, we shall apply the result of Theorem 3.1 to stochastic Hamiltonian systems by some examples. Let $H(x)$ be a Hamiltonian function given by

$$H(x) := H(z,u) := \frac{1}{2}|u|^2 + V(z),$$

where $x = (z,u) \in \mathbb{R}^d \times \mathbb{R}^d$, z and u denote the position and velocity of the motion of a particle in the statistical physics, respectively. $V(z)$ is a potential function bounded from below 1. Consider the following stochastic Hamiltonian system in the phase space $\mathbb{R}^{2d}$:

$$\begin{cases} \mathrm{d}Z_t = \partial_u H(Z_t,U_t)\mathrm{d}t = U_t\mathrm{d}t, & Z_0 = z, \\ \mathrm{d}U_t = -[\partial_z H(Z_t,U_t) + \Phi_t(Z_t,U_t)]\mathrm{d}t + \Theta_t(Z_t,U_t)\mathrm{d}W_t & \\ \quad\;\; = -[\nabla V(Z_t) + \Phi_t(Z_t,U_t)]\mathrm{d}t + \Theta_t(Z_t,U_t)\mathrm{d}W_t, & U_0 = u, \end{cases}$$

where $\Phi_t : \mathbb{R}_+ \times \mathbb{R}^{2d} \to \mathbb{R}^d$ and $\Theta_t : \mathbb{R}_+ \times \mathbb{R}^{2d} \to \mathbb{R}^d \times \mathbb{R}^m$. A Lyapunov function $\mathscr{V}$ is given by

$$\begin{aligned} \mathscr{V}(x) :=& H(x) \\ (x = (z,u)) =& \frac{1}{2}|u|^2 + V(z). \end{aligned}$$

With the convention, we use $\partial_u H(x)$ and $\nabla_u H(x)$ as synonymy.

We make the following assumptions on V, Φ and Θ: for some $\alpha \in (0,1]$,

(A1)

$$|\nabla V(z) - \nabla V(\tilde{z})| \leqslant C_V \cdot (V^\alpha(z) + V^\alpha(\tilde{z})) \cdot |z - \tilde{z}|. \tag{4.1}$$

(A2)

$$|\Phi_t(x)| \leqslant C_\Phi \cdot \mathscr{V}^{\frac{1}{2}}(x), \tag{4.2}$$
$$|\Phi_t(x) - \Phi_t(\tilde{x})| \leqslant C_\Phi \cdot (\mathscr{V}^\alpha(x) + \mathscr{V}^\alpha(\tilde{x})) \cdot |x - \tilde{x}|, \tag{4.3}$$
$$|\nabla\Phi_t(x) - \nabla\Phi_t(\tilde{x})| \leqslant C_\Phi \cdot (\mathscr{V}^\alpha(x) + \mathscr{V}^\alpha(\tilde{x})) \cdot |x - \tilde{x}|. \tag{4.4}$$

(A3)

$$|\Theta_t(x)|^2 \leqslant C_\Theta \cdot \mathscr{V}^{1-\alpha}(x), \tag{4.5}$$
$$|\nabla_u\Theta_t(x)|^2 \leqslant C_\Theta \cdot \mathscr{V}^\alpha(x), \tag{4.6}$$
$$|\Theta_t(x) - \Theta_t(\tilde{x})|^2 \leqslant C_\Theta \cdot (\mathscr{V}^\alpha(x) + \mathscr{V}^\alpha(\tilde{x})) \cdot |x - \tilde{x}|^2, \tag{4.7}$$
$$|\nabla_u\Theta_t(x) - \nabla_u\Theta_t(\tilde{x})|^2 \leqslant C_\Theta \cdot (\mathscr{V}^{(3\alpha-1)}(x) + \mathscr{V}^{(3\alpha-1)}(\tilde{x})) \cdot |x - \tilde{x}|^2. \tag{4.8}$$

Now we apply our result to stochastic Hamiltonian systems.
The corresponding perturbation equation is as follows:

$$\begin{cases} \mathrm{d}Z_t = U_t \mathrm{d}t, & Z_0 = z, \\ \mathrm{d}U_t = -[\partial_z V(Z_t) + \Phi_t(Z_t, U_t)]\mathrm{d}t + \epsilon\Theta_t(Z_t, U_t)\mathrm{d}W_t, & U_0 = u. \end{cases}$$

Theorem 4.1. *Assume that V, Φ and Θ satisfy* (A1)-(A3). Then $\{X^\varepsilon\}_{\epsilon>0}$ satisfies the MDP with the rate function $I(f)$ and with the speed $h^2(\epsilon)$.

Proof. Applying Theorem 3.1 to the following SDEs:

$$\mathrm{d}X_t(x) = b_t(X_t(x))\mathrm{d}t + \sigma_t(X_t(x))\mathrm{d}W_t, X_0 = (z, u),$$

where

$$b_t(x) := b_t(z, u) := (u, -\nabla V(z) - \Phi_t(z, u))$$

and

$$\sigma_t(x) := \sigma_t(z, u) := \begin{pmatrix} 0, & 0 \\ 0, & \Theta_t(z, u) \end{pmatrix}.$$

By equations (4.2), (4.5) and $u \leqslant C\mathscr{V}^{\frac{1}{2}}(x)$, we have

$$(b(x), \nabla\mathscr{V}(x)) = -\Phi_t(x) \cdot u \leqslant C\Phi_t(x) \cdot \mathscr{V}^{\frac{1}{2}}(x) \leqslant \tilde{C}_5 \cdot \mathscr{V}(x),$$

$$|\sigma_t^*(x) \cdot \nabla\mathscr{V}(x)|^2 = |\Theta_t^*(x) \cdot u|^2 \leqslant C|\Theta_t^*(x)|^2 \cdot \mathscr{V}(x) \leqslant \tilde{C}_1 \cdot \mathscr{V}^{2-\alpha}(x).$$

By equation (4.5) and $\mathscr{V}(x) \geqslant 1$, we get

$$\sum_{i,j=1}^{2d}\sum_{k=1}^{d+m}(\sigma_t^{ik}\sigma_t^{jk})(x)\partial_i\partial_j\mathscr{V}(x) = \sum_{i,j=1}^{d}(\Theta\Theta^*)^{ij}(x) \leqslant \sum_{i,j=1}^{d}(\Theta\Theta^*)^{ij}(x)\mathscr{V}^\alpha(x) \leqslant \tilde{C}_6 \cdot \mathscr{V}(x).$$

Combining equations (4.1), (4.3), (4.4) and (4.7) together, we verify (b, σ) satisfies (H1) and (H2).

By equations (4.1), (4.2), (4.3) and the definition of $\mathscr{V}(x)$, we have

$$\begin{aligned}|b_t(x)| &\leqslant |b_t(x)-b_t(0)|+|b_t(0)|\\ &\leqslant \tilde{C}_4\cdot(\mathscr{V}^\alpha(x)+\mathscr{V}^\alpha(0))\cdot|x|+|\nabla V(0)-\Phi_t(0)|\\ &\leqslant \tilde{C}_7(1+\mathscr{V}^\alpha(x)|x|).\end{aligned}$$

By equations (4.2), (4.3), (4.4) and (4.5), we can check (b,σ) satisfies (H3). The proof is complete. □

Example 4.2. For specificity, consider the following stochastic nonlinear oscillator equation:

$$\ddot{Z}_t=[C_0\dot{Z}_t-\nabla V(Z_t)]\mathrm{d}t+\Theta(Z_t)\mathrm{d}W_t,\quad (Z_0,\dot{Z}_0)=(z,u)\in\mathbb{R}^2,$$

where $H(z,u):=\frac{1}{2}|u|^2+V(z)$ for $|z|>1$. Let $x:=(z,u)$, assume that V and Θ satisfy (A1) and (A3), respectively. Then $\{X^\epsilon\}_{\epsilon>0}$ satisfies the MDP with the rate function $I(f)$ and with the speed $h^2(\epsilon)$.

Note that by (A1) it is easy to derive that $\frac{|\nabla V(z)-\nabla V(\tilde{z})|}{|z-\tilde{z}|}\leqslant C_V\cdot(V^\alpha(z)+V^\alpha(\tilde{z}))$ if $|z|>1,|\tilde{z}|>1,$ and $z\neq\tilde{z}$. Moreover, if V is second-order differentiable, we can get $|V''(u)|\leqslant C_V\cdot V^\alpha(u)$ from the above inequality for $|u|>1$. This is basically a bound of the second derivative of V and implies more or less sub-exponential growth of V. Therefore, the condition to verify whether (A1) is tenable to a certain extent is transformed into the condition to verify whether the inequality $|V''(u)|\leqslant C_V\cdot V^\alpha(u)$ is tenable $|u|>1$. Below we give some examples of $V(z)$ satisfying condition (A1).

(1) If $V(z)=|z|^p$ for $|z|>1$ and $p>2$, then the second derivative is $|V''(z)|\leqslant C|z|^{p-2}=C_V\cdot V^{\frac{p-2}{p}}(z)$ for $|z|>1$ and $p>2$. Therefore, (A1) holds with $\alpha\in[(p-2)/p,1]$. Notice that if p is large then α must be close to 1.

(2) If $V(z)=e^{|z|}$ for $|z|>1$, then (A1) holds with $\alpha=1$.

(3) If $V(z)=|z|^p+e^{|z|}$ for $|z|>1$, then the second derivative is $|V''(z)|\leqslant C(|z|^{p-2}+e^{|z|})\leqslant C(|z|^p+e^{|z|})=C_V\cdot V(z)$ for $|z|>1$. Thus, (A1) holds with $\alpha=1$.

(4) If $V(z)=|z|^pe^{|z|}$ for $|z|>1$, then the second derivative is $|V''(z)|\leqslant C(|z|^{p-2}+|z|^{p-1}+|z|^p)e^{|z|})\leqslant C|z|^pe^{|z|}=C_V\cdot V(z)$ for $|z|>1$. Thus, (A1) holds with $\alpha=1$.

(5) If $V(z)=e^{\beta|z|}+|\sin e^{\gamma|z|}|$ for $\beta>0>\gamma$ and $|z|>1$, then the second derivative is $|V''(z)|\leqslant C(e^{\beta|z|}+e^{\gamma|z|}+e^{2\gamma|z|}|\sin e^{\gamma|z|}|)\leqslant C(e^{\beta|z|}+|\sin e^{\gamma|z|}|)=C_V\cdot V(z)$ for $|z|>1$. So, (A1) holds with $\alpha=1$.

(6) If $V(z)=|z|^p+|\sin|z|^q|$ for $q\leqslant 1$ and $|z|>1$, then the second derivative is $|V''(z)|\leqslant C(|z|^{p-2}+|z|^{q-2}|\cos|z|^q|+|z|^{2q-2}|\sin|z|^q|)\leqslant C(|z|^p+|\sin|z|^q|)=C_V\cdot V(z)$ for $|z|>1$. So, (A1) holds with $\alpha=1$.

(7) If $V(z)=e^{|z|}+|\sin|z|^q|$ for $q\leqslant 1$ and $|z|>1$, then the second derivative is $|V''(z)|\leqslant C(e^{|z|}+|z|^{q-2}|\cos|z|^q|+|z|^{2q-2}|\sin|z|^q|)\leqslant C(e^{|z|}+|\sin|z|^q|)=C_V\cdot V(z)$ for $|z|>1$. Thus, (A1) holds with $\alpha=1$.

Acknowledgements. The authors would like to thank Editor-in-Chief, two anonymous reviewers and Professor Jicheng Liu for their careful reading of manuscript, correcting errors and making very helping suggestions, which improve the quality of this paper. The authors also acknowledge the support provided by NSFs of China Nos. 11971154 and 11901154.

References

[1] A. Budhiraja, J. Chen and P. Dupuis, Large deviations for stochastic partial differential equations driven by a Poisson random measure. *Stochastic Process Appl.* **123** (2013) 523–560.

[2] A. Budhiraja and P. Dupuis, A variational representation for positive functionals of infinite dimensional Brownian motion. *Probab. Math. Stat.* **20** (2000) 39–61.

[3] A. Budhiraja, P. Dupuis and V. Maroulas, Large deviations for infinite dimensional stochastic dynamical systems. *Ann. Probab.* **36** (2008) 1390–1420.

[4] J. Bao and C. Yuan, Large deviations for neutral functional SDEs with jumps. *Stochastics* **87** (2015) 48–70.

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

[6] M. Boué, P. Dupuis and R.S. Ellis, Large deviations for small noise diffusions with discontinuous statistics. *Probab. Theory Related Fields* **116** (2000) 125–149.
[7] P.L. Chow, Some parabolic Itô equations. *Commun. Pure Appl. Math.* **45** (1992) 97–120.
[8] E.A. Coddington and N. Levinson, Theory of Ordinary Differential Equations. McGraw-Hill, New York (1955).
[9] P. Dupuis and R.S. Ellis, A Weak Convergence Approach to the Theory of Large Deviations. Wiley, New York (1997).
[10] M.I. Freidlin and A.D. Wentzell, On small random perturbations of dynamical system. *Russian Math. Surveys* **25** (1970) 1–55.
[11] G. Kallianpur and J. Xiong, Stochastic differential equations in infinite dimensional spaces. IMS Lecture Notes- Monograph Series, vol. 26. *Institute Math. Stat.* (1995).
[12] I. Karatzas and S.E. Shreve, Brownian Motion and Stochastic Calculus, 2nd ed. Graduate Texts in Mathematics. Springer-Verlag, New York (1991).
[13] A. Kolmogorov and S. Fomin, Elements of the Theory of Functions and Functional Analysis. Graylock, Rochester (1957).
[14] Y. Ma, R. Wang and L. Wu, Moderate Deviation Principle for dynamical systems with small random perturbation. *J. Math.* **32** (2012) 395–401.
[15] V. Maroulas, Large deviations for infinite-dimensional stochastic systems with jumps. *Mathematika* **57** (2011) 175–192.
[16] J. Ren, Large deviation principle for homeomorphism flows of stochastic Hamiltonian systems. *Acta Math. Sin. (Chin. Ser.)* **61** (2018) 383–402.
[17] J. Ren and X. Zhang, Freidlin–Wentzell's large deviations for stochastic evolution equations. *J. Funct. Anal.* **254** (2008) 3148–3172.
[18] J. Ren and X. Zhang, Freidlin–Wentzell's large deviations for homemorphism flows of non-Lipschitz SDEs. *Bull. Sci. Math.* **129** (2005) 643–655.
[19] J. Ren and X. Zhang, Schilder theorem for the Brownian motion on the diffeomorphism group of the circle. *J. Funct. Anal.* **224** (2005) 107–133.
[20] R. Sowers, Large deviations for a reaction diffusion equation with non-Gaussian perturbations. *Ann. Probab.* **20** (1992) 504–537.
[21] Y. Suo, J. Tao and W. Zhang, Moderate deviation and central limit theorem for stochastic differential delay equations with polynomial growth. *Front. Math. China* **013** (2018) 913–933.
[22] R. Wang, J. Zhai and T. Zhang, A moderate deviation principle for 2-D stochastic Navier–Stokes equations. *J. Differ. Equ.* **258** (2015) 3363–3390.
[23] R. Wang and T. Zhang, Moderate deviations for stochastic reaction-diffusion equations with multiplicative noise. *Potential Anal.* **42** (2015) 99–113.
[24] J. Wu, Uniform large deviations for multivalued stochastic differential equations with Poisson jumps. *Kyoto J Math.* **51** (2011) 535–559.
[25] L. Wu, Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stochastic Process. Appl.* **91** (2001) 205–238.
[26] L. Xie and X. Zhang, Ergodicity of stochastic differential equations with jumps and singular coefficients. *Ann. I. H. Poincare* **56** (2020) 175–229.
[27] T. Yamada and S. Watanabe, On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.* 11 (1971) 155–167.
[28] F. X, C. Zhou and F. Wu, On strong Feller property, exponential ergodicity and large deviations principle for stochastic damping Hamiltonian systems with state-dependent switching. *J. Differ. Equ.* **286** (2021) 856–891.
[29] X. Zhang, Stochastic flows and Bismut formulas for stochastic Hamiltonian systems. *Stochastic Process Appl.* **120** (2010) 1929–1949.
[30] X. Zhang, Stochastic Volterra equations in Banach spaces and stochastic partial differential equation. *J. Funct. Anal.* **258** (2010) 1361–1425.
[31] X. Zhang, Well-posedness and large deviation for degenerate SDEs with Sobolev coefficients. *Rev. Mat. Iberoamericana* **29** (2013) 25–52.
[32] X. Zhang, A variational representation for random functionals on abstract Wiener spaces. *J. Math. Kyoto Univ.* **49** (2009) 475–490.