Large Deviations Principle for SDEs with Dini Continuous Drifts *

Lingyan Cheng[a)], **Xing Huang**[b)]

[a)]School of Science, Nanjing University of Science and Technology, Nanjing 210094, China

[b)]Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

xinghuang@tju.edu.cn

chengly@amss.ac.cn

Abstract

In this paper, using Zvonkin type transform, the large deviation principle is proved for stochastic differential equations with Dini continuous drifts, where the existed methods for large deviation principle are unavailable. The method and result are new in related fields. Moreover, the result is also extended to a class of degenerate stochastic differential equations with Dini continuous drifts.

AMS subject Classification: 60H35, 60H10, 60C30
Keywords: Dini continuity, Zvonkin type transform, Large deviation, Degenerate SDEs

1 Introduction

The large deviation principle (LDP for short) is proved for various stochastic differential equations (SDEs) with Lipschitz continuous drift. For instance, Freidlin and Wentzell [10] firstly studied the LDP for the finite dimensional setting, where the SDE is driven by finitely many Brownian motions and its coefficients satisfy suitable regularity properties. Peszat [20] (also the references therein) investigated the LDP for stochastic partial differential equations (SPDEs) under global Lipschitz condition on the nonlinear term. Cerrai and Röckner [6] obtained the LDP for stochastic reaction-diffusion systems with multiplicative noise under local Lipschitz conditions. Moreover, the LDP for semilinear parabolic equations on a Gelfand triple was proved by Chow in [7]. Röckner, Wang and Wu [23] established the LDP for stochastic porous media equations within the variational framework. All these papers mainly

*Supported in part by NNSFC (11801406).

used the classical ideas of discretization approximations and the contraction principle, which was firstly developed by Freidlin and Wentzell.

Budhiraja, Dupuis and Maroulas [3] also get the LDP of the infinite dimensional setting by the weak convergence method (see [1]). This approach is now a powerful tool which has been extensively used to prove LDP for various stochastic dynamical systems. For instance, Cerrai and Freidlin [5] established the LDP for the langevin equation, see also [2, 17, 18, 19, 22, 24, 25, 28, 30] and the references therein for more works. There are also some results with non-Lipschitz coefficients, for instance, [8, 14, 15].

Recently, pathwise uniqueness of SDEs/SPDEs with singular drifts is proved. The main idea is to construct Zvonkin's transform ([31]) which is a homeomorphism map to transform the original one to a new one, where the singular drift is killed and the pathwise uniqueness can be obtained. This technique strongly depends on the regularity of the solution to PDE like (2.4) below with singular coefficients. Wang [26] proved the pathwise uniqueness for semi-linear SPDEs with Dini continuous drift and non-degenerate noise. In [27], Wang and Zhang studied existence and uniqueness for stochastic Hamiltonian system with Hölder-Dini continuous drifts, where the noise is degenerate. There are also many other results on this topic, see [9, 12, 13, 21, 29] and references therein.

So far, there are no results on LDP for SDEs with Hölder continuous or more singular drifts, where the Gronwall lemma, which is crucial in both discretization approximations and weak convergence, can not been used directly. The aim of this paper is to solve this problem. To this end, we need to search for new technique and Zvonkin's transform offers an effective method to regularize the singular drifts. The idea is to use Zvonkin's transform to change the SDEs with singular drifts as a new one with Lipschitz continuous coefficients, where the LDP holds. Then we can obtain the LDP for the original SDE by the inverse of Zvonkin's transform and the definition of LDP.

Throughout the paper, the following notations will be used. For $T > 0$, $d \in \mathbb{N}^+$, let $C([0,T], \mathbb{R}^d)$ be all $\mathbb{R}^d$-valued and continuous functions on $[0,T]$. For a function f from $\mathbb{R}^m$ to $\mathbb{R}^n$, set $\|f\|_\infty := \sup_{x\in\mathbb{R}^m} |f(x)|$.

Before moving on, let us recall some knowledge on LDP.

Definition 1.1. Let S be a Polish space. A function $I : S \to \mathbb{R}^1$ is called a rate function if it is lower semicontinuous. If for any constant $c > 0$, the level set $\{f; I(f) \le c\}$ is compact in S, then I is called a good rate function.

Definition 1.2. Let S be a Polish space. We call a family of S-valued random variables $\{Z^\varepsilon\}_{\varepsilon\in(0,1)}$ satisfies an LDP with speed function ε^{-1} and rate function $I : S \to [0, \infty)$, if the following conditions hold.

(1) For any closed subset $F \subset S$,

$$\limsup_{\varepsilon\to 0^+} \varepsilon \log \mathbb{P}(Z^\varepsilon \in F) \le - \inf_{f\in F} I(f).$$

(2) For any open subset $G \subset S$,

$$\liminf_{\varepsilon\to 0^+} \varepsilon \log \mathbb{P}(Z^\varepsilon \in G) \ge - \inf_{f\in G} I(f).$$

From now on, we fix $T > 0$. Next, we give a known result in Lemma 1.2 which will be used in the sequel, see [10] or the introduction in [11]. Consider an SDE on $\mathbb{R}^n$:

$$\mathrm{d}\tilde{X}_t^\varepsilon = b_1^\varepsilon(\tilde{X}_t^\varepsilon) + \sqrt{\varepsilon}\sigma(\tilde{X}_t^\varepsilon)\mathrm{d}W_t, \quad t \in [0,T], \quad \tilde{X}_0^\varepsilon = x_0 \in \mathbb{R}^n, \tag{1.1}$$

where $\varepsilon \in (0,1)$, $b_1^\varepsilon : \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : \mathbb{R}^n \to \mathbb{R}^n \otimes \mathbb{R}^n$, and $(W_t)_{t\in[0,T]}$ is an n-dimensional Brownian motion defined on a complete filtration probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in[0,T]}, \mathbb{P})$. Without loss of generality, we assume $x_0 = 0$.

(A1) There exists a constant $L > 0$ such that for any $\varepsilon \in (0,1)$,

$$\|\sigma(x) - \sigma(y)\| + |b_1^\varepsilon(x) - b_1^\varepsilon(y)| \le L|x-y|, \quad x, y \in \mathbb{R}^n. \tag{1.2}$$

Moreover, there exists a Lipschitz continuous function $b_1^0 : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$\lim_{\varepsilon\to 0}\left\{\sup_{x\in\mathbb{R}^n} |b_1^\varepsilon(x) - b_1^0(x)|\right\} = 0. \tag{1.3}$$

Let $C([0,T],\mathbb{R}^n)$ be equipped with sup-norm, and define rate function $I : C([0,T],\mathbb{R}^n) \to [0,\infty)$ as

$$I(f) = \frac{1}{2}\inf_{f=g(h),h\in\mathcal{H}} \|h\|_H^2, \quad f \in C([0,T],\mathbb{R}^n), \tag{1.4}$$

where

$$\mathcal{H} = \left\{h \in C([0,T],\mathbb{R}^n); \|h\|_H^2 := \int_0^T |\dot{h}_t|^2 dt < \infty\right\}$$

and for any $h \in \mathcal{H}$, $g(h) \in C([0,T],\mathbb{R}^n)$ satisfies

$$(g(h))_t = \int_0^t b_1^0((g(h))_s)\mathrm{d}s + \int_0^t \sigma((g(h))_s)\dot{h}_s\mathrm{d}s, \quad t \in [0,T]. \tag{1.5}$$

Remark 1.1. *Under* **(A1)**, *for any $\varepsilon \in (0,1)$, (1.1) has a unique strong solution denoted by $\{(\tilde{X}_t^\varepsilon)_{t\in[0,T]}\}$. Furthermore,* **(A1)** *also implies that for any $h \in \mathcal{H}$, $g(h)$ defined above is the unique solution to the following deterministic differential equation:*

$$\mathrm{d}Z_t = b_1^0(Z_t)\mathrm{d}t + \sigma(Z_t)\dot{h}_t\mathrm{d}t, \quad t \in [0,T], Z_0 = 0. \tag{1.6}$$

Lemma 1.2. *Under* **(A1)**, *the family $\{(\tilde{X}_t^\varepsilon)_{t\in[0,T]}\}_{\varepsilon\in(0,1)}$ obeys an LDP on $C([0,T];\mathbb{R}^n)$ with the speed function ε^{-1} and the good rate function I given by (1.4).*

The outline of this paper is organized as follows: In Section 2, we study the LDP for non-degenerate SDEs with singular drift; In Section 3, we investigate an LDP for degenerate SDEs with singular drift.

2 LDP for Non-degenerate SDEs

In this section, we add a small singular interruption in (1.1), i.e., consider the following SDE on $\mathbb{R}^n$:

$$\mathrm{d}X_t^\varepsilon = b_1^\varepsilon(X_t^\varepsilon) + \varepsilon b_2(X_t^\varepsilon)\mathrm{d}t + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)\mathrm{d}W_t, \quad t \in [0,T], \quad X_0^\varepsilon = x_0, \tag{2.1}$$

where $\varepsilon, \sigma, b_1^\varepsilon$ and $(W_t)_{t\in[0,T]}$ are introduced in Section 1, and $b_2 : \mathbb{R}^n \to \mathbb{R}^n$ is the singular drift. Without loss of generality, we assume $x_0 = 0$.

To characterize the singularity of b_2, we introduce some definitions which are taken from [4] and [27].

Definition 2.1. (1) An increasing function $\phi : [0,\infty) \to [0,\infty)$ is called a Dini function if

$$\int_0^1 \frac{\phi(s)}{s}\mathrm{d}s < \infty.$$

(2) A function f defined on the Euclidean space is called Dini continuous if

$$|f(x) - f(y)| \le \phi(|x-y|)$$

holds for some Dini function ϕ.

(3) A measurable function $\phi : [0,\infty) \to [0,\infty)$ is called a *slowly varying* function at zero (see [4]) if for any $\delta > 0$,

$$\lim_{t\to 0} \frac{\phi(\delta t)}{\phi(t)} = 1.$$

Let $\mathscr{D}_0$ be the set of all Dini functions, and $\mathscr{T}_0$ the set of all slowly varying functions at zero that are bounded away from 0 and ∞ on $[\varepsilon,\infty)$ for any $\varepsilon > 0$. Notice that the typical examples for functions contained in $\mathscr{D}_0 \cap \mathscr{T}_0$ are $\phi(t) := (\log(1+t^{-1}))^{-\beta}$ for $\beta > 1$.

To obtain the LDP for (2.1), we make the following assumptions.

(A1') Besides **(A1)**, there exists a constant $K > 1$ such that

$$\sup_{\varepsilon\in(0,1)} \|b_1^\varepsilon\|_\infty + \|b_2\|_\infty \le K$$

and

$$K^{-1}I \le \sigma\sigma^* \le KI. \tag{2.2}$$

(A2) There exists $\phi \in \mathscr{D}_0 \cap \mathscr{T}_0$ such that

$$|b_2(x) - b_2(y)| \le \phi(|x-y|), \quad x, y \in \mathbb{R}^n. \tag{2.3}$$

Under **(A1')** and **(A2)**, (2.1) admits a unique non-explosive strong solution $(X_t^\varepsilon)_{t\in[0,T]}$; see, e.g., [27, Corollary 1.5]. In fact, by Zvonkin's transform, we can kill b_2, see (2.8) below for more details.

Our main result is

Theorem 2.1. *Assume* **(A1')-(A2)***, then* $\{(X_t^\varepsilon)_{t\in[0,T]}\}_{\varepsilon\in(0,1)}$ *obeys LDP on* $C([0,T];\mathbb{R}^n)$ *with the speed function* ε^{-1} *and the good rate function* I *given by* (1.4).

Remark 2.2. *Due to the singularity of* b_2*, we need to give stronger condition* **(A1')** *in Theorem 2.1 than* **(A1)** *in Lemma 1.2, see the proof of Theorem 2.1 for more details.*

2.1 Proof of Theorem 2.1

In order to obtain the LDP for (2.1), we adopt Zvonkin type transform to change (2.1) to a new equation with Lipschitz continuous coefficients, where the Freidlin-Wentzell theorem ([10]) can be available. Let $(e_i)_{i\geq 1}$ be an orthogonal basis of $\mathbb{R}^n$. For any $\lambda > 0$, consider the following $\mathbb{R}^n$-valued PDE:

$$\mathscr{L}u_\lambda + b_2 + \nabla_{b_2}u_\lambda = \lambda u_\lambda, \tag{2.4}$$

where

$$\mathscr{L} := \frac{1}{2}\sum_{i,j=1}^n \langle(\sigma\sigma^*)e_i, e_j\rangle\nabla_{e_i}\nabla_{e_j}.$$

By [27, Theorem 3.10] with $d_1 = 0$, $d_2 = n$, there exists a constant $\lambda_0 > 0$ such that for any $\lambda \geq \lambda_0$, the equation (2.4) has a unique solution u_λ satisfying

$$\|u_\lambda\|_\infty + \|\nabla u_\lambda\|_\infty + \|\nabla^2 u_\lambda\|_\infty \leq \frac{1}{2}. \tag{2.5}$$

For any $\lambda \geq \lambda_0$, let $\theta_\lambda : \mathbb{R}^n \to \mathbb{R}^n$ be defined by $\theta_\lambda(x) := x + u_\lambda(x), x \in \mathbb{R}^n$. By (2.5), θ_λ is a homeomorphism on $\mathbb{R}^n$. Let θ_λ^{-1} denote the inverse of θ_λ, then it holds that $\nabla\theta_\lambda^{-1} = (\nabla\theta_\lambda)^{-1}$.

We are now in a position to complete the Proof of Theorem 2.1.

Proof of Theorem 2.1. Throughout the whole proof, we assume $\lambda \geq \lambda_0$. Since

$$\mathrm{d}X_t^\varepsilon = b_1^\varepsilon(X_t^\varepsilon) + \varepsilon b_2(X_t^\varepsilon)\mathrm{d}t + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)\mathrm{d}W_t, \quad t \in [0,T], \quad X_0^\varepsilon = x_0, \tag{2.6}$$

applying Itô's formula to $\theta_\lambda(X_t^\varepsilon)$, we deduce from (2.4) that

$$\mathrm{d}\theta_\lambda(X_t^\varepsilon) = \varepsilon\lambda u_\lambda(X_t^\varepsilon)\mathrm{d}t + (\nabla\theta_\lambda b_1^\varepsilon)(X_t^\varepsilon)\mathrm{d}t + \sqrt{\varepsilon}(\nabla\theta_\lambda\sigma)(X_t^\varepsilon)\mathrm{d}W_t, \quad t \in [0,T]. \tag{2.7}$$

Denote $Y_t^\varepsilon := \theta_\lambda(X_t^\varepsilon)$, then (2.7) becomes

$$\begin{aligned}\mathrm{d}Y_t^\varepsilon &= \varepsilon\lambda u_\lambda(\theta_\lambda^{-1}(Y_t^\varepsilon))\mathrm{d}t + (\nabla\theta_\lambda b_1^\varepsilon)(\theta_\lambda^{-1}(Y_t^\varepsilon))\mathrm{d}t + \sqrt{\varepsilon}(\nabla\theta_\lambda\sigma)(\theta_\lambda^{-1}(Y_t^\varepsilon))\mathrm{d}W_t \\ &=: \tilde{b}^\varepsilon(Y_t^\varepsilon)\mathrm{d}t + \sqrt{\varepsilon}\tilde{\sigma}(Y_t^\varepsilon)\mathrm{d}W_t, \quad t \in [0,T], \quad Y_0^\varepsilon = \theta_\lambda(x_0),\end{aligned} \tag{2.8}$$

where

$$\tilde{b}^\varepsilon(x) = \varepsilon\lambda u_\lambda(\theta_\lambda^{-1}(x)) + (\nabla\theta_\lambda b_1^\varepsilon)(\theta_\lambda^{-1}(x)), \quad \tilde{\sigma}(x) = (\nabla\theta_\lambda\sigma)(\theta_\lambda^{-1}(x)), \quad x \in \mathbb{R}^n.$$

Since θ_λ is a diffeomorphic operator, by **(A1')** and (2.5), $\tilde{b}^\varepsilon$ and $\tilde{\sigma}$ satisfy the following conditions:

(1) for some constant $\tilde{K} > 1$, we have

$$\|\tilde{\sigma}(x) - \tilde{\sigma}(y)\| + |\tilde{b}^\varepsilon(x) - \tilde{b}^\varepsilon(y)| \le \tilde{K}|x-y|, \quad x, y \in \mathbb{R}^n.$$

(2) Let $\tilde{b}^0 := (\nabla\theta_\lambda b_1^0) \circ \theta_\lambda^{-1}$, then

$$\lim_{\varepsilon\to 0} \|\tilde{b}^\varepsilon - \tilde{b}^0\|_\infty = 0.$$

By Lemma 1.2, $\{Y_t^\varepsilon, t \in [0,T]\}_{\varepsilon\in(0,1)}$ satisfies the LDP in $C([0,T],\mathbb{R}^n)$ with the speed function ε^{-1} and the good rate function given by

$$I^Y(f) := \frac{1}{2} \inf_{f=g^Y(h), h\in\mathcal{H}} \|h\|_H^2 \tag{2.9}$$

with

$$(g^Y(h))_t = \int_0^t \tilde{b}^0((g^Y(h))_s)\mathrm{d}s + \int_0^t \tilde{\sigma}((g^Y(h))_s)\dot{h}_s\mathrm{d}s, \quad t \in [0,T].$$

This implies that

(i) for any constant $c > 0$, the level set $\{f; I^Y(f) \le c\}$ is compact in $C([0,T];\mathbb{R}^n)$;

(ii) for any closed subset $F \subset C([0,T];\mathbb{R}^n)$,

$$\limsup_{\varepsilon\to 0^+} \varepsilon \log \mathbb{P}(Y^\varepsilon \in F) \le -\inf_{f\in F} I^Y(f);$$

(iii) for any open subset $G \subset C([0,T];\mathbb{R}^n)$,

$$\liminf_{\varepsilon\to 0^+} \varepsilon \log \mathbb{P}(Y^\varepsilon \in G) \ge -\inf_{f\in G} I^Y(f).$$

Define

$$I^X(f) := \frac{1}{2} \inf_{f=g^X(h), h\in\mathcal{H}} \|h\|_H^2 \tag{2.10}$$

with

$$(g^X(h))_t = \int_0^t b_1^0((g^X(h))_s)\mathrm{d}s + \int_0^t \sigma((g^X(h))_s)\dot{h}_s\mathrm{d}s, \quad t \in [0,T].$$

In the following, we will prove that $\{X_t^\varepsilon, t \in [0,T]\}_{\varepsilon\in(0,1)}$ satisfies the LDP in $C([0,T],\mathbb{R}^n)$ with the speed function ε^{-1} and the good rate function I^X. This will be completed in Lemma 2.3. □

Lemma 2.3. *Assume* **(A1')** *and* **(A2)**, *then* $\{X^\varepsilon_t, t \in [0,T]\}_{\varepsilon\in(0,1)}$ *satisfies the LDP in* $C([0,T],\mathbb{R}^n)$ *with the speed function* ε^{-1} *and the good rate function* I^X.

Proof. We only need to prove that (i)-(iii) hold with notation Y replaced by notation X. For any $\lambda \geq \lambda_0$, define Θ_λ on $C([0,T];\mathbb{R}^n)$ as

$$(\Theta_\lambda(\xi))_t = \theta_\lambda(\xi_t), \quad t \in [0,T], \xi \in C([0,T];\mathbb{R}^n).$$

For any $\xi, \tilde{\xi} \in C([0,T];\mathbb{R}^n)$ and $s,t \in [0,T]$,

$$|(\Theta_\lambda(\xi))_t - (\Theta_\lambda(\tilde{\xi}))_s| = \theta_\lambda(\xi_t) - \theta_\lambda(\tilde{\xi}_s) \leq \|\nabla\theta_\lambda\|_\infty |\xi_t - \tilde{\xi}_s|, \tag{2.11}$$

which means that $\Theta_\lambda(\xi) \in C([0,T];\mathbb{R}^n)$ by taking $\xi = \tilde{\xi}$. Moreover, for any $\xi \in C([0,T];\mathbb{R}^n)$, let $\eta \in C([0,T];\mathbb{R}^n)$ be defined as $\eta_s = \theta_\lambda^{-1}(\xi_s),\ s \in [0,T]$. Then $\Theta_\lambda(\eta) = \xi$. On the other hand, for any $\xi, \bar{\xi} \in C([0,T];\mathbb{R}^n)$ satisfying $\Theta_\lambda(\xi) = \Theta_\lambda(\bar{\xi})$, i.e., $\theta_\lambda(\xi_s) = \theta_\lambda(\bar{\xi}_s), s \in [0,T]$, we have $\xi = \bar{\xi}$. So, Θ_λ is a bijection on $C([0,T];\mathbb{R}^n)$. Moreover, taking $t = s$ in (2.11) implies that Θ_λ is a continuous map. Similarly, Θ_λ^{-1} is also a continuous map. Thus, Θ_λ is a homeomorphism.

(i) We firstly prove that I^X is a good rate function. $I^X = I^Y(\Theta_\lambda(\cdot))$. By chain rule, we have

$$\begin{aligned}
\theta_\lambda((g^X(h))_t) &= \int_0^t [(\nabla\theta_\lambda b_1^0) \circ \theta_\lambda^{-1}](\theta_\lambda((g^X(h))_s))\mathrm{d}s \\
&\quad + \int_0^t [(\nabla\theta_\lambda \sigma) \circ \theta_\lambda^{-1}](\theta_\lambda((g^X(h))_s))\dot{h}_s \mathrm{d}s \\
&= \int_0^t \tilde{b}^0(\theta_\lambda((g^X(h))_s))\mathrm{d}s \\
&\quad + \int_0^t \tilde{\sigma}(\theta_\lambda((g^X(h))_s))\dot{h}_s \mathrm{d}s, \quad t \in [0,T].
\end{aligned}$$

By the uniqueness of solution, we have $\theta_\lambda((g^X(h))_t) = (g^Y(h))_t, t \in [0,T]$, i.e., $\Theta_\lambda(g^X(h)) = g^Y(h)$. Combining the definition of I^X and I^Y, it is easy to see that $I^X = I^Y(\Theta_\lambda(\cdot))$. Thus, for any $c > 0$, $\{f; I^X(f) \leq c\} = \{f; I^Y(\Theta_\lambda(f)) \leq c\} = \Theta_\lambda^{-1}\{f; I^Y(f) \leq c\}$. Since $\{f; I^Y(f) \leq c\}$ is a compact set, and Θ_λ is a homeomorphism, we conclude that $\{f; I^X(f) \leq c\}$ is a compact set.

(ii) For any closed subset $F \subset C([0,T];\mathbb{R}^n)$,

$$\begin{aligned}
&\limsup_{\varepsilon\to0^+} \varepsilon \log \mathbb{P}(X^\varepsilon \in F) \\
&= \limsup_{\varepsilon\to0^+} \varepsilon \log \mathbb{P}(Y^\varepsilon \in \Theta_\lambda(F)) \\
&\leq -\inf_{f\in\Theta_\lambda(F)} I^Y(f) \\
&= -\inf_{f\in F} I^Y(\Theta_\lambda(f)) = -\inf_{f\in F} I^X(f).
\end{aligned}$$

Similarly, for any open subset $G \subset C([0,T];\mathbb{R}^n)$,

$$\liminf_{\varepsilon\to 0^+} \varepsilon \log \mathbb{P}(X^\varepsilon \in G) \geq -\inf_{f\in G} I^X(f).$$

Thus, (iii) holds.

We finish the proof. $\square$

3 LDP for Degenerate SDEs

Consider the following degenerate SDEs on $\mathbb{R}^{d_1+d_2}$:

$$\begin{cases} \mathrm{d}X_t = \bar{b}^\varepsilon(X_t, Y_t)\mathrm{d}t, \\ \mathrm{d}Y_t = \bar{B}^\varepsilon(X_t, Y_t)\mathrm{d}t + \varepsilon b(Y_t)\mathrm{d}t + \sqrt{\varepsilon}\sigma(Y_t)\mathrm{d}W_t, \\ (X_0, Y_0) = (x_0, y_0) \in \mathbb{R}^{d_1+d_2}, \end{cases} \tag{3.1}$$

where $\varepsilon \in (0,1)$, $W = (W_t)_{t\geq 0}$ is a d_2-dimensional standard Brownian motion with respect to a complete filtration probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$, $\bar{b}^\varepsilon : \mathbb{R}^{d_1+d_2} \to \mathbb{R}^{d_1}$, $\bar{B}^\varepsilon : \mathbb{R}^{d_1+d_2} \to \mathbb{R}^{d_2}$, $b : \mathbb{R}^{d_2} \to \mathbb{R}^{d_2}$ and $\sigma : \mathbb{R}^{d_2} \to \mathbb{R}^{d_2} \otimes \mathbb{R}^{d_2}$ are measurable and locally bounded (bounded on bounded sets). Again we assume $(x_0, y_0) = 0$.

Suppose that there exist $\phi \in \mathscr{D}_0 \cap \mathscr{T}_0$ and a constant $K > 1$ such that the following conditions hold.

(H1) $\|\bar{B}^\varepsilon\|_\infty + \|b\|_\infty \leq K$,

$$\|\sigma(y_1) - \sigma(y_2)\| \leq K|y_1 - y_2|, \quad y_1, y_2 \in \mathbb{R}^{d_2}, \tag{3.2}$$

and

$$|\bar{b}^\varepsilon(z_1) - \bar{b}^\varepsilon(z_2)| + |\bar{B}^\varepsilon(z_1) - \bar{B}^\varepsilon(z_2)| \leq K|z_1 - z_2|, \quad z_1, z_2 \in \mathbb{R}^{d_1+d_2}.$$

Moreover,

$$K^{-1} I_{d_2\times d_2} \leq \sigma\sigma^* \leq K I_{d_2\times d_2}. \tag{3.3}$$

(H2) There exist Lipschitz continuous functions $\bar{b}^0 : \mathbb{R}^{d_1+d_2} \to \mathbb{R}^{d_1}$ and $\bar{B}^0 : \mathbb{R}^{d_1+d_2} \to \mathbb{R}^{d_2}$ such that

$$\lim_{\varepsilon\to 0} \left\{\|\bar{b}^\varepsilon - \bar{b}^0\|_\infty\right\} = 0, \tag{3.4}$$

and

$$\lim_{\varepsilon\to 0} \left\{\|\bar{B}^\varepsilon - \bar{B}^0\|_\infty\right\} = 0. \tag{3.5}$$

(H3) (Regularity of b_2)

$$|b(y_1) - b(y_2)| \leq \phi(|y_1 - y_2|), \quad y_1, y_2 \in \mathbb{R}^{d_2}. \tag{3.6}$$

Under **(H1)** and **(H3)**, for any $\varepsilon \in (0,1)$, (3.1) admits a unique non-explosive strong solution $(X_t^\varepsilon, Y_t^\varepsilon)_{t\in[0,T]}$; see, e.g., [27, Theorem 1.1].

Let $C([0,T],\mathbb{R}^{d_2})$ be equipped with sup-norm, and define rate function $I: C([0,T],\mathbb{R}^{d_2}) \to [0,\infty)$ as

$$I(f) = \frac{1}{2}\inf_{f=g(h),h\in\tilde{\mathcal{H}}} \|h\|_{\tilde{H}}^2, \tag{3.7}$$

where

$$\tilde{\mathcal{H}} = \left\{ h \in C([0,T],\mathbb{R}^{d_2}); \|h\|_{\tilde{H}}^2 := \int_0^T |\dot{h}_t|^2 dt < \infty \right\}$$

and for any $h \in \tilde{\mathcal{H}}$, $g(h) \in C([0,T],\mathbb{R}^{d_1+d_2})$ satisfies

$$(g(h))_t = \int_0^t (\bar{b}^0((g(h))_s), \bar{B}^0((g(h))_s)\mathrm{d}s + \int_0^t (0, \sigma((g(h))_s)\dot{h}_s)\mathrm{d}s, \quad t\in[0,T].$$

3.1 Main results

The main result of this section is the following theorem.

Theorem 3.1. *Assume* **(H1)**-**(H3)**. *The family* $\{(X_t^\varepsilon, Y_t^\varepsilon))_{t\in[0,T]}\}_{\varepsilon\in(0,1)}$ *obeys the LDP on* $C([0,T];\mathbb{R}^{d_1+d_2})$ *with the speed function* ε^{-1} *and the good rate function* I *given by* (3.7).

3.2 Proof of Theorem 3.1

Similarly to the proof of Theorem 2.1, let $(e_i)_{i\ge1}$ be an orthogonal basis of $\mathbb{R}^{d_2}$. For any $\lambda > 0$, consider the following $\mathbb{R}^{d_2}$-valued PDE:

$$\tilde{\mathscr{L}}u_\lambda + b + \nabla_b u_\lambda = \lambda u_\lambda, \tag{3.8}$$

where

$$\tilde{\mathscr{L}} := \frac{1}{2}\sum_{i,j=1}^{d_2} \langle (\sigma\sigma^*)e_i, e_j\rangle \nabla_{e_i}\nabla_{e_j}.$$

Then by [27, Theorem 3.10], there exists a constant $\lambda_0 > 0$ such that for any $\lambda \ge \lambda_0$, the equation (3.8) has a unique solution u_λ satisfying

$$\|u_\lambda\|_\infty + \|\nabla u_\lambda\|_\infty + \|\nabla^2 u_\lambda\|_\infty \le \frac{1}{2}. \tag{3.9}$$

For any $\lambda \ge \lambda_0$, let $\theta_\lambda : \mathbb{R}^{d_2} \to \mathbb{R}^{d_2}$ be defined by $\theta_\lambda(x) := x + u_\lambda(x), x \in \mathbb{R}^{d_2}$. By (3.9), θ_λ is a homeomorphism on $\mathbb{R}^{d_2}$. Let θ_λ^{-1} denote the inverse of θ_λ, then it holds that $\nabla\theta_\lambda^{-1} = (\nabla\theta_\lambda)^{-1}$. Throughout the whole proof, we assume $\lambda \ge \lambda_0$. Since

$$\begin{cases} \mathrm{d}X_t^\varepsilon = \bar{b}^\varepsilon(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t, \\ \mathrm{d}Y_t^\varepsilon = \bar{B}^\varepsilon(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + \varepsilon b(Y_t^\varepsilon)\mathrm{d}t + \sqrt{\varepsilon}\sigma(Y_t^\varepsilon)\mathrm{d}W_t, \\ (X_0, Y_0) = (x_0, y_0) \in \mathbb{R}^{d_1+d_2}, \end{cases} \tag{3.10}$$

it follows from Itô's formula and (2.4) that

$$\begin{cases} \mathrm{d}X_t^\varepsilon = \bar{b}^\varepsilon(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t, \\ \mathrm{d}\theta_\lambda(Y_t^\varepsilon) = \varepsilon\lambda u_\lambda(Y_t^\varepsilon)\mathrm{d}t + \nabla\theta_\lambda(Y_t^\varepsilon)\bar{B}^\varepsilon(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + \sqrt{\varepsilon}(\nabla\theta_\lambda\sigma)(Y_t^\varepsilon)\mathrm{d}W_t. \end{cases} \tag{3.11}$$

Denote $\tilde{Y}_t^\varepsilon := \theta_\lambda(Y_t^\varepsilon)$, then (3.11) can be written as

$$\begin{cases} \mathrm{d}X_t^\varepsilon = \tilde{b}^\varepsilon(X_t^\varepsilon, \tilde{Y}_t^\varepsilon)\mathrm{d}t, \\ \mathrm{d}\tilde{Y}_t^\varepsilon = \tilde{B}^\varepsilon(X_t^\varepsilon, \tilde{Y}_t^\varepsilon)\mathrm{d}t + \sqrt{\varepsilon}\tilde{\sigma}(\tilde{Y}_t^\varepsilon)\mathrm{d}W_t, \end{cases} \tag{3.12}$$

where

$$\tilde{B}^\varepsilon(x,y) = \varepsilon\lambda u_\lambda(\theta_\lambda^{-1}(y)) + \nabla\theta_\lambda(\theta_\lambda^{-1}(y))\bar{B}^\varepsilon(x, \theta_\lambda^{-1}(y)),$$

and

$$\tilde{b}^\varepsilon(x,y) = \bar{b}^\varepsilon(x, \theta_\lambda^{-1}(y)), \quad \tilde{\sigma}(y) = (\nabla\theta_\lambda\sigma)(\theta_\lambda^{-1}(y)), \quad (x,y) \in \mathbb{R}^{d_1+d_2}.$$

Since θ_λ is a diffeomorphic operator, by **(H1)**, **(H2)** and (3.9), $\tilde{B}^\varepsilon, \tilde{b}^\varepsilon$ and $\tilde{\sigma}$ satisfy the following conditions:

(1) There exists a constant $\tilde{K} > 1$ such that for any $z_1 = (x_1, y_1), z_2 = (x_2, y_2) \in \mathbb{R}^{d_1+d_2}$,

$$\|\tilde{\sigma}(y_1) - \tilde{\sigma}(y_2)\| + |\tilde{b}^\varepsilon(x_1,y_1) - \tilde{b}^\varepsilon(x_2,y_2)| + |\tilde{B}^\varepsilon(x_1,y_1) - \tilde{B}^\varepsilon(x_2,y_2)| \le \tilde{K}|z_1 - z_2|.$$

(2) Let $\tilde{b}^0(x,y) = \bar{b}^0(x, \theta_\lambda^{-1}(y))$ and $\tilde{B}^0(x,y) := \nabla\theta_\lambda(\theta_\lambda^{-1}(y))\bar{B}^0(x, \theta_\lambda^{-1}(y))$, $(x,y) \in \mathbb{R}^{d_1+d_2}$, then it holds that

$$\lim_{\varepsilon\to 0}\left\{\|\tilde{b}^\varepsilon - \tilde{b}^0\|_\infty\right\} = 0,$$

and

$$\lim_{\varepsilon\to 0}\left\{\|\tilde{B}^\varepsilon - \tilde{B}^0\|_\infty\right\} = 0.$$

Again by Lemma 1.2, $\{(X_t^\varepsilon, \tilde{Y}_t^\varepsilon), t \in [0,T]\}_{\varepsilon\in(0,1)}$ satisfies the LDP in $C([0,T], \mathbb{R}^{d_1+d_2})$ with the speed function ε^{-1} and the good rate function $\tilde{I}$ given by

$$\tilde{I}(f) := \frac{1}{2}\inf_{f=\tilde{g}(h), h\in\tilde{\mathcal{H}}}\|h\|_{\tilde{H}}^2 \tag{3.13}$$

with

$$(\tilde{g}(h))_t = \int_0^t (\tilde{b}^0((\tilde{g}(h))_s), \tilde{B}^0((\tilde{g}(h))_s))\mathrm{d}s + \int_0^t (0, \tilde{\sigma}((\tilde{g}(h))_s)\dot{h}_s\mathrm{d}s), \quad t \in [0,T].$$

This implies that

(i') for any constant $c > 0$, the level set $\{f; \tilde{I}(f) \le c\}$ is compact in $C([0,T]; \mathbb{R}^{d_1+d_2})$;

(ii') for any closed subset $F \subset C([0,T];\mathbb{R}^{d_1+d_2})$,

$$\limsup_{\varepsilon\to 0^+} \varepsilon \log \mathbb{P}((X^\varepsilon, \tilde{Y}^\varepsilon) \in F) \le -\inf_{f\in F} \tilde{I}(f);$$

(iii') for any open subset $G \subset C([0,T];\mathbb{R}^{d_1+d_2})$,

$$\liminf_{\varepsilon\to 0^+} \varepsilon \log \mathbb{P}((X^\varepsilon, \tilde{Y}^\varepsilon) \in G) \ge -\inf_{f\in G} \tilde{I}(f).$$

Next, we will prove that $\{(X_t^\varepsilon, Y_t^\varepsilon), t \in [0,T]\}_{\varepsilon\in(0,1)}$ satisfies the LDP in $C([0,T],\mathbb{R}^{d_1+d_2})$ with the speed function ε^{-1} and the good rate function I defined by

$$I(f) := \frac{1}{2} \inf_{f=g(h), h\in\tilde{\mathcal{H}}} \|h\|_{\tilde{H}}^2 \tag{3.14}$$

with

$$(g(h))_t = \int_0^t (\bar{b}^0((g(h))_s), \bar{B}^0((g(h))_s))\mathrm{d}s + \int_0^t (0, \sigma((g(h))_s)\dot{h}_s \mathrm{d}s), \quad t \in [0,T].$$

This will be completed in Lemma 3.2.

Lemma 3.2. *Assume* **(H1)-(H3)**, *then* $\{(X_t^\varepsilon, Y_t^\varepsilon), t \in [0,T]\}_{\varepsilon\in(0,1)}$ *satisfies the LDP in* $C([0,T],\mathbb{R}^{d_1+d_2})$ *with the speed function* ε^{-1} *and the good rate function* I *given in* (3.14).

Proof. We only need to prove that (i')-(iii') hold with $\tilde{Y}$ replaced by Y and the good rate function $\tilde{I}$ replaced by I. For any $\lambda \ge \lambda_0$, $\xi = (\xi^1, \xi^2) \in C([0,T];\mathbb{R}^{d_1+d_2})$, let

$$(\Theta_\lambda(\xi))_t = (\xi_t^1, \theta_\lambda(\xi_t^2)), \quad t \in [0,T].$$

Then it is easy to see that Θ_λ is a homeomorphism on $C([0,T];\mathbb{R}^{d_1+d_2})$. In fact, for any $\xi \in C([0,T];\mathbb{R}^{d_1+d_2})$,

$$|(\Theta_\lambda(\xi))_t - (\Theta_\lambda(\xi))_s| \le (\|\nabla\theta_\lambda\|_\infty \vee 1)|\xi_t - \xi_s|,$$

which means $\Theta_\lambda(\xi) \in C([0,T];\mathbb{R}^{d_1+d_2})$. Moreover, for any $\xi \in C([0,T];\mathbb{R}^{d_1+d_2})$, let $\eta \in C([0,T];\mathbb{R}^{d_1+d_2})$ be defined as $\eta_s = (\xi_s^1, \theta_\lambda^{-1}(\xi_s^2))$, $s \in [0,T]$. Then $\Theta_\lambda(\eta) = \xi$. On the other hand, for any $\xi, \bar{\xi} \in C([0,T];\mathbb{R}^{d_1+d_2})$ satisfying $\Theta_\lambda(\xi) = \Theta_\lambda(\bar{\xi})$, i.e., $\xi_s^1 = \bar{\xi}_s^1$ and $\theta_\lambda(\xi_s^2) = \theta_\lambda(\bar{\xi}_s^2), s \in [0,T]$, we have $\xi = \bar{\xi}$. So, Θ_λ is a bijection. Moreover, for any $\xi, \bar{\xi} \in C([0,T];\mathbb{R}^{d_1+d_2})$, we have

$$\|\Theta_\lambda(\xi) - \Theta_\lambda(\bar{\xi})\|_\infty \le (\|\nabla\theta_\lambda\|_\infty \vee 1) \sup_{t\in[0,T]} |\xi_t - \bar{\xi}_t| = (\|\nabla\theta_\lambda\|_\infty \vee 1)\|\xi - \bar{\xi}\|_\infty,$$

which means that Θ_λ is a continuous map. Similarly, Θ_λ^{-1} is also a continuous map. Thus, Θ_λ is a homeomorphism on $C([0,T];\mathbb{R}^{d_1+d_2})$.

(i') We firstly prove that $I = \tilde{I}(\Theta_\lambda(\cdot))$. By chain rule and the definition of $\tilde{B}^0, \tilde{b}^0$, $\tilde{\sigma}$ and Θ_λ, it is not difficult to see that

$$\begin{aligned}(\Theta_\lambda(g(h)))_t = &\int_0^t (\tilde{b}^0((\Theta_\lambda(g(h)))_s), \tilde{B}^0((\Theta_\lambda(g(h)))_s))\mathrm{d}s \\ &+ \int_0^t (0, \tilde{\sigma}((\Theta_\lambda(g(h)))_s)\dot{h}_s \mathrm{d}s), \quad t \in [0,T].\end{aligned}$$

By the uniqueness of solution, we have $\Theta_\lambda(g(h)) = \tilde{g}(h)$. Combining the definition of I and $\tilde{I}$, we arrive at $I = \tilde{I}(\Theta_\lambda(\cdot))$. Thus, for any $c > 0$, $\{f; I(f) \le c\} = \{f; \tilde{I}(\Theta_\lambda(f)) \le c\} = \Theta_\lambda^{-1}\{f; \tilde{I}(f) \le c\}$. Since $\{f; \tilde{I}(f) \le c\}$ is a compact set and Θ_λ is a homeomorphism, we conclude that $\{f; I(f) \le c\}$ is a compact set.

(ii') for any closed subset $F \subset C([0,T]; \mathbb{R}^{d_1+d_2})$,

$$\begin{aligned}&\limsup_{\varepsilon \to 0^+} \varepsilon \log \mathbb{P}((X^\varepsilon, Y^\varepsilon) \in F) \\ &= \limsup_{\varepsilon \to 0^+} \varepsilon \log \mathbb{P}((X^\varepsilon, \tilde{Y}^\varepsilon) \in \Theta_\lambda(F)) \\ &\le - \inf_{f \in \Theta_\lambda(F)} \tilde{I}(f) \\ &= - \inf_{f \in F} \tilde{I}(\Theta_\lambda(f)) = - \inf_{f \in F} I(f).\end{aligned}$$

Similarly, for any open subset $G \subset C([0,T]; \mathbb{R}^{d_1+d_2})$,

$$\liminf_{\varepsilon \to 0^+} \varepsilon \log \mathbb{P}((X^\varepsilon, Y^\varepsilon) \in G) \ge - \inf_{f \in G} I(f).$$

Thus, (iii') holds.

We finish the proof. □

Remark 3.3. *By [16, Lemma 3.2], we know that* (2.5) *and* (3.9) *also hold if we assume* **(A2)** *and* **(H3)** *for $\phi(x) = x^\alpha$ with $\alpha \in (0,1)$. Thus, the assertions in Theorem 2.1 and Theorem 3.1 still hold by replacing* (2.3) *and* (3.6) *with $\phi(x) = x^\alpha$ for some $\alpha \in (0,1)$.*

Acknowledgement. The authors would like to thank Professor Feng-Yu Wang for helpful comments.

References

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

[2] A. Budhiraja, P. Dupuis, Large deviations for stochastic partial differential equations driven by a Poisson random measure, *Stochastic Processes* & *Their Applications*, **123**(2013), 523–560.

[3] A. Budhiraja, P. Dupuis, V. Maroulas, Large deviations for infinite dimensional stochastic dynamical systems continuous time processes, *Annals of Probability*, **36**(2008), 1390–1420.

[4] N. H. Bingham, C. M. Goldie, J. L. Teugels, *Regular variation,* Cambridge University Press, Cambridge, 1987.

[5] S. Cerrai, M. Freidlin, Mark Large deviations for the Langevin equation with strong damping. *J. Stat. Phys.*, **4**(2015), 859–875.

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

[7] P. L. Chow, Large deviation problem for some parabolic Itô equations, *Commun. Pure Appl. Math.*, **45**(1992), 97–120.

[8] S. Fang, T. Zhang, A study of a class of stochastic differential equations with non-Lipschitzian coefficients, *Probability Theory and Related Fields*, **132**(2005), 356-390.

[9] E. Fedrizzi, F. Flandoli, Pathwise uniqueness and continuous dependence for SDEs with non-irregular drift, *Stochastics* **83**(2011), 241–257.

[10] M. I. Freidlin, A. D. Wentzell, Random Perturbations of Dynamical Systems, *New York: Springer Science and Business Media*, vol. 260, 1984.

[11] F. Gao, S. Wang, Asymptotic behaviors for functionals of random dynamical systems, *Stoch. Anal. Appl.*, **34**(2016), 258-277.

[12] I. Gyöngy, T. Martinez, On stochastic differential equations with locally unbounded drift, *Czechoslovak Math.J.*, **51**(2001), 763–783.

[13] N. V. Krylov, M. Röckner, Strong solutions of stochastic equations with singular time dependent drift, *Probability Theory and Related Fields*, **131**(2005), 154-196.

[14] A. Kulik, D. Sobolieva, Large deviation principle for one-dimensional SDEs with discontinuous coefficients, *Modern Stochastics: Theory and Applications*, **3**(2016), 145-164.

[15] G. Lan, Large deviation principle of stochastic differential equations with non-Lipschitzian coefficients, *Frontiers of Mathematics in China*, **8**(2013), 1307-1321.

[16] H. Li, D. Luo, J. Wang, Harnack inequalities for SDEs with multiplicative noise and non-regular drift, *Stoch. Dyn.* **15**(2015), 18pp.

[17] W. Liu, Large deviations for stochastic evolution equations with small multiplicative noise, *App. Math. Opt.*, **61**(2010), 27–56.

[18] YanLv and A.J. Roberts, Large Deviation Principle for Singularly Perturbed Stochastic Damped Wave Equations, *Stochastic Analysis* & *Applications*, **32**(2014), 50–60.

[19] U. Manna, S. S. Sritharan, P. Sundar, Large deviations for the stochastic shell model of turbulence, *Nonlinear Differential Equations and Applications*, **16**(2009), 493–521.

[20] S. Peszat, Large deviation principle for stochastic evolution equations, *Probability Theory and Related Fields*, **98**(1994), 113–136.

[21] P. E. C. D. Raynal, Strong existence and uniqueness for stochastic differential equation with Hölder drift and degenerate noise, to appear in Ann. Inst. Henri Poincaré Probab. Stat..

[22] J. Ren, X. Zhang, Freidlin-Wentzell large deviations for homeomorphism flows of non-Lipschitz SDE, *Bull. Sci.*, **129**(2005), 643–655.

[23] M. Röckner, F.-Y. Wang, L. Wu, Large deviations for stochastic Generalized Porous Media Equations, *Stoch. Proc. Appl.*, **116**(2006), 1677–1689.

[24] M. Röckner, T. Zhang, X. Zhang, Large deviations for stochastic tamed 3D Navier- Stokes equations, *App. Math. Opt.*, **61**(2010), 267–285.

[25] C. Rovira, M. Sanz-solé, Large deviations for stochastic Volterra equations in the plane, *Potential Anal.*, **12**(2000), 359–383.

[26] F.-Y. Wang, Gradient estimates and applications for SDEs in Hilbert space with multiplicative noise and Dini continuous drift, *J. Differential Equations*, **260**(2016), 2792–2829.

[27] F.-Y. Wang, X. Zhang, Degenerate SDE with Hölder-Dini Drift and Non-Lipschitz Noise Coefficient, *SIAM J. Math. Anal.*, **48**(2016), 2189–2226.

[28] T. Xu, T. Zhang, White noise driven SPDEs with reflection: Existence, uniqueness and large deviation principles, *Stochastic Processes & Their Applications*, **119**(2017), 3453–3470.

[29] X. Zhang, Strong solutions of SDEs with singural drift and Sobolev diffusion coefficients, *Stoch. Proc. Appl.*, **115**(2005), 1805–1818.

[30] X. Zhang, Stochastic Volterra equations in Banach spaces and stochastic partial differential equation, *Journal of Functional Analysis*, **258**(2017), 1361–1425.

[31] A. K. Zvonkin, A transformation of the phase space of a diffusion process that removes the drift, *Math. Sb.* , **93**(1974), 129–149.