# Large Deviations for Stochastic Evolution Equations with Small Multiplicative Noise *

**Wei Liu** †

Fakultät Für Mathematik, Universität Bielefeld, D-33501 Bielefeld, Germany

School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China

**Abstract**

The Freidlin-Wentzell large deviation principle is established for the distributions of stochastic evolution equations with general monotone drift and small multiplicative noise. Roughly speaking, besides the assumptions for existence and uniqueness of the solution, one only need assume some additional assumptions on diffusion coefficient in order to obtain Large deviation principle for the distribution of solution. As applications we can apply the main result to different type examples of SPDEs (e.g. stochastic reaction-diffusion equation, stochastic porous media and fast diffusion equations, stochastic p-Laplacian equation) in Hilbert space. The weak convergence approach is employed to verify the Laplace principle, which is equivalent to large deviation principle in our framework.

AMS subject Classification: 60F10, 60H15.
Keywords: Stochastic evolution equation, large deviation principle, Laplace principle, variational approach, weak convergence approach, reaction-diffusion equation, porous media equation, fast diffusion equation, p-Laplacian equation.

## 1 Introduction

There basically exist three different approaches to analyze stochastic partial differential equations(SPDE) in literature. The "martingale measure approach" initiated by J. Walsh

*Supported in part by the DFG through the Internationales Graduiertenkolleg "Stochastics and Real World Models" and NNSFC(10121101).

†E-mail: weiliu0402@yahoo.com.cn

in [41]. The “variational approach” was first used by Bensoussan and Temam in [3, 4] to study SPDE, then this approach was further developed in the works of Pardoux [26], Krylov and Rozovoskii [24] for more general case. Concerning the “semigroup (or mild solution) approach” we can refer to the classical book by Da Prato and Zabcyzk [15]. In this paper we will use the variational approach because we mainly treat the nonlinear SPDEs of evolutionary type. All kinds of dynamics with stochastic influence in nature or man-made complex systems can be modeled by such equations. This type of SPDEs have been studied intensively also in recent years, we refer to [5, 12, 13, 14, 23, 29, 32, 33, 42, 45] for many different generalizations and applications.

Concerning the large deviation principle(LDP), there also exist fruitful results under the different framework for SPDE. The LDP was first formulated by Varadhan [38] in 1966. For its validity to stochastic differential equations in the finite dimensional case we mainly refer to the well known Freidlin-Wentzell's LDP ([22]). And the same problem was also treated by Varadhan in [40] and Stroock in [37] by a different approach, which followed the large deviation theory developed by Azencott [2], Donsker-Varadhan [16] and Varadhan [38]. In the classical paper [21] Freidlin studied large deviations for the small noise limit of stochastic reaction-diffusion equations. Subsequently, many authors have obtained Large deviations results under less and less restrictive conditions. For the extensions to infinite dimensional diffusions or stochastic PDE under global Lipschitz condition on the nonlinear term, we refer the reader to Da Prato and Zabczyk [15] and Peszat [27] (also see the references therein). For the case of local Lipschitz conditions we refer to [10] where also multiplicative and degenerate noise is handled. The LDP for semilinear parabolic equations on a Gelfand triple was studied by Chow in [11]. Recently, Röckner, Wang and Wu obtained the LDP in [35] for the distributions of stochastic porous media equations. All these papers mainly used the classical ideas of discretization approximations and Contraction principle which first developed by Wentzell and Freidlin. But the situation became much dispersive in infinite dimensional case since each type of stochastic nonlinear PDE needs different specific techniques and estimates. Another alternative approach for LDP has been developed in [20], which mainly used nonlinear semigroup theory and infinite dimensional Hamilton-Jacobi equation.

In this paper we will study the Large deviation principle for stochastic evolution equations with general monotone drift and multiplicative noise, which is more general and complicated than semilinear case studied in [11] and additive noise case in [35]. It's quite difficult here to follow the classical discretization approach that we mentioned above. The reason is many technical difficulties appear since the coefficients of SPDE in our framework live on a Gelfand triple(three different spaces are involved). So we would use the stochastic control and weak convergence approach in this paper, which is mainly based on a variational representation for certain functionals of infinite dimensional Brownian Motion in [7]. The main advantage of this approach is one can avoid all exponential probability estimates which are very difficult to derive for infinite dimensional models(especially for our case). This approach is used to obtain the Large deviations for homeomorphism flows of non-Lipschitz SDEs in [30], for two-dimensional stochastic Navier-Stokes equations in [36] and reaction-diffusion type SPDEs in [8]. For more references on this approach we refer to [17, 31].

We first recall some standard definitions and results from the large deviation theory. Let $\{X^\varepsilon\}$ be a family of random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and taking values in some Polish space $E$. The large deviations theory is mainly concerned with events $A$ for which probabilities $\mathbf{P}(X^\varepsilon \in A)$ converge to zero exponentially fast as $\varepsilon \to 0$. The rate of such exponential decay is expressed by the "rate function".

**Definition 1.1.** (Rate function) A function $I : E \to [0, +\infty]$ is called a rate function if $I$ is lower semicontinuous. A rate function $I$ is called a good rate function if the level set $\{x \in E : I(x) \le K\}$ is compact for all $K < \infty$.

**Definition 1.2.** (Large deviation principle) The sequence $\{X^\varepsilon\}$ is said to satisfy the *large deviation principle with rate function* $I$ if for each Borel subset $A$ of $E$

$$-\inf_{x\in A^o} I(x) \le \liminf_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{P}(X^\varepsilon \in A) \le \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{P}(X^\varepsilon \in A) \le -\inf_{x\in \bar{A}} I(x),$$

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

If one is interested in obtaining the exponential estimates on general functions instead of indicator functions of Borel subsets of $E$ above, then one can study the following Laplace principle.

**Definition 1.3.** (Laplace principle) The sequence $\{X^\varepsilon\}$ is said to satisfy the *Laplace principle (LP in short) with rate function* $I$ if for each real-valued, bounded continuous function $h$ defined on $E$,

$$\lim_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{E} \left\{ \exp\left[ -\frac{1}{\varepsilon^2} h(X^\varepsilon) \right] \right\} = -\inf_{x\in E} \{h(x) + I(x)\}.$$

The starting point for the weak convergence approach is the equivalence of LDP and LP if $E$ is Polish space and the rate function is good, which was first formulated in [28]. This equivalence is essentially a consequence of Varadhan's Lemma [38] and Bryc's converse to Varadhan's Lemma [9]. We refer to [17, 18] for an elementary proof. In view of this equivalence, we only need to concern with the study of the Laplace principle.

Let $\{W_t\}_{t\ge 0}$ is a cylindrical Wiener process on a separable Hilbert space $U$ w.r.t a complete filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ (i.e. the path of $W$ take values in $C([0,T]; U_1)$, where $U_1$ is another Hilbert space such that the embedding $U \subset U_1$ is Hilbert-Schmidt). Suppose $g^\varepsilon : C([0,T]; U_1) \to E$ is a measurable map and $X^\varepsilon = g^\varepsilon(W_\cdot)$. Let

$$\mathcal{A} = \left\{ v : v \text{ is } U\text{-valued } \mathcal{F}_t\text{-predictable process s.t. } \int_0^T \|v_s(\omega)\|_U^2 \mathrm{d}s < \infty \ a.s. \right\},$$

$$S_N = \left\{ \phi \in L^2([0,T],U) : \int_0^T \|\phi_s\|_U^2 \mathrm{d}s \le N \right\}.$$

The set $S_N$ endowed with the weak topology is a Polish space (we will always refer to weak topology on $S_N$ in this paper if we don't state explicitly). Define

$$\mathcal{A}_N = \{v \in \mathcal{A} : v(\omega) \in S_N, \mathbf{P} - a.s.\}.$$

Then the crucial step in the proof of Laplace principle is based on the following variational representation for certain functionals of Brownian motion obtained in [7].

$$-\log \mathbf{E} \exp\{-f(W)\} = \inf_{v \in \mathcal{A}} \mathbf{E} \left( \frac{1}{2} \int_0^T \|v_s\|_U^2 \mathrm{d}s + f\left( W + \int_0^\cdot v_s \mathrm{d}s \right) \right) \tag{1.1}$$

where $f$ is any bounded Borel measurable function from $C([0,T];U_1)$ to $\mathbf{R}$. The connection between exponential functionals and variational representations appeared to be first exploited by Fleming in [19]. The formula (1.1) for finite dimensional Brownian motion case is first obtained in [6]. Now we can formulate the following sufficient condition of the Laplace principle (equivalently, Large deviation principle) in [7] for $X^\varepsilon$ as $\varepsilon \to 0$.

**(A)** There exists a measurable map

$$g^0 : C([0,T];U_1) \to E$$

such that following two conditions hold:

(i) Let $\{v^\varepsilon : \varepsilon > 0\} \subset \mathcal{A}_N$ for some $N < \infty$. If $v^\varepsilon$ converge to $v$ in distribution as $S_N$-valued random elements, then

$$g^\varepsilon \left( W_\cdot + \frac{1}{\varepsilon} \int_0^\cdot v_s^\varepsilon \mathrm{d}s \right) \to g^0 \left( \int_0^\cdot v_s \mathrm{d}s \right)$$

in distribution as $\varepsilon \to 0$.

(ii) For each $N < \infty$, the set

$$K_N = \left\{ g^0 \left( \int_0^\cdot \phi_s \mathrm{d}s \right) : \phi \in S_N \right\}$$

is a compact subset of $E$.

In this case, for each $f \in E$ we define

$$I(f) = \inf_{\{\phi \in L^2([0,T];U):\ f = g^0(\int_0^\cdot \phi_s \mathrm{d}s)\}} \left\{ \frac{1}{2} \int_0^T \|\phi(s)\|_U^2 \mathrm{d}s \right\} \tag{1.2}$$

where the infimum over an empty set is taken as $\infty$.

**Lemma 1.1.** *[7, Theorem 4.4] If $\{g^\varepsilon\}$ satisfies $(\mathbf{A})$, then the family $\{X^\varepsilon\}$ satisfies the Laplace principle (hence Large deviation principle) on $E$ with the good rate function $I$ given by (1.2).*

## 2 Main framework and result

Suppose

$$V \subset H \equiv H^* \subset V^*$$

is a Gelfand triple, i.e. $V$ is a reflexive and separable Banach space, $H$ is a separable Hilbert space and identified with it's dual space by Riesz isomorphism, $V$ is continuously and densely embedded in $H$. The inner product in $H$ is denoted by $\langle \cdot, \cdot \rangle_H$, the dual between $V^*$ and $V$ is denoted by ${}_{V^*}\langle \cdot, \cdot \rangle_V$. It's obvious that

$${}_{V^*}\langle u, v \rangle_V = \langle u, v \rangle_H, \ u \in H, v \in V.$$

$\{W_t\}_{t \geq 0}$ is a cylindrical Wiener process on a separable Hilbert space $U$ w.r.t a complete filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$. $(L_2(U; H), \|\cdot\|_2)$ and $L(U; H)$ denote the space of all Hilbert-Schmidt operators and bounded linear operators from $U$ to $H$ respectively.

Consider the following stochastic evolution equation

$$\mathrm{d}X_t = A(t, X_t)\mathrm{d}t + B(t, X_t)\mathrm{d}W_t, \tag{2.1}$$

where $A : [0, T] \times V \to V^*$ and $B : [0, T] \times V \to L_2(U; H)$ are measurable. In order to obtain Large deviation principle, we assume the following conditions which are slightly stronger than those in [24] for the existence and uniqueness of strong solution to (2.1).

For a fixed $\alpha > 1$, there exist constants $\delta > 0$ and $K$ such that the following conditions hold all $v, v_1, v_2 \in V$ and $t \in [0, T]$.

$(A1)$ Semicontinuity of $A$: The map

$$\mathbb{R} \ni s \mapsto {}_{V^*}\langle A(t, v_1 + sv_2), v \rangle_V$$

is continuous.

$(A2)$ Strong Monotonicity of $(A, B)$:

$$2{}_{V^*}\langle A(t, v_1) - A(t, v_2), v_1 - v_2 \rangle_V + \|B(t, v_1) - B(t, v_2)\|_2^2 \leq -\delta \|v_1 - v_2\|_V^\alpha + K\|v_1 - v_2\|_H^2.$$

$(A3)$ Boundedness of growth $A$:

$$\|A(t, v)\|_{V^*}^{\frac{\alpha}{\alpha-1}} \leq K(1 + \|v\|_V^\alpha).$$

$(A4)$ Local Lipschitz and growth property of $B$: There exists constant $K_N$

$$\|B(t, v_1) - B(t, v_2)\|_2^2 \leq K_N \|v_1 - v_2\|_H^2, \ \|v_i\|_H \leq N, \ i = 1, 2.$$
$$\|B(t, v)\|_2^2 \leq K(1 + \|v\|_H^2).$$

If $1 < \alpha < 2$, we assume also

$$\|B(t, v)\|_{L(U, V^*)} \leq K(1 + \|v\|_V^{\alpha-1}).$$

$(A5)$ Approximated property of $B$: Suppose $B(\cdot, 0)$ is continuous on $[0, T]$ and there exist

$$H_n \subseteq H_{n+1},\ H_n \hookrightarrow V \text{compact},\ \bigcup_{n=1}^{\infty} H_n \subseteq H \text{ dense}$$

such that for any $M > 0$

$$\sup_{(t,v)\in[0,T]\times S_M} \|P_n B(t,v) - B(t,v)\|_2 \to 0\ (n \to \infty) \tag{2.2}$$

where $P_n : H \to H_n$ is standard projection operator and

$$S_M = \{v \in V : \|v\|_H \leq M\}.$$

*Remark* 2.1. (i) Notice $(A1) - (A3)$ is mainly for the existence and uniqueness of strong solution to (2.1). Additional assumptions $(A4) - (A5)$ on diffusion coefficient are used in order to establish large deviation principle. We will verify that many examples in [32, 24] (e.g. stochastic reaction-diffusion equation, stochastic porous media equation, stochastic p-Laplacian equation) satisfy $(A1) - (A5)$.
(ii) By $(A2)$ and $(A3)$ we can easily obtain the Coercivity of $(A, B)$:

$$2_{V^*}\langle A(t,v), v\rangle_V + \|B(t,v)\|_2^2 + \frac{\delta}{2}\|v\|_V^{\alpha} \leq C(1 + \|v\|_H^2).$$

(iii) Since for all $(t, v) \in [0, T] \times V$

$$\|P_n B(t,v) - B(t,v)\|_2 \to 0\ (n \to \infty).$$

Hence a simple sufficient condition for (2.2) hold is to assume

$$\{B(t,v) : (t,v) \in [0,T] \times S\}$$

is a relatively compact set in $L_2(U; H)$. For example, we can take

$$B(t,v) = \sum_{i=1}^{N} b_i(v) B_i(t),$$

where $b_i(\cdot) : V \to \mathbb{R}$ are Lipschitz functions and $B_i(\cdot) : [0, T] \to L_2(U; H)$ are contiunous.

If one take $n$-dimensional subspace of $V$ as $H_n$, denote $\{u_i\}$ and $\{e_i\}(\subseteq V)$ are ONB of $U$ and $H$ respectively. Suppose

$$B(t,v) = \sum_{i,j=1}^{\infty} b_{i,j}(t,v) u_i \otimes e_j.$$

Then (2.2) holds if

$$|b_{i,j}(t,v)| \leq b_{i,j}(1+\|v\|_H^{\gamma})$$

where $\gamma, b_{i,j} > 0$ and $\sum_{i,j=1}^{\infty} b_{i,j}^2 < \infty$.
(iv) If there exists a Hilbert space $H_0$ such that the embedding $H_0 \subseteq H$ is compact, $\{e_i\} \subseteq H_0 \cap V$ is an ONB in $H_0$ and also orthogonal set in $H$. Suppose for all $M > 0$

$$\sup_{(t,v)\in[0,T]\times S_M} \|B(t,v)\|_{L_2(U;H_0)} < \infty.$$

Then (2.2) holds. Because $B(t,v) = \sum_{i,j=1}^{\infty} b_{i,j}(t,v)u_i \otimes e_j$, by assumptions we know $\|e_j\|_H^2 \to 0$ and

$$\sup_{(t,v)\in[0,T]\times S_M} \sum_{i,j=1}^{\infty} b_{i,j}^2(t,v) < \infty.$$

then

$$\|P_n B(t,v) - B(t,v)\|_2^2 = \sum_{i=1}^{\infty} \sum_{j=n+1}^{\infty} b_{i,j}^2(t,v)\|e_j\|_H^2.$$

hence (2.2) holds. □

According to [24, Theorem II2.1] and the remark above, if $(A1)-(A3)$ hold, then for any $X_0 \in L^2(\Omega \to H; \mathcal{F}_0; \mathbf{P})$, (2.1) has an unique solution $\{X_t\}_{t\in[0,T]}$ which is an adapted continuous process on $H$ such that $\mathbf{E}\int_0^T (\|X_t\|_V^{\alpha} + \|X_t\|_H^2)\,\mathrm{d}t < \infty$ and

$$\langle X_t, v\rangle_H = \langle X_0, v\rangle_H + \int_0^t {}_{V^*}\langle A(s,X_s), v\rangle_V ds + \int_0^t \langle B(s,X_s)\mathrm{d}W_s, v\rangle_H, \ \mathbf{P}-a.s.$$

holds for all $v \in V$ and $t \in [0,T]$. Moreover, we have

$$\mathbf{E} \sup_{t\in[0,T]} \|X_t\|_H^2 \mathrm{d}t < \infty$$

and the following crucial Itô formula

$$\|X_t\|_H^2 = \|X_0\|_H^2 + \int_0^t \left(2{}_{V^*}\langle A(s,X_s), X_s\rangle_V + \|B(s,X_s)\|_2^2\right) \mathrm{d}s + 2\int_0^t \langle X_s, B(s,X_s)\mathrm{d}W_s\rangle_H.$$

Let us consider the stochastic evolution equation (2.1) with small noise:

$$\mathrm{d}X_t^{\varepsilon} = A(t,X_t^{\varepsilon})\mathrm{d}t + \varepsilon B(t,X_t^{\varepsilon})\mathrm{d}W_t, \quad \varepsilon > 0, \ X_0^{\varepsilon} = x \in H. \tag{2.3}$$

Let $T > 0$ and $x$ be fixed. By assumptions we know there exists an unique strong solution $\{X^{\varepsilon}\}$ of (2.3) with values in $C([0,T];H) \cap L^{\alpha}([0,T];V)$. The metric on $C([0,T];H) \cap L^{\alpha}([0,T];V)$ is defined as

$$\rho(f,g) := \sup_{t\in[0,T]} \|f_t - g_t\|_H + \left(\int_0^T \|f_t - g_t\|_V^{\alpha}\mathrm{d}t\right)^{\frac{1}{\alpha}}. \tag{2.4}$$

Obviously, $(C([0,T];H)\cap L^\alpha([0,T];V),\rho)$ is a Polish space. It follows (from infinite dimensional version of Yamada-Watanabe theorem in [32, Appendix E] or [34]) that there exists a Borel-measurable function

$$g^\varepsilon : C([0,T];U_1)\to C([0,T];H)\cap L^\alpha([0,T];V)$$

such that $X^\varepsilon = g^\varepsilon(W)$ *a.s.*. To state our main result, let us introduce the skeleton equation associated to (2.3):

$$\frac{\mathrm{d}z_t^\phi}{\mathrm{d}t} = A(t,z_t^\phi) + B(t,z_t^\phi)\phi_t, \quad z_0^\phi = x \tag{2.5}$$

where $\phi\in L^2([0,T];U)$. An element $z^\phi\in C([0,T];H)\cap L^\alpha([0,T];V)$ is called a solution to (2.5) if for any $v\in V$,

$$\langle z_t^\phi, v\rangle_H = \langle x,v\rangle_H + \int_0^t {}_{V^*}\langle A(s,z_s^\phi) + B(t,z_t^\phi)\phi_t, v\rangle_V \mathrm{d}s, \quad t\in[0,T]. \tag{2.6}$$

We will prove (see Lemma 3.1) that $(A1)-(A4)$ imply the existence and the uniqueness of the solution to (2.5) for any $\phi\in L^2([0,T];U)$.

Define $g^0 : C([0,T];U_1)\to C([0,T];H)\cap L^\alpha([0,T];V)$ by

$$g^0(h) := \begin{cases} z^\phi, & \text{if } h=\int_0^\cdot \phi_s \mathrm{d}s \ \text{ for some } \phi\in L^2([0,T];U),\\ 0, & \text{otherwise.}\end{cases}$$

Then it's obvious that the rate function in (1.2) can be written as

$$I(z) = \inf\{\frac{1}{2}\int_0^T \|\phi_s\|_U^2 \mathrm{d}s :\ z=z^\phi,\ \phi\in L^2([0,T],U)\}, \tag{2.7}$$

where $z\in C([0,T];H)\cap L^\alpha([0,T];V)$.

Now we can formulate the main result which is a well-known Freidlin-Wentzell type estimate:

**Theorem 2.1.** *Assume* $(A1)-(A5)$ *hold. For each* $\varepsilon>0$*, let* $X^\varepsilon = (X_t^\varepsilon)_{t\in[0,T]}$ *be the solution to* (2.3)*. Then as* $\varepsilon\to 0$*,* $\{X^\varepsilon\}$ *satisfies the LDP on* $C([0,T];H)\cap L^\alpha([0,T];V)$ *with the good rate function* $I$ *which is given by* (2.7)*.*

*Remark* 2.2. (i) According to [8, Theorem 5], we can also prove uniform Laplace principle by using the same arguments but with more cumbersome notation. Hence we omit the details here for simplicity.

(ii)The LDP is obtained on the space $C([0,T];H)\cap L^\alpha([0,T];V)$, which is stronger than the same LDP on $C([0,T];H)$ since the embedding

$$C([0,T];H)\cap L^\alpha([0,T];V)\subseteq C([0,T];H)$$

is obviously continuous.

(iii)This theorem can't be applied to stochastic fast-diffusion equation in [25, 29] since $(A2)$ fails to satisfy. However, if we replace $(A2)$ by usual monotone and coercive condition as in [24], then LDP can be obtained on $C([0,T];H)$(not $C([0,T];H)\cap L^\alpha([0,T];V)$) by similar proof. This result will be formulate in section 6.

The organization of this paper as follows. In section 3, we prove Large deviation principle by using weak convergence approach under some additional assumptions on $B$. Section 4 and 5 are devoted to drop those superfluous assumptions by some approximation arguments. In section 6 we prove the LDP under a slightly weaker assumptions than in Theorem 2.1, which can also be applied to stochastic fast diffusion equations. In Section 7 we apply the main results to different class of SPDEs in Hilbert space as applications.

# 3 Proof of Theorem 2.1 under additional assumptions on $B$

In order to prove large deviations principle for $X^\varepsilon$, we only need to verify **(A)** according to Lemma 1.1. Concerning the proof of (i) in **(A)**, we need to assume additionally

$(A6)$ $B:[0,T]\times V\to L(U;V_0)$, $V_0\subseteq V$ is compact embedding such that

$$\|B(t,v)\|^2_{L(U;V_0)}\le C(1+\|v\|_V^\alpha+\|v\|_H^2).$$

$(A7)$ $B$ is globally Lipschitz and bounded:

$$\|B(t,v_1)-B(t,v_2)\|_2^2\le K\|v_1-v_2\|_H^2;$$
$$\sup_{(t,v)\in[0,T]\times V}\|B(t,v)\|_2^2<\infty.$$

Then **(A)** can be verified by a series of Lemmas.

For reader's convenience, we recall two well-known inequalities which used quite often in the proof. Throughout the paper, the generic constants may be different from line to line. If it is essential, we will written the dependence of the constant on parameters explicitly.

**Young's inequality:** Given $p,q>1$ satisfy $\frac{1}{p}+\frac{1}{q}=1$, for any positive number $\sigma,a,b$

$$ab\le\sigma\frac{a^p}{p}+\sigma^{-\frac{q}{p}}\frac{b^q}{q}.$$

**Gronwall's inequality:** Let $W,\Phi,\Psi:[0,T]\to\mathbf{R}^+$ be Lebesgue measurable. Suppose $\Psi$ is locally integrable and $\int_0^T\Psi(s)W(s)\mathrm{d}s<\infty$. If

$$\begin{aligned}&W(t)\le\Phi(t)+\int_0^t\Psi(s)W(s)\mathrm{d}s,\ t\in[0,T]\ \text{ or}\\ &\frac{\mathrm{d}}{\mathrm{d}t}W(t)\le\frac{\mathrm{d}}{\mathrm{d}t}\Phi(t)+\Psi(t)W(t),\ t\in[0,T),\ W(0)\le\Phi(0).\end{aligned}\tag{3.1}$$

Then

$$W(t) \le \Phi(t) + \int_0^t \exp\left[\int_s^t \Psi(u)\mathrm{d}u\right] \Psi(s)\Phi(s)\mathrm{d}s,\ \ t \in [0,T]. \tag{3.2}$$

**Lemma 3.1.** *Assume* $(A1)-(A4)$ *hold. Let* $\|z\| := \sup_{t\in[0,T]} \|z_t\|_H$ *for* $z \in C([0,T];H)$. *For all* $x \in H$ *and* $\phi \in L^2([0,T];U)$ *there exists a unique solution* $z^\phi$ *to* (2.5) *and*

$$\begin{aligned} &\|z^\phi - z^\psi\|^2 + \delta \int_0^T \|z_t^\phi - z_t^\psi\|_V^\alpha \mathrm{d}t \\ &\le \exp\left\{\int_0^T \left(K + \|\phi_t\|_U^2 + \|B(t, z_t^\psi)\|_2^2\right)\mathrm{d}t\right\} \int_0^T \|\phi_t - \psi_t\|_U^2 \mathrm{d}t \end{aligned} \tag{3.3}$$

*hold for some constant* $K$ *and all* $\phi, \psi \in L^2([0,T];U)$.

*Proof.* To verify the existence of the solution, we make use of [24, Theorem II.2.1](see e.g. $(B1)-(B4)$ in section 6). First we assume $\phi \in L^\infty([0,T];U)$ and

$$\tilde{A}(s,v) := A(s,v) + B(s,v)\phi_s.$$

Then, due to $(A1)-(A4)$, it's easy to verify that $\tilde{A}$ satisfies Assumptions $A_i)(i = 1,..,5)$ on page 1252 of [24].

(i) Semicontinuity of $\tilde{A}$ follows from $(A1)$ and $(A2)$(or $(A4)$).

(ii) Monotonicity and Coercivity of $\tilde{A}$ follows from $(A2)$ and $(A4)$.

(iii) Boundedness of $\tilde{A}$ follows from $(A3)$ and $(A4)$ if $1 < \alpha < 2$.

If $\alpha \ge 2$, then by $(A3)$ and $(A4)$

$$\begin{aligned} \|\tilde{A}(s,v)\|_{V^*}^{\frac{\alpha}{\alpha-1}} &\le C\|A(s,v)\|_{V^*}^{\frac{\alpha}{\alpha-1}} + C\|B(t,v)\phi_t\|_{V^*}^{\frac{\alpha}{\alpha-1}} \\ &\le C(1+\|v\|_V^\alpha) + C\|B(t,v)\|_{L(U,V^*)}^{\frac{\alpha}{\alpha-1}} \\ &\le C(1+\|v\|_V^\alpha) + C\|B(t,v)\|_2^{\frac{\alpha}{\alpha-1}} \\ &\le C(1+\|v\|_V^\alpha) + C(1+\|v\|_H^{\frac{\alpha}{\alpha-1}}) \\ &\le C(1+\|v\|_V^\alpha). \end{aligned}$$

By [24, Theorems II.2.1 and II.2.2](see also [44, Theorem 30.A])we know (2.5) has an unique solution.

For general $\phi \in L^2([0,T];U)$, we can find a sequence of $\phi^n \in L^\infty([0,T];U)$ such that

$$\phi_n \to \phi \ \text{ in } \ L^2([0,T];U).$$

Let $z^n$ be the unique solution to (2.5) for $\phi^n$, we will show $\{z^n\}$ is a Cauchy sequence in $C([0,T];H)\cap L^\alpha([0,T];V)$. By Itô's formula due to [24, Theorem I.3.2] and $(A2)$ we have

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|z_t^n - z_t^m\|_H^2 =& 2_{V^*}\langle A(t,z_t^n) - A(t,z_t^m), z_t^n - z_t^m\rangle_V \\
&+ 2\langle B(t,z_t^n)\phi_t^n - B(t,z_t^m)\phi_t^m, z_t^n - z_t^m\rangle_H \\
\le& 2_{V^*}\langle A(t,z_t^n) - A(t,z_t^m), z_t^n - z_t^m\rangle_V + \|B(t,z_t^n) - B(t,z_t^m)\|_2^2 \\
&+ \|\phi_t^n\|_U^2\|z_t^n - z_t^m\|_H^2 + 2\langle z_t^n - z_t^m, B(t,z_t^m)\phi_t^n - B(t,z_t^m)\phi_t^m\rangle_H \\
\le& -\delta\|z_t^n - z_t^m\|_V^\alpha + (K + \|\phi_t^n\|_U^2)\|z_t^n - z_t^m\|_H^2 \\
&+ 2\|B^*(t,z_t^m)\left(z_t^n - z_t^m\right)\|_U\|\phi_t^n - \phi_t^m\|_U \\
\le& -\delta\|z_t^n - z_t^m\|_V^\alpha + \|\phi_t^n - \phi_t^m\|_U^2 \\
&+ \left(K + \|\phi_t^n\|_U^2 + \|B(t,z_t^m)\|_2^2\right)\|z_t^n - z_t^m\|_H^2.
\end{aligned}
\tag{3.4}
$$

where $B^*$ denote the adjoint operator of $B$ and we also use the fact

$$\|B^*\|_{L(H;U)} = \|B\|_{L(U;H)} \le \|B\|_2.$$

Then by Gronwall lemma we have

$$
\begin{aligned}
&\|z^n - z^m\|^2 + \delta\int_0^T \|z_t^n - z_t\|_V^\alpha \mathrm{d}t \\
&\le \exp\left\{\int_0^T \left(K + \|\phi_t^n\|_U^2 + \|B(t,z_t^m)\|_2^2\right)\mathrm{d}t\right\}\int_0^T \|\phi_t^n - \phi_t^m\|_U^2\mathrm{d}t
\end{aligned}
\tag{3.5}
$$

By the similar argument we have

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|z_t^n\|_H^2 =& 2_{V^*}\langle A(t,z_t^n), z_t^n\rangle_V + 2\langle B(t,z_t^n)\phi_t^n, z_t^n\rangle_H \\
\le& -\frac{\delta}{2}\|z_t^n\|_V^\alpha + C(1 + \|z_t^n\|_H^2) + \|\phi_t^n\|_U^2\|z_t^n\|_H^2 \\
\le& -\frac{\delta}{2}\|z_t^n\|_V^\alpha + C + \left(C + \|\phi_t^n\|_U^2\right)\|z_t^n\|_H^2
\end{aligned}
\tag{3.6}
$$

Then by Gronwall lemma and boundedness of $\phi^n$ in $L^2([0,T];U)$

$$
\|z^n\|^2 + \int_0^T \|z_t^n\|_V^\alpha \mathrm{d}t \le C\exp\left\{\int_0^T \left(C + \|\phi_t^n\|_U^2\right)\mathrm{d}t\right\}\left(\|x\|_H^2 + T\right) \le \textbf{Constant} < \infty.
\tag{3.7}
$$

Hence we have

$$
\int_0^T \|B(t,z_t^m)\|_2^2\mathrm{d}t \le c\int_0^T \left(1 + \|z_t^m\|_H^2 + \|z_t^m\|_V^\alpha\right)\mathrm{d}t \le \textbf{Constant} < \infty.
\tag{3.8}
$$

Combining (3.5),(3.8) and $\phi^n \to \phi$, we can conclude that $\{z^n\}$ is a Cauchy sequence in $C([0,T];H) \cap L^\alpha([0,T];V)$, the limit is denoted by $z^\phi$. Similarly, from (3.7), $(A3)$ and $(A4)$ we know there exist $\eta$ such that

$$\begin{aligned} A(\cdot, z^n_\cdot) \to \eta \ \text{ weakly in } L^{\frac{\alpha}{\alpha-1}}([0,T];V^*); \\ B(\cdot, z^n_\cdot)\phi^n_\cdot \to B(\cdot, z_\cdot)\phi_\cdot \text{ strongly in } \ L^2([0,T];H). \end{aligned}$$

Then one can verify $\eta = A(\cdot, z^\phi_\cdot)$ by using $(A2)$ and standard monotonicity argument(e.g.[44, Theorem 30.A]). Hence $z^\phi$ is the solution of (2.5) corresponding to $\phi$.

And (3.3) can be derived by the same argument for (3.5). So the proof is complete. □

The following Lemma shows that $I$ defined by (2.7) is a good rate function.

**Lemma 3.2.** *Assume* $(A1) - (A4)$ *hold. For every* $N < \infty$*, the set*

$$K_N = \{g^0(\int_0^\cdot \phi_s \mathrm{d}s) : \phi \in S_N\}$$

*is a compact subset of* $C([0,T];H) \cap L^\alpha([0,T];V)$*.*

*Proof.* **Step 1**: we first assume $B$ also satisfy $(A6)$. By definition we know

$$K_N = \{z^\phi : \phi \in L^2([0,T];U), \ \int_0^T \|\phi_s\|_U^2 \mathrm{d}s \le N\}.$$

For any sequence $\phi^n \subset S_N$, we may assume $\phi^n \to \phi$ weakly in $L^2([0,T];U)$ since $L^2([0,T];U)$ is weakly compact. Denote $z^n, z$ are the solutions of (2.5) corresponding to $\phi^n, \phi$ respectively. Now it sufficient to show $z^n \to z$ strongly in $C([0,T];H) \cap L^\alpha([0,T];V)$.

From (3.4) we have

$$\begin{aligned} &\|z^n_t - z_t\|_H^2 + \delta \int_0^t \|z^n_s - z_s\|_V^\alpha \mathrm{d}s \\ &\le \int_0^t (K + \|\phi^n_s\|_U^2)\|z^n_s - z_s\|_H^2 \mathrm{d}s + 2\int_0^t \langle z^n_s - z_s, B(s,z_s)(\phi^n_s - \phi_s)\rangle_H \mathrm{d}s \end{aligned} \tag{3.9}$$

Define

$$h^n_t = \int_0^t B(s,z_s)(\phi^n_s - \phi_s)\mathrm{d}s.$$

By $(A5)$ and (3.8) we know $h^n \in C([0,T];V_0)$ and

$$\begin{aligned} \sup_{t\in[0,T]} \|h^n_t\|_{V_0} &\le \int_0^T \|B(s,z_s)(\phi^n_s - \phi_s)\|_{V_0} \mathrm{d}s \\ &\le \left(\int_0^T \|B(s,z_s)\|_{L(U,V_0)}^2 \mathrm{d}s\right)^{1/2} \left(\int_0^T \|\phi^n_s - \phi_s\|_U^2 \mathrm{d}s\right)^{1/2} \\ &\le \textbf{Constant} < \infty. \end{aligned} \tag{3.10}$$

Since the embedding $V_0 \subseteq V$ is compact and $\phi^n \to \phi$ weakly in $L^2([0,T];U)$, it's easy to show that $h^n \to 0$ in $C([0,T];V)$ by using the Arzèla-Ascoli theorem(also see e.g.[7, Lemma 3.2]). In particular, $h^n \to 0$ in $C([0,T];H) \cap L^\alpha([0,T];V)$.

Moreover the derivative (w.r.t. time variable) is given by

$$(h_s^n)' = B(s,z_s)(\phi_s^n - \phi_s).$$

If $\alpha \geq 2$ we have

$$\begin{aligned}\int_0^T \|(h_s^n)'\|_{V^*}^{\frac{\alpha}{\alpha-1}} \mathrm{d}s &\leq \int_0^T \|B(s,z_s)(\phi_s^n - \phi_s)\|_{V^*}^{\frac{\alpha}{\alpha-1}} \mathrm{d}s \\ &\leq C\left(1 + \sup_{t\in[0,T]} \|z_t\|_H^2\right)^{\frac{\alpha}{2(\alpha-1)}} \left(\int_0^T \|\phi_s^n - \phi_s\|_U^{\frac{\alpha}{\alpha-1}} \mathrm{d}s\right) \\ &\leq \textbf{Constant} < \infty.\end{aligned} \tag{3.11}$$

Hence $(h^n_\cdot)'$ is an element in $L^{\frac{\alpha}{\alpha-1}}([0,T];V^*)$. If $1 < \alpha < 2$, we need to approximate $\phi^n, \phi$ by element in $L^\infty([0,T];U)$ as in Lemma 3.1. Since Lemma 3.1 shows that the convergence of the corresponding solution $z^\phi$ is uniformly on $S_N$ w.r.t. this approximation, the conclusion on case $1 < \alpha < 2$ can de derived by the following proof for the case $\alpha \geq 2$ and standard $3\varepsilon$-argument. We omit the details here.

By [44, Proposition 23.23] we have the following integration by parts formula

$$\langle z_t^n - z_t, h_t^n\rangle_H = \int_0^t {}_{V^*}\langle (z_s^n - z_s)', h_s^n\rangle_V \mathrm{d}s + \int_0^t {}_{V^*}\langle (h_s^n)', z_s^n - z_s\rangle_V \mathrm{d}s.$$

Since $z_s^n - z_s$ and $(h_s^n)'$ all take values in $H$, we know

$${}_{V^*}\langle (h_s^n)', z_s^n - z_s\rangle_V = \langle z_s^n - z_s, B(s,z_s)(\phi_s^n - \phi_s)\rangle_H.$$

Hence one has

$$\begin{aligned}&\int_0^t \langle z_s^n - z_s, B(s,z_s)(\phi_s^n - \phi_s)\rangle_H \mathrm{d}s \\ =&\langle z_t^n - z_t, h_t^n\rangle_H - \int_0^t {}_{V^*}\langle (z_s^n - z_s)', h_s^n\rangle_V \mathrm{d}s \\ =&\langle z_t^n - z_t, h_t^n\rangle_H - \int_0^t {}_{V^*}\langle A(s,z_s^n) - A(s,z_s), h_s^n\rangle_V \mathrm{d}s \\ &- \int_0^t \langle B(s,z_s^n)\phi_s^n - B(s,z_s)\phi_s, h_s^n\rangle_H \mathrm{d}s \\ =:&I_1 + I_2 + I_3\end{aligned} \tag{3.12}$$

By using the Hölder inequality, $(A3)$ and (3.7) we have

$$
\begin{aligned}
I_1 &\leq \|z_t^n - z_t\|_H \cdot \|h_t^n\|_H \leq \frac{1}{4}\|z_t^n - z_t\|_H^2 + \|h_t^n\|_H^2. \\
I_2 &\leq \int_0^t \|A(s, z_s^n) - A(s, z_s)\|_{V^*} \|h_s^n\|_V \mathrm{d}s \\
&\leq \left(\int_0^t \|A(s, z_s^n) - A(s, z_s)\|_{V^*}^{\frac{\alpha}{\alpha-1}} \mathrm{d}s\right)^{\frac{\alpha-1}{\alpha}} \left(\int_0^t \|h_s^n\|_V^\alpha \mathrm{d}s\right)^{\frac{1}{\alpha}} \\
&\leq \left(\int_0^t C\left(1 + \|z_s\|_V^\alpha + \|z_s^n\|_V^\alpha\right) \mathrm{d}s\right)^{\frac{\alpha-1}{\alpha}} \left(\int_0^t \|h_s^n\|_V^\alpha \mathrm{d}s\right)^{\frac{1}{\alpha}} \\
&\leq C\left(\int_0^t \|h_s^n\|_V^\alpha \mathrm{d}s\right)^{\frac{1}{\alpha}}. \qquad (3.13) \\
I_3 &\leq \int_0^t \|B(s, z_s^n)\phi_s^n - B(s, z_s)\phi_s\|_H \cdot \|h_s^n\|_H \mathrm{d}s \\
&\leq \sup_{s\in[0,t]} \|h_s^n\|_H \int_0^t \|B(s, z_s^n)\phi_s^n - B(s, z_s)\phi_s\|_H \mathrm{d}s \\
&\leq \sup_{s\in[0,t]} \|h_s^n\|_H \left\{N^{1/2}\left(\int_0^t \|B(s, z_s^n)\|_2^2 \mathrm{d}s\right)^{1/2} + N^{1/2}\left(\int_0^t \|B(s, z_s)\|_2^2 \mathrm{d}s\right)^{1/2}\right\} \\
&\leq C \sup_{s\in[0,t]} \|h_s^n\|_H.
\end{aligned}
$$

where $C$ is a constant which come from the following estimate

$$
\int_0^t \|B(s, z_s^n)\|_2^2 \mathrm{d}s \leq C\int_0^t \left(1 + \|z_s^n\|_H^2 + \|z_s^n\|_V^\alpha\right) \mathrm{d}s \leq \textbf{Constant} < \infty
$$

Combining (3.9) and (3.12)-(3.13) we have

(3.14)

$$
\begin{aligned}
&\|z_t^n - z_t\|_H^2 + \delta\int_0^T \|z_t^n - z_t\|_V^\alpha \mathrm{d}t \\
&\leq C\int_0^t (1 + \|\phi_s^n\|_U^2)\|z_s^n - z_s\|_H^2 \mathrm{d}s + C\left(\sup_{s\in[0,t]} \|h_s^n\|_H + \sup_{s\in[0,t]} \|h_s^n\|_H^2 + \left(\int_0^t \|h_s^n\|_V^\alpha \mathrm{d}s\right)^{\frac{1}{\alpha}}\right)
\end{aligned}
$$

Then by Gronwall Lemma and $L^2$-boundedness of $\phi^n$, there exists a constant $C$ such that

$$
\|z^n - z\|^2 + \delta\int_0^T \|z_s^n - z_s\|_V^\alpha \mathrm{d}s \leq C\left(\sup_{s\in[0,T]} \|h_s^n\|_H + \sup_{s\in[0,T]} \|h_s^n\|_H^2 + \left(\int_0^T \|h_s^n\|_V^\alpha \mathrm{d}s\right)^{\frac{1}{\alpha}}\right).
$$

Since $h^n \to 0$ in $C([0,T];H)\cap L^\alpha([0,T];V)$, we know

$$z^n \to z \text{ strongly } \text{ in } C([0,T];H)\cap L^\alpha([0,T];V)$$

as $n\to\infty$.

**Step 2**: we prove the conclusion for general $B$ without condition $(A6)$. Denote $z^\phi_{t,n}$ the solution of the following equation

$$\frac{\mathrm{d}z^\phi_{t,n}}{\mathrm{d}t} = A(t,z^\phi_{t,n}) + P_n B(t,z^\phi_{t,n})\phi_t, \quad z^\phi_{0,n} = x$$

Where $P_n$ is the standard projection(see (A5) and section 4 for details). By using the same argument in Lemma 3.1 we can prove

$$\begin{aligned}&\|z^\phi_n - z^\phi\|^2 + \delta\int_0^T \|z^\phi_{s,n} - z^\phi_s\|_V^\alpha \mathrm{d}s\\ &\le \exp\left\{\int_0^T (K + 2\|\phi_s\|_U^2)\mathrm{d}s\right\}\int_0^T \|(I-P_n)B(s,z^\phi_s)\|_2^2\mathrm{d}s\end{aligned} \tag{3.15}$$

Since $B(\cdot,\cdot)$ are Hilbert-Schmidt (hence compact) operators, then by dominated convergence theorem we know

$$\int_0^T \|(I-P_n)B(s,z^\phi_s)\|_2^2\mathrm{d}s \to 0 \text{ as } n\to\infty.$$

Hence $z^\phi_n \to z^\phi$ in $C([0,T];H)\cap L^\alpha([0,T];V)$ as $n\to\infty$. Moreover, this convergence is uniformly (w.r.t $\phi$) on bounded set of $L^2([0,T];U)$ which follows from (3.15) and (3.8). Notice $P_nB$ satisfy $(A6)$, combining with **Step 1** and standard $3\varepsilon$-argument we can conclude that $z^n \to z$ strongly in $C([0,T];H)\cap L^\alpha([0,T];V)$ also for general $B$. Now the proof is complete. ☐

**Lemma 3.3.** *Assume* $(A1)-(A4)$ *and* $(A6)-(A7)$ *hold. Let* $\{v^\varepsilon\}_{\varepsilon>0}\subset \mathcal{A}_N$ *for some* $N<\infty$. *Assume* $v^\varepsilon$ *converge to* $v$ *in distribution as* $S_N$*-valued random elements, then*

$$g^\varepsilon\left(W_\cdot + \frac{1}{\varepsilon}\int_0^\cdot v^\varepsilon_s \mathrm{d}s\right) \to g^0\left(\int_0^\cdot v_s\mathrm{d}s\right)$$

*in distribution as* $\varepsilon\to 0$.

*Proof.* By Girsanov theorem and uniqueness of solution to (2.3), it's easy to see that $X^\varepsilon := g^\varepsilon\left(W_\cdot + \frac{1}{\varepsilon}\int_0^\cdot v^\varepsilon_s\mathrm{d}s\right)$ (the abuse of notation here is for simplicity) is the unique solution of the following equation

$$\mathrm{d}X^\varepsilon_t = (A(t,X^\varepsilon_t) + B(t,X^\varepsilon_t)v^\varepsilon_t)\,\mathrm{d}t + \varepsilon B(t,X^\varepsilon_t)\mathrm{d}W_t,\ X^\varepsilon_0 = x. \tag{3.16}$$

Now we only need to show $X^\varepsilon \to z^v$ in distribution as $\varepsilon \to 0$. We may assume $\varepsilon \le \frac{1}{2}$, by using Itô formula, the Young inequality and (A2) we have

$$
\begin{aligned}
\mathrm{d}\|X_t^\varepsilon - z_t^v\|_H^2 =& 2_{V^*}\langle A(t,X_t^\varepsilon) - A(t,z_t^v), X_t^\varepsilon - z_t^v\rangle_V \mathrm{d}t \\
&+ 2\langle X_t^\varepsilon - z_t^v, (B(t,X_t^\varepsilon) - B(t,z_t^v))v_t^\varepsilon + B(t,z_t^v)(v_t^\varepsilon - v_t)\rangle_H \mathrm{d}t \\
&+ \varepsilon^2 \|B(t,X_t^\varepsilon)\|_2^2 \mathrm{d}t + 2\varepsilon\langle X_t^\varepsilon - z_t^v, B(t,X_t^\varepsilon)\mathrm{d}W_t\rangle_H \\
\le& \left(2_{V^*}\langle A(t,X_t^\varepsilon) - A(t,z_t^v), X_t^\varepsilon - z_t^v\rangle_V + \|B(t,X_t^\varepsilon) - B(t,z_t^v)\|_2^2\right)\mathrm{d}t \\
&+ 2\|v_t^\varepsilon\|_U^2 \|X_t^\varepsilon - z_t^v\|_H^2 \mathrm{d}t + 2\langle X_t^\varepsilon - z_t^v, B(t,z_t^v)(v_t^\varepsilon - v_t)\rangle_H \mathrm{d}t \\
&+ 2\varepsilon^2 \|B(t,z_t^v)\|_2^2 \mathrm{d}t + 2\varepsilon\langle X_t^\varepsilon - z_t^v, B(t,X_t^\varepsilon)\mathrm{d}W_t\rangle_H \\
\le& -\delta\|X_t^\varepsilon - z_t^v\|_V^\alpha + \left[C(1+\|v_t^\varepsilon\|_U^2)\|X_t^\varepsilon - z_t^v\|_H^2 + 2\varepsilon^2\|B(t,z_t^v)\|_2^2\right]\mathrm{d}t \\
&+ 2\langle X_t^\varepsilon - z_t^v, B(t,z_t^v)(v_t^\varepsilon - v_t)\rangle_H \mathrm{d}t + 2\varepsilon\langle X_t^\varepsilon - z_t^v, B(t,X_t^\varepsilon)\mathrm{d}W_t\rangle_H
\end{aligned} \tag{3.17}
$$

Similarly we define

$$
h_t^\varepsilon = \int_0^t B(s,z_s^v)(v_s^\varepsilon - v_s)\mathrm{d}s,
$$

then we can show $h^\varepsilon \to 0$ in distribution as $C([0,T];V)$-valued random element(see e.g.[7, Lemma 3.2]), consequently also in $C([0,T];H) \cap L^\alpha([0,T];V)$. Notice that

$$
2\langle X_t^\varepsilon - z_t^v, h_t^\varepsilon\rangle_H = \|X_t^\varepsilon - z_t^v + h_t^\varepsilon\|_H^2 - \|X_t^\varepsilon - z_t^v\|_H^2 - \|h_t^\varepsilon\|_H^2.
$$

By using the Itô formula for corresponding square norm we can conclude that

$$
\begin{aligned}
&\int_0^t \langle X_s^\varepsilon - z_s^v, B(s,z_s^v)(v_s^\varepsilon - v_s)\rangle_H \mathrm{d}s \\
=& \langle X_t^\varepsilon - z_t^v, h_t^\varepsilon\rangle_H - \int_0^t {}_{V^*}\langle A(s,X_s^\varepsilon) - A(s,z_s^v), h_s^\varepsilon\rangle_V \mathrm{d}s \\
&- \int_0^t \langle B(s,X_s^\varepsilon)v_s^\varepsilon - B(s,z_s^v)v_s, h_s^\varepsilon\rangle_H \mathrm{d}s - \varepsilon\int_0^t \langle B(s,X_s^\varepsilon)\mathrm{d}W_s, h_s^\varepsilon\rangle_H
\end{aligned} \tag{3.18}
$$

By using the same argument in (3.13) we have

$$
\begin{aligned}
&\int_0^t \langle X_s^\varepsilon - z_s^v, B(s,z_s^v)(v_s^\varepsilon - v_s)\rangle_H \mathrm{d}s \\
\le& \frac{1}{4}\|X_t^\varepsilon - z_t^v\|_H^2 + \sup_{s\in[0,t]}\|h_s^\varepsilon\|_H^2 - \varepsilon\int_0^t \langle B(s,X_s^\varepsilon)\mathrm{d}W_s, h_s^\varepsilon\rangle_H \\
&+ C\left(\int_0^t (1+\|z_s^v\|_V^\alpha + \|X_s^\varepsilon\|_V^\alpha)\,\mathrm{d}s\right)^{\frac{\alpha-1}{\alpha}} \cdot \left(\int_0^t \|h_s^\varepsilon\|_V^\alpha \mathrm{d}s\right)^{\frac{1}{\alpha}} \\
&+ C\sup_{s\in[0,t]}\|h_s^\varepsilon\|_H \left\{\left(\int_0^t \|B(s,X_s^\varepsilon)\|_2^2 \mathrm{d}s\right)^{1/2} + \left(\int_0^t \|B(s,z_s^v)\|_2^2 \mathrm{d}s\right)^{1/2}\right\}
\end{aligned} \tag{3.19}
$$

Hence from (3.17)-(3.19) we have

$$
\begin{aligned}
&\|X_t^\varepsilon - z_t^v\|_H^2 + \delta \int_0^t \|X_t^\varepsilon - z_t^v\|_V^\alpha \mathrm{d}s \\
&\le c_1 \int_0^t (1 + \|v_s^\varepsilon\|_U^2)\|X_t^\varepsilon - z_t^v\|_H^2 \mathrm{d}s + c_2(\varepsilon^2 + \sup_{s\in[0,t]} \|h_s^\varepsilon\|_H^2) \\
&+ c_3 \left(1 + \int_0^t \|X_s^\varepsilon\|_V^\alpha \mathrm{d}s\right)^{\frac{\alpha-1}{\alpha}} \cdot \left(\int_0^t \|h_s^\varepsilon\|_V^\alpha \mathrm{d}s\right)^{\frac{1}{\alpha}} \\
&+ c_4 \sup_{s\in[0,t]} \|h_s^\varepsilon\|_H \left\{1 + \left(\int_0^t \|X_s^\varepsilon\|_H^2 \mathrm{d}s\right)^{1/2}\right\} \\
&+ 4\varepsilon \int_0^t \langle X_s^\varepsilon - z_s^v - h_s^\varepsilon, B(s, X_s^\varepsilon)\mathrm{d}W_s\rangle_H
\end{aligned}
\tag{3.20}
$$

where we used the estimate (see (3.6)-(3.8)) that there exists constant $C$ such that

$$\int_0^T \|B(s, z_s^v)\|_2^2 \mathrm{d}s + \int_0^T \|z_s^v\|_V^\alpha \mathrm{d}s \le C, \quad a.s.$$

By applying Gronwall Lemma we have

(3.21)

$$
\begin{aligned}
&\sup_{s\in[0,t]} \|X_s^\varepsilon - z_s^v\|_H^2 + \delta \int_0^t \|X_s^\varepsilon - z_s^v\|_V^\alpha \mathrm{d}s \\
&\le C\big[\varepsilon^2 + \sup_{s\in[0,t]} \|h_s^\varepsilon\|_H^2 + \left(1 + \int_0^t \|X_s^\varepsilon\|_V^\alpha \mathrm{d}s\right)^{\frac{\alpha-1}{\alpha}} \left(\int_0^t \|h_s^\varepsilon\|_V^\alpha \mathrm{d}s\right)^{\frac{1}{\alpha}} \\
&+ \sup_{s\in[0,t]} \|h_s^\varepsilon\|_H \left\{1 + \left(\int_0^t \|X_s^\varepsilon\|_H^2 \mathrm{d}s\right)^{1/2}\right\} + \sup_{u\in[0,t]} \left|\varepsilon \int_0^u \langle X_s^\varepsilon - z_s^v - h_s^\varepsilon, B(s, X_s^\varepsilon)\mathrm{d}W_s\rangle_H\right| \big]
\end{aligned}
$$

Define stopping time

$$\tau^{M,\varepsilon} = \inf\left\{t \le T : \sup_{s\in[0,t]} \|X_s^\varepsilon\|_H^2 + \int_0^t \|X_s^\varepsilon\|_V^\alpha \mathrm{d}s > M\right\}.$$

By Burkhölder-Davis-Gundy inequality one has

$$
\begin{aligned}
&\varepsilon\mathbf{E}\sup_{t\in[0,\tau^{M,\varepsilon}]}|\int_0^t\langle X_s^\varepsilon-z_s^v-h_s^\varepsilon,B(s,X_s^\varepsilon)\mathrm{d}W_s\rangle_H|\\
\le&3\varepsilon\mathbf{E}\left\{\int_0^{\tau^{M,\varepsilon}}\|X_s^\varepsilon-z_s^v-h_s^\varepsilon\|_H^2\|B(s,X_s^\varepsilon)\|_2^2\mathrm{d}s\right\}^{1/2}\\
\le&3\varepsilon\mathbf{E}\left\{\sup_{s\in[0,\tau^{M,\varepsilon}]}\|X_s^\varepsilon-z_s^v-h_s^\varepsilon\|_H^2+C\int_0^{\tau^{M,\varepsilon}}\left(1+\|X_s^\varepsilon\|_H^2\right)\mathrm{d}s\right\}\\
\le&C\varepsilon\to0\ (\varepsilon\to0).
\end{aligned}
\tag{3.22}
$$

Since $B$ is bounded, by using the similar argument in (3.17) we have

$$
\mathrm{d}\|X_t^\varepsilon\|_H^2\le-\frac{\delta}{2}\|X_t^\varepsilon\|_V^\alpha+C(1+\|v_t^\varepsilon\|_U^2+\|X_t^\varepsilon\|_H^2)\mathrm{d}t+2\varepsilon\langle X_t^\varepsilon,B(t,X_t^\varepsilon)\mathrm{d}W_t\rangle_H,
$$

where $C$ is a constant. Repeat the same argument in [24, Theorem 3.10] we can prove

$$
\sup_{\varepsilon\in[0,1)}\mathbf{E}\left\{\sup_{t\in[0,T]}\|X_t^\varepsilon\|_H^2+\int_0^T\|X_t^\varepsilon\|_V^\alpha\mathrm{d}t\right\}\le\mathbf{constant}<\infty
$$

Hence it's easy to show that for a suitable constant $C$

$$
\liminf_{\varepsilon\to0}\mathbf{P}\{\tau^{M,\varepsilon}=T\}\ge1-\frac{C}{M}
\tag{3.23}
$$

Recall that $h^\varepsilon\to0$ in distribution in $C([0,T];H)\cap L^\alpha([0,T];V)$, combining with (3.21)-(3.23) one can conclude

$$
\sup_{t\in[0,T]}\|X_t^\varepsilon-z_t^v\|_H^2+\int_0^T\|X_t^\varepsilon-z_t^v\|_V^\alpha\mathrm{d}t\to0\ (\varepsilon\to0)
$$

in distribution. Hence the proof is complete. □

*Remark* 3.1. According to Lemma 1.1, Lemma 3.2 and Lemma 3.3, $\{X^\varepsilon\}$ satisfy LDP provided $(A1)-(A4)$ and $(A6)-(A7)$ hold. The next two section is to drop $(A6)$ and $(A7)$ by using some approximate argument.

# 4 From $L(U;V_0)$ to $L_2(U;H)$: drop $(A6)$

For any fixed $n\ge1$, let $H_n\subseteq V$ compact and $P_n:H\to H_n$ be the orthogonal projection. Let $X_t^{\varepsilon,n}$ be the solution of

$$
\mathrm{d}X_t^{\varepsilon,n}=A(t,X_t^{\varepsilon,n})\mathrm{d}t+\varepsilon P_nB(t,X_t^{\varepsilon,n})\mathrm{d}W_t,\quad X_0^{\varepsilon,n}=x.
\tag{4.1}
$$

Since $P_nB$ satisfy $(A6)$, according to section 3 and [7, Theorem 4.4] we know $\{X^{\varepsilon,n}\}$ satisfy (LDP) provided $(A1)-(A4)$ and $(A7)$ hold. Now we prove that $\{X^{\varepsilon,n}\}$ are the exponential good approximation to $\{X^\varepsilon\}$.

**Lemma 4.1.** *If $(A1)-(A5)$ and $(A7)$ hold, then $\forall \sigma > 0$*

$$\limsup_{n\to\infty}\limsup_{\varepsilon\to 0}\varepsilon^2\log \mathbf{P}\left(\rho(X^\varepsilon, X^{\varepsilon,n}) > \sigma\right) = -\infty, \tag{4.2}$$

*where $\rho$ is the metric on $C([0,T];H)\cap L^\alpha([0,T];V)$ defined in (2.4).*

*Proof.* For $\varepsilon < \frac{1}{2}$, by using Itô's formula and $(A2)$ we have

$$\begin{aligned}
&\mathrm{d}\|X_t^\varepsilon - X_t^{\varepsilon,n}\|_H^2\\
=&\left(2_{V^*}\langle A(t,X_t^\varepsilon) - A(t,X_t^{\varepsilon,n}), X_t^\varepsilon - X_t^{\varepsilon,n}\rangle_V + \varepsilon^2\|B(t,X_t^\varepsilon) - P_nB(t,X_t^{\varepsilon,n})\|_2^2\right)\mathrm{d}t\\
&+2\varepsilon\langle X_t^\varepsilon - X_t^{\varepsilon,n}, (B(t,X_t^\varepsilon) - P_nB(t,X_t^{\varepsilon,n}))\mathrm{d}W_t\rangle_H\\
\le& -\delta\|X_t^\varepsilon - X_t^{\varepsilon,n}\|_V^\alpha\mathrm{d}t + C(\|X_t^\varepsilon - X_t^{\varepsilon,n}\|_H^2 + \varepsilon^2\|(I-P_n)B(t,X_t^\varepsilon)\|_2^2)\mathrm{d}t + 2\varepsilon\mathrm{d}M_t^{(n)}
\end{aligned}$$

where $\mathrm{d}M_t^{(n)} := \langle X_t^\varepsilon - X_t^{\varepsilon,n}, (B(t,X_t^\varepsilon) - P_nB(t,X_t^{\varepsilon,n}))\mathrm{d}W_t\rangle_H$ and $C$ is a constant. Define

$$\|X_t^\varepsilon - X_t^{\varepsilon,n}\| = \|X_t^\varepsilon - X_t^{\varepsilon,n}\|_H^2 + \delta\int_0^t\|X_s^\varepsilon - X_s^{\varepsilon,n}\|_V^\alpha\mathrm{d}s;\quad a_n = \sup_{(t,v)\in[0,T]\times V}\|(I-P_n)B(t,v)\|_2^2.$$

By the standard truncation argument (see section 5 for details) we can replace $B$ by truncated map $B_N$ here, hence by $(A5)$ and monotone convergence theorem

$$a_n = \max\left\{\sup_{(t,v)\in[0,T]\times S_{2N}}\|(I-P_n)B(t,v)\|_2^2,\ \sup_{t\in[0,T]}\|(I-P_n)B(t,0)\|_2^2\right\}\to 0(n\to\infty).$$

Therefore

$$\|X_t^\varepsilon - X_t^{\varepsilon,n}\| \le C\int_0^t(\|X_s^\varepsilon - X_s^{\varepsilon,n}\|_H^2 + \varepsilon^2 a_n)\mathrm{d}s + 2\varepsilon M_t^{(n)}$$

From $(A7)$ we know the quadratic variation process of the local martingale $M^{(n)}$ satisfies

$$\mathrm{d}\langle M^{(n)}\rangle_t \le C\|X_t^\varepsilon - X_t^{\varepsilon,n}\|_H^2(\|X_t^\varepsilon - X_t^{\varepsilon,n}\|_H^2 + a_n)\mathrm{d}t.$$

Define $\varphi_\theta(y) = (a_n + y)^\theta$ for some $\theta > 0$, then

$$\begin{aligned}
&\mathrm{d}\varphi_\theta(\|X_t^\varepsilon - X_t^{\varepsilon,n}\|)\\
\le&\theta(a_n + \|X_t^\varepsilon - X_t^{\varepsilon,n}\|)^{\theta-1}\left(C(\|X_t^\varepsilon - X_t^{\varepsilon,n}\|_H^2 + \varepsilon^2 a_n)\mathrm{d}t + 2\varepsilon\mathrm{d}M_t^{(n)}\right)\\
&+2C\varepsilon^2\theta(\theta-1)(a_n + \|X_t^\varepsilon - X_t^{\varepsilon,n}\|)^{\theta-2}\|X_t^\varepsilon - X_t^{\varepsilon,n}\|_H^2\left(\|X_t^\varepsilon - X_t^{\varepsilon,n}\|_H^2 + a_n\right)\mathrm{d}t\\
\le&C(\theta + \theta^2\varepsilon^2)\varphi_\theta(\|X_t^\varepsilon - X_t^{\varepsilon,n}\|)\mathrm{d}t + \mathrm{d}\beta_t
\end{aligned}\tag{4.3}$$

where $\beta_t$-term is a local martingale. By standard localization argument we may assume $\beta_t$ is a martingale here for simplicity. Let $\theta = \frac{1}{\varepsilon^2}$ we know

$$N_t := \mathrm{e}^{-\frac{2C}{\varepsilon^2}t}\varphi_{\frac{1}{\varepsilon^2}}(\|X_t^\varepsilon - X_t^{\varepsilon,n}\|)$$

is a supermartingale. Hence

$$\begin{aligned}
&\mathbf{P}\left(\rho(X^\varepsilon, X^{\varepsilon,n}) > 2\sigma\right)\\
\leq&\mathbf{P}\left(\sup_{t\in[0,T]}\|X_t^\varepsilon - X_t^{\varepsilon,n}\|_H > \sigma\right) + \mathbf{P}\left(\int_0^T \|X_t^\varepsilon - X_t^{\varepsilon,n}\|_V^\alpha \mathrm{d}t > \sigma^\alpha\right)\\
\leq&\mathbf{P}\left(\sup_{t\in[0,T]} N_t > \exp\{-\frac{2C}{\varepsilon^2}T\}(\sigma^2 + a_n)^{\frac{1}{\varepsilon^2}}\right) + \mathbf{P}\left(\sup_{t\in[0,T]} N_t > \exp\{-\frac{2C}{\varepsilon^2}T\}(\delta\sigma^\alpha + a_n)^{\frac{1}{\varepsilon^2}}\right)\\
\leq&\exp\{\frac{2C}{\varepsilon^2}T\}(\sigma^2 + a_n)^{-\frac{1}{\varepsilon^2}}\mathbf{E}N_0 + \exp\{\frac{2C}{\varepsilon^2}T\}(\delta\sigma^\alpha + a_n)^{-\frac{1}{\varepsilon^2}}\mathbf{E}N_0\\
=&\exp\{\frac{2C}{\varepsilon^2}T\}\left[\left(\frac{a_n}{\sigma^2 + a_n}\right)^{\frac{1}{\varepsilon^2}} + \left(\frac{a_n}{\sigma^\alpha + a_n}\right)^{\frac{1}{\varepsilon^2}}\right].
\end{aligned}$$

Hence we have

$$\begin{aligned}
&\limsup_{\varepsilon\to 0}\varepsilon^2 \log \mathbf{P}\left(\sup_{t\in[0,T]}\|X_t^\varepsilon - X_t^{\varepsilon,n}\| > 2\sigma\right)\\
&\leq 2CT + \max\left\{\log\frac{a_n}{\sigma^2 + a_n}, \log\frac{a_n}{\delta\sigma^\alpha + a_n}\right\}.
\end{aligned}$$

Since $a_n \to 0$ as $n \to \infty$, the proof is complete.

□

**Corollary 4.2.** *If (A1) – (A5) and (A7) hold, then $\{X^\varepsilon\}$ satisfy LDPin $C([0,T];H) \cap L^\alpha([0,T];V)$ with rate function (2.7).*

*Proof.* According to [43, Theorem 2.1] and section 3 one can conclude $\{X^\varepsilon\}$ satisfy $LDP$ with the following rate function

$$\tilde{I}(f) := \sup_{r>0}\liminf_{n\to\infty}\inf_{g\in S_r(f)} I^n(g) = \sup_{r>0}\limsup_{n\to\infty}\inf_{g\in S_r(f)} I^n(g).$$

where $S_r(f)$ is the closed ball in $C([0,T];H) \cap L^\alpha([0,T];V)$ centered at $f$ with radius $r$ and $I^n$ is given by

$$I^n(z) := \inf\{\frac{1}{2}\int_0^T \|\phi_s\|_U^2 \mathrm{d}s :\ z = z^{n,\phi},\ \phi \in L^2([0,T],U)\}, \tag{4.4}$$

where $z^{n,\phi}$ is the unique solution of following equation

$$\frac{\mathrm{d}z_t^n}{\mathrm{d}t} = A(t, z_t^n) + P_n B(t, z_t^n)\phi_t,\ z_0^n = x.$$

Now we only need to prove $\tilde{I} = I$, i.e.

$$I(f) = \sup_{r>0} \liminf_{n\to\infty} \inf_{g\in S_r(f)} I^n(g).$$

We will first show that for any $r > 0$

$$I(f) \geq \liminf_{n\to\infty} \inf_{g\in S_r(f)} I^n(g).$$

We assume $I(f) < \infty$, then by Lemma 3.2 there exists $\phi$ such that

$$f = z^\phi \quad \text{and} \quad I(f) = \frac{1}{2}\int_0^T \|\phi_s\|_U^2 \mathrm{d}s.$$

Since $z^{n,\phi} \to z^\phi$, for $n$ large enough we have

$$f_n := z^{n,\phi} \in S_r(f).$$

Notice $I^n(f_n) \leq \frac{1}{2}\int_0^T \|\phi_s\|_U^2 \mathrm{d}s$, hence we have

$$\liminf_{n\to\infty} \inf_{g\in S_r(f)} I^n(g) \leq \liminf_{n\to\infty} I^n(f_n) \leq I(f).$$

Since $r$ is arbitrary we have proved the lower bound

$$I(f) \geq \sup_{r>0} \liminf_{n\to\infty} \inf_{g\in S_r(f)} I^n(g).$$

For the upper bound we can proceed as in finite dimensional case in [37, Lemma 4.6] to show

$$\limsup_{n\to\infty} \inf_{g\in S_r(f)} I^n(g) \geq \inf_{g\in S_r(f)} I(g)$$

Hence we have

$$\sup_{r>0} \limsup_{n\to\infty} \inf_{g\in S_r(f)} I^n(g) \geq \sup_{r>0} \inf_{g\in S_r(f)} I(g) \geq I(f).$$

Hence the proof is complete. □

# 5 From boundedness to linear growth: drop $(A7)$

In order to drop the additional assumption $(A7)$ on $B$, we need to use some truncation techniques. This part is a modification of the argument in [37, 11].

**Lemma 5.1.** *Assume* $(A1)-(A4)$ *hold, then*

$$\lim_{R\to\infty} \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{P}\Big( \sup_{t\in[0,T]} \|X_t^\varepsilon\|_H^2 + \delta \int_0^T \|X_t^\varepsilon\|_V^\alpha \mathrm{d}t > R\Big) = -\infty. \tag{5.1}$$

*Proof.* By using Itô's formula and $(A2)-(A3)$, for $\varepsilon < \frac{1}{2}$ we have

$$
\begin{aligned}
\mathrm{d}\|X_t^\varepsilon\|_H^2 &= \left(2_{V^*}\langle A(t,X_t^\varepsilon), X_t^\varepsilon\rangle_V + \varepsilon^2 \|B(t,X_t^\varepsilon)\|_2^2\right)\mathrm{d}t + 2\varepsilon\langle X_t^\varepsilon, (B(t,X_t^\varepsilon)\mathrm{d}W_t\rangle_H \\
&\le -\frac{\delta}{2}\|X_t^\varepsilon\|_V^\alpha \mathrm{d}t + C(1+\|X_t^\varepsilon\|_H^2)\mathrm{d}t + 2\varepsilon \mathrm{d}M_t^{(n)}
\end{aligned}
$$

where $M_t^{(n)} := \int_0^t \langle X_s^\varepsilon, B(s,X_s^\varepsilon)\mathrm{d}W_s\rangle_H$ is a local martingale. Then

$$
\|X_t^\varepsilon\| \le C\int_0^t (1+\|X_s^\varepsilon\|_H^2)\mathrm{d}s + 4\varepsilon M_t^{(n)}
$$

where $\|X_t^\varepsilon\| := \|X_t^\varepsilon\|_H^2 + \delta\int_0^t \|X_s^\varepsilon\|_V^\alpha \mathrm{d}s$. The quadratic variation process of $M^{(n)}$ satisfies

$$
d\langle M^{(n)}\rangle_t \le C\|X_t^\varepsilon\|_H^2(1+\|X_t^\varepsilon\|_H^2)\mathrm{d}t.
$$

Define $\varphi_\theta(y) = (1+y)^\theta$ for some $\theta > 0$, then

$$
\begin{aligned}
\mathrm{d}\varphi_\theta(\|X_t^\varepsilon\|) \le&\theta(1+\|X_t^\varepsilon\|)^{\theta-1}\left(C(1+\|X_t^\varepsilon\|_H^2)\mathrm{d}t + 2\varepsilon \mathrm{d}M_t^{(n)}\right) \\
&+ 8C\varepsilon^2\theta(\theta-1)(1+\|X_t^\varepsilon\|)^{\theta-2}\|X_t^\varepsilon\|_H^2\left(1+\|X_t^\varepsilon\|_H^2\right)\mathrm{d}t \\
\le& C(\theta+\theta^2\varepsilon^2)\varphi_\theta(\|X_t^\varepsilon\|)\mathrm{d}t + \mathrm{d}\beta_t
\end{aligned}
\tag{5.2}
$$

where $\beta_t$-term is a local martingale. We also omit the standard localization procedure here for simplicity. Let $\theta = \frac{1}{\varepsilon^2}$ we know

$$
N_t := \mathrm{e}^{-\frac{2C}{\varepsilon^2}t}\varphi_{\frac{1}{\varepsilon^2}}(\|X_t^\varepsilon\|)
$$

is a supermartingale. Hence

$$
\begin{aligned}
&\mathbf{P}\left(\sup_{t\in[0,T]}\|X_t^\varepsilon\|_H^2 + \delta\int_0^T \|X_t^\varepsilon\|_V^\alpha \mathrm{d}t > R\right) \\
\le& \mathbf{P}\left(\sup_{t\in[0,T]} N_t > \exp\{-\frac{2C}{\varepsilon^2}T\}(1+R)^{\frac{1}{\varepsilon^2}}\right) \\
\le& \exp\{\frac{2C}{\varepsilon^2}T\}(1+R)^{-\frac{1}{\varepsilon^2}}\mathbf{E}N_0 \\
=& \exp\{\frac{2C}{\varepsilon^2}T\}\left(\frac{1}{1+R}\right)^{\frac{1}{\varepsilon^2}}.
\end{aligned}
$$

Hence we have

$$
\limsup_{\varepsilon\to 0}\varepsilon^2\log\mathbf{P}\left(\sup_{t\in[0,T]}\|X_t^\varepsilon\| > R\right) \le \log\frac{1}{1+R} + 2CT.
$$

Then the proof is complete. □

Now we can finish the proof of Theorem 2.1.

**Proof of Theorem 2.1**: The proof here is a slight modification of [37, Theorem 4.13]. Define $\xi : V \to [0,1]$ be a $C_0^\infty$-function such that

$$\xi(v) := \begin{cases} 0, & \text{if } \|v\|_H > 2, \\ 1, & \text{if } \|v\|_H \le 1. \end{cases}$$

Let $\xi_N(v) = \xi(\frac{v}{N})$ and

$$B_N(t,v) = \xi_N(v)B(t,v) + (1-\xi_N(v))B(t,0).$$

Consider the mollified problem for equation (2.3):

$$\mathrm{d}X^\varepsilon_{t,N} = A(t, X^\varepsilon_{t,N})\mathrm{d}t + \varepsilon B_N(t, X^\varepsilon_{t,N})\mathrm{d}W_t,\ X_0 = x. \tag{5.3}$$

It's easily to see that $A, B_N$ satisfy $(A1)-(A5)$ and $(A7)$. Hence by Corollary 4.2 we know $\{X^\varepsilon_N\}_{\varepsilon>0}$ satisfy large deviation principle on $C([0,T];H)\cap L^\alpha([0,T];V)$ with the following mollified rate function

$$I_N(z) := \inf\{\frac{1}{2}\int_0^T \|\phi_s\|_U^2 \mathrm{d}s :\ z = z_N^\phi,\ \phi \in L^2([0,T],U)\}, \tag{5.4}$$

where $z_N^\phi$ is the unique solution of following equation

$$\frac{\mathrm{d}z_{t,N}}{\mathrm{d}t} = A(t, z_{t,N}) + B_N(t, z_{t,N})\phi_t,\ z_{0,N} = x.$$

Let $N \to \infty$, then the LDP for $\{X^\varepsilon\}$ can be derived as in the finite dimensional case. According to Lemma 3.2, $I$ defined in (2.7) is a (good) rate function. Notice $I_N(z) = I(z)$ for any $z \in C([0,T];H)\cap L^\alpha([0,T];V)$ satisfy

$$\|z\|_T := \sup_{t\in[0,T]} \|z_t\|_H \le N.$$

We now first show that for any open set $G \subseteq C([0,T];H)\cap L^\alpha([0,T];V)$

$$\liminf_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{P}\left(X^\varepsilon \in G\right) \ge -\inf_{z\in G} I(z).$$

Obviously, we only need to prove that for all $\overline{z} \in G$ with $\overline{z}_0 = x$

$$\liminf_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{P}\left(X^\varepsilon \in G\right) \ge -I(\overline{z}).$$

Choose $R > 0$ such that $\|\overline{z}\|_T < R$ and set

$$N_R = \{z \in C([0,T];H)\cap L^\alpha([0,T];V) : \|z\|_T < R\}.$$

Then we have

$$\begin{aligned}\liminf_{\varepsilon\to 0}\varepsilon^2\log\mathbf{P}\left(X^\varepsilon\in G\right) &\geq \liminf_{\varepsilon\to 0}\varepsilon^2\log\mathbf{P}\left(X^\varepsilon\in G\cap N_R\right)\\ &= \liminf_{\varepsilon\to 0}\varepsilon^2\log\mathbf{P}\left(X_N^\varepsilon\in G\cap N_R\right)\\ &\geq -\inf_{z\in G\cap N_R} I_N(z)\\ &\geq -I(\overline{z}).\end{aligned}$$

Finally, given a closed set $F$ and an $L<\infty$, by Lemma 5.1 there exists $R$ such that

$$\begin{aligned}\limsup_{\varepsilon\to 0}\varepsilon^2\log\mathbf{P}\left(X^\varepsilon\in F\right) &\leq \limsup_{\varepsilon\to 0}\varepsilon^2\log\left(\mathbf{P}(X^\varepsilon\in F\cap\overline{N_R})+\mathbf{P}(X^\varepsilon\in N_R^c)\right)\\ &\leq \left(-\inf_{z\in F\cap\overline{N_R}} I_N(z)\right)\vee(-L)\\ &\leq -\left[\inf_{z\in F} I(z)\wedge L\right].\end{aligned}$$

Let $L\to\infty$, we obtain

$$\limsup_{\varepsilon\to 0}\varepsilon^2\log\mathbf{P}\left(X^\varepsilon\in F\right)\leq -\inf_{z\in F} I(z).$$

Now the proof of Theorem 2.1 is complete. □

# 6 LDP on $C([0,T];H)$ under weaker assumptions

As applications of Theorem 2.1, many examples of SPDE are studied in the next section. But Theorem 2.1 can't be applied to conclude LDP for stochastic fast diffusion equation in [29, 25] since strong Monotonicity condition $(A2)$ doesn't satisfy. Notice $(A2)$ mainly used to prove the additional convergence in $L^\alpha([0,T];V)$. Hence, if we only interested in LDP on $C([0,T];H)$, then we can prove the analogue of Theorem 2.1 under the following weaker assumptions than $(A1)-(A5)$, which can be applied to stochastic fast diffusion equation also.

For a fixed $\alpha>1$, there exist constants $\delta>0$ and $K$ such that the following conditions hold all $v,v_1,v_2\in V$ and all $t\in[0,T]$.

$(B1)$ Semicontinuity of $A$: The following map

$$\mathbb{R}\ni s\mapsto {}_{V^*}\langle A(t,v_1+sv_2),v\rangle_V$$

is continuous.

$(B2)$ Monotonicity of $(A,B)$:

$$2_{V^*}\langle A(t,v_1)-A(t,v_2),v_1-v_2\rangle_V+\|B(t,v_1)-B(t,v_2)\|_2^2\leq K\|v_1-v_2\|_H^2.$$

$(B3)$ Coercivity of $(A, B)$:

$$2_{V^*}\langle A(t,v), v\rangle_V + \|B(t,v)\|_2^2 + \delta\|v\|_V^\alpha \le K(1 + \|v\|_H^2),$$

$(B4)$ Boundedness of growth $A$:

$$\|A(t,v)\|_{V^*}^{\frac{\alpha}{\alpha-1}} \le K(1 + \|v\|_V^\alpha).$$

$(B5)$ Local Lipschitz and growth property of $B$:

$$\|B(t,v_1) - B(t,v_2)\|_2^2 \le K_N\|v_1 - v_2\|_H^2,\ \|v_i\|_H \le N,\ i = 1, 2.$$
$$\|B(t,v)\|_2^2 \le K(1 + \|v\|_H^2).$$

If $1 < \alpha < 2$, we assume also

$$\|B(t,v)\|_{L(U,V^*)} \le K(1 + \|v\|_V^{\alpha-1}).$$

$(B6)$ Approximated property of $B$: Suppose $B(\cdot, 0)$ is continuous on $[0, T]$ and there exist

$$H_n \subseteq H_{n+1},\ H_n \hookrightarrow V \text{compact},\ \bigcup_{n=1}^\infty H_n \subseteq H \text{ dense}$$

such that for any $M > 0$

$$\sup_{(t,v)\in[0,T]\times S_M} \|P_n B(t,v) - B(t,v)\|_2 \to 0\ (n \to \infty) \tag{6.1}$$

where $P_n : H \to H_n$ is standard projection operator and

$$S_M = \{v \in V : \|v\|_H \le M\}.$$

*Remark* 6.1. If we replace constant 1 in $(B3)$ and $(B4)$ by a integrable function, then the condition $(B1) - (B4)$ is exactly the same conditions in [24, 32] for the existence and uniqueness of the solution to (2.1).

**Theorem 6.1.** *Assume $(B1) - (B6)$ hold. For each $\varepsilon > 0$, let $X^\varepsilon = (X_t^\varepsilon)_{t\in[0,T]}$ be the solution to* (2.3)*. Then as $\varepsilon \to 0$, $\{X^\varepsilon\}$ satisfies the LDP on $C([0,T];H)$ with the good rate function $I$ which is given by* (2.7)*.*

The proof is a small modification (only consider the convergence in $C([0,T];H)$) of the argument for Theorem 2.1. We omit the details here.

# 7 Examples

Now we can apply the result to many stochastic evolution equations as applications. As a preparation we prove the following lemma first.

**Lemma 7.1.** *Let $(H, \langle\cdot,\cdot\rangle)$ is a Hilbert space, then for any $r \geq 0$ we have*

$$\langle \|a\|^r a - \|b\|^r b, a-b\rangle \geq 2^{-r}\|a-b\|^{r+2}, a, b \in H. \tag{7.1}$$

$$\|\|a\|^{r-1}a - \|b\|^{r-1}b\| \leq \max\{r,1\}\|a-b\|(\|a\|^{r-1} + \|b\|^{r-1}), a, b \in H. \tag{7.2}$$

*If $0 < r < 1$, then there exists a constant $C > 0$ such that*

$$||a|^{r-1}a - |b|^{r-1}b| \leq C|a-b|^r, \ \ a, b \in \mathbb{R}. \tag{7.3}$$

*Proof.* (i)If $\|a\| = \|b\|$, then (7.1) holds obviously.
If $\|a\| \neq \|b\|$, we may assume $\|a\| > \|b\|$, then

$$\begin{aligned}
& \langle \|a\|^r a - \|b\|^r b, a-b\rangle \\
= \ & \|b\|^r\|a-b\|^2 + (\|a\|^r - \|b\|^r)\langle a, a-b\rangle \\
= \ & \|b\|^r\|a-b\|^2 + (\|a\|^r - \|b\|^r)\cdot\frac{1}{2}(\|a\|^2 + \|a-b\|^2 - \|b\|^2) \\
> \ & \|b\|^r\|a-b\|^2 + \frac{1}{2}(\|a\|^r - \|b\|^r)\|a-b\|^2 \\
= \ & \frac{1}{2}(\|a\|^r + \|b\|^r)\|a-b\|^2 \\
\geq \ & 2^{-r}\|a-b\|^{r+2}
\end{aligned}$$

since $\|a-b\|^r \leq 2^{r-1}(\|a\|^r + \|b\|^r)$.
(ii) The proof of (7.2) and (7.3) is similar.

□

The first example is to obtain $LDP$ for a class of reaction-diffusion type SPDEs, which improve the result in [11].

**Example 7.2.** *Let $\Lambda$ is an open bounded domain in $\mathbb{R}^d$ with smooth boundary, $L$ is a negative definite self-adjoint operator on $H := L^2(\Lambda)$. Suppose*

$$V := \mathscr{D}(\sqrt{-L}), \quad \|v\|_V := \|\sqrt{-L}v\|_H.$$

*is a Banach space, $V \subseteq H$ is dense and compact, and $L$ can be extended as a continuous operator from $V$ to it's dual space $V^*$. Consider the following semilinear stochastic equation*

$$\mathrm{d}X_t^\varepsilon = (LX_t^\varepsilon + F(t, X_t^\varepsilon))\mathrm{d}t + \varepsilon B(t, X_t^\varepsilon)\mathrm{d}W_t, \ \ X_0^\varepsilon = x \in H, \tag{7.4}$$

*where $W_t$ is cylindrical Wiener process on another separable Hilbert space $U$ and*

$$F : [0,T] \times V \to V^*, \quad B : [0,T] \times V \to L_2(U;V).$$

*If $F$ and $B$ satisfy the following conditions:*

$$\begin{aligned} 2{}_{V^*}\langle F(t,v_1) - F(t,v_2), v_1 - v_2\rangle_V &\leq C\|v_1 - v_2\|_H^2, \\ \|B(t,v_1) - B(t,v_2)\|_2^2 &\leq C\|v_1 - v_2\|_H^2, \\ \|F(t,v)\|_{V^*}^2 + \|B(t,v)\|_2^2 &\leq C(1 + \|v\|_H^2), \ \ v, v_1, v_2 \in V. \end{aligned} \tag{7.5}$$

*where $C$ is a constant and $B(\cdot,0)$ is continuous on $[0,T]$, then $\{X^\varepsilon\}$ satisfy Large deviation principle on $C([0,T];H) \cap L^2([0,T];V)$.*

*Proof.* From the assumptions (7.5), it's easy to show that $(A1) - (A5)$ hold for $\alpha = 2$. Hence the conclusion follows from Theorem 2.1. □

*Remark* 7.1. (i) We can simply take $L$ as Laplace operator with Dirichlet boundary condition and $F(t,X_t) = -|X_t|^{p-2}X_t (1 \leq p \leq 2)$ as a concrete example.

(ii) Compare with the result in [11, Theorem 4.2](only time homogeneous case), the author in [11] need to assume $F$ is local Lipschitz and have more restricted range conditions:

$$F : [0,T] \times V \to H.$$

In our example we can allow $F$ is monotone and take values in $V^*$. Another difference is we also drop the non-degenerated condition $(A.4)$ in [11] on $B$.

(iii) Notice one can also take $B : V \to L_2(U;H)$ with locally compact range, which seems not allowed in [11, Theorem 4.2].

Second example is the stochastic generalized porous media equation, which have been studied quite intensively in recent years, see e.g.[5, 14, 29, 35, 42]. We use the same framework as in [35, 42].

**Example 7.3.** *(Stochastic porous media equation)*
*Let $(E, \mathcal{M}, \mathbf{m})$ be a separable probability space and $(L, \mathscr{D}(L))$ a negative definite self-adjoint linear operator on $(L^2(\mathbf{m}), \langle\cdot,\cdot\rangle)$ with spectrum contained in $(-\infty, -\lambda_0]$ for some $\lambda_0 > 0$. Then the embedding*

$$H^1 := \mathscr{D}(\sqrt{-L}) \subseteq L^2(\mathbf{m})$$

*is dense and continuous. Define $H$ is the dual Hilbert space of $H^1$ realized through this embedding. Assume $L^{-1}$ is continuous on $L^{r+1}(\mathbf{m})$.*

*For fixed $r > 1$, we consider the following triple*

$$V := L^{r+1}(\mathbf{m}) \subseteq H \subseteq V^*$$

*and the stochastic porous media equation*

$$\mathrm{d}X_t^\varepsilon=(L\Psi(t,X_t^\varepsilon)+\Phi(t,X_t^\varepsilon))\mathrm{d}t+\varepsilon B(t,X_t^\varepsilon)\mathrm{d}W_t,\ X_0^\varepsilon=x. \tag{7.6}$$

*where $W_t$ is cylindrical Wiener process on $L^2(\mathbf{m})$ and*

$$\Psi,\Phi:[0,T]\times\mathbb{R}\to\mathbb{R},\ B:[0,T]\times V\to L_2(L^2(\mathbf{m});H)$$

*are measurable and continuous in the second variable. Suppose $L^2(\mathbf{m})\subseteq\mathbf{H}$ is compact and $B:[0,T]\times V\to L_2(L^2(\mathbf{m}))$. If $B(\cdot,0)$ is continuous on $[0,T]$ and there exist two constants $\delta>0$ and $K$ such that*

$$\begin{aligned}&|\Psi(t,x)|+|\Phi(t,x)|\le K(1+|x|^r),\quad t\in[0,T],x\in\mathbb{R};\\&-\langle\Psi(t,u)-\Psi(t,v),u-v\rangle-\langle\Phi(t,u)-\Phi(t,v),L^{-1}(u-v)\rangle\\&\qquad\le-\delta\|u-v\|_V^{r+1}+K\|u-v\|_H^2;\\&\|B(t,u)-B(t,v)\|_2^2\le K\|u-v\|_H^2,\quad\|B(t,0)\|_2\le K.\ \ t\in[0,T],u,v\in V.\end{aligned}\tag{7.7}$$

*Then $\{X^\varepsilon\}$ satisfy Large deviation principle on $C([0,T];H)\cap L^{r+1}([0,T];V)$.*

*Proof.* From the assumptions and the relation

$${}_{V^*}\langle L\Phi(t,u)+\Phi(t,u),u\rangle_V=-\langle\Phi(t,u),u\rangle-\langle\Phi(t,u),L^{-1}u\rangle,$$

it's easy to show that $(A1)-(A5)$ holds for $\alpha=r+1$ from (7.7). We refer to [33, Example 4.1.11] for details, see also [14, 35, 32, 42]. Hence the conclusion follows from Theorem 2.1. □

*Remark* 7.2. (i) If we take $L$ the Laplace operator on a smooth bounded domain in a complete Riemannian manifold with Dirichlet boundary condition. A simple example for $\Psi$ and $\Phi$ satisfy 7.7 is given by

$$\Psi(t,x)=f(t)sgn(x)|x|^r,\quad\Phi(t,x)=g(t)x$$

for some strictly positive continuous function $f$ and bounded function $g$ on $[0,T]$.

(ii) This example generalized the main result in [35, Theorem 1.1] which obtain $LDP$ for stochastic porous media equations with additive noise. In [35] the authors mainly used the piecewise linear approximation to the path of Wiener process and generalized contraction principle, which are totally different with the weak convergence approach in this paper.

If we assume $0<r<1$ in the above example, then the equation turns into the stochastic version of classical fast diffusion equation. The behavior of the solutions to these two type of PDEs are essentially different, see e.g.[1].

**Example 7.4.** *(Stochastic fast diffusion equation)*
*Assume the same framework as Example 7.3 for $0 < r < 1$, i.e. assume the embedding $V := L^{r+1}(\mathbf{m}) \subseteq H$ is continuous and dense. We consider the equation*

$$\mathrm{d}X_t^\varepsilon = \left\{L\Psi(t, X_t^\varepsilon) + \gamma_t X_t^\varepsilon\right\}\mathrm{d}t + \varepsilon B(t, X_t^\varepsilon)\mathrm{d}W_t, \tag{7.8}$$

*where $\gamma : [0, T] \to \mathbb{R}$ is locally bounded and measurable and*

$$\Psi : [0, T] \times \mathbb{R} \to \mathbb{R}$$

*be measurable and continuous in the second variable.*

$$B : [0, T] \times V \to L_2(L^2(\mathbf{m}))$$

*are measurable and $W_t$ is cylindrical Wiener process on $L^2(\mathbf{m})$.*

*For a fixed number $r \in (0, 1)$, we assume $B(\cdot, 0)$ is continuous on $[0, T]$ and there exist constants $\delta > 0$ and $K$ such that for all $x, x_1, x_2 \in \mathbb{R}, t \in [0, T]$ and $u, v \in V$*

$$\begin{aligned}
&|\Psi(t, x)| \leq K(1 + |x|^r);\\
&(\Psi(t, x_1) - \Psi(t, x_2))(x_1 - x_2) \geq \delta|x_1 - x_2|^2(|x_1| \vee |x_2|)^{r-1};\\
&\|B(t, u) - B(t, v)\|_2^2 \leq K\|u - v\|_H^2, \quad \|B(t, 0)\|_2^2 \leq K;\\
&\|B(t, u)\|_{L(L^2(\mathbf{m}), V^*)} \leq K(1 + \|u\|_V^r).
\end{aligned} \tag{7.9}$$

*Then $\{X^\varepsilon\}$ satisfy Large deviation principle on $C([0, T]; H)$.*

*Proof.* Notice

$${}_{V^*}\langle L\Psi(t, u) + \gamma_t u, u\rangle_V = -\langle \Psi(t, u), u\rangle_{L^2} + \langle \gamma_t u, u\rangle_H$$

it's easy to show $(B1) - (B6)$ hold for $\alpha = r + 1$ under assumptions (7.9). Then the conclusion follows from Theorem 6.1. □

*Remark* 7.3. (i)In particular, if $\gamma = 0, B = 0$ and $\Psi(t, s) = s^r := |s|^{r-1}s$ for some $r \in (0, 1)$, then (7.8) reduces back to the classical fast-diffusion equation (see e.g. [1]).

(ii)Here we assume the embedding $L^{r+1}(\mathbf{m}) \subseteq H$ is continuous and dense for simplicity, see e.g. [33, Remark 4.1.15] and [25] for some sufficient conditions of this assumption. But in general $L^{r+1}(\mathbf{m})$ and $H$ are incomparable, then one need to consider the more general framework in [29].

**Example 7.5.** *(Stochastic "p-Laplacian" equation)*
*Let $\Lambda$ is an open bounded domain in $\mathbb{R}^d$ with smooth boundary, consider the following triple*

$$V := H_0^{1,p}(\Lambda) \subseteq H := L^2(\Lambda) \subseteq (H_0^{1,p}(\Lambda))^*$$

*and the stochastic "p-Laplacian" equation*

$$\mathrm{d}X_t^\varepsilon = \left[\mathbf{div}(|\nabla X_t^\varepsilon|^{p-2}\nabla X_t^\varepsilon) - \eta_t|X_t^\varepsilon|^{\tilde{p}-2}X_t^\varepsilon\right]\mathrm{d}t + \varepsilon B(t, X_t^\varepsilon)\mathrm{d}W_t, X_0^\varepsilon = x, \tag{7.10}$$

*where $2 \le p < \infty, 1 \le \tilde{p} \le p$, $B : [0,T] \times V \to L_2(H)$ and $W_t$ is a cylindrical Wiener process on $H$. Assume*

$$B(t,v) = \sum_{i=1}^{N} b_i(v) B_i(t),$$

*where $b_i(\cdot) : V \to \mathbb{R}$ and $B_i(\cdot) : [0,T] \to L_2(U;H)$. If $B_i(\cdot)$ are contiunous for $1 \le i \le n$, $\eta$ are continuous function such that*

$$0 \le \eta_t \le K, \quad t \in [0,T].$$
$$|b_i(u) - b_i(v)|^2 \le K \|u - v\|_H^2, \ u, v \in V.$$

*where $K$ is a constant. Then $\{X^\varepsilon\}$ satisfy Large deviation principle on $C([0,T];H) \cap L^p([0,T];V)$.*

*Proof.* The assumptions for existence and uniqueness of the solution was verified in [33, Example 4.1.9] for $\alpha = p$. Hence we only need to prove $(A2)$ hold here. By using (7.1) in Lemma 7.1 we have

$$\begin{aligned}
& {}_{V^*}\langle \mathbf{div}(|\nabla u|^{p-2}\nabla u) - \mathbf{div}(|\nabla v|^{p-2}\nabla v), u - v\rangle_V \\
= \; & -\int_\Lambda \langle |\nabla u(x)|^{p-2}\nabla u(x) - |\nabla v(x)|^{p-2}\nabla v(x), \nabla u(x) - \nabla v(x)\rangle \mathrm{d}x \\
\le \; & -2^{p-2} \int_\Lambda |\nabla u(x) - \nabla v(x)|^p \mathrm{d}x \\
\le \; & -c\|u - v\|_V^p.
\end{aligned}$$

where c is a positive constant and follows from Poincaré inequality.

By the monotonicity of function $|x|^{\tilde{p}-2}x$ we know

$${}_{V^*}\langle |u|^{\tilde{p}-2}u - |v|^{\tilde{p}-2}v, u - v\rangle_V \ge 0.$$

Hence $(A2)$ holds. Then the conclusion follows from Theorem 2.1. □

The following equation is taken from [24]. The main part of drift is a high order generalization of Laplacian operator.

**Example 7.6.** *Let $\Lambda$ is an open bounded domain in $\mathbb{R}^1$ and $m \in \mathbb{N}_+$, consider the following triple*

$$V := H_0^{m,p}(\Lambda) \subseteq H := L^2(\Lambda) \subseteq (H_0^{m,p}(\Lambda))^*$$

*and the stochastic evolution equation*

$$\begin{aligned}
\mathrm{d}X_t^\varepsilon(x) = & \big[(-1)^{m+1}\frac{\partial}{\partial x^m}(|\frac{\partial^m}{\partial x^m}X_t^\varepsilon(x)|^{p-2}\frac{\partial^m}{\partial x^m}X_t^\varepsilon(x)) + F(t, X_t(x))\big]\mathrm{d}t \\
& + \varepsilon B(t, X_t^\varepsilon(x))\mathrm{d}W_t, \qquad X_0^\varepsilon = x
\end{aligned} \tag{7.11}$$

*where $2 \le p < \infty$ and $W_t$ is cylindrical Wiener process on $H$. Suppose*

$$F : [0,T] \times V \to V^*, \quad B(t,v) = \sum_{i=1}^{N} b_i(v) B_i(t)$$

*where $b_i(\cdot) : V \to \mathbb{R}$ and $B_i(\cdot) : [0,T] \to L_2(U;H)$. Assume $B_i$ are continuous for $1 \le i \le N$ and*

$$\begin{aligned} 2_{V^*}\langle F(t,v_1) - F(t,v_2), v_1 - v_2\rangle_V &\le C\|v_1 - v_2\|_H^2, \\ |b_i(v_1) - b_i(v_2)|^2 &\le C\|v_1 - v_2\|_H^2, \\ \|F(t,v)\|_{V^*}^{\frac{p}{p-1}} &\le C(1 + \|v\|_H^2), \ v, v_1, v_2 \in V. \end{aligned}$$

*where $K$ is a constant. Then $\{X^\varepsilon\}$ satisfy Large deviation principle on $C([0,T];H) \cap L^p([0,T];V)$.*

*Proof.* By using Lemma 7.1, the conclusions follow from Theorem 2.1 by the same arguments as in Example 7.5.

□


## Acknowledgements

The author would like to thank Professor Michael Röckner for his valuable discussions and suggestions, and also thank Professors Paul Dupuis and Amarjit Budhiraja for stimulating communications on their works. The useful comments from Professors Fengyu Wang and Xicheng Zhang are also acknowledged.


## References


[1] D.G. Aronson, *The porous medium equation,* Lecture Notes in Mathematics 1224, Springer, Berlin, 1–46, 1986.

[2] R.G. Azencott, *Grandes deviations et applications, Ecole d'Eté de Probabilités de Saint-Flour VII,* Lecture Notes in Mathematics 774, 1980.

[3] A. Bensoussan, *Filtrage optimale des systemes linéaires,* Dunod, Paris, 1971.

[4] A. Bensoussan and R. Temam, *Equations aux derives partielles stochastiques non linéaires,* Isr. J. Math. 11(1972), 95–129.

[5] V.I. Bogachev, G. Da Prato and M. Röckner, *Invariant measures of stochastic generalized porous medium equations,* Doklady Math. Russian Acad. Sci. 396(2004), no. 1, English translation: Russ. Math. Dokl.

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

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

[8] A. Budhiraja, P. Dupuis and V. Maroulas, *Large deviations for infinite dimensional stochastic dynamical systems,* to appear in Ann. Probab.

[9] W. Bryc, *Large deviations by the asymptotic value method,* In M. Pinsky, editor, Diffusion Processes and Related Problems in Analysis, vol.1, 447–472. Birkhäuser, Boston, 1990.

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

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

[12] G. Da Prato and M. Röckner, *Weak solutions to stochastic porous media equations,* J. Evolution Equ. 4(2004), 249–271.

[13] G. Da Prato and M. Röckner, *Invariant measures for a stochastic porous medium equation,* Stochstic Analysis and Related Topics (Hiroshi Kunita, Yoichiro Takanashi, and Shinzo Watanabe, eds.), Advanced Studies in Pure Mathematics, 41(2004), 13-29.

[14] G. Da Prato, M. Röckner, Rozovskii and F.-Y. Wang, *Strong solutions to stochastic generalized porous media equations: existence, uniqueness and ergodicity,* Comm. Part. Diff. Equat. 31(2006), no.2, 277-291.

[15] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions,* Encyclopedia of Mathematics and its Applications, Cambridge University Press. 1992.

[16] M.D. Donsker and S.R.S. Varadhan, *Asymptotic evalution of certain Markov process expectations for large time, I, II, III,* Comm. Pure Appl. Math. 28(1975), 1–47; 28(1975), 279–301; 29(1977), 389–461.

[17] P. Dupuis and R. Ellis, *A weak convergence approach to the theory of large deviations,* Wiley, New York. 1997.

[18] A. Dembo and O. Zeitouni, *Large deviations techniques and applications,* Springer-Verlag, New York. 2000.

[19] W.H. Fleming, *Exit probabilities and optimal stochastic control,* Appl. Math. Optim. 4(1978), 329–346.

[20] J. Feng and T.G. Kurtz, *Large Deviations of Stochastic Processes,* vol. 131 of *Mathematical Surveys and Monographs,* American Mathematical Society, Providence, RI, 2006.

[21] M.I. Freidlin, *Random perturbations of reaction-diffusion equations: the quasi-deterministic approximations,* Trans. Amer. Math. Soc. 305 (1988), 665–697.

[22] M.I. Freidlin and A.D. Wentzell, *Random perturbations of dynamical systems,* Translated from the Russian by Joseph Szu"cs. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 260. Springer-Verlag, New York, 1984.

[23] I. Gyöngy and A. Millet, *On discretization schemes for stochastic evolution equations,* Pot. Anal. 23(2005), 99–134.

[24] N.V. Krylov and B.L. Rozovskii, *Stochastic evolution equations,* Translated from Itogi Naukii Tekhniki, Seriya Sovremennye Problemy Matematiki 14(1979), 71–146, Plenum Publishing Corp. 1981.

[25] W. Liu and F.-Y. Wang, *Harnack inequality and strong Feller property for stochastic fast diffusion equations,* to appear in J. Math. Anal. Appl.

[26] E. Pardoux, *Equations aux dérivées partielles stochastiques non linéaires monotones,* Thesis, Université Paris XI, 1975.

[27] S. Peszat, *Large deviation principle for stochastic evolution equations,* Probab. Theory Relat. Fields. 98(1994), 113–136.

[28] A.A. Pukhalskii, *On the theory of large deviations,* Theory probab. Appl. 38(1993), 490–497.

[29] J. Ren, M. Röckner and F.-Y. Wang, *Stochastic generalized porous media and fast diffusion equations,* J. Diff. Equat. 238(2007), 118–152.

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

[31] 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.

[32] M. Röckner, *Introduction to Stochastic Partial Differential Equations,* Lecture Notes in Purdue University, 2006.

[33] C.Prévôt and M. Röckner, *A Concise Course on Stochastic Partial Differential Equations,* Lecture Notes in Mathematics 1905, Springer, 2007.

[34] M. Röckner, B. Schmuland and X. Zhang, *Yamada-Watanabe Theorem for stochastic evolution equations in infinite dimensions,* BiBos-Preprint.

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

[36] S.S. Sritharan and P. Sundar, *Large deviations for the two-dimensional Navier-Stokes equations with multiplicative noise,* Stoc. Proc. Appl. 116(2006), 1636–1659.

[37] D.W. Stroock, *An Introduction to the Theory of Large Deviations,* Spring-Verlag, New York, 1984.

[38] S.R.S. Varadhan, *Asymptotic probabilities and differential equations,* Comm. Pure Appl. Math. 19 (1966), 261–286.

[39] S.R.S. Varadhan, *Diffusion processes in a small time interval,* Comm. Pure Appl. Math. 20 (1967), 659–685.

[40] S.R.S. Varadhan, *Large Deviations and Applications,* CBMS 46, SIAM, Philadelphia, 1984.

[41] J.B. Walsh, *An introduction to stochastic partial differential equations,* Ecole d'Ete de Probabilite de Saint-Flour XIV (1984), P.L. Hennequin editor, Lecture Notes in Mathematics 1180 , 265–439.

[42] F.-Y. Wang, *Harnack Inequality and Applications for Stochastic Generalized Porous Media equations,* Ann. Probab. 35(2007), 1333-1350.

[43] L. Wu, *On large deviations for moving average processes,* In Probability, Finance and Insurance, pp.15-49, the proceeding of a Workshop at the University of Hong-Kong (15-17 July 2002), Eds: T.L. Lai, H.L. Yang and S.P. Yung. World Scientific 2004, Singapour.

[44] E. Zeidler, *Nonlinear Functional Analysis and its Applications, II/B, Nonlinear Monotone Operators,* Springer-Verlag, New York: 1990.

[45] X. Zhang, *On Stochastic evolution equations with non-Lipschitz coefficients,* Preprint.