J Stat Phys (2018) 170:845–861
https://doi.org/10.1007/s10955-018-1958-4

Moderate Deviations for the Langevin Equation with Strong Damping

Lingyan Cheng[1] · **Ruinan Li**[2] · **Wei Liu**[3,4]

Received: 8 March 2017 / Accepted: 9 January 2018 / Published online: 25 January 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract In this paper, we establish a moderate deviation principle for the Langevin dynamics with strong damping. The weak convergence approach plays an important role in the proof.

Keywords Stochastic Langevin equation · Large deviations · Moderate deviations

Mathematics Subject Classification 60H10 · 60F10

1 Introduction

For every $\varepsilon > 0$, consider the following Langevin equation with strong damping

$$\begin{cases} \ddot{q}^{\varepsilon}(t) = b(q^{\varepsilon}(t)) - \frac{\alpha(q^{\varepsilon}(t))}{\varepsilon}\dot{q}^{\varepsilon}(t) + \sigma(q^{\varepsilon}(t))\dot{B}(t), \\ q^{\varepsilon}(0) = q \in \mathbb{R}^d, \quad \dot{q}^{\varepsilon}(0) = p \in \mathbb{R}^d. \end{cases} \tag{1.1}$$

Here $B(t)$ is a d-dimensional standard Wiener process, defined on some complete stochastic basis $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$. The coefficients b, α and σ satisfy some regularity conditions (see

✉ Ruinan Li
ruinanli@amss.ac.cn

Lingyan Cheng
lycheng@tju.edu.cn

Wei Liu
wliu.math@whu.edu.cn

1 Center of Applied Mathematics, Tianjin University, Tianjin 300072, People's Republic of China

2 School of Statistics and Information, Shanghai University of International Business and Economics, Shanghai 201620, People's Republic of China

3 School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, People's Republic of China

4 Computational Science Hubei Key Laboratory, Wuhan University, Wuhan, Hubei 430072, People's Republic of China

Sect. 2 for details) such that for any fixed $\varepsilon > 0$, $T > 0$ and $k \geq 1$, Eq.(1.1) admits a unique solution q^ε in $L^k(\Omega; C([0,T];\mathbb{R}^d))$. Let $q_\varepsilon(t) := q^\varepsilon(t/\varepsilon)$, $t \geq 0$, then Eq. (1.1) becomes

$$\begin{cases} \varepsilon^2 \ddot{q}_\varepsilon(t) = b(q_\varepsilon(t)) - \alpha(q_\varepsilon(t))\dot{q}_\varepsilon(t) + \sqrt{\varepsilon}\sigma(q_\varepsilon(t))\dot{w}(t), \\ q_\varepsilon(0) = q \in \mathbb{R}^d, \quad \dot{q}_\varepsilon(0) = \frac{p}{\varepsilon} \in \mathbb{R}^d, \end{cases} \tag{1.2}$$

where $w(t) := \sqrt{\varepsilon}B(t/\varepsilon)$, $t \geq 0$, is also a $\mathbb{R}^d$-valued Wiener process.

In [3], Cerrai and Freidlin established a large deviation principle (LDP for short) for Eq. (1.2) as $\varepsilon \to 0+$. More precisely, for any $T > 0$, they proved that the family $\{q_\varepsilon\}_{\varepsilon>0}$ satisfies the LDP in the space $C([0,T];\mathbb{R}^d)$, with the same rate function I and the same speed function ε^{-1} that describe the LDP of the first order equation

$$\dot{g}_\varepsilon(t) = \frac{b(g_\varepsilon(t))}{\alpha(g_\varepsilon(t))} + \sqrt{\varepsilon}\frac{\sigma(g_\varepsilon(t))}{\alpha(g_\varepsilon(t))}\dot{w}(t), \quad g_\varepsilon(0) = q \in \mathbb{R}^d. \tag{1.3}$$

Explicitly, this means that

(1) for any constant $c > 0$, the level set $\{f; I(f) \leq c\}$ is compact in $C([0,T];\mathbb{R}^d)$;
(2) for any closed subset $F \subset C([0,T];\mathbb{R}^d)$,

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

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

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

The dynamics system (1.3) can be regarded as the random perturbation of the following deterministic differential equation

$$\dot{q}_0(t) = \frac{b(q_0(t))}{\alpha(q_0(t))}, \quad q_0(0) = q \in \mathbb{R}^d. \tag{1.4}$$

Roughly speaking, the LDP result in [3] shows that the asymptotic probability of $\mathbb{P}(\|q_\varepsilon - q_0\| \geq \delta)$ converges exponentially to 0 as $\varepsilon \to 0$ for any $\delta > 0$, where $\|\cdot\|$ is the sup-norm on $C([0,T];\mathbb{R}^d)$.

Similarly to the large deviations, the moderate deviations arise in the theory of statistical inference quite naturally. The moderate deviation principle (MDP for short) can provide us with the rate of convergence and a useful method for constructing asymptotic confidence intervals (see, e.g., recent works [6,8,9,11] and references therein). Usually, the quadratic form of the rate function corresponding to the MDP allows for the explicit minimization, and particularly it allows one to obtain an asymptotic evaluation for the exit time (see [10]). Recently, the study of the MDP estimates for stochastic (partial) differential equation has been carried out as well, see e.g. [1,7,12,13] and so on.

In this paper, we shall investigate the MDP problem for the family $\{q_\varepsilon\}_{\varepsilon>0}$ on $C([0,T];\mathbb{R}^d)$. That is, the asymptotic behavior of the trajectory

$$X_\varepsilon(t) = \frac{1}{\sqrt{\varepsilon}h(\varepsilon)}\left(q_\varepsilon(t) - q_0(t)\right), \quad t \in [0,T]. \tag{1.5}$$

Here the deviation scale satisfies

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

Due to the complexity of q_ε, we mainly use the weak convergence approach to deal with this problem. Comparing with the approximating method used in Gao and Wang [5], our method is simpler since we only need the moment estimation rather than the exponential moment estimation of the solution.

The organization of this paper is as follows. In Sect. 2, we present the framework of the Langevin equation, and then state our main results. Section 3 is devoted to proving the MDP.

2 Framework and Main Results

Let $|\cdot|$ be the Euclidean norm of a vector in $\mathbb{R}^d$, $\langle\cdot,\cdot\rangle$ the inner production in $\mathbb{R}^d$, and $\|\cdot\|_{\mathrm{HS}}$ the Hilbert-Schmidt norm in $\mathbb{R}^{d\times d}$ (the space of $d\times d$ matrices). For a function $b:\mathbb{R}^d\to\mathbb{R}^d$, $Db=\left(\frac{\partial}{\partial x_j}b^i\right)_{1\le i,j\le d}$ is the Jacobian matrix of b. Recall that $\|\cdot\|$ is the sup-norm on $C([0,T];\mathbb{R}^d)$. Throughout this paper, $T>0$ is some fixed constant, $C(\cdot)$ is a positive constant depending on the parameters in the bracket and independent of ε. The value of $C(\cdot)$ may be different from line to line.

Assume that the coefficients b, α and σ in (1.2) satisfy the following hypothesis.

Hypothesis 2.1 *(a) The mappings $b:\mathbb{R}^d\to\mathbb{R}^d$ and $\sigma:\mathbb{R}^d\to\mathbb{R}^{d\times d}$ are continuously differentiable, and there exists some constant $K>0$ such that for all $x,y\in\mathbb{R}^d$,*

$$|b(x)-b(y)|\le K|x-y|, \tag{2.1}$$

and

$$\|\sigma(x)-\sigma(y)\|_{\mathrm{HS}}\le K|x-y|,\ \|\sigma(x)\|_{\mathrm{HS}}\le K.$$

Moreover, the matrix $\sigma(q)$ is invertible for any $q\in\mathbb{R}^d$, and $\sigma^{-1}:\mathbb{R}^d\to\mathbb{R}^{d\times d}$ is bounded.

(b) The mapping $\alpha:\mathbb{R}^d\to\mathbb{R}$ belongs to $C_b^1(\mathbb{R}^d)$ and there exist some constants $0<\alpha_0\le\alpha_1$ and $K>0$ such that

$$\alpha_0=\inf_{x\in\mathbb{R}^d}\alpha(x),\ \alpha_1=\sup_{x\in\mathbb{R}^d}\alpha(x) \text{ and } \sup_{x\in\mathbb{R}^d}|\nabla\alpha(x)|\le K.$$

Notice that:

(1) $\|Db\|_{\mathrm{HS}}\le K$ since b is continuously differentiable and satisfies (2.1);
(2) σ/α is Lipschitz continuous and bounded due to the Lipschitz-continuity and the boundness of the functions σ and $1/\alpha$.

Under Hypothesis 2.1, according to [5, Theorem 2.2], we know that the family $\left\{(g_\varepsilon-q_0)/[\sqrt{\varepsilon}h(\varepsilon)]\right\}_{\varepsilon>0}$ satisfies the LDP on $C([0,T];\mathbb{R}^d)$ with speed $h^2(\varepsilon)$ and a good rate function I given by

$$I(\psi)=\frac{1}{2}\inf_{h\in\mathcal{H};\psi=\Gamma_0(h)}\|h\|_{\mathcal{H}}^2, \tag{2.2}$$

where

$$\mathcal{H}:=\left\{h\in C([0,T];\mathbb{R}^d);\ h(t)=\int_0^t\dot{h}(s)ds,\ \|h\|_{\mathcal{H}}^2:=\int_0^T\left|\dot{h}(t)\right|^2dt<\infty\right\} \tag{2.3}$$

and

$$\Gamma_0(h(t))=\int_0^t D\left(\frac{b(q_0(s))}{\alpha(q_0(s))}\right)\Gamma_0(h(s))ds+\int_0^t\frac{\sigma(q_0(s))}{\alpha(q_0(s))}\dot{h}(s)ds, \tag{2.4}$$

with the convention $\inf\emptyset=\infty$. This special kind of LDP is just the MDP for the family $\{g_\varepsilon\}_{\varepsilon>0}$ (see [4]).

The main goal of this paper is to prove that the family $\{q_\varepsilon\}_{\varepsilon>0}$ satisfies the same MDP as the family $\{g_\varepsilon\}_{\varepsilon>0}$ on $C([0,T];\mathbb{R}^d)$.

Theorem 2.2 *Under Hypothesis 2.1, the family* $\{(q_\varepsilon - q_0)/[\sqrt{\varepsilon}h(\varepsilon)]\}_{\varepsilon>0}$ *obeys an LDP on* $C([0,T];\mathbb{R}^d)$ *with the speed function* $h^2(\varepsilon)$ *and the rate function* I *given by (2.2).*

3 Proof of MDP

3.1 Weak Convergence Approach in LDP

In this subsection, we will give the general criteria for the LDP given in [2].

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with an increasing family $\{\mathcal{F}_t\}_{0\le t\le T}$ of the sub-σ-fields of $\mathcal{F}$ satisfying the usual conditions. Let $\mathcal{E}$ be a Polish space with the Borel σ-field $\mathcal{B}(\mathcal{E})$. The Cameron-Martin space associated with the Wiener process $\{w(t)\}_{0\le t\le T}$ (defined on the filtered probability space given above) is given by (2.3). See [4]. The space $\mathcal{H}$ is a Hilbert space with inner product

$$\langle h_1, h_2\rangle_{\mathcal{H}} := \int_0^T \left\langle \dot{h}_1(s), \dot{h}_2(s)\right\rangle ds.$$

Let $\mathcal{A}$ denote the class of all $\{\mathcal{F}_t\}_{0\le t\le T}$-predictable processes belonging to $\mathcal{H}$ a.s.. Define for any $N \in \mathbb{N}$,

$$S_N := \left\{h \in \mathcal{H};\ \int_0^T \left|\dot{h}(s)\right|^2 ds \le N\right\}.$$

Consider the weak convergence topology on $\mathcal{H}$, i.e., for any $h_n, h \in \mathcal{H}, n \ge 1$, h_n converges weakly to h as $n \to +\infty$ if

$$\langle h_n - h, g\rangle_{\mathcal{H}} \to 0,\ \text{as}\ n \to +\infty,\ \forall g \in \mathcal{H}.$$

It is easy to check that S_N is a compact set in $\mathcal{H}$ under the weak convergence topology. Define

$$\mathcal{A}_N := \{\phi \in \mathcal{A};\ \phi(\omega) \in S_N,\ \mathbb{P}\text{-a.s.}\}.$$

We present the following result from Budhiraja et al. [2].

Theorem 3.1 *([2]) Let* $\mathcal{E}$ *be a Polish space with the Borel* σ*-field* $\mathcal{B}(\mathcal{E})$*. For any* $\varepsilon > 0$*, let* Γ_ε *be a measurable mapping from* $C([0,T];\mathbb{R}^d)$ *into* $\mathcal{E}$*. Let* $X_\varepsilon(\cdot) := \Gamma_\varepsilon(w(\cdot))$*. Suppose there exists a measurable mapping* $\Gamma_0 : C([0,T];\mathbb{R}^d) \to \mathcal{E}$ *such that*

(a) for every $N < +\infty$*, the set*

$$\left\{\Gamma_0\left(\int_0^{\cdot} \dot{h}(s)ds\right);\ h \in S_N\right\}$$

is a compact subset of $\mathcal{E}$*;*

(b) for every $N < +\infty$ *and any family* $\{h^\varepsilon\}_{\varepsilon>0} \subset \mathcal{A}_N$ *satisfying that* h^ε *(as* S_N*-valued random elements) converges in distribution to* $h \in \mathcal{A}_N$ *as* $\varepsilon \to 0$*,*

$$\Gamma_\varepsilon\left(w(\cdot) + \frac{1}{\sqrt{\varepsilon}}\int_0^{\cdot} \dot{h}^\varepsilon(s)ds\right)\ \textit{converges to}\ \Gamma_0\left(\int_0^{\cdot} \dot{h}(s)ds\right)$$

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

Then the family $\{X_\varepsilon\}_{\varepsilon>0}$ *satisfies the LDP on* $\mathcal{E}$ *with the rate function* I *given by*

$$I(g) := \inf_{h\in\mathcal{H}; g=\Gamma_0\left(\int_0^\cdot \dot{h}(s)ds\right)} \left\{\frac{1}{2}\int_0^T \left|\dot{h}(s)\right|^2 ds\right\}, \; g\in\mathcal{E}, \tag{3.1}$$

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

3.2 Reduction to the Bounded Case

Under Hypothesis 2.1, for every fixed $\varepsilon > 0$, Eq. (1.2) admits a unique solution q_ε in $L^k(\Omega; C([0,T];\mathbb{R}^d))$. According to the proof of Theorem 3.3 in [3], we know that the solution q_ε of Eq. (1.2) can be expressed in the following form:

$$q_\varepsilon(t) = q + \int_0^t \frac{b(q_\varepsilon(s))}{\alpha(q_\varepsilon(s))}ds + \sqrt{\varepsilon}\int_0^t \frac{\sigma(q_\varepsilon(s))}{\alpha(q_\varepsilon(s))}dw(s) + R_\varepsilon(t), \tag{3.2}$$

where

$$\begin{aligned} R_\varepsilon(t) &:= \frac{p}{\varepsilon}\int_0^t e^{-A_\varepsilon(s)}ds - \frac{1}{\alpha(q_\varepsilon(t))}\int_0^t e^{-A_\varepsilon(t,s)}b(q_\varepsilon(s))ds \\ &\quad + \int_0^t \left(\int_0^s e^{-A_\varepsilon(s,r)}b(q_\varepsilon(r))dr\right)\frac{1}{\alpha^2(q_\varepsilon(s))}\langle\nabla\alpha(q_\varepsilon(s)), \dot{q}_\varepsilon(s)\rangle ds \\ &\quad - \frac{1}{\alpha(q_\varepsilon(t))}H_\varepsilon(t) + \int_0^t \frac{1}{\alpha^2(q_\varepsilon(s))}H_\varepsilon(s)\langle\nabla\alpha(q_\varepsilon(s)), \dot{q}_\varepsilon(s)\rangle ds \\ &=: \sum_{k=1}^5 I_\varepsilon^k(t), \end{aligned} \tag{3.3}$$

with

$$\begin{aligned} A_\varepsilon(t,s) &:= \frac{1}{\varepsilon^2}\int_s^t \alpha(q_\varepsilon(r))dr, \quad A_\varepsilon(t) := A_\varepsilon(t,0), \\ H_\varepsilon(t) &:= \sqrt{\varepsilon}e^{-A_\varepsilon(t)}\int_0^t e^{A_\varepsilon(s)}\sigma(q_\varepsilon(s))dw(s). \end{aligned}$$

We denote the solution functional from $C([0,T];\mathbb{R}^d)$ into $C([0,T];\mathbb{R}^d)$ by $\mathcal{G}_\varepsilon$, i.e.,

$$\mathcal{G}_\varepsilon(w(t)) := q_\varepsilon(t), \; \forall t\in[0,T]. \tag{3.4}$$

Let

$$X_\varepsilon(t) := \Gamma_\varepsilon(w(t)) := \frac{\mathcal{G}_\varepsilon(w(t)) - q_0(t)}{\sqrt{\varepsilon}h(\varepsilon)}, \; \forall t\in[0,T]. \tag{3.5}$$

Then X_ε solves the following equation

$$\begin{aligned} X_\varepsilon(t) &= \frac{1}{\sqrt{\varepsilon}h(\varepsilon)}\int_0^t \left[\frac{b(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon(s))}{\alpha(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon(s))} - \frac{b(q_0(s))}{\alpha(q_0(s))}\right]ds \\ &\quad + \frac{1}{h(\varepsilon)}\int_0^t \frac{\sigma(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon(s))}{\alpha(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon(s))}dw(s) + \frac{R_\varepsilon(t)}{\sqrt{\varepsilon}h(\varepsilon)}, t\in[0,T]. \end{aligned} \tag{3.6}$$

We shall prove that $\{X_\varepsilon\}_{\varepsilon>0}$ obeys an LDP on $C([0,T];\mathbb{R}^d)$ with speed function $h^2(\varepsilon)$ and the rate function I given by (2.2).

Since the family $\{q_\varepsilon\}_{\varepsilon>0}$ satisfies the LDP in the space $C([0,T];\mathbb{R}^d)$ with the rate function I and the speed function ε^{-1} under Hypothesis 2.1 (see Cerrai and Freidlin [3]), there exist some positive constants R, C such that

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\left(\|q_\varepsilon\| \geq R\right) \leq -C.$$

Noticing (1.6), we have

$$\limsup_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\left(\|q_\varepsilon\| \geq R\right) = -\infty. \tag{3.7}$$

For any fixed constant $M > R$, define

$$b^M(x) := \begin{cases} b(x), & |x| < M; \\ g(x), & M \leq |x| < M+1; \\ 0, & |x| \geq M+1, \end{cases}$$

where $g(x)$ is some infinitely differentiable function on $\mathbb{R}^d$ such that $b^M(x)$ is continuous differentiable on $\mathbb{R}^d$. Then for all $t \in [0,T]$, we denote

$$\begin{aligned} q_0^M(t) &:= q + \int_0^t \frac{b^M(q_0^M(s))}{\alpha(q_0^M(s))} ds; \\ q_\varepsilon^M(t) &:= q + \int_0^t \frac{b^M(q_\varepsilon^M(s))}{\alpha(q_\varepsilon^M(s))} ds + \sqrt{\varepsilon} \int_0^t \frac{\sigma(q_\varepsilon^M(s))}{\alpha(q_\varepsilon^M(s))} dw(s) + R_\varepsilon^M(t); \\ X_\varepsilon^M(t) &:= \frac{1}{\sqrt{\varepsilon}h(\varepsilon)} \int_0^t \left[\frac{b^M(q_0^M(s) + \sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^M(s))}{\alpha(q_0^M(s) + \sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^M(s))} - \frac{b^M(q_0^M(s))}{\alpha(q_0^M(s))} \right] ds \\ &\quad + \frac{1}{h(\varepsilon)} \int_0^t \frac{\sigma(q_0^M(s) + \sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^M(s))}{\alpha(q_0^M(s) + \sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^M(s))} dw(s) + \frac{R_\varepsilon^M(t)}{\sqrt{\varepsilon}h(\varepsilon)}, \end{aligned}$$

where the expression of $R_\varepsilon^M(t)$ is similar to Eq. (3.3) with b^M, q_ε^M in place of b, q_ε.

Notice that $\|q_0\|$ is finite by the continuity of b and α. Hence, we can choose M large enough such that $q_0(t) = q_0^M(t)$, for all $t \in [0,T]$. Then for some M large enough, by Eq. (3.7), for all $\delta > 0$, we have

$$\begin{aligned} &\limsup_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\left(\left\| X_\varepsilon - X_\varepsilon^M \right\| > \delta \right) \\ =&\limsup_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\left(\left\| \frac{q_\varepsilon - q_\varepsilon^M}{\sqrt{\varepsilon}h(\varepsilon)} \right\| > \delta \right) \\ \leq&\limsup_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}\left(\left\| q_\varepsilon - q_\varepsilon^M \right\| > 0 \right) \\ \leq&\limsup_{\varepsilon\to 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{P}(\|q_\varepsilon\| \geq M) = -\infty, \end{aligned} \tag{3.8}$$

which means that X_ε is $h^2(\varepsilon)$-exponentially equivalent to X_ε^M. Hence, to prove the LDP for $\{X_\varepsilon\}_{\varepsilon>0}$ on $C([0,T];\mathbb{R}^d)$, it is enough to prove that for $\{X_\varepsilon^M\}_{\varepsilon>0}$, which is the task of the next part.

3.3 The LDP for $\{X_\varepsilon^M\}_{\varepsilon>0}$

In this subsection, we prove that for some fixed constant M large enough , $\{X_\varepsilon^M\}_{\varepsilon>0}$ obeys an LDP on $C([0,T];\mathbb{R}^d)$ with speed function $h^2(\varepsilon)$ and the rate function I given by (2.2). Without loss of generality, we assume that b is bounded, i.e., $|b| \le K$ for some positive constant K. Then $\frac{b}{\alpha}$ is also Lipschitz continuous and bounded, and by the differentiability of $\frac{b}{\alpha}$, $D(\frac{b}{\alpha})$ is also bounded. From now on, we can drop the M in the notations for the sake of simplicity.

3.3.1 Skeleton Equations

For any $h \in \mathcal{H}$, consider the deterministic equation:

$$g^h(t) = \int_0^t D\left(\frac{b(q_0(s))}{\alpha(q_0(s))}\right) g^h(s)ds + \int_0^t \frac{\sigma(q_0(s))}{\alpha(q_0(s))}\dot{h}(s)ds. \tag{3.9}$$

Lemma 3.2 *Under Hypothesis 2.1, for any $h \in \mathcal{H}$, Eq. (3.9) admits a unique solution g^h in $C([0,T];\mathbb{R}^d)$, denoted by $g^h(\cdot)=:\Gamma_0\left(\int_0^\cdot \dot{h}(s)ds\right)$. Moreover, for any $N > 0$, there exists some positive constant $C(K, N, T, \alpha_0, \alpha_1)$ such that*

$$\sup_{h\in S_N} \left\| g^h \right\| \le C(K, N, T, \alpha_0, \alpha_1). \tag{3.10}$$

Proof The existence and uniqueness of the solution can be proved similarly to the case of stochastic differential equation (1.3), but much more simply. (3.10) follows from the boundness conditions of the coefficient functions and Gronwall's inequality. Here we omit the relative proof. □

Proposition 3.3 *Under Hypothesis 2.1, for every positive number $N < +\infty$, the family*

$$K_N := \left\{\Gamma_0\left(\int_0^\cdot \dot{h}(s)ds\right); h \in S_N\right\}$$

is compact in $C([0,T];\mathbb{R}^d)$.

Proof To prove this proposition, it is sufficient to prove that the mapping Γ_0 defined in Lemma 3.2 is continuous from S_N to $C([0,T];\mathbb{R}^d)$, since the fact that K_N is compact follows from the compactness of S_N under the weak topology and the continuity of the mapping Γ_0 from S_N to $C([0,T];\mathbb{R}^d)$.

Assume that $h_n \to h$ weakly in S_N as $n \to \infty$. We consider the following equation

$$\begin{aligned}
&g^{h_n}(t) - g^h(t) \\
&= \int_0^t D\left(\frac{b(q_0(s))}{\alpha(q_0(s))}\right)\left(g^{h_n}(s) - g^h(s)\right)ds + \int_0^t \frac{\sigma(q_0(s))}{\alpha(q_0(s))}\left(\dot{h}_n(s) - \dot{h}(s)\right)ds \\
&=: I_1^n(t) + I_2^n(t).
\end{aligned}$$

Due to Cauchy-Schwartz inequality and the boundness of functions σ, α, we know that for any $0 \le t_1 \le t_2 \le T$,

$$\begin{aligned} |I_2^n(t_2) - I_2^n(t_1)| =& \left| \int_{t_1}^{t_2} \frac{\sigma(q_0(s))}{\alpha(q_0(s))} \left(\dot{h}_n(s) - \dot{h}(s)\right) ds \right| \\ \le& \left(\int_{t_1}^{t_2} \left\| \frac{\sigma(q_0(s))}{\alpha(q_0(s))} \right\|_{HS}^2 ds \right)^{\frac{1}{2}} \cdot \left(\int_{t_1}^{t_2} \left| \dot{h}_n(s) - \dot{h}(s) \right|^2 ds \right)^{\frac{1}{2}} \\ \le& C(K, \alpha_0) N^{\frac{1}{2}} (t_2 - t_1)^{\frac{1}{2}}. \end{aligned} \tag{3.11}$$

Hence, the family of functions $\{I_2^n\}_{n \ge 1}$ is equiv-continuous in $C([0, T]; \mathbb{R}^d)$. Particularly, taking $t_1 = 0$, we obtain that

$$\left\| I_2^n \right\| \le C(K, N, T, \alpha_0) < \infty, \tag{3.12}$$

where $C(K, N, T, \alpha_0)$ is independent of n. Thus, by the Ascoli-Arzelá theorem, the set $\{I_2^n\}_{n \ge 1}$ is compact in $C([0, T]; \mathbb{R}^d)$.

On the other hand, for any $v \in \mathbb{R}^d$, by the boundness of σ/α, we know that the function $\frac{\sigma(q_0)}{\alpha(q_0)} v$ belongs to $L^2([0, T]; \mathbb{R}^d)$. Since $\dot{h}_n \to \dot{h}$ weakly in $L^2([0, T]; \mathbb{R}^d)$ as $n \to +\infty$, we know that

$$\left\langle I_2^n(t), v \right\rangle = \int_0^t \frac{\sigma(q_0(s))}{\alpha(q_0(s))} \left(\dot{h}_n(s) - \dot{h}(s)\right) v ds \to 0, \quad \text{as } n \to \infty. \tag{3.13}$$

Then by the compactness of $\{I_2^n\}_{n \ge 1}$, we have

$$\lim_{n \to \infty} \left\| I_2^n \right\| = 0. \tag{3.14}$$

Set $\zeta^n(t) = \sup_{0 \le s \le t} \left| g^{h_n}(s) - g^h(s) \right|$. By the boundness of $D(b/\alpha)$, we have

$$\zeta^n(t) \le C(K, \alpha_0, \alpha_1) \int_0^t \zeta^n(s) ds + \left\| I_2^n \right\|.$$

By Gronwall's inequality and (3.14), we have

$$\left\| g^{h_n} - g^h \right\| \le e^{C(K, \alpha_0, \alpha_1)T} \cdot \left\| I_2^n \right\| \to 0, \text{ as } n \to \infty,$$

which completes the proof. □

3.3.2 MDP

For any predictable process $\dot{u}$ taking values in $L^2([0, T]; \mathbb{R}^d)$, we denote by $q_\varepsilon^u(t)$ the solution of the following equation

$$\begin{cases} \varepsilon^2 \ddot{q}_\varepsilon^u(t) = b(q_\varepsilon^u(t)) - \alpha(q_\varepsilon^u(t)) \dot{q}_\varepsilon^u(t) + \sqrt{\varepsilon} \sigma(q_\varepsilon^u(t)) \dot{w}(t) \\ + \sqrt{\varepsilon} h(\varepsilon) \sigma(q_\varepsilon^u(t)) \dot{u}(t), \ t \in [0, T], \\ q_\varepsilon^u(0) = q \in \mathbb{R}^d, \quad \dot{q}_\varepsilon^u(0) = \frac{p}{\varepsilon} \in \mathbb{R}^d. \end{cases} \tag{3.15}$$

As is well known, for any fixed $\varepsilon > 0$, $T > 0$ and $k \ge 1$, this equation admits a unique solution q_ε^u in $L^k(\Omega; C([0, T]; \mathbb{R}^d))$ as follows

$$q_\varepsilon^u(t) = \mathcal{G}_\varepsilon \left(w(t) + h(\varepsilon) \int_0^t \dot{u}(s) ds \right),$$

where $\mathcal{G}_\varepsilon$ is defined by (3.4).

Lemma 3.4 *Under Hypothesis 2.1, for every fixed $N \in \mathbb{N}$ and $\varepsilon > 0$, let $u^\varepsilon \in \mathcal{A}_N$ and Γ_ε be given by (3.5). Then $X_\varepsilon^{u^\varepsilon}(\cdot) := \Gamma_\varepsilon\left(w(\cdot) + h(\varepsilon)\int_0^\cdot \dot{u}^\varepsilon(s)ds\right)$ is the unique solution of the following equation*

$$
\begin{aligned}
X_\varepsilon^{u^\varepsilon}(t) = &\int_0^t \frac{1}{\sqrt{\varepsilon}h(\varepsilon)}\left[\frac{b(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))}{\alpha(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))} - \frac{b(q_0(s))}{\alpha(q_0(s))}\right]ds \\
&+ \int_0^t \frac{\sigma(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))}{\alpha(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))}\dot{u}^\varepsilon(s)ds \\
&+ \frac{1}{h(\varepsilon)}\int_0^t \frac{\sigma(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))}{\alpha(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))}dw(s) + \frac{R_\varepsilon^{u^\varepsilon}(t)}{\sqrt{\varepsilon}h(\varepsilon)}, \quad t \in [0,T],
\end{aligned}
\tag{3.16}
$$

where

$$
\begin{aligned}
R_\varepsilon^{u^\varepsilon}(t) = &\frac{p}{\varepsilon}\int_0^t e^{-A_\varepsilon^{u^\varepsilon}(s)}ds - \frac{1}{\alpha(q_\varepsilon^{u^\varepsilon}(t))}\int_0^t e^{-A_\varepsilon^{u^\varepsilon}(t,s)}b(q_\varepsilon^{u^\varepsilon}(s))ds \\
&+ \int_0^t\left(\int_0^s e^{-A_\varepsilon^{u^\varepsilon}(s,r)}b(q_\varepsilon^{u^\varepsilon}(r))dr\right)\frac{1}{\alpha^2(q_\varepsilon^{u^\varepsilon}(s))}\left\langle\nabla\alpha(q_\varepsilon^{u^\varepsilon}(s)), \dot{q}_\varepsilon^{u^\varepsilon}(s)\right\rangle ds \\
&- \frac{1}{\alpha(q_\varepsilon^{u^\varepsilon}(t))}H_\varepsilon^{1,u^\varepsilon}(t) + \int_0^t \frac{1}{\alpha^2(q_\varepsilon^{u^\varepsilon}(s))}H_\varepsilon^{1,u^\varepsilon}(s)\left\langle\nabla\alpha(q_\varepsilon^{u^\varepsilon}(s)), \dot{q}_\varepsilon^{u^\varepsilon}(s)\right\rangle ds \\
&- \frac{1}{\alpha(q_\varepsilon^{u^\varepsilon}(t))}H_\varepsilon^{2,u^\varepsilon}(t) + \int_0^t \frac{1}{\alpha^2(q_\varepsilon^{u^\varepsilon}(s))}H_\varepsilon^{2,u^\varepsilon}(s)\left\langle\nabla\alpha(q_\varepsilon^{u^\varepsilon}(s)), \dot{q}_\varepsilon^{u^\varepsilon}(s)\right\rangle ds \\
=: &\sum_{k=1}^7 I_\varepsilon^{k,u^\varepsilon},
\end{aligned}
$$

with

$$
\begin{aligned}
A_\varepsilon^{u^\varepsilon}(t,s) &:= \frac{1}{\varepsilon^2}\int_s^t \alpha(q_\varepsilon^{u^\varepsilon}(r))dr, \quad A_\varepsilon^{u^\varepsilon}(t) = A_\varepsilon^{u^\varepsilon}(t,0), \\
H_\varepsilon^{1,u^\varepsilon}(t) &:= \sqrt{\varepsilon}e^{-A_\varepsilon^{u^\varepsilon}(t)}\int_0^t e^{A_\varepsilon^{u^\varepsilon}(s)}\sigma(q_\varepsilon^{u^\varepsilon}(s))dw(s), \\
H_\varepsilon^{2,u^\varepsilon}(t) &:= \sqrt{\varepsilon}h(\varepsilon)e^{-A_\varepsilon^{u^\varepsilon}(t)}\int_0^t e^{A_\varepsilon^{u^\varepsilon}(s)}\sigma(q_\varepsilon^{u^\varepsilon}(s))\dot{u}^\varepsilon(s)ds.
\end{aligned}
\tag{3.17}
$$

Furthermore, there exists a positive constant $\varepsilon_0 > 0$ such that for any $\varepsilon \in (0, \varepsilon_0]$,

$$
\mathbb{E}\left[\int_0^T \left|X_\varepsilon^{u^\varepsilon}(t)\right|^2 dt\right] \le C(K, N, T, \alpha_0, \alpha_1, |p|, |q|). \tag{3.18}
$$

Moveover, we have

$$
\mathbb{E}\left[\left\|X_\varepsilon^{u^\varepsilon}\right\|^2\right] \le C(K, N, T, \alpha_0, \alpha_1, |p|, |q|). \tag{3.19}
$$

To prove Lemma 3.4 and our main result, we present the following three lemmas. The first lemma is similar to [3, Lemma 3.1].

Lemma 3.5 *Under Hypothesis 2.1, for any $T > 0$, $k \geq 1$ and $N > 0$, there exists some constant $\varepsilon_0 > 0$ such that for any $u^\varepsilon \in \mathcal{A}_N$ and $\varepsilon \in (0, \varepsilon_0]$, we have*

$$\sup_{t\in[0,T]} \mathbb{E}\left[\left|H_\varepsilon^{1,u^\varepsilon}(t)\right|^k\right] \leq C(k, K, N, T, \alpha_0, \alpha_1)\left(|q|^k + |p|^k + 1\right)\varepsilon^{\frac{3k}{2}} + C(k, K)\varepsilon^{\frac{k}{2}} t^{\frac{k}{2}} e^{-\frac{k\alpha_0 t}{\varepsilon^2}}. \tag{3.20}$$

Moveover, we have

$$\mathbb{E}\left\|H_\varepsilon^{1,u^\varepsilon}\right\| \leq \sqrt{\varepsilon} C(K, N, T, \alpha_0, \alpha_1)(1 + |q| + |p|). \tag{3.21}$$

Proof Notice that Eq. (3.15) can be rewritten as the following equation: for all $t \in [0, T]$,

$$\begin{cases} \dot{q}_\varepsilon^{u^\varepsilon}(t) = p_\varepsilon^{u^\varepsilon}(t), \\ \varepsilon^2 \dot{p}_\varepsilon^{u^\varepsilon}(t) = b(q_\varepsilon^{u^\varepsilon}(t)) - \alpha(q_\varepsilon^{u^\varepsilon}(t)) p_\varepsilon^{u^\varepsilon}(t) + \sqrt{\varepsilon}\sigma(q_\varepsilon^{u^\varepsilon}(t))\dot{w}(t) + \sqrt{\varepsilon} h(\varepsilon)\sigma(q_\varepsilon^{u^\varepsilon}(t))\dot{u}^\varepsilon(t), \\ q_\varepsilon^{u^\varepsilon}(0) = q \in \mathbb{R}^d, \;\; p_\varepsilon^{u^\varepsilon}(0) = \frac{p}{\varepsilon} \in \mathbb{R}^d. \end{cases}$$

From the notation given in Eq. (3.17), we have

$$\dot{q}_\varepsilon^{u^\varepsilon}(t) = p_\varepsilon^{u^\varepsilon}(t) = \frac{1}{\varepsilon} e^{-A_\varepsilon^{u^\varepsilon}(t)} p + \frac{1}{\varepsilon^2}\int_0^t e^{-A_\varepsilon^{u^\varepsilon}(t,s)} b(q_\varepsilon^{u^\varepsilon}(s))ds + \frac{1}{\varepsilon^2} H_\varepsilon^{2,u^\varepsilon}(t) + \frac{1}{\varepsilon^2} H_\varepsilon^{1,u^\varepsilon}(t). \tag{3.22}$$

Integrating with respect to t, we obtain that

$$q_\varepsilon^{u^\varepsilon}(t) = q + \frac{1}{\varepsilon}\int_0^t e^{-A_\varepsilon^{u^\varepsilon}(s)} p ds + \frac{1}{\varepsilon^2}\int_0^t \int_0^s e^{-A_\varepsilon^{u^\varepsilon}(s,r)} b(q_\varepsilon^{u^\varepsilon}(r)) dr ds + \frac{1}{\varepsilon^2}\int_0^t H_\varepsilon^{2,u^\varepsilon}(s)ds + \frac{1}{\varepsilon^2}\int_0^t H_\varepsilon^{1,u^\varepsilon}(s)ds.$$

By Hypothesis 2.1 and Young's inequality for integral operators, we have

$$\begin{aligned} \left|q_\varepsilon^{u^\varepsilon}(t)\right| &\leq |q| + \frac{\varepsilon}{\alpha_0}|p| + C(K, T, \alpha_0)\int_0^t (1 + \left|q_\varepsilon^{u^\varepsilon}(s)\right|)ds \\ &\quad + C(K, \alpha_0)\sqrt{\varepsilon} h(\varepsilon)\int_0^t \left|\dot{u}^\varepsilon(s)\right| ds + \frac{1}{\varepsilon^2}\int_0^t \left|H_\varepsilon^{1,u^\varepsilon}(s)\right| ds \\ &\leq C(K, N, T, \alpha_0)\left(|q| + \varepsilon|p| + \sqrt{\varepsilon} h(\varepsilon)\right) \\ &\quad + \frac{1}{\varepsilon^2}\int_0^t \left|H_\varepsilon^{1,u^\varepsilon}(s)\right| ds + C(K, T, \alpha_0)\int_0^t \left|q_\varepsilon^{u^\varepsilon}(s)\right| ds. \end{aligned}$$

Since $\lim_{\varepsilon\to 0}\sqrt{\varepsilon} h(\varepsilon) = 0$, for ε small enough, by Gronwall's inequality,

$$\left|q_\varepsilon^{u^\varepsilon}(t)\right| \leq C(K, N, T, \alpha_0)(|q| + |p| + 1) + C(K, T, \alpha_0)\frac{1}{\varepsilon^2}\int_0^t \left|H_\varepsilon^{1,u^\varepsilon}(s)\right| ds. \tag{3.23}$$

Hence by the similar proof to that in [3, Lemma 3.1], we obtain (3.20) and (3.21). □

For $H_\varepsilon^{2,u^\varepsilon}(t)$, we have the following estimation.

Lemma 3.6 *Under Hypothesis 2.1, for any $T > 0$, $k \geq 1$ and $N \in \mathbb{N}$, there exists some constant $\varepsilon_0 > 0$ such that for any $u^\varepsilon \in \mathcal{A}_N$ and $\varepsilon \in (0, \varepsilon_0]$, we have*

$$\mathbb{E}\left[\left\|H_\varepsilon^{2,u^\varepsilon}\right\|^k\right] \leq C(K, N, \alpha_0)\varepsilon^{\frac{3k}{2}} h^k(\varepsilon). \tag{3.24}$$

Proof For any $t \in [0, T]$ and $u^\varepsilon \in \mathcal{A}_N$, by the boundness of σ and Cauchy-Schwarz inequality, we have

$$
\begin{aligned}
\left|H_\varepsilon^{2,u^\varepsilon}(t)\right| &= \left|\sqrt{\varepsilon}h(\varepsilon)e^{-A_\varepsilon^{u^\varepsilon}(t)}\int_0^t e^{A_\varepsilon^{u^\varepsilon}(s)}\sigma(q_\varepsilon^{u^\varepsilon}(s))\dot{u}^\varepsilon(s)ds\right| \\
&\le K\sqrt{\varepsilon}h(\varepsilon)e^{-A_\varepsilon^{u^\varepsilon}(t)}\int_0^t e^{A_\varepsilon^{u^\varepsilon}(s)}\left|\dot{u}^\varepsilon(s)\right|ds \\
&\le K\sqrt{\varepsilon}h(\varepsilon)e^{-A_\varepsilon^{u^\varepsilon}(t)}\left(\int_0^t e^{2A_\varepsilon^{u^\varepsilon}(s)}ds\right)^{\frac{1}{2}}\left(\int_0^T \left|\dot{u}^\varepsilon(s)\right|^2 ds\right)^{\frac{1}{2}} \\
&\le KN^{\frac{1}{2}}\sqrt{\varepsilon}h(\varepsilon)e^{-A_\varepsilon^{u^\varepsilon}(t)}\left(\int_0^t e^{2A_\varepsilon^{u^\varepsilon}(s)}ds\right)^{\frac{1}{2}}.
\end{aligned}
$$

Since $A_\varepsilon^{u^\varepsilon}(t) = \frac{1}{\varepsilon^2}\int_0^t \alpha(q_\varepsilon^{u^\varepsilon}(r))dr$, we have

$$
\begin{aligned}
\int_0^t e^{2A_\varepsilon^{u^\varepsilon}(s)}ds &= \int_0^t \frac{\varepsilon^2}{2\alpha(q_\varepsilon^{u^\varepsilon}(s))}de^{\frac{2}{\varepsilon^2}\int_0^s \alpha(q_\varepsilon^{u^\varepsilon}(r))dr} \\
&\le \frac{\varepsilon^2}{2\alpha_0}\int_0^t de^{\frac{2}{\varepsilon^2}\int_0^s \alpha(q_\varepsilon^{u^\varepsilon}(r))dr} \\
&= \frac{\varepsilon^2}{2\alpha_0}\left(e^{2A_\varepsilon^{u^\varepsilon}(t)} - 1\right).
\end{aligned}
$$

Hence

$$
\begin{aligned}
\left|H_\varepsilon^{2,u^\varepsilon}(t)\right| &\le KN^{\frac{1}{2}}\frac{\varepsilon^{\frac{3}{2}}h(\varepsilon)}{\sqrt{2\alpha_0}}e^{-A_\varepsilon^{u^\varepsilon}(t)}\left(e^{2A_\varepsilon^{u^\varepsilon}(t)} - 1\right)^{\frac{1}{2}} \\
&\le C(K, N, \alpha_0)\varepsilon^{\frac{3}{2}}h(\varepsilon)e^{-A_\varepsilon^{u^\varepsilon}(t)}e^{A_\varepsilon^{u^\varepsilon}(t)} \\
&= C(K, N, \alpha_0)\varepsilon^{\frac{3}{2}}h(\varepsilon),
\end{aligned}
$$

and furthermore

$$
\mathbb{E}\left\|H_\varepsilon^{2,u^\varepsilon}\right\|^k \le C(K, N, \alpha_0)\varepsilon^{\frac{3k}{2}}h^k(\varepsilon),
$$

which completes the proof. □

Lemma 3.7 *Under Hypothesis 2.1, for any $T > 0$ and any $u^\varepsilon \in \mathcal{A}_N$, we have*

$$
\mathbb{E}\left\|\frac{R_\varepsilon}{\sqrt{\varepsilon}h(\varepsilon)}\right\| \to 0, \quad as \quad \varepsilon \to 0. \tag{3.25}
$$

Moreover, we have

$$
\mathbb{E}\left[\left\|\frac{R_\varepsilon}{\sqrt{\varepsilon}h(\varepsilon)}\right\|^2\right] \to 0, \quad as \quad \varepsilon \to 0. \tag{3.26}
$$

Proof Similarly to the proof [3, (3.17)], under Hypothesis 2.1, we have

$$
\mathbb{E}\left\|\frac{\sum_{k=1}^5 I_\varepsilon^{k,u^\varepsilon}}{\sqrt{\varepsilon}h(\varepsilon)}\right\| \le \frac{1}{h(\varepsilon)}C(K, N, T, \alpha_0, \alpha_1, |p|, |q|) \to 0, \quad \text{as } \varepsilon \to 0. \tag{3.27}
$$

Next, we will estimate $\mathbb{E}\left\|\frac{I_\varepsilon^{6,u^\varepsilon}}{\sqrt{\varepsilon}h(\varepsilon)}\right\|$ and $\mathbb{E}\left\|\frac{I_\varepsilon^{7,u^\varepsilon}}{\sqrt{\varepsilon}h(\varepsilon)}\right\|$. By Lemma 3.6, we have

$$\mathbb{E}\left\|\frac{I_\varepsilon^{6,u^\varepsilon}}{\sqrt{\varepsilon}h(\varepsilon)}\right\| \le \frac{1}{\sqrt{\varepsilon}h(\varepsilon)\alpha_0}\mathbb{E}\left\|H_\varepsilon^{2,u^\varepsilon}\right\| \le \varepsilon C(K,N,\alpha_0) \to 0, \quad \text{as } \varepsilon \to 0. \tag{3.28}$$

By Cauchy-Schwarz inequality, we have

$$\begin{aligned}\mathbb{E}\left\|\frac{I_\varepsilon^{7,u^\varepsilon}}{\sqrt{\varepsilon}h(\varepsilon)}\right\| &\le \frac{C(K,\alpha_0)}{\sqrt{\varepsilon}h(\varepsilon)}\mathbb{E}\left[\sup_{t\in[0,T]}\int_0^t \left|H_\varepsilon^{2,u^\varepsilon}(s)\right|\cdot\left|\dot{q}_\varepsilon^{u^\varepsilon}(s)\right|ds\right]\\ &\le \frac{C(K,\alpha_0)}{\sqrt{\varepsilon}h(\varepsilon)}\left[\int_0^T \mathbb{E}\left[\left|H_\varepsilon^{2,u^\varepsilon}(s)\right|^2\right]ds\right]^{\frac{1}{2}}\cdot\left[\int_0^T \mathbb{E}\left[\left|\dot{q}_\varepsilon^{u^\varepsilon}(s)\right|^2\right]ds\right]^{\frac{1}{2}}.\end{aligned}$$

By (3.23), we have for all $\varepsilon > 0$ small enough,

$$\int_0^T \left|\dot{q}_\varepsilon^{u^\varepsilon}(s)\right|^2 ds \le C(K,N,T,\alpha_0,|p|,|q|) + \frac{C(K,T,\alpha_0)}{\varepsilon^4}\int_0^T \left|H_\varepsilon^{1,u^\varepsilon}(s)\right|^2 ds.$$

Hence, by (3.20) and Lemma 3.6, we have

$$\begin{aligned}&\mathbb{E}\left\|\frac{I_\varepsilon^{7,u^\varepsilon}}{\sqrt{\varepsilon}h(\varepsilon)}\right\|\\ &\le \frac{C(K,N,T,\alpha_0,|p|,|q|)}{\sqrt{\varepsilon}h(\varepsilon)}\left[\left(\int_0^T \mathbb{E}\left[\left|H_\varepsilon^{2,u^\varepsilon}(s)\right|^2\right]ds\right)^{\frac{1}{2}}\right.\\ &\quad\left.+\frac{C(K,N,T,\alpha_0)}{\varepsilon^{\frac{5}{2}}h(\varepsilon)}\left(\int_0^T \mathbb{E}\left[\left|H_\varepsilon^{2,u^\varepsilon}(s)\right|^2\right]ds\right)^{\frac{1}{2}}\cdot\left(\int_0^T \mathbb{E}\left[\left|H_\varepsilon^{1,u^\varepsilon}(s)\right|^2\right]ds\right)^{\frac{1}{2}}\right.\\ &\le \sqrt{\varepsilon}C(K,N,T,\alpha_0,\alpha_1,|p|,|q|) \to 0, \quad \text{as } \varepsilon \to 0.\end{aligned} \tag{3.29}$$

This together with (3.27) and (3.28) implies (3.25).

(3.26) can be easily obtained by applying the similar estimation process for

$$\mathbb{E}\left[\left\|\frac{I_\varepsilon^{i,u^\varepsilon}}{\sqrt{\varepsilon}h(\varepsilon)}\right\|^2\right], \; i = 1,2,3,\cdots,7,$$

as given above. Hence we omit the proof. □

Now we prove Lemma 3.4.

The proof of Lemma 3.4 For any $\varepsilon > 0$ and $u^\varepsilon \in \mathcal{A}_N$, define

$$d\mathbb{Q}^{u^\varepsilon} := \exp\left\{-h(\varepsilon)\int_0^t \dot{u}^\varepsilon(s)dw(s) - \frac{h^2(\varepsilon)}{2}\int_0^t \left|\dot{u}^\varepsilon(s)\right|^2 ds\right\}d\mathbb{P}.$$

Since $\frac{d\mathbb{Q}^{u^\varepsilon}}{d\mathbb{P}}$ is an exponential martingale, $\mathbb{Q}^{u^\varepsilon}$ is a probability measure on Ω. By Girsanov theorem, the process

$$\tilde{w}^\varepsilon(t) = w(t) + h(\varepsilon)\int_0^t \dot{u}^\varepsilon(s)ds$$

is a $\mathbb{R}^d$-valued Wiener process under the probability measure $\mathbb{Q}^{u^\varepsilon}$. Rewriting Eq. (3.16) with $\tilde{w}^\varepsilon(t)$, we obtain Eq. (3.6) with $\tilde{w}^\varepsilon(t)$ in place of $w(t)$. Let $X_\varepsilon^{u^\varepsilon}$ be the unique solution of Eq. (3.6) with $\tilde{w}^\varepsilon(t)$ on the space $(\Omega, \mathcal{F}, \mathbb{Q}^{u^\varepsilon})$. Then $X_\varepsilon^{u^\varepsilon}$ satisfies Eq. (3.16), $\mathbb{Q}^{u^\varepsilon}$-a.s.. By the equivalence of probability measures, $X_\varepsilon^{u^\varepsilon}$ satisfies Eq. (3.16), $\mathbb{P}$-a.s..

Now we prove (3.18). By (3.26), there exists some constant $\varepsilon_0 > 0$ such that for any $\varepsilon \in (0, \varepsilon_0]$,

$$\mathbb{E}\left[\left\|\frac{R_\varepsilon^{u^\varepsilon}}{\sqrt{\varepsilon}h(\varepsilon)}\right\|^2\right] \le C(K, N, T, \alpha_0, \alpha_1, |p|, |q|). \tag{3.30}$$

Notice that b/α is Lipschitz continuous and σ/α is bounded, then we have

$$\begin{aligned}\left|X_\varepsilon^{u^\varepsilon}(t)\right|^2 &\le C(K, \alpha_0, \alpha_1)\int_0^t \left|X_\varepsilon^{u^\varepsilon}(s)\right|^2 ds + C(K, N, T, \alpha_0)\\ &\quad + \frac{C(K, \alpha_0)}{h^2(\varepsilon)} w^2(t) + C\left|\frac{R_\varepsilon^{u^\varepsilon}(t)}{\sqrt{\varepsilon}h(\varepsilon)}\right|^2.\end{aligned} \tag{3.31}$$

Hence by (1.6) and (3.30), for any $\varepsilon \in (0, \varepsilon_0]$, taking expectation in both sides in (3.31), we have

$$\mathbb{E}\left[\left|X_\varepsilon^{u^\varepsilon}(t)\right|^2\right] \le C(K, \alpha_0, \alpha_1)\int_0^T \mathbb{E}\left[\left|X_\varepsilon^{u^\varepsilon}(s)\right|^2\right] ds + C(K, N, T, \alpha_0, \alpha_1, |p|, |q|).$$

By Gronwall's inequality, we get

$$\mathbb{E}\left[\left|X_\varepsilon^{u^\varepsilon}(t)\right|^2\right] \le C(K, N, T, \alpha_0, \alpha_1, |p|, |q|), \tag{3.32}$$

then by Fubini's theorem,

$$\mathbb{E}\left[\int_0^T \left|X_\varepsilon^{u^\varepsilon}(s)\right|^2 ds\right] \le C(K, N, T, \alpha_0, \alpha_1, |p|, |q|). \tag{3.33}$$

First taking supremum with respect to $t \in [0, T]$ in (3.31), and then taking expectation in both sides, for any $\varepsilon \in (0, \varepsilon_0]$, by BDG inequality, (1.6), (3.30) and (3.33), we obtain that

$$\begin{aligned}\mathbb{E}\left[\left\|X_\varepsilon^{u^\varepsilon}\right\|^2\right] &\le C(K, \alpha_0, \alpha_1)\mathbb{E}\left[\int_0^T \left|X_\varepsilon^{u^\varepsilon}(s)\right|^2 ds\right] + C(K, N, T, \alpha_0, \alpha_1, |p|, |q|)\\ &\le C(K, N, T, \alpha_0, \alpha_1, |p|, |q|),\end{aligned}$$

which completes the proof. □

Proposition 3.8 *Under Hypothesis 2.1, for every fixed $N \in \mathbb{N}$, let $\{u^\varepsilon\}_{\varepsilon>0}$ be a family of processes in $\mathcal{A}_N$ that converges in distribution to some $u \in \mathcal{A}_N$ as $\varepsilon \to 0$, as random variables taking values in the space S_N, endowed with the weak topology. Then*

$$\Gamma_\varepsilon\left(w(\cdot) + h(\varepsilon)\int_0^\cdot \dot{u}^\varepsilon(s)ds\right) \to \Gamma_0\left(\int_0^\cdot \dot{u}(s)ds\right),$$

in distribution in $C([0, T]; \mathbb{R}^d)$ as $\varepsilon \to 0$.

Proof By the Skorokhod representation theorem, there exists a probability basis $(\bar{\Omega}, \bar{\mathcal{F}}, (\bar{\mathcal{F}}_t), \bar{\mathbb{P}})$, and on this basis, a Brownian motion $\bar{w}$ and a family of $\bar{\mathcal{F}}_t$-predictable processes $\{\bar{u}^\varepsilon\}_{\varepsilon>0}, \bar{u}$ taking values in S_N, $\bar{\mathbb{P}}$-a.s., such that the joint law of (u^ε, u, w) under $\mathbb{P}$ coincides with that of $(\bar{u}^\varepsilon, \bar{u}, \bar{w})$ under $\bar{\mathbb{P}}$ and

$$\lim_{\varepsilon\to 0}\langle \bar{u}^\varepsilon - \bar{u}, g\rangle_{\mathcal{H}} = 0,\ \ \forall g\in\mathcal{H},\ \bar{\mathbb{P}}\text{-}a.s..$$

Let $\bar{X}_\varepsilon^{\bar{u}^\varepsilon}$ be the solution of a similar equation to (3.16) with u^ε replaced by $\bar{u}^\varepsilon$ and w by $\bar{w}$, and let $\bar{X}^{\bar{u}}$ be the solution of a similar equation to (3.9) with h replaced by $\bar{u}$. Thus, to prove this proposition, it is sufficient to prove that

$$\lim_{\varepsilon\to 0}\left\|\bar{X}_\varepsilon^{\bar{u}^\varepsilon} - \bar{X}^{\bar{u}}\right\| = 0, \quad \text{in probability.} \tag{3.34}$$

From now on, we drop the bars in the notation for the sake of simplicity.

Notice that, for any $t\in[0,T]$,

$$\begin{aligned}
&X_\varepsilon^{u^\varepsilon}(t) - X^u(t)\\
&= \int_0^t \left\{\frac{1}{\sqrt{\varepsilon}h(\varepsilon)}\left[\frac{b(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))}{\alpha(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))} - \frac{b(q_0(s))}{\alpha(q_0(s))}\right] - D\left(\frac{b(q_0(s))}{\alpha(q_0(s))}\right)X^u(s)\right\}ds\\
&\quad+\int_0^t\left[\frac{\sigma(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))}{\alpha(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))}\dot{u}^\varepsilon(s) - \frac{\sigma(q_0(s))}{\alpha(q_0(s))}\dot{u}(s)\right]ds\\
&\quad+\frac{1}{h(\varepsilon)}\int_0^t\frac{\sigma(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))}{\alpha(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))}dw(s) + \frac{R_\varepsilon^{u^\varepsilon}(t)}{\sqrt{\varepsilon}h(\varepsilon)}\\
&=:\sum_{k=1}^4 Y_\varepsilon^{k,u^\varepsilon}(t).
\end{aligned}\tag{3.35}$$

We shall prove this proposition in the following four steps.

Step 1: For the first term $Y_\varepsilon^{1,u^\varepsilon}$, denote $x_\varepsilon(t) := \sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(t)$, by Taylor's formula, there exists a random variable η_ε taking values in $(0,1)$ such that

$$\begin{aligned}
&\left|Y_\varepsilon^{1,u^\varepsilon}(t)\right|\\
&=\left|\int_0^t\left[D\left(\frac{b(q_0(s)+\eta_\varepsilon(s)x_\varepsilon(s))}{\alpha(q_0(s)+\eta_\varepsilon(s)x_\varepsilon(s))}\right)X_\varepsilon^{u^\varepsilon}(s) - D\left(\frac{b(q_0(s))}{\alpha(q_0(s))}\right)X^u(s)\right]ds\right|\\
&\le\left|\int_0^t D\left(\frac{b(q_0(s)+\eta_\varepsilon(s)x_\varepsilon(s))}{\alpha(q_0(s)+\eta_\varepsilon(s)x_\varepsilon(s))}\right)\cdot\left(X_\varepsilon^{u^\varepsilon}(s)-X^u(s)\right)ds\right|\\
&\quad+\left|\int_0^t\left[D\left(\frac{b(q_0(s)+\eta_\varepsilon(s)x_\varepsilon(s))}{\alpha(q_0(s)+\eta_\varepsilon(s)x_\varepsilon(s))}\right) - D\left(\frac{b(q_0(s))}{\alpha(q_0(s))}\right)\right]\cdot X^u(s)ds\right|\\
&=: y_\varepsilon^{11}(t) + y_\varepsilon^{12}(t).
\end{aligned}$$

For the first term y_ε^{11}, by the boundness of $D\left(\frac{b}{\alpha}\right)$, we have

$$y_\varepsilon^{11}(t)\le C(K,\alpha_0,\alpha_1)\int_0^t\left|X_\varepsilon^{u^\varepsilon}(s) - X^u(s)\right|ds. \tag{3.36}$$

Next we deal with the second term y_ε^{12}. For each $R>\|q_0\|$ and $\rho\in(0,1)$, set

$$\eta_{R,\rho} := \sup_{|x|\le R,|y|\le R,|x-y|\le\rho}\left|D\left(\frac{b}{\alpha}\right)(x) - D\left(\frac{b}{\alpha}\right)(y)\right|.$$

Then by the continuous differentiability of $\frac{b}{\alpha}$, we know that for any fixed $R > 0$,

$$\lim_{\rho\to 0}\eta_{R,\rho} = 0.$$

Since $\sqrt{\varepsilon}h(\varepsilon) \to 0$ as $\varepsilon \to 0$, there exists some $\varepsilon_0 > 0$ small enough such that for all $0 < \varepsilon \le \varepsilon_0$,

$$\sup_{\|q_0\|\le R, \sqrt{\varepsilon}h(\varepsilon)\|X_\varepsilon^{u^\varepsilon}\|\le\rho} \left\| \left(D\left(\frac{b}{\alpha}\right)(q_0+\eta_\varepsilon\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}) - D\left(\frac{b}{\alpha}\right)(q_0)\right) X^u \right\| \le \eta_{R+1,\rho}\left\|X^u\right\|$$

for any $\rho \in (0, 1)$.

Thus, we obtain that for any $r > 0$, $R > \|q_0\|$,

$$\begin{aligned} &\mathbb{P}\left(\left\|y_\varepsilon^{12}\right\| > r\right) \\ \le&\mathbb{P}\left(\sqrt{\varepsilon}h(\varepsilon)\left\|X_\varepsilon^{u^\varepsilon}\right\| > \rho\right) + \mathbb{P}\left(\eta_{R+1,\rho}\left\|X^u\right\| > \frac{r}{T}\right) \\ \le&\frac{\varepsilon h^2(\varepsilon)}{\rho^2}\mathbb{E}\left[\left\|X_\varepsilon^{u^\varepsilon}\right\|^2\right] + \frac{T^2\eta_{R+1,\rho}^2}{r^2}\mathbb{E}\left[\left\|X^u\right\|^2\right]. \end{aligned} \tag{3.37}$$

By (3.10) and (3.19), letting $\varepsilon \to 0$ and then $\rho \to 0$ in (3.37), we can prove that

$$\lim_{\varepsilon\to 0}\mathbb{P}\left(\left\|y_\varepsilon^{12}\right\| > r\right) = 0, \quad \text{for any } r > 0. \tag{3.38}$$

Step 2: For the second term $Y_\varepsilon^{2,u^\varepsilon}$ we have

$$\begin{aligned} &\left|Y_\varepsilon^{2,u^\varepsilon}(t)\right| \\ \le&\left|\int_0^t \frac{\sigma(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))}{\alpha(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))}\left(\dot{u}^\varepsilon(s)-\dot{u}(s)\right)ds\right| \\ &+\left|\int_0^t\left[\frac{\sigma(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))}{\alpha(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_\varepsilon^{u^\varepsilon}(s))} - \frac{\sigma(q_0(s))}{\alpha(q_0(s))}\right]\dot{u}(s)ds\right| \\ =:&\left|Y_\varepsilon^{2,u^\varepsilon,1}(t)\right| + \left|Y_\varepsilon^{2,u^\varepsilon,2}(t)\right|. \end{aligned}$$

Using the same argument as that in the proof of (3.14), we obtain that

$$\lim_{\varepsilon\to 0}\left\|Y_\varepsilon^{2,u^\varepsilon,1}\right\| = 0, \quad \text{a.s..} \tag{3.39}$$

Since $\left\|Y_\varepsilon^{2,u^\varepsilon,1}\right\| \le C(K, N, T, \alpha_0)$, by the dominated convergence theorem, Eq. (3.39) implies that

$$\lim_{\varepsilon\to 0}\mathbb{E}\left\|Y_\varepsilon^{2,u^\varepsilon,1}\right\| = 0.$$

Due to the Lipschitz continuity of σ/α, we have

$$\left\|Y_\varepsilon^{2,u^\varepsilon,2}\right\| \le C(K, \alpha_0, \alpha_1)\int_0^T \sqrt{\varepsilon}h(\varepsilon)\left|X_\varepsilon^{u^\varepsilon}(t)\right|\cdot|\dot{u}(t)|\,ds. \tag{3.40}$$

By (3.18) and Hölder's inequality, we get

$$\mathbb{E}\left[\int_0^T\left|X_\varepsilon^{u^\varepsilon}(t)\right|\cdot|\dot{u}(t)|\,dt\right] \le C(K, N, T, \alpha_0, \alpha_1, |p|, |q|).$$

Hence by (1.6), we obtain that

$$\mathbb{E}\left\|Y_{\varepsilon}^{2,u^{\varepsilon}}\right\| \to 0, \quad \text{as } \varepsilon \to 0. \tag{3.41}$$

Step 3: For the third term $Y_{\varepsilon}^{3,u^{\varepsilon}}$, by BDG inequality and (1.6), we have

$$\begin{aligned}
\mathbb{E}\left\|Y_{\varepsilon}^{3,u^{\varepsilon}}\right\| &= \frac{1}{h(\varepsilon)}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \frac{\sigma(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_{\varepsilon}^{u^{\varepsilon}}(s))}{\alpha(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_{\varepsilon}^{u^{\varepsilon}}(s))}dw(s)\right|\right] \\
&\le \frac{C}{h(\varepsilon)}\mathbb{E}\left(\int_0^T \left\|\frac{(\sigma*\sigma^T)(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_{\varepsilon}^{u^{\varepsilon}}(s))}{\alpha^2(q_0(s)+\sqrt{\varepsilon}h(\varepsilon)X_{\varepsilon}^{u^{\varepsilon}}(s))}\right\|_{HS} ds\right)^{\frac{1}{2}} \\
&\le \frac{C(K,T,\alpha_0)}{h(\varepsilon)} \to 0, \quad \text{as } \varepsilon \to 0.
\end{aligned} \tag{3.42}$$

Step 4: For the last term $Y_{\varepsilon}^{4,u^{\varepsilon}}$, by Lemma 3.7, we have

$$\mathbb{E}\left\|Y_{\varepsilon}^{4,u^{\varepsilon}}\right\| \to 0, \quad \text{as} \quad \varepsilon \to 0. \tag{3.43}$$

By Eq. (3.35) and (3.36), we obtain that

$$\begin{aligned}
&\sup_{0\le s\le t}\left|X_{\varepsilon}^{u^{\varepsilon}}(s) - X^u(s)\right| \\
&\le C(K,\alpha_0,\alpha_1)\int_0^t \sup_{0\le v\le s}\left|X_{\varepsilon}^{u^{\varepsilon}}(v) - X^u(v)\right| ds + \sup_{0\le s\le t} y_{\varepsilon}^{12}(s) \\
&\quad + \sup_{0\le s\le t}\left|Y_{\varepsilon}^{2,u^{\varepsilon}}(s)\right| + \sup_{0\le s\le t}\left|Y_{\varepsilon}^{3,u^{\varepsilon}}(s)\right| + \sup_{0\le s\le t}\left|Y_{\varepsilon}^{4,u^{\varepsilon}}(s)\right|.
\end{aligned} \tag{3.44}$$

Using Gronwall's inequality, we have that

$$\left\|X_{\varepsilon}^{u^{\varepsilon}} - X^u\right\| \le C\left(\|y_{\varepsilon}^{12}\| + \sum_{l=2,3,4}\left\|Y_{\varepsilon}^{l,u^{\varepsilon}}\right\|\right).$$

This, together with (3.38), (3.41), (3.42) and (3.43), implies that

$$\lim_{\varepsilon\to 0}\left\|X_{\varepsilon}^{u^{\varepsilon}} - X^u\right\| = 0, \quad \text{in probability},$$

which completes the proof. □

According to Theorem 3.1, the MDP of $\{X_{\varepsilon}^M\}_{\varepsilon>0}$ follows from Proposition 3.3 and Proposition 3.8, which completes the proof of our main result Theorem 2.2.

Acknowledgements We thank the anonymous referees for their valuable comments and suggestions which help us improve the quality of this paper. Liu W. is supported by Natural Science Foundation of China (11571262, 11731009).

References

1. Budhiraja, A., Dupuis, P., Ganguly, A.: Moderate deviations principles for stochastic differential equations with jumps. Ann. Probab. **44**, 1723–1775 (2016)
2. Budhiraja, A., Dupuis, P., Maroulas, V.: Large deviations for infinite dimensional stochastic dynamical systems. Ann. Probab. **36**, 1390–1420 (2008)

3. Cerrai, S., Freidlin, M.: Mark large deviations for the Langevin equation with strong damping. J. Stat. Phys. **4**(161), 859–875 (2015)
4. Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications. Applications of Mathematics, 2nd edn. Springer, Berlin (1998)
5. Gao, F.Q., Wang, S.: Asymptotic behaviors for functionals of random dynamical systems. Stoch. Anal. Appl. **34**(2), 258–277 (2016)
6. Gao, F.Q., Zhao, X.Q.: Delta method in large deviations and moderate deviations for estimators. Ann. Stat. **39**, 1211–1240 (2011)
7. Guillin, A., Liptser, R.: Examples of moderate deviations principle for diffusion processes. Discret. Contin. Dyn. Syst. Ser. B **6**, 803–828 (2006)
8. Hall, P., Schimek, M.: Moderate-deviations-based inference for random degeneration in paired rank lists. J. Am. Stat. Assoc. **107**, 661–672 (2012)
9. Kallenberg, W.: On moderate deviations theory in estimation. Ann. Stat. **11**, 498–504 (1983)
10. Klebaner, F., Liptser, R.: Moderate deviations for randomly perturbed dynamical systems. Stoch. Process. Appl. **80**, 157–176 (1999)
11. Miao, Y., Shen, S.: Moderate deviations principle for autoregressive processes. J. Multivar. Anal. **100**, 1952–1961 (2009)
12. Wang, R., Zhai, J., Zhang, T.: A moderate deviations principle for 2-D stochastic Navier-Stokes equations. J. Differ. Equ. **258**, 3363–3390 (2015)
13. Wang, R., Zhang, T.: Moderate deviations for stochastic reaction-diffusion equations with multiplicative noise. Potential Anal. **42**, 99–113 (2015)