Chin. Ann. Math. Ser. B
40(2), 2019, 285–308
DOI: 10.1007/s11401-019-0133-9

Chinese Annals of
Mathematics, Series B

On Uniform Large Deviations Principle for Multi-valued SDEs via the Viscosity Solution Approach*

Jiagang REN[1] Jing WU[1]

Abstract This paper deals with the uniform large deviations for multivalued stochastic differential equations (MSDEs for short) by applying a stability result of the viscosity solutions of second order Hamilton-Jacobi-Belleman equations with multivalued operators. Moreover, the large deviation principle is uniform in time and in starting point.

Keywords Multivalued stochastic differential equation, Large deviation principle, Viscosity solution, Exponential tightness, Laplace limit

2000 MR Subject Classification 60H10, 60F10, 49L25

1 Introduction

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ be a probability space satisfying the usual conditions and $(W_t)_{t\geq 0}$ be a standard d_1-dimensional Brownian motion. In this paper we are concerned with a large deviation principle (LDP for short) for a family of diffusions generated by the following multivalued stochastic differential equation (MSDE for short) perturbed with small noises:

$$\mathrm{d}X_n(t) \in b(t, X_n(t))\mathrm{d}t + \frac{1}{\sqrt{n}}\sigma(t, X_n(t))\mathrm{d}W(t) - A(X_n(t))\mathrm{d}t, \quad X_n(0) = x \in \overline{D(A)}, \qquad (1.1)$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d_1}$ are measurable functions, A is a multivalued map $\mathbb{R}^d \to 2^{\mathbb{R}^d}$ with domain $D(A) := \{x \in \mathbb{R}^d; A(x) \neq \emptyset\}$ and graph $\mathrm{Gr}(A) := \{(x, y) \in \mathbb{R}^{2d} : x \in \mathbb{R}^d,\ y \in A(x)\}$, maximal monotone.

Let us introduce the mathematical formulation of (1.1). Consider the case with $n = 1$. By a solution we mean a pair (X, K) of $(\mathscr{F}_t)$-adapted continuous processes satisfying:

(1) $X(0) = x$ and for all $t\geq 0$, $X(t) \in \overline{D(A)}$ a.s.,

(2) K is of locally finite variation and $K(0) = 0$ a.s.,

(3) $\mathrm{d}X(t) = b(X(t))\mathrm{d}t + \sigma(X(t))\mathrm{d}W(t) - \mathrm{d}K(t)$, $0\leq t < \infty$ a.s.,

(4) for any pair of continuous functions (α, β) such that $(\alpha(t), \beta(t)) \in \mathrm{Gr}(A)$, $\forall t \in [0, +\infty)$, $\langle X(t) - \alpha(t), \mathrm{d}K(t) - \beta(t)\mathrm{d}t\rangle \geq 0$ a.s.

Manuscript received November 22, 2015. Revised March 7, 2017.
[1]School of Mathematics, Sun Yat-Sen University, Guangzhou 510275, China.
E-mail: renjg@mail.sysu.edu.cn wjjosie@hotmail.com
*This work was supported by the National Natural Science Foundation of China (Nos. 11471340, 11671408, 11871484) and the Pearl River Nova Program of Guangzhou (No. 201710010045).

This type of stochastic equations has been solved by Cépa in [1] by using a fixed point method and standard iterations. Recently Zălinescu [2] investigated the optimal control problem and established existence and comparison principle results of viscosity solutions for the corresponding Hamilton-Jacobi-Bellman (HJB for short) equations. Based on this result, we aim to present a PDE approach to the large deviations for diffusions perturbed with small random noises like (1.1). This principle was proved for the first time in [3–4] by the weak convergence approach. Compared with the probabilistic proof in [4], the main goal of this paper is to present a viscosity approach, which contributes to exploring possible applications of the recent theory of Hamilton-Jacobi-Bellman equations with multivalued maximal monotone operators developed in [2], and to studying in a more general context the connection between viscosity solution theory and the LDP theory.

Using viscosity method in large deviations is not new and among the early works we mention [5], where the analytic method is applied to asymptotic behaviors of perturbed diffusions. There is a large literature of large deviations results using this method (see, to name a few, [6–11]). In the book [6], Feng and Kurtz expanded the viscosity solution tool in a remarkably general setting. We aim at carrying out the method to establish a uniform LDP for the corresponding MSDEs in the space $\mathcal{C}([0,T] \times \overline{D(A)}; \overline{D(A)})$, which, to the best of our knowledge, has never been done in this way before.

To prove the LDP in the path space level, according to the well developed procedure, one needs to prove that the finite dimensional distributions satisfy an LDP and that the laws in the path space are exponentially tight. It is in the first step that the viscosity solution comes to play a role. Roughly speaking, the Laplace integrals of the solution processes at a single time turn out to satisfy some nonlinear second order HJB equations in the viscosity sense. According to the well-known equivalence between LDP and Laplace limit, to establish the LDP at a single time of the sequence of the generated diffusions, we need to prove a result concerning the stability of the viscosity solutions of the associated HJB equations. With the LDP at a single time established, the finite dimensional LDP will be a consequence of the Markov property and some convergence result (see Section 4 and Subsection 5.1). To obtain the exponential tightness we need to prove an exponential compact containment result, which can be obtained through using an exponential moment estimate (see Propositions 2.3–2.4). With the LDP in finite dimensional case well established, we finally prove the LDP of the solutions processes uniform in time and starting points in $\mathcal{C}(\overline{D(A)} \times [0,T]; \overline{D(A)})$.

Compared with the usual SDEs, there is an additional multivalued maximal monotone operator A in (1.1). A remarkable example of A is the subdifferential of a proper, convex lower semicontinuous function φ defined by

$$\partial\varphi(x) := \{z \in \mathbb{R}^d;\ \langle y - x, z\rangle + \varphi(x) \le \varphi(y),\ \forall y \in \mathbb{R}^d\},$$

in which case (1.1) becomes a stochastic variational inequality. The singularity and unboundedness of A cause quite a few tricky problems. For example, in proving exponential tightness of the solution process, one usually uses uniform Hölder continuity modulus estimates to find compact sets in the case of usual SDEs, but in respect of multivalued case we cannot get any Hölder continuity result for the solutions. To overcome this difficulty, we will turn to Puhalskii's criterion for exponential tightness, an analogue of Aldous's condition for tightness. Correspondingly, due to the existence of A, the viscosity solution results of the HJB equations cannot be obtained as easy extensions of those in SDEs' case. For example, in order to prove the stability, we need to use relaxed semi-limits of the viscosity solutions.

Compared with [12], where diffusions generated by stochastic equations with nonlinear, m-dissipative operators are considered, here the nonlinear operator A is generally multivalued and we do not assume that $\overline{D(A)} = \mathbb{R}^d$, and the method of [12] does not apply to our case.

We now give an outline of the paper. After presenting some notations and standing assumptions in Section 2, we present some estimates concerning the solutions of multivalued SDEs in Section 3 and a stability result for viscosity solutions of second order HJB equations carrying multivalued operators in Section 4, and by using these two results, the LDP of path level is established. In Section 5 the C-exponential tightness result is established for the solution processes starting from a fixed point and finally, with the help of this intermediate result, we obtain the LDP uniformly in initial values. We also give a detailed proof of the comparison principle in Section 6 for corresponding HJB equations, which have different form from that in [2].

2 Preliminaries and Estimates

2.1 Notations about A

Let us recall some definitions and properties of maximal monotone operators. For more general details we refer the readers to [13].

A multivalued operator $A : \mathbb{R}^d \to 2^{\mathbb{R}^d}$ is called monotone if

$$\langle y_1 - y_2, x_1 - x_2 \rangle \geq 0, \quad \forall (x_1, y_1), (x_2, y_2) \in \mathrm{Gr}(A).$$

A monotone operator A is called maximal monotone if and only if

$$(x_1, y_1) \in \mathrm{Gr}(A) \Leftrightarrow \langle y_1 - y_2, x_1 - x_2 \rangle \geq 0, \quad \forall (x_2, y_2) \in \mathrm{Gr}(A).$$

Proposition 2.1 *For a maximal monotone operator A on $\mathbb{R}^d$, the following hold*:

(1) $\overline{D(A)}$ *and* $\mathrm{Int}(D(A))$ *are convex subsets of* $\mathbb{R}^d$, *and moreover,* $\mathrm{Int}(D(A)) = \mathrm{Int}(\overline{D(A)})$.

(2) *A is locally bounded on* $\mathrm{Int}(D(A))$, *i.e., for every compact subset Γ of* $\mathrm{Int}(D(A))$, $\bigcup_{x \in \Gamma} A(x)$ *is bounded.*

Proposition 2.2 (see [1, Proposition 4.1, Proposition 4.4]) *Suppose that (X, K) is a pair of continuous and $(\mathscr{F}_t)$-adapted processes satisfying* (4) *in Section 1, i.e., for any continuous* *functions* (α, β) *with*

$$(\alpha(t), \beta(t)) \in \mathrm{Gr}(A), \quad \forall t \in [0, +\infty),$$

then the measure

$$\langle X(t) - \alpha(t), \mathrm{d}K(t) - \beta(t)\mathrm{d}t\rangle \geq 0, \quad a.s.$$

Moreover, there exist $\gamma > 0$, $\mu \geq 0$ *such that for all* $0 \leq s < t \leq T$,

$$\int_s^t \langle X(r), \mathrm{d}K(r)\rangle \geq \gamma |K|_t^s - \mu \int_s^t |X(r)|\mathrm{d}r - \gamma\mu(t-s),$$

where $|K|_t^s$ *denotes the total variation of* K *on* $[s, t]$.

Also, suppose if $(\widetilde{X}, \widetilde{K})$ *is another such pair, then*

$$\langle X(t) - \widetilde{X}(t), \mathrm{d}K(t) - \mathrm{d}\widetilde{K}(t)\rangle \geq 0, \quad a.s.$$

2.2 Large deviations

For reader's convenience we also recall the definitions of the large deviation principle and exponential tightness here (see [6, Chapter 4]). Let $(\mathcal{X}, d)$ be a separable metric space and (Y_n) a sequence of random variables on a probability space $(\Omega, \mathscr{F}, \mathbf{P})$ with values in $\mathcal{X}$.

Definition 2.1 *A function* $I : \mathcal{X} \to [0, \infty]$ *with compact level subsets* $\{x; I(x) \leq r\}$ *for every* $r \in [0, +\infty)$ *is a rate function.*

(Y_n) *is said to satisfy the large deviation principle with a rate function* I *if there exists a measurable map* $I : \mathcal{X} \to [0, \infty]$ *such that for any Borel measurable set* F,

$$-\inf_{x \in F^\circ} I(x) \leq \liminf_{n\to\infty} \frac{1}{n} \log \mathbf{P}(Y_n \in F) \leq \limsup_{n\to\infty} \frac{1}{n} \log \mathbf{P}(Y_n \in F) \leq -\inf_{x \in \overline{F}} I(x).$$

Definition 2.2 $\{Y_n\}$ *is said to be exponentially tight if for every* $r < \infty$, *there exists a compact subset* Γ_r *of* $\mathcal{X}$ *such that*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbf{P}(Y_n \in \Gamma_r^c) \leq -r.$$

Definition 2.3 *A sequence of stochastic processes* X_n *that is exponentially tight in* $\mathcal{D}([0, T); \mathcal{X})$ (*the Skorohod space of càdlàg functions over* $[0, T)$) *is called C-exponentially tight if for each* $\epsilon > 0$ *and* $T > 0$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbf{P}(\sup_{t \leq T} d(X_n(t), X_n(t-)) \geq \epsilon) = -\infty,$$

where d *is the metric on* $\mathcal{X}$.

2.3 Assumptions

Throughout the paper we make the following assumptions:

(H1) σ and b are continuous and satisfy that for all $x, y \in \mathbb{R}^d$, $t, s \geq 0$,

$$|\sigma(t, x) - \sigma(s, y)| \vee |b(t, x) - b(s, y)| \leq L(|t-s| + |x-y|),$$

$$|\sigma(t,x)|\vee|b(t,x)|\leq L'(1+|x|),$$

where $L, L' > 0$ are constants.

(H2) $0\in \mathrm{Int}(D(A))$.

Note that (H2) is not a real restriction since the general case $\mathrm{Int}(D(A))\neq\emptyset$ can be reduced to this one by a linear transformation (see [2]).

2.4 Some estimates about the solution

As is noted in Section 1, to pass from finite dimensional distribution LDP to the path space level LDP, we need to prove the exponential tightness. To this aim we first establish an exponential moment estimate for the solutions, which may be of independent interest.

Proposition 2.3 *Let $X_n(t,x;s)$ be the solution to* (1.1) *and set $X_n(s):=X_n(0,x;s)$. Then there exist constants $c_1, c_2>0$ independent of n such that for all n,*

$$\mathbf{E}\sup_{s\leq T}\mathrm{e}^{nc_1\left(\log(\mathrm{e}+|X_n(s)|^2)\right)^2}\leq \mathrm{e}^{nc_2}.$$

Proof Take $\beta>0$, $\alpha>0$ and

$$g(x):=(\log(\mathrm{e}+x^2))^2.$$

Set

$$\tau_R:=\inf\{t\leq T; |X_n(t)|>R\}.$$

Then by Itô's formula and Proposition 2.2,

$$\begin{aligned}
&\mathrm{e}^{n\beta\mathrm{e}^{-\alpha(s\wedge\tau_R)}g(|X_n(s\wedge\tau_R)|)}\\
&=\mathrm{e}^{n\beta g(|x|)}-\alpha\int_0^{s\wedge\tau_R}n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}\mathrm{d}r\\
&\quad+4\int_0^{s\wedge\tau_R}n\beta\mathrm{e}^{-\alpha r}\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}\log(\mathrm{e}+|X_n(r)|^2)\frac{1}{\mathrm{e}+|X_n(r)|^2}\\
&\quad\cdot\Big(\langle X_n(r), b(r,X_n(r))\rangle\mathrm{d}r+\frac{1}{\sqrt{n}}\langle X_n(r),\sigma(r,X_n(r))\rangle\mathrm{d}W(r)-\langle X_n(r),\mathrm{d}K_n(r)\rangle\Big)\\
&\quad+8\int_0^{s\wedge\tau_R}n\beta^2\mathrm{e}^{-2\alpha r}\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}g(|X_n(r)|)\frac{1}{\mathrm{e}+|X_n(r)|^2}\mathrm{tr}(\sigma\sigma^*(r,X_n(r)))\mathrm{d}r\\
&\quad+6\int_0^{s\wedge\tau_R}\beta\mathrm{e}^{-\alpha r}\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}\mathrm{tr}(\sigma\sigma^*(r,X_n(r)))\Big(\frac{X_n(r)\otimes X_n(r)}{(\mathrm{e}+|X_n(r)|^2)^2}\\
&\quad+\log(\mathrm{e}+|X_n(r)|^2)\frac{\mathrm{e}+|X_n(r)|^2-2X_n(r)\otimes X_n(r)}{2(\mathrm{e}+|X_n(r)|^2)^2}\Big)\mathrm{d}r\\
&\leq\mathrm{e}^{n\beta g(|x|)}-\alpha\int_0^{s\wedge\tau_R}n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}\mathrm{d}r\\
&\quad+6L'\int_0^{s\wedge\tau_R}n\beta\mathrm{e}^{-\alpha r}\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}\log(\mathrm{e}+|X_n(r)|^2)\mathrm{d}r
\end{aligned}$$

$$+4\int_0^{s\wedge\tau_R}\sqrt{n}\beta \mathrm{e}^{-\alpha r}\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}\frac{\log(\mathrm{e}+|X_n(r)|^2)}{\mathrm{e}+|X_n(r)|^2}\langle X_n(r),\sigma(r,X_n(r))\rangle\mathrm{d}W(r)$$

$$+C(\gamma,\mu)\int_0^{s\wedge\tau_R}n\beta\mathrm{e}^{-\alpha r}\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}\log(\mathrm{e}+|X_n(r)|^2)\mathrm{d}r$$

$$+8L'^2\int_0^{s\wedge\tau_R}n\beta^2\mathrm{e}^{-2\alpha r}\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}g(|X_n(r)|)\mathrm{d}r$$

$$+6L'^2\int_0^{s\wedge\tau_R}\beta\mathrm{e}^{-\alpha r}\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}[1+\log(\mathrm{e}+|X_n(r)|^2)]\mathrm{d}r$$

$$=:\sum_{i=1}^{6}\mathrm{I}_i(s).$$

Note that

$$\mathrm{I}_1(s)=-\alpha\int_0^{s\wedge\tau_R}n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}\mathrm{d}r,$$

$$\mathrm{I}_2(s)\le 6L'\int_0^{s\wedge\tau_R}n\beta\mathrm{e}^{-\alpha r}\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}g(|X_n(r)|)\mathrm{d}r,$$

$$\mathrm{I}_4(s)\le C(\gamma,\mu)\int_0^{s\wedge\tau_R}n\beta\mathrm{e}^{-\alpha r}\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}g(|X_n(r)|)\mathrm{d}r,$$

$$\mathrm{I}_5(s)\le 8L'^2\int_0^{s\wedge\tau_R}n\beta^2\mathrm{e}^{-2\alpha r}\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}g(|X_n(r)|)\mathrm{d}r,$$

$$\mathrm{I}_6(s)\le 12L'^2\int_0^{s\wedge\tau_R}\beta\mathrm{e}^{-\alpha r}\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}g(|X_n(r)|)\mathrm{d}r.$$

Thus by taking expectations we get

$$\begin{aligned}&\mathbf{E}\mathrm{e}^{n\beta\mathrm{e}^{-\alpha(s\wedge\tau_R)}g(|X_n(s\wedge\tau_R)|)}\\ &\le \mathrm{e}^{n\beta g(|x|)}-\alpha\mathbf{E}\int_0^{s\wedge\tau_R}n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}\mathrm{d}r\\ &\quad+\Big(6L'+8\beta L'^2+C(\gamma,\mu)+\frac{12L'^2}{n}\Big)\mathbf{E}\int_0^{s\wedge\tau_R}n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}\mathrm{d}r.\end{aligned}$$

Taking

$$\alpha=(6+8\beta L'+12L')L'+C(\gamma,\mu)+1,$$

we have

$$\mathbf{E}\mathrm{e}^{n\beta\mathrm{e}^{-\alpha(s\wedge\tau_R)}g(|X_n(s\wedge\tau_R)|)}+\mathbf{E}\int_0^{s\wedge\tau_R}n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)\mathrm{e}^{n\beta\mathrm{e}^{-\alpha r}g(|X_n(r)|)}\mathrm{d}r\le \mathrm{e}^{n\beta g(|x|)}. \quad (2.1)$$

Now take α satisfying (2.1) for $\beta=2$. With this α it is trivial to see that (2.1) holds for $\beta=1$ as well. Applying BDG's inequality to $\mathrm{I}_3(s)$ with the same α and $\beta=1$ gives

$$\begin{aligned}\mathbf{E}\sup_{s\le T}\mathrm{I}_3(s)&\le\frac{8L'}{\sqrt{2}}\Big(\mathbf{E}\int_0^{T\wedge\tau_R}2n\mathrm{e}^{-2\alpha r}\mathrm{e}^{2n\mathrm{e}^{-\alpha r}g(|X_n(r)|)}g(|X_n(r)|)\mathrm{d}r\Big)^{\frac{1}{2}}\\ &\le 8L'\mathrm{e}^{ng(|x|)},\end{aligned}$$

where the last inequality follows from (2.1) with $\beta = 2$.

Hence we have

$$\mathbf{E}\sup_{s\le T} \mathrm{e}^{n\mathrm{e}^{-\alpha(s\wedge\tau_R)}g(|X_n(s\wedge\tau_R)|)} \le (8L'+1)\mathrm{e}^{n(\log(\mathrm{e}+|x|^2))^2}.$$

Letting $R\to\infty$ gives

$$\mathbf{E}\sup_{s\le T} \mathrm{e}^{n\mathrm{e}^{-\alpha(s)}g(|X_n(s)|)} \le (8L'+1)\mathrm{e}^{n(\log(\mathrm{e}+|x|^2))^2}.$$

Now by taking $c_1 = \mathrm{e}^{-\alpha T}$ and c_2 sufficiently large we complete the proof.

From this result we deduce easily the following proposition.

Proposition 2.4 *For every $M>0$, there exists a compact set $\Gamma_M \subset \overline{D(A)}$ such that*

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbf{P}(\exists t\in[0,T],\ s.t.\ X_n(t)\notin\Gamma_M)\le -M.$$

Proof Take

$$\Gamma_M := \{|x|\le r\}\cap\overline{D(A)}$$

where r will be specified later. Note that for every t, $X_n(t)\in\overline{D(A)}$ a.e. Applying the above proposition we get

$$\begin{aligned}
&\frac{1}{n}\log\mathbf{P}(\exists t\in[0,T],\ s.t.\ X_n(t)\notin\Gamma_M)\\
\le&\frac{1}{n}\log\mathbf{P}\Big(\sup_{t\in[0,T]}|X_n(t)|>r\Big)\\
\le&\frac{1}{n}\log\mathbf{P}\Big(\sup_{t\in[0,T]}\mathrm{e}^{nc_1(\log(\mathrm{e}+|X_n(t)|^2))^2}>\mathrm{e}^{nc_1(\log(\mathrm{e}+r^2))^2}\Big)\\
\le&\frac{1}{n}\log\mathrm{e}^{-nc_1(\log(\mathrm{e}+r^2))^2}\mathrm{e}^{nc_2}\\
=&\,c_2-c_1(\log(\mathrm{e}+r^2))^2\le -M
\end{aligned}$$

provided that

$$r=\exp\Big\{\frac{1}{2}\sqrt{\frac{M+c_2}{c_1}}\Big\}.$$

The next result follows from standard calculations and we omit the proof.

Proposition 2.5 *Suppose that $X_n(t,x;s)$ and $X_n(t,y;s)$ are solutions to* (1.1) *with initial values x, y at t respectively. Then under* (H1)–(H2)*, there exists a constant $C>0$ independent of n such that*

$$\mathbf{E}\sup_{s\in[t,T]}|X_n(t,x;s)-X_n(t,y;s)|^2\le C|x-y|^2.$$

3 HJB Equations and Viscosity Solutions

Now we prepare some materials about the viscosity solution of second order HJB equations with multivalued maximal monotone operators. Given a multivalued maximal monotone operator A with domain $\overline{D(A)}$ and a Hamiltonian

$$H : [0,1] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathscr{S}^d \mapsto \mathbb{R},$$

suppose that H is elliptic, i.e.,

$$M, N \in \mathscr{S}^d, \quad M \geq N \Rightarrow H(t,x,r,p,M) \leq H(t,x,r,p,N), \quad \forall (t,x,p),$$

where $\mathscr{S}^d$ stands for the space of all real positively definite matrices. Consider the Cauchy problem

$$\begin{cases} \dfrac{\partial u}{\partial t} + H(t,x,u,Du,D^2u) \in \langle A(x), Du\rangle & \text{in } (0,T) \times \overline{D(A)}, \\ u(T,\cdot) = h(\cdot) & \text{on } \overline{D(A)}. \end{cases} \tag{3.1}$$

Identify A_* (resp. A^*) as the lower (resp. upper) semi-continuous envelope of A. The following definition is adjusted from [2].

Definition 3.1 *Let A be a multivalued maximal monotone operator with domain $D(A)$. Define for any $x \in \overline{D(A)}$ and $y \in \mathbb{R}^d$,*

$$A^*(x,y) := \lim_{\varepsilon \downarrow 0} \sup_{(z,w) \in D_\varepsilon(x,y)} \langle z, w\rangle,$$
$$A_*(x,y) := \lim_{\varepsilon \downarrow 0} \inf_{(z,w) \in D_\varepsilon(x,y)} \langle z, w\rangle,$$

where $D_\varepsilon(x,y) := \{(z,w);\ z \in A(x'),\ x' \in D(A),\ |x-x'| < \varepsilon,\ |y-w| < \varepsilon\}$.

Definition 3.2 (1) $u \in \mathrm{USC}([0,T] \times \overline{D(A)})$ (*the space of upper semicontinuous functions on* $[0,T] \times \overline{D(A)}$) *is called a viscosity subsolution of* (3.1) *if* $u(T,\cdot) \leq h(\cdot)$ *on* $\overline{D(A)}$ *and for any* $\varphi \in \mathcal{C}^{1,2}([0,T) \times \overline{D(A)})$,

$$\frac{\partial \varphi}{\partial t}(t,x) + H(t,x,u,D\varphi(t,x),D^2\varphi(t,x)) \geq A_*(x, D\varphi(t,x))$$

provided that $(t,x) \in (0,T) \times \overline{D(A)}$ *is a local maximizer of* $u - \varphi$.

(2) $u \in \mathrm{LSC}([0,T] \times \overline{D(A)})$ (*the space of lower semicontinuous functions on* $[0,T] \times \overline{D(A)}$) *is called a viscosity supersolution of* (3.1) *if* $u(T,\cdot) \geq h(\cdot)$ *on* $\overline{D(A)}$ *and for any* $\varphi \in \mathcal{C}^{1,2}((0,T) \times \overline{D(A)})$,

$$\frac{\partial \varphi}{\partial t}(t,x) + H(t,x,u,D\varphi(t,x),D^2\varphi(t,x)) \leq A^*(x, D\varphi(t,x))$$

provided that $(t,x) \in (0,T) \times \overline{D(A)}$ *is a local minimizer of* $u - \varphi$.

(3) *u is called a viscosity solution if it is both a subsolution and a supersolution.*

Remark 3.1 As usual, the term local maximizer (resp. local minimizer) in the above definition can be replaced by global strict maximizer (resp. global strict minimizer).

We will need the following stability result.

Theorem 3.1 *Suppose that* $H_n, H \in \mathcal{C}([0,T] \times \overline{D(A)} \times \mathbb{R} \times \mathbb{R}^d \times \mathscr{S}^d)$, $h_n, h \in \mathcal{C}(\overline{D(A)})$, $n \in \mathbb{N}$, *and* u_n *is a subsolution (resp. supersolution) of* (3.1) *with the Hamiltonian* H_n, *and final value* h_n, *with* $\{u_n(t,x), (t,x) \in [0,T] \times \overline{D(A)}, n \in \mathbb{N}\}$ *being uniformly bounded. Suppose that*

$$\lim_{n\to\infty} H_n = H \quad \text{in } \mathcal{C}([0,T] \times \overline{D(A)} \times \mathbb{R} \times \mathbb{R}^d \times \mathscr{S}^d).$$

Let u^* *and* u_* *be the superior and inferior relaxed semi-limit of* u *respectively, i.e.,*

$$u^*(t,x) := \lim_{n\to\infty} \sup\{u_m(s,y) : |s-t| + |y-x| < n^{-1},\ (s,y) \in [0,T] \times \overline{D(A)},\ m \geq n\}$$
$$u_*(t,x) := \lim_{n\to\infty} \inf\{u_m(s,y) : |s-t| + |y-x| < n^{-1},\ (s,y) \in [0,T] \times \overline{D(A)},\ m \geq n\}.$$

Similarly we define h^* *and* h_*. *Then* u^* *(resp.* u_**) is a subsolution (resp. supersolution) of* (3.1) *with Hamiltonian* H *and final value function* h^* *(resp.* h_**).*

Moreover, if the comparison principle holds for the limit equation (i.e., any subsolution is less than any supersolution) and u_n *is a solution of* (3.1), *then* $u^* = u_* =: u$ *is continuous and is a solution of* (3.1).

Proof Suppose that $\varphi \in \mathcal{C}^{1,2}((0,T) \times \overline{D(A)})$ and $(t_0, x_0) \in (0,T) \times \overline{D(A)}$ are such that

$$0 = (u^* - \varphi)(t_0, x_0) > (u^* - \varphi)(t,x), \quad \forall (t,x) \in (0,T) \times \overline{D(A)} \setminus (t_0, x_0).$$

Then it is clear by [14, Lemma 2.10] that there exists a sequence (t_n, x_n) in a compact subset $U_r(t_0, x_0) \subset [0,T] \times \overline{D(A)}$ such that

$$\lim_{n\to\infty} (t_n, x_n) = (t_0, x_0), \quad \lim_{n\to\infty} u_n(t_n, x_n) = u^*(t_0, x_0),$$

and for all n,

$$(u_n - \varphi)(t_n, x_n) = \max_{(t,x)\in U_r(t_0,x_0)} (u_n - \varphi)(t,x).$$

Since

$$\begin{cases} \dfrac{\partial \varphi}{\partial t}(t_n, x_n) + H_n(t_n, x_n, u_n(t_n, x_n), D\varphi(t_n, x_n), D^2\varphi(t_n, x_n)) \geq A_*(x_n, D\varphi(t_n, x_n)), \\ u_n(T, \cdot) \leq h_n(\cdot) \quad \text{on } \overline{D(A)}. \end{cases} \tag{3.2}$$

By sending $n \to \infty$, we get

$$\frac{\partial \varphi}{\partial t}(t_0, x_0) + H(t_0, x_0, u^*(t_0, x_0), D\varphi(t_0, x_0), D^2\varphi(t_0, x_0)) \geq A_*(x_0, D\varphi(t_0, x_0)) \tag{3.3}$$

and

$$u^*(T, \cdot) \leq h^*(\cdot) \quad \text{on } \overline{D(A)}.$$

This proves that u^* is a subsolution with final value function h^*. A similar argument shows that u_* is a supersolution. If moreover the comparison principle holds, then $u^* \leq u_*$. But $u^* \geq u_*$ by definition. Hence $u^* = u_* = u$ and the proof is completed.

4 LDP in Time

Consider the following equation

$$\begin{cases} \mathrm{d}X(s) \in b(s, X(s))\mathrm{d}s + \sigma(s, X(s))\mathrm{d}W(s) - A(X(s))\mathrm{d}s, & s \in [t, T], \\ X(t) = x \in \overline{D(A)}. \end{cases} \tag{4.1}$$

We denote by $(X(t, x; s), K(t, x; s))$ its unique solution. Let h be a bounded continuous function on $\overline{D(A)}$. Then by [2, Theorem 4] the function

$$u(t, x) := \mathbf{E}[h(X(t, x; T))], \quad (t, x) \in [0, T] \times \overline{D(A)}$$

is the unique viscosity solution of the following PDE:

$$\begin{cases} \dfrac{\partial u}{\partial t} + H(t, x, Du, D^2u) \in \langle A(x), Du \rangle & \text{in } (0, T) \times \overline{D(A)}, \\ u(T, \cdot) = h(\cdot) & \text{on } \overline{D(A)} \end{cases} \tag{4.2}$$

with

$$H(t, x, q, M) := \frac{1}{2}\mathrm{tr}(\sigma\sigma^*(t, x)M) + \langle b(t, x), q \rangle.$$

Consequently if $h > 0$ then the logarithmic transformation of u,

$$v := c \log u,$$

is a viscosity solution of (4.2) with

$$H(t, x, q, M) := \frac{1}{2}\mathrm{tr}(\sigma\sigma^*(t, x)M) + \frac{1}{2c}|\sigma^* q|^2 + \langle b(t, x), q \rangle$$

and the final value condition $v(T, \cdot) = c \log h(\cdot)$.

Applying the above observation to the solution $X_n(t, x; s)$ of

$$\begin{cases} \mathrm{d}X_n(s) \in b(s, X_n(s))\mathrm{d}s + \dfrac{1}{\sqrt{n}}\sigma(s, X_n(s))\mathrm{d}W(s) - A(X_n(s))\mathrm{d}s, & s \in [t, T], \\ X_n(t) = x \in \overline{D(A)}, \end{cases} \tag{4.3}$$

we see that if h is a bounded continuous function defined on $\overline{D(A)}$ then the function

$$u_n(t, x) := -\frac{1}{n} \log \mathbf{E}[\exp\{-nh(X_n(t, x; T))\}] \tag{4.4}$$

is the unique viscosity solution to

$$\begin{cases} \dfrac{\partial u}{\partial t} + H_n(t, x, Du, D^2u) \in \langle A(x), Du \rangle & \text{in } (0, T) \times \overline{D(A)}, \\ u(T, \cdot) = h(\cdot) & \text{on } \overline{D(A)}, \end{cases} \tag{4.5}$$

where

$$H_n(t, x, q, M) := \frac{1}{2n}\mathrm{tr}(\sigma\sigma^*(t, x)M) - \frac{1}{2}|\sigma^*(t, x)q|^2 + \langle b(t, x), q \rangle.$$

Applying Proposition 2.5, we get that for every $n\geq 1$ and every $(t,x)\in[0,T]\times\overline{D(A)}$, there exists a constant $L>0$ satisfying

$$|u_n(t,x)-u_n(t,y)|\leq L|x-y|. \tag{4.6}$$

Moreover, it is obvious that for each n, u_n is a viscosity supersolution of (4.5) with n replaced by $n+1$. Thus $u_{n+1}\leq u_n$ by the comparison principle of (4.5) (see Section 6) when σ is bounded. Similarly, $u_n\geq u$ for all n where u is the viscosity solution of

$$\begin{cases}\dfrac{\partial u}{\partial t}+H(t,x,Du)\in\langle A(x),Du\rangle & \text{in } (0,T)\times\overline{D(A)},\\ u(T,\cdot)=h(\cdot) & \text{on } \overline{D(A)}\end{cases} \tag{4.7}$$

with

$$H(t,x,q):=-\frac{1}{2}|\sigma^*(t,x)q|^2+\langle b(t,x),q\rangle.$$

Hence the function $\widetilde{u}=\lim\limits_{n\to\infty}u_n$ is well defined.

Note that the comparison principle holds for the equation when σ is bounded ($c\equiv 0$ in Section 6). By Theorem 3.1 $\widetilde{u}$ is a viscosity solution of (4.7). Since the viscosity solution to (4.7) is unique, we have $u=\widetilde{u}$ and therefore have proved the following proposition.

Proposition 4.1 *Suppose* $\|\sigma\|\leq M$ *for some constant* $M>0$. *Let* u_n *be defined in* (4.4) *and* u *be the unique viscosity solution of* (4.7). *Then for every* $(t,x)\in[0,T]\times\overline{D(A)}$,

$$\lim_{n\to\infty}u_n(t,x)=u(t,x).$$

Moreover, since $\{u_n\}$ *is a sequence of decreasing functions, the convergence holds uniformly on compact subsets of* $\overline{D(A)}$ *with respect to* x.

Since

$$-\frac{1}{2}|\sigma^*(x)q|^2=\inf_{z\in\mathbb{R}^d}\Big\{\langle q,\sigma(x)z\rangle+\frac{1}{2}|z|^2\Big\},$$

by [15, Theorem 4] (note that a time reversible is necessary to use this result), we have

$$\begin{aligned}u(t,x)&=\inf\Big\{\frac{1}{2}\int_t^T|z_s|^2\mathrm{d}s+h(X^z_{t,x}(T)),\ z\in L^2([0,T],\mathbb{R}^{d_1})\Big\}\\&=\inf_{y\in\mathbb{R}^d}\Big\{\inf\Big\{\frac{1}{2}\int_t^T|z_s|^2\mathrm{d}s;\ z\in L^2([0,T],\mathbb{R}^{d_1})\ s.t.\ X^z_{t,x}(T)=y\Big\}+h(y)\Big\},\end{aligned}$$

where X^z is the unique solution of

$$\begin{cases}\mathrm{d}X^z_{t,x}(s)\in b(s,X^z_{t,x}(s))\mathrm{d}s+\sigma(s,X^z_{t,x}(s))z_s\mathrm{d}s-A(X^z_{t,x}(s))\mathrm{d}s, & s\in[t,T],\\ X^z_{t,x}(t)=x\in\overline{D(A)}.\end{cases} \tag{4.8}$$

Theorem 4.1 *Let* $0<t<t_2$ *and* $x\in\overline{D(A)}$ *be fixed. Then under* (H1)–(H2), $\{X_n(t,x;t_2)\}$ *satisfies the LDP with a good rate function* I *given by*

$$I_{t,t_2}(x;y):=\inf\Big\{\frac{1}{2}\int_t^{t_2}|z_s|^2\mathrm{d}s;\ z\in L^2([0,T],\mathbb{R}^{d_1})\ s.t.\ X^z_{t,x}(t_2)=y\Big\}.$$

Here $I_{t,t_2}(x;y)$ *is considered as a function of* y, *which depends on the parameters* t,t_2 *and* x.

Proof It will be convenient to use the notations

$$u_n(t,x)=V_n(T-t,h;x),\quad u(t,x)=V(T-t,h;x).$$

By Proposition 2.4, for every $t\in[0,T]$, $\{X_n(t)\}$ is exponentially tight. Hence by Bryc formula (see, e.g., [16, Corollary 1.2.5, 6, Proposition 3.8]), the theorem is proved when $\|\sigma\|\le M$ for some constant $M>0$.

Now we aim to prove the theorem without the assumption $\|\sigma\|\le M$. It is clear that for $|x|\le r,\ r>0$, $\|\sigma(t,x)\|\le m(r)$ for some $m(r)>0$. Denote by $X_n^{(m)}$ the solution to (4.3) in this case, and respectively by $u_n^{(m)}, u^{(m)}$ in place of u_n, u.

Trivially, for every $(t,x)\in[0,T]\times\overline{D(A)}$,

$$\sup_{s\in[t,T]}|X^z(t,x;s)|^2\le C(1+|x|^2).$$

Then as we have seen in Proposition 4.1, for every $r>0$, $\exists m_0(r)$ such that for $m\ge m_0(r)$,

$$\lim_{n\to\infty}u_n^{(m)}(t,x)=u^{(m)}(t,x)=u(t,x)\quad\text{uniformly on compact subsets of }\overline{D(A)}.$$

Thus it suffices to prove

$$\lim_{n\to\infty}\sup_{|x|\le r}|u_n(t,x)-u_n^{(m)}(t,x)|=0. \tag{4.9}$$

Set $\Omega_1:=\big\{\omega;\ \sup_{s\in[t,T]}|X_n(t,x;s)|>m\big\}$. Then on $\Omega-\Omega_1$,

$$X_n(t,x;\cdot)=X_n^{(m)}(t,x;\cdot)\quad\text{on }[t,T]$$

and thus for $0<t<t_2\le T$,

$$\mathbf{E}[\mathrm{e}^{-nh(X_n^{(m)}(t,x;t_2))}]=\mathbf{E}[\mathrm{e}^{-nh(X_n(t,x;t_2))}]-\mathbf{E}[\mathrm{e}^{-nh(X_n(t,x;t_2))}\chi_{\Omega_1}]+\mathbf{E}[\mathrm{e}^{-nh(X_n^{(m)}(t,x;t_2))}\chi_{\Omega_1}].$$

This yields that

$$\begin{aligned}&-\frac{1}{n}\log(\mathbf{E}[\mathrm{e}^{-nh(X_n(t,x;t_2))}]+\mathbf{E}[\mathrm{e}^{-nh(X_n^{(m)}(t,x;t_2))}\chi_{\Omega_1}])\\ \le&-\frac{1}{n}\log\mathbf{E}[\mathrm{e}^{-nh(X_n^{(m)}(t,x;t_2))}]=u_n^{(m)}(t,x)\\ \le&-\frac{1}{n}\log(\mathbf{E}[\mathrm{e}^{-nh(X_n(t,x;t_2))}]-\mathbf{E}[\mathrm{e}^{-nh(X_n(t,x;t_2))}\chi_{\Omega_1}])\end{aligned}$$

and further,

$$\begin{aligned}&V_n(t_2-t,h;x)-\frac{1}{n}\log\Big(1+\frac{\mathbf{E}[\mathrm{e}^{-nh(X_n^{(m)}(t,x;t_2))}\chi_{\Omega_1}]}{\mathbf{E}[\mathrm{e}^{-nh(X_n(t,x;t_2))}]}\Big)\\ \le&u_n^{(m)}(t,x)\\ \le&V_n(t_2-t,h;x)-\frac{1}{n}\log\Big(1-\frac{\mathbf{E}[\mathrm{e}^{-nh(X_n(t,x;t_2))}\chi_{\Omega_1}]}{\mathbf{E}[\mathrm{e}^{-nh(X_n(t,x;t_2))}]}\Big).\end{aligned}$$

But by Proposition 2.3, for every $r>0$, there exists $m\geq m_0(r)$ sufficiently large such that

$$\mathbf{P}(\Omega_1)\leq e^{-3n\|h\|},$$

which implies that for such m,

$$\frac{\mathbf{E}[e^{-nh(X_n^{(m)}(t,x;t_2))}\chi_{\Omega_1}]}{\mathbf{E}[e^{-nh(X_n(t,x;t_2))}]}\leq e^{-n\|h\|},\quad \frac{\mathbf{E}[e^{-nh(X_n(t,x;t_2))}\chi_{\Omega_1}]}{\mathbf{E}[e^{-nh(X_n(t,x;t_2))}]}\leq e^{-n\|h\|}.$$

Therefore for small $s>0$,

$$V_n(t_2-t,h;x)-\frac{1}{n}e^{-n\|h\|}\leq u_n^{(m)}(t,x)\leq V_n(t_2-t,h;x)+e^{-2n\|h\|},$$

and (4.9) follows.

Before approaching to finite dimensional case, we now state the following proposition, which, similar to [11, Proposition 7.7], can be proved by using exponential tightness and Proposition 4.1.

Proposition 4.2 *For every $(t,x)\in[0,T]\times\overline{D(A)}$ and every $n\geq 1$, $0\leq t<t_2\leq T$, $V_n(t_2-t,h;\cdot)$ is continuous on $\overline{D(A)}$. Moreover, suppose that $\{h_n\}$ is a sequence of uniformly bounded continuous functions converging to a continuous bounded function h. Then*

$$\lim_n V_n(t_2-t,h_n;x)=V(t_2-t,h;x)$$

uniformly on compact subset of $\overline{D(A)}$.

Starting from the above results, we can go by Markovian property from one dimensional distributions to finite dimensional distributions. More precisely we have the following theorem.

Theorem 4.2 *Assume that* (H1)–(H2) *hold. Let m be an integer and $0<t_1<t_2<\cdots<t_m$ be arbitrarily fixed numbers. Then the sequence $\{X_n^x(t_1),X_n^x(t_2),\cdots,X_n^x(t_m)\}$, where X_n^x is the solution of* (4.3) *with initial value x starting at* 0*, satisfies the large deviation principle with a good rate function given by*

$$\begin{aligned}&I_{t_1,t_2,\cdots,t_m}(x;y_1,y_2,\cdots,y_m)\\&:=\inf\Big\{\frac{1}{2}\int_0^{t_m}|z_s|^2\mathrm{d}s;\ z_s\in\mathbb{R}^d\ s.t.\ X^z_{t_i,y_i}(t_{i+1})=y_{i+1},\ i=0,1,\cdots,m-1\Big\}.\end{aligned}$$

Here we have set $t_0=0$ and $y_0=x$.

Proof Let $f(x_1,x_2,\cdots,x_m)$ be a bounded Lipschitz function on $(\mathbb{R}^d)^m$. For every n define inductively by

$$\begin{aligned}&f_{n,m}(x_1,\cdots,x_m)=f(x_1,x_2,\cdots,x_m),\\&f_{n,k}(x_1,\cdots,x_k):=-\frac{1}{n}\log\mathbf{E}[e^{-nf_{n,k+1}(x_1,\cdots x_k,X_n^x(t_{k+1}))}\mid X_n^x(t_k)=x_k],\quad k=1,\cdots,m-1.\end{aligned}$$

By the Markov property we have

$$-\frac{1}{n}\log \mathbf{E}(\mathrm{e}^{-n[f(X_n^x(t_1),\cdots,X_n^x(t_m))]}) = V_n(t_1, f_{n,1}, x).$$

Define again f_k successively by

$$f_m(x_1,\cdots,x_m) = f(x_1,x_2,\cdots,x_m),$$
$$f_k(x_1,\cdots,x_k) := V(t_{k+1}-t_k, f_{k+1}(x_1,\cdots,x_k,\cdot),x_k), \quad k=m-1,m-2,\cdots,1.$$

Since $f_{n,m-1}$ and f_{m-1} are respectively viscosity solutions to (4.5) and (4.7) with final value functions $f_{n,m}$ and f_m, by the comparison principle we know that $\{f_{n,m-1}\}_n$ is a decreasing sequence of uniformly bounded and continuous functions, and converges uniformly on compact subsets of $\overline{D(A)}$ to f_{m-1} according to Proposition 4.2. This in turn implies that $\{f_{n,m-2}\}_n$ is a decreasing sequence of uniformly bounded, continuous functions, and that for every n, $f_{n,m-2}$ is the viscosity solution to (4.5) with final value function $f_{n,m-1}$. Thus again by Proposition 4.2, $\{f_{n,m-2}\}$ converges to f_{m-2} uniformly on compact subsets of $\overline{D(A)}$. Continuing this process we finally have that

$$\lim_{n\to\infty} -\frac{1}{n}\log \mathbf{E}(\mathrm{e}^{-n[f(X_n^x(t_1),\cdots,X_n^x(t_m))]}) = V(t,f_1;x) =: f_0(x),$$

from which we deduce easily

$$f_0(x) = \inf_{x_1,x_2,\cdots,x_m}\Big\{\inf_z\Big\{\frac{1}{2}\int_0^{t_m}|z_s|^2\mathrm{d}s; X_x^z(t_i)=x_i\Big\} + f(x_1,\cdots,x_m)\Big\}.$$

Now we can obtain the conclusion by [16, Theorem 1.2.3].

To prove the LDP of $\{X_n\}$ in $\mathcal{D}([0,T],\overline{D(A)})$, we need to prove the following C-exponential tightness result. This is obtained through the use of the exponential compact containment proved in Section 3.

Theorem 4.3 *The sequence $\{X_n\}$ is C-exponentially tight in $\mathcal{D}([0,T],\overline{D(A)})$, the Skorohod space over $[0,T]$.*

Proof Since almost all sample paths of X_n are continuous, it suffices to prove that $\{X_n\}$ is tight in $\mathcal{D}([0,T],\overline{D(A)})$. By the above theorem and [6, Theorem 4.1], it remains to prove that for $\lambda\in\mathbb{R}$ and $s>0$, there exist random variables $\xi_n(s,\lambda)$ satisfying that for $0\le t\le t+s\le T$,

$$\mathbf{E}[\mathrm{e}^{n\lambda(|X_n^x(t+s)-X_n^x(t)|^2\wedge 1)}|\mathscr{F}_t]\le\mathbf{E}[\mathrm{e}^{\xi_n(s,\lambda)}|\mathscr{F}_t] \tag{4.10}$$

and

$$\lim_{s\to 0}\limsup_{n\to\infty}\frac{1}{n}\log\mathbf{E}[\mathrm{e}^{\xi_n(s,\lambda)}] = 0. \tag{4.11}$$

Let $r>0$ be large enough such that by setting $\Omega_1 := \big\{\omega : \sup_{s\in[0,T]}|X_n^x(s)|>r\big\}$, we have according to Proposition 2.3,

$$\mathbf{P}(\Omega_1)\le\mathrm{e}^{-n|\lambda|}.$$

By Itô's formula,

$$
\begin{aligned}
&\mathbf{E}[\mathrm{e}^{\lambda n(|X_n^x(t+s)-X_n^x(t)|^2\wedge 1)}|\mathscr{F}_t] \\
&= \mathbf{E}[\mathrm{e}^{\lambda n(|X_n^x(t+s)-X_n^x(t)|^2\wedge 1)}\cdot\chi_{\Omega_1}|\mathscr{F}_t^n] + \mathbf{E}[\mathrm{e}^{\lambda n(|X_n^x(t+s)-X_n^x(t)|^2\wedge 1)}\cdot\chi_{\Omega-\Omega_1}|\mathscr{F}_t] \\
&\le \mathbf{E}[\mathrm{e}^{n|\lambda|}\cdot\chi_{\Omega_1}|\mathscr{F}_t] + \mathbf{E}[\mathrm{e}^{\lambda n|X_n^x(t+s)-X_n^x(t)|^2}\cdot\chi_{\Omega-\Omega_1}|\mathscr{F}_t] \\
&= \mathbf{E}[\mathrm{e}^{n|\lambda|}\cdot\chi_{\Omega_1}|\mathscr{F}_t] + \mathbf{E}\Big[\exp\Big(\lambda n\Big\{\frac{2}{\sqrt{n}}\int_t^{t+s}\langle X_n^x(l)-X_n^x(t),\sigma(l,X_n^x(l))\rangle\mathrm{d}W(l) \\
&\quad + \int_t^{t+s}\Big(2\langle X_n^x(l)-X_n^x(t),b(l,X_n^x(l))\rangle + \frac{1}{n}\mathrm{tr}(\sigma\sigma^*)(X_n^x(l))\Big)\mathrm{d}l \\
&\quad - 2\int_t^{t+s}\langle X_n^x(l)-X_n^x(t),\mathrm{d}K_n^x(l)\rangle\Big\}\Big)\cdot\chi_{\Omega-\Omega_1}\Big|\mathscr{F}_t\Big] \\
&\le \mathbf{E}[\mathrm{e}^{n|\lambda|}\cdot\chi_{\Omega_1}|\mathscr{F}_t] + \mathrm{e}^{nC(\lambda,r,L')s}\mathbf{E}\Big[\Big(\exp\Big\{4\lambda\sqrt{n}\int_t^{t+s}\langle X_n^x(l)-X_n^x(t),\sigma(l,X_n^x(l))\rangle\mathrm{d}W(l)\Big\} \\
&\quad + \exp\Big\{4\lambda n\int_t^{t+s}\langle X_n^x(t)-X_n^x(l),\mathrm{d}K_n^x(l)\rangle\Big\}\Big)\cdot\chi_{\Omega-\Omega_1}\Big|\mathscr{F}_t\Big] \\
&=: \mathrm{I}_1+\mathrm{I}_2+\mathrm{I}_3.
\end{aligned}
$$

Note that since

$$
\begin{aligned}
\varsigma(u) := \exp\Big(&\lambda\sqrt{n}\int_t^u\langle X_n^x(l)-X_n^x(t),\sigma(l,X_n^x(l))\rangle\mathrm{d}W(l) \\
&- \frac{n\lambda^2}{2}\int_t^u|\sigma^*(l,X_n^x(l))(X_n^x(l)-X_n^x(t))|^2\mathrm{d}l\Big)
\end{aligned}
$$

is an exponential martingale with respect to $\mathscr{F}_t$, we have $\mathbf{E}[\varsigma(t+s)|\mathscr{F}_t]=1$ and furthermore,

$$\mathrm{I}_2\le\mathrm{e}^{nC(\lambda,r,L')s}.$$

Now for $r>0$ and ε sufficiently small we set

$$\mathbb{D}_r^\varepsilon := \Big\{ax,\ 0\le a\le 1-\frac{\varepsilon}{2},\ x\in\partial(\overline{D(A)}\cap\overline{B(0,r)})\Big\}.$$

Then $\mathbb{D}_r^\varepsilon$ is a nonempty compact convex subset of $\mathrm{Int}(D(A))$ satisfying $d(x,\mathbb{D}_r^\varepsilon)<\varepsilon r$ for every $x\in\overline{D(A)}\cap\overline{B(0,r)}$.

Set

$$l_r(\varepsilon) := \sup_{x\in\mathbb{D}_r^\varepsilon}\sup_{y\in A(x)}|y|.$$

By Proposition 2.1, A is locally bounded on $\mathrm{Int}(D(A))$. Thus the function

$$(0,\varepsilon_0)\ni\varepsilon\to l_r(\varepsilon)$$

is well defined for some $\varepsilon_0>0$ and decreasing.

Define

$$q_r(s) := \inf\{\varepsilon\in(0,\varepsilon_0): l_r(\varepsilon)\le s^{-\frac{1}{2}}\}.$$

Then

$$\lim_{s\to 0} q_r(s) = 0, \quad l_r(s+q_r(s)) \le s^{-\frac{1}{2}} \quad \text{for every } s>0.$$

Now fix s sufficiently small such that $s+q_r(s)\le\varepsilon_0$. Let Y_n^x be the projection of $X_n^x(t)$ on $\mathbb{D}_r^{s+q_r(s)}$. Then $Y_n^x \in \mathrm{Int}(D(A))$ and on the set $\Omega-\Omega_1$,

$$\begin{aligned}
&\int_t^{t+s} \langle X_n^x(t) - X_n^x(l), \mathrm{d}K_n^x(l)\rangle \\
&= \int_t^{t+s} \langle X_n^x(t) - Y_n^x, \mathrm{d}K_n^x(l)\rangle + \int_t^{t+s} \langle Y_n^x - X_n^x(l), \mathrm{d}K_n^x(l)\rangle \\
&\le |X_n^x(t) - Y_n^x||K_n^x|_t^{t+s} + \int_t^{t+s} \langle Y_n^x - X_n^x(l), \mathrm{d}K_n^x(l)\rangle \\
&\le (s+q_r(s))r|K_n^x|_t^{t+s} + \int_t^{t+s} \langle Y_n^x - X_n^x(l), \mathrm{d}K_n^x(l) - z\mathrm{d}l\rangle + \int_t^{t+s} \langle Y_n^x - X_n^x(l), z\mathrm{d}l\rangle \\
&\le (s+q_r(s))r|K_n^x|_t^{t+s} + 2rs^{\frac{1}{2}},
\end{aligned}$$

where $z\in A(Y_n^x)$ is arbitrarily fixed and we used the fact $|z|\le s^{-\frac{1}{2}}$. Thus by taking

$$\xi_n(s,\lambda) := \log(\mathrm{e}^{nC(\lambda,r,L')s} + \mathrm{e}^{n|\lambda|}\cdot\chi_{\Omega_1} + \exp\{4n|\lambda|[(s+q_r(s))r|K_n^x|_T^0 + 2rs^{\frac{1}{2}}] + nC(\lambda,r,L')s\}),$$

we get (4.10).

To prove (4.11), observe that by Proposition 2.2 there exist constants γ and μ independent of n such that on $\Omega-\Omega_1$,

$$|K_n^x|_T^0 \le C(r,\mu,\gamma,L') + \frac{2}{\gamma\sqrt{n}} \int_0^T \langle X_n^x(l), \sigma(l, X_n^x(l))\rangle \mathrm{d}W(l).$$

Therefore

$$\begin{aligned}
\mathbf{E}\mathrm{e}^{\xi_n(s,\lambda)} &\le 1 + \mathrm{e}^{nC(\lambda,r,L')s} + \mathbf{E}\exp\{4n|\lambda|[(s+q_r(s))r|K_n^x|_T^0 + 2rs^{\frac{1}{2}}] + nC(\lambda,r,L')s\} \\
&\le 1 + \mathrm{e}^{nC(\lambda,r,L')s} + \mathbf{E}\Big[\exp\Big(\frac{8\sqrt{n}}{\gamma}|\lambda|(s+q_r(s))r\int_0^T \langle X_n^x(l), \sigma(l, X_n^x(l))\rangle \mathrm{d}W(l)\Big)\chi_{\Omega-\Omega_1}\Big] \\
&\quad \cdot \exp(4n|\lambda|[(s+q_r(s))C(r,\mu,\gamma,L') + 2rs^{\frac{1}{2}}] + nC(\lambda,r,L')s) \\
&\le 1 + \mathrm{e}^{nC(\lambda,r,L')s} + \exp\{4n[C(r,\mu,\lambda,\gamma,L')(s+q_r(s)+2rs^{\frac{1}{2}})]\},
\end{aligned}$$

where the last inequality follows from the fact that for every constant c and $\alpha_l := \chi_{(\sup_{u\le l}|X_n^x(u)|\le r)}$, the process

$$\exp\Big\{c\int_0^r \langle X_n^x(l)\alpha_l, \sigma(l, X_n^x(l))\rangle \mathrm{d}W(l) - \frac{c^2}{2}\int_0^r \alpha_l |X_n^x(l)|^2 \|\sigma(l, X_n^x(l))\|^2 \mathrm{d}l\Big\}_{r\in(0,T]}$$

is an exponential martingale with respect to $\mathscr{F}_r$ and that according to the locality of stochastic integrals (see e.g. [17, Chapter 4, Lemma 2.11]), the following holds:

$$\exp\Big\{c\int_0^r \langle X_n^x(l), \sigma(l, X_n^x(l))\rangle \mathrm{d}W(l) - \frac{c^2}{2}\int_0^r |X_n^x(l)|^2 \|\sigma(l, X_n^x(l))\|^2 \mathrm{d}l\Big\}\chi_{\Omega-\Omega_1}$$

$$= \exp\Big\{c\int_0^r \langle X_n^x(l)\alpha_l, \sigma(l, X_n^x(l))\rangle \mathrm{d}W(l) - \frac{c^2}{2}\int_0^r \alpha_l |X_n^x(l)|^2 \|\sigma(l, X_n^x(l))\|^2 \mathrm{d}l\Big\}\chi_{\Omega-\Omega_1}.$$

Hence with this estimate it is easy to see that

$$\lim_{s\to 0}\limsup_{n\to\infty}\frac{1}{n}\log \mathbf{E}\mathrm{e}^{\xi_n(s,\lambda)}$$
$$\leq \lim_{s\to 0}(C(\lambda, r, L')s + 4C(r,\mu,\lambda,\gamma,L')(s + q_r(s) + rs^{\frac{1}{2}})) = 0.$$

Combining Theorems 4.2–4.3, and using [6, Sections 4.4, 4.7] we arrive at the following theorem.

Theorem 4.4 *For each $x \in \overline{D(A)}$, let X_n^x denote the solution of* (4.3) *with $t = 0$. Then $\{X_n^x\}$ satisfies the LDP in $\mathcal{C}([0,T], \overline{D(A)})$.*

Remark 4.1 Using Theorem 4.2 and a standard procedure as explained in [18, Chapter 5] it is possible to write down the good rate function as

$$I(f) := \frac{1}{2}\inf\Big\{\int_0^T |z_s|^2 \mathrm{d}s;\ f = X_x^z\Big\},$$

where X_x^z is the solution to (4.8), starting from x at 0, and $f \in \mathcal{C}([0,T], \overline{D(A)})$.

5 Uniform Large Deviation Principle

Our final purpose is an LDP uniform both in time and in initial conditions. To this end we need to consider the multi-point motion. Let m be an arbitrary positive integer and $x_i \in \mathbb{R}^d$, $i = 1, \cdots, m$. Set $\mathbf{x} = (x_1, \cdots, x_m)$ and

$$\sigma^{(m)}(t,\mathbf{x}) := \begin{pmatrix}\sigma(t.x_1)\\ \sigma(t,x_2)\\ \vdots \\ \sigma(t,x_m)\end{pmatrix}, \quad b^{(m)}(t,\mathbf{x}) := \begin{pmatrix}b(t,x_1)\\ b(t,x_2)\\ \vdots \\ b(t,x_m)\end{pmatrix}, \quad A^{(m)}(\mathbf{x}) := \begin{pmatrix}A(x_1)\\ A(x_2)\\ \vdots \\ A(x_m)\end{pmatrix}.$$

Then $\sigma^{(m)}$ and $b^{(m)}$ are functions defined on $(\mathbb{R}^d)^m$ and $A^{(m)}$ is a multivalued maximal operator with domain $D(A)^m$ and they satisfy the assumptions (H1)–(H2) in Section 2.

We consider the following MSDE:

$$\begin{cases}\mathrm{d}X_n^{(m)}(t) \in b^{(m)}(t, X_n^{(m)}(t))\mathrm{d}t + \dfrac{1}{\sqrt{n}}\sigma^{(m)}(t, X_n^{(m)}(t))\mathrm{d}W(t) - A(X^{(m)}(t))\mathrm{d}t,\\ X_n^{(m)}(0) = \mathbf{x} \in \overline{D(A)^m}.\end{cases} \tag{5.1}$$

Denote its solution by $X_n^{(m)}(\mathbf{x},\cdot)$. Then using Theorem 4.4, we have the following result.

Theorem 5.1 $\{X_n^{(m)}(\mathbf{x},\cdot)\}$ *satisfies the LDP.*

As what we have just done in Section 5, to pass from the finite multi-point motion LDP to the LDP on $\mathcal{C}([0,T]\times\overline{D(A)}, \overline{D(A)})$ we have to prove the exponential tightness in this latter space. This will be done with the help of the following lemma.

Lemma 5.1 *Let $\mathbb{E}$ be a separable Banach space and suppose that for every $x \in \overline{D(A)}$, $\{\xi_n(x)\} \subset \mathbb{E}$ is exponentially tight and there is a constant C independent of n such that*

$$\mathbf{E}|\xi_n(x) - \xi_n(y)|_{\mathbb{E}}^n \leq C^n|x - y|^n.$$

Then $\{\xi_n(\cdot)\} \subset \mathcal{C}(\overline{D(A)}, \mathbb{E})$ is exponentially tight. Here $\mathcal{C}(\overline{D(A)}, \mathbb{E})$ is endowed with the locally uniform topology.

Proof For an integer $r \geq 0$ set

$$L_r := \overline{D(A)} \cap \{x, |x| \leq r\}.$$

By [19, Chapter 1, Theorem 2.1], for every $r > 0$ there exists a constant C_r independent of n such that

$$\mathbf{E}\Big[\sup_{\substack{x \neq y \\ x,y \in L_r}} \frac{|\xi_n(x) - \xi_n(y)|^n}{|x - y|^{\frac{n}{2}}}\Big] \leq C_r^n$$

for all n. Thus by Chebeyshev's inequality we have for all $\varepsilon > 0$,

$$\mathbf{P}\Big(\sup_{\substack{0 < |x-y| < \delta \\ x,y \in L_r}} |\xi_n(x) - \xi_n(y)| \geq \varepsilon\Big) \leq \varepsilon^{-n} \delta^{\frac{n}{2}} C_r^n, \quad \forall n. \tag{5.2}$$

Take a dense subset $G = \{x_i\}_{i=1}^{\infty} \subset \overline{D(A)}$ and set

$$G_r := G \cap L_r.$$

Let $M \geq 1$ be arbitrarily fixed. For each i take a compact $E_i \subset \mathbb{E}$ such that

$$\mathbf{P}(\xi_n(x_i) \notin E_i) \leq \mathrm{e}^{-2Mni}.$$

Taking $\delta_{r,i} = 2^{-2i(4Mr+1)} C_r^{-2}$ we have by (5.2)

$$\mathbf{P}\Big(\sup_{\substack{|x-y| \leq \delta_{r,i} \\ x,y \in L_r}} |\xi_n(x) - \xi_n(y)|_{\mathbb{E}} \geq 2^{-i}\Big) \leq 2^{-4Mnir}.$$

Let

$$\Gamma := \Big\{f \in \mathcal{C}(\overline{D(A)}, \mathbb{E}) : \forall r, i,\ f(x_i) \in E_i \text{ and } \sup_{\substack{|x-y| \leq \delta_{r,i} \\ x,y \in L_r \setminus L_{r-1}}} |f(x) - f(y)|_{\mathbb{E}} < 2^{-i}\Big\}.$$

We claim $\Gamma \subset \mathcal{C}(\overline{D(A)}, \mathbb{E})$ is relatively compact.

Taking the claim for granted we have

$$\mathbf{P}(\xi_n \notin \Gamma) \leq \mathbf{P}(\exists i,\ \xi_n(x_i) \notin E_i) + \mathbf{P}\Big(\exists r, i, \sup_{\substack{|x-y| \leq \delta_{r,i} \\ x,y \in L_r \setminus L_{r-1}}} |\xi_n(x) - \xi_n(y)| \geq 2^{-i}\Big)$$

$$\leq \sum_{i=1}^{\infty} 2^{-2Mni} + \sum_{r,i=1}^{\infty} 2^{-4Mnri} \leq 2^{-Mn},$$

from which we deduce

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbf{P}(\xi_n\notin\Gamma)\leq -M\log 2.$$

Hence we are done by the arbitrariness of M.

Now we prove the claim. For this we have to prove that Γ_r is relatively compact in $\mathcal{C}(L_r,\mathbb{E})$ for every r where Γ_r is the set consisting of the restriction to L_r of the elements of Γ. Obviously, Γ_r is an equicontinuous family and hence it only remains to prove that $\{f(x),\ f\in\Gamma_r\}$ is relatively compact in $\mathbb{E}$ for every $x\in L_r$. In fact, let $x\in L_r$ be fixed. Write

$$G_r=\{x_{r_1},x_{r_2},\cdots\}$$

with $r_1<r_2<\cdots$. To simplify notations we set $y_k:=x_{r_k}$ and $F_k:=E_{r_k}$. For $\varepsilon>0$, take first an integer $k_1>2\varepsilon^{-1}$ and then an integer k_2 such that $|x-y_{k_2}|<\delta_{r,k_1}$. Since F_{k_2} is totally bounded, there exists a finite $\frac{\varepsilon}{2}$-net $\big\{B\big(e_i,\frac{\varepsilon}{2}\big)\big\}_{i=1}^N$ containing F_{k_2}. It follows that $\{B(e_i,\varepsilon)\}_{i=1}^N$ is a finite ε-net containing $\{f(x),\ f\in\Gamma_r\}$. Hence $\{f(x),f\in\Gamma_r\}$ is relatively compact. Now the relative compactness of Γ_r follows by Arzelà-Ascoli Theorem and the proof is complete.

Next we prove the following lemma.

Lemma 5.2 *For the sequence of solutions of* (4.3), $\{X_n(\cdot,\cdot)\}_{n=1}^\infty$ *is exponentially tight in* $\mathcal{C}(\overline{D(A)};\mathcal{C}([0,T],\overline{D(A)}))$.

Proof By Itô's formula and Proposition 2.2, we have

$$|X_n(x,t)-X_n(y,t)|^n\leq|x-y|^n+M_{n,t}+Cn\int_0^t|X_n(x,s)-X_n(y,s)|^n\mathrm{d}s, \tag{5.3}$$

where

$$\begin{aligned}M_{n,t}=n^{\frac12}\int_0^t&|X_n(x,s)-X_n(y,s)|^{n-1}\mathrm{sgn(X_n(x,s)-X_n(y,s))}\\&\cdot(\sigma(X_n(x,s))-\sigma(X_n(y,s)))\mathrm{d}W_s.\end{aligned} \tag{5.4}$$

Taking expectation gives

$$\mathbf{E}[|X_n(x,t)-X_n(y,t)|^n]\leq|x-y|^n+Cn\int_0^t\mathbf{E}[|X_n(x,t)-X_n(y,t)|^n]\mathrm{d}s.$$

By Gronwall's lemma,

$$\mathbf{E}[|X_n(x,t)-X_n(y,t)|^n]\leq|x-y|^n\mathrm{e}^{Cnt}.$$

Using (5.4), Doob's inequality and BDG inequality, we have

$$\begin{aligned}\mathbf{E}\Big[\sup_{0\leq s\leq t}|M_{n,s}|^2\Big]&\leq Cn\int_0^t\mathbf{E}[|X_n(x,s)-X_n(y,s)|^{2n}]\mathrm{d}s\\&\leq Cn\mathrm{e}^{Cnt}|x-y|^{2n}.\end{aligned}$$

Hence

$$\mathbf{E}\Big[\sup_{0\leq s\leq t}|M_{n,s}|\Big]\leq Cn^{\frac12}\mathrm{e}^{Cnt}|x-y|^n.$$

By (5.3) we have

$$\begin{aligned}&\sup_{0\le s\le t}|X_n(x,s)-X_n(y,s)|^n\\ \le&|x-y|^n+\sup_{0\le s\le t}|M_{n,s}|+Cn\int_0^t\Big[\sup_{0\le u\le s}|X_n(x,u)-X_n(y,u)|^n\Big]\mathrm{d}s.\end{aligned}$$

Taking expectation we have

$$\begin{aligned}&\mathbf{E}\Big[\sup_{0\le s\le t}|X_n(x,s)-X_n(y,s)|^n\Big]\\ \le&\mathrm{e}^{Cn}|x-y|^n+Cn\int_0^t\mathbf{E}\Big[\sup_{0\le u\le s}|X_n(x,u)-X_n(y,u)|^n\Big]\mathrm{d}s.\end{aligned}$$

By Gronwall's lemma we have

$$\mathbf{E}\Big[\sup_{0\le s\le t}|X_n(x,s)-X_n(y,s)|^n\Big]\le\mathrm{e}^{Cn}|x-y|^n.$$

Hence by Lemma 5.1, the family $\{X_n\}_{n=1}^\infty$ is exponentially tight in $\mathcal{C}(\overline{D(A)},\mathcal{C}([0,T],\overline{D(A)}))$.

Lemma 5.2 combined with Theorem 5.1 shows that $\{X_n(\cdot,\cdot)\}_{n=1}^\infty$ satisfies the LDP in $\mathcal{C}(\overline{D(A)},\mathcal{C}([0,T],\overline{D(A)}))$. But

$$\mathcal{C}(\overline{D(A)},\mathcal{C}([0,T],\overline{D(A)}))\cong\mathcal{C}(\overline{D(A)}\times[0,T],\overline{D(A)}),$$

so we have completed the proof of the following theorem, which is the main result of the present paper.

Theorem 5.2 $\{X_n\}$ *satisfies the LDP in* $\mathcal{C}(\overline{D(A)}\times[0,T],\overline{D(A)})$.

Remark 5.1 We have a similar remark as Remark 4.1 and, consequently, have a variational expression of the rate function as

$$I(f):=\inf\Big\{\frac{1}{2}\int_0^T|z(s)|^2\mathrm{d}s,\ f_x=X_x^z,\ \forall x\in\overline{D(A)}\Big\}$$

for $f\in\mathcal{C}(\overline{D(A)}\times[0,T],\overline{D(A)})$.

6 Appendix

Proposition 6.1 *Suppose that u and v are respectively viscosity subsolution and supersolution to HJB equation:*

$$\begin{cases}\dfrac{\partial u}{\partial t}+H(t,x,Du,D^2u)\in\langle A(x),Du\rangle & \text{in } (0,T)\times\overline{D(A)},\\ u(T,\cdot)=h(\cdot) & \text{on } \overline{D(A)}\end{cases}\tag{6.1}$$

with

$$H(t,x,q,M):=-\frac{1}{2}|\sigma^*(t,x)q|^2+\langle b(t,x),q\rangle+\frac{c}{2}\mathrm{tr}(\sigma\sigma^*(t,x)M),\quad 0\le c<\infty.$$

Assume that $\|\sigma(t,x)\|\leq M$ *for some* $M>0$, u,v *are bounded from above and either* u *or* v *is Lipscthiz continuous with respect to* x. *Then* $u\leq v$ *on* $[0,T]\times\overline{D(A)}$.

Proof Through a transformation $u\to \mathrm{e}^{\lambda t}u$ we consider the following HJB equation:

$$\begin{cases}-\lambda u+\dfrac{\partial u}{\partial t}+H(t,x,Du,D^2u)\in\langle A(x),Du\rangle,\\ u(T,\cdot)=\mathrm{e}^{\lambda T}h(\cdot)\quad\text{on } \overline{D(A)}.\end{cases}\tag{6.2}$$

Here $\lambda>0$ is a constant.

Assume that u and v are subsolution and supersolution of (6.2) respectively. Suppose there exists $(t_0,x_0)\in(0,T]\times\overline{D(A)}$ such that $l_0:=u(t_0,x_0)-v(t_0,x_0)>0$. For $\varepsilon>0$, define

$$u^\varepsilon(t,x):=u(t,x)-\varepsilon|x|^2,\quad v^\varepsilon(t,x):=v(t,x)+\varepsilon|x|^2.$$

For $\alpha>0$ and $(t,x,s,y)\in((0,T]\times\overline{D(A)})^2$, set

$$\begin{aligned}&\Phi_{\alpha,\varepsilon}:=u^\varepsilon(t,x)-v^\varepsilon(s,y)-\psi_\alpha(t,x,s,y),\\ &\Psi(t,x,s,y):=\frac{\alpha}{2}(|x-y|^2+|t-s|^2)+\frac{l_0t_0}{8}\Big(\frac{1}{t}+\frac{1}{s}\Big).\end{aligned}$$

Note that for $\varepsilon\leq\frac{l_0}{4|x_0|^2}$ and every $\alpha>0$,

$$\Phi_{\alpha,\varepsilon}(t_0,x_0,t_0,x_0)>0,\qquad \lim_{\frac{1}{t}\vee\frac{1}{s}\vee|x|\vee|y|\to\infty}\Phi_{\alpha,\varepsilon}(t,x,s,y)\leq 0.$$

Thus there exists $\overline{\zeta}:=(\overline{t},\overline{x},\overline{s},\overline{y})\in((0,T]\times\overline{D(A)})^2$ (here and in what follows ζ and $\overline{t},\overline{x},\overline{s},\overline{y}$ depend on α,ε but for simplicity we drop the subscripts) such that

$$N_{\alpha,\varepsilon}:=\Phi_{\alpha,\varepsilon}(\overline{\zeta})=\sup_{((0,T]\times\overline{D(A)})^2}\Phi_{\alpha,\varepsilon}(t,x,s,y)<+\infty.$$

One then gets

$$N_{\alpha,\varepsilon}\geq u(t_0,x_0)-v(t_0,x_0)-2\varepsilon|x_0|^2-\frac{l_0}{4}\geq\frac{l_0}{2}>0,\quad\forall\varepsilon\leq\frac{l_0}{8|x_0|^2}.\tag{6.3}$$

Note that $\alpha\to N_{\alpha,\varepsilon}$ is decreasing. Then the limit $\lim\limits_{\alpha\to+\infty}M_{\alpha,\varepsilon}$ exists and is finite. Moreover,

$$N_{\frac{\alpha}{2},\varepsilon}\geq\Phi_{\frac{\alpha}{2},\varepsilon}(\overline{\zeta})=N_{\alpha,\varepsilon}+\frac{\alpha}{4}(|\overline{x}-\overline{y}|^2+|\overline{t}-\overline{s}|^2),$$

which yields that

$$\lim_{\alpha\to+\infty}\alpha(|\overline{x}-\overline{y}|^2+|\overline{t}-\overline{s}|^2)=0.\tag{6.4}$$

Since

$$\Phi_{\alpha,\varepsilon}(\overline{\zeta})\geq\Phi_{\alpha,\varepsilon}(\overline{t},\overline{x},\overline{s},\overline{x}),$$

we get

$$\frac{\alpha}{2}|\overline{x}-\overline{y}|^2\leq v(\overline{s},\overline{x})-v(\overline{s},\overline{y})+\varepsilon(|\overline{x}|^2-|\overline{y}|^2)$$

$$\leq L|\overline{x}-\overline{y}|+\varepsilon(|\overline{x}|+|\overline{y}|)|\overline{x}-\overline{y}|,$$

which yields that

$$\alpha|\overline{x}-\overline{y}|\leq 2L+2\varepsilon(|\overline{x}|+|\overline{y}|). \tag{6.5}$$

If for some fixed $\varepsilon\leq\varepsilon_0$, there exists an increasing unbounded sequence α_n such that $t_{\alpha_n,\varepsilon}\vee s_{\alpha_n,\varepsilon}=T$. Then one can find $\widetilde{x}\in\overline{D(A)}$ such that $(t_{\alpha_n,\varepsilon},x_{\alpha_n,\varepsilon},s_{\alpha_n,\varepsilon},y_{\alpha_n,\varepsilon})\to(T,\widetilde{x},T,\widetilde{x})$. By (6.3),

$$\frac{l_0}{2}\leq\lim_n N_{\alpha_n,\varepsilon}\leq\limsup_n(u(t_{\alpha_n,\varepsilon},x_{\alpha_n,\varepsilon})-v(s_{\alpha_n,\varepsilon},y_{\alpha_n,\varepsilon}))\leq u(T,\widetilde{x})-v(T,\widetilde{x})\leq 0.$$

This contradicts the assumption $l_0>0$. Therefore for $\varepsilon\leq\varepsilon_0$, there exists $\alpha_\varepsilon>0$ such that for $\alpha>\alpha_\varepsilon$, $\overline{\zeta}\in((0,T)\times\overline{D(A)})^2$.

Take $\varepsilon\leq\frac{l_0}{8|x_0|^2}$. By applying [20, Theorem 3.2] at the point $\overline{\zeta}$, we can find matrices $Q,R\in\mathcal{S}^d$ such that

$$\Big(\frac{\partial\Psi}{\partial t}(\overline{\zeta}),D_x\Psi(\overline{\zeta})+2\varepsilon\overline{x},Q+2\varepsilon\Big(\frac{\overline{x}\otimes\overline{x}}{|\overline{x}|^2}+I\Big)\Big)\in\overline{\mathcal{P}}^{1,2,+}_{\overline{D(A)}}u(\overline{t},\overline{x}),$$
$$\Big(-\frac{\partial\Psi}{\partial s}(\overline{\zeta}),-D_y\Psi(\overline{\zeta})-2\varepsilon\overline{y},R-2\varepsilon\Big(\frac{\overline{y}\otimes\overline{y}}{|\overline{y}|^2}+I\Big)\Big)\in\overline{\mathcal{P}}^{1,2,-}_{\overline{D(A)}}v(\overline{s},\overline{y}),$$

where $\mathcal{P}^{1,2,\pm}_{\overline{D(A)}}$ is the same as in [2, 20] and

$$\begin{pmatrix}Q & 0\\ 0 & -R\end{pmatrix}\leq 3\alpha\begin{pmatrix}I & -I\\ -I & I\end{pmatrix}. \tag{6.6}$$

Then since u (resp. v) ia a viscosity subsolution (resp. supersolution), we have

$$-\lambda u(\overline{t},\overline{x})+\frac{\partial\Psi}{\partial t}(\overline{\zeta})+H\Big(\overline{t},\overline{x},\alpha(\overline{x}-\overline{y})+2\varepsilon\overline{x},Q+2\varepsilon\Big(\frac{\overline{x}\otimes\overline{x}}{|\overline{x}|^2}+I\Big)\Big)\geq A_*(\overline{x},\alpha(\overline{x}-\overline{y})+2\varepsilon\overline{x}),$$
$$-\lambda v(\overline{s},\overline{y})-\frac{\partial\Psi}{\partial s}(\overline{\zeta})+H\Big(\overline{s},\overline{y},\alpha(\overline{x}-\overline{y})-2\varepsilon\overline{y},R-2\varepsilon\Big(\frac{\overline{y}\otimes\overline{y}}{|\overline{y}|^2}+I\Big)\Big)\leq A^*(\overline{y},\alpha(\overline{x}-\overline{y})-2\varepsilon\overline{y}).$$

Note that for $\overline{x},\overline{y}\in\overline{D(A)}$ (see [2, Lemma 2]),

$$A_*(\overline{x},\alpha(\overline{x}-\overline{y})+2\varepsilon\overline{x})=\inf_{x^*\in A(\overline{x})}\langle x^*,\alpha(\overline{x}-\overline{y})+2\varepsilon\overline{x}\rangle,$$
$$A^*(\overline{y},\alpha(\overline{x}-\overline{y})-2\varepsilon\overline{y})=\sup_{y^*\in A(\overline{y})}\langle y^*,\alpha(\overline{x}-\overline{y})-2\varepsilon\overline{y}\rangle.$$

Then by simple calculation one gets

$$A_*(\overline{x},\alpha(\overline{x}-\overline{y})+2\varepsilon\overline{x})\geq A^*(\overline{y},\alpha(\overline{x}-\overline{y})-2\varepsilon\overline{y}).$$

By subtraction we get

$$-\lambda(u(\overline{t},\overline{x})-v(\overline{s},\overline{y}))-\frac{l_0t_0}{8}\Big(\frac{1}{\overline{t}^2}+\frac{1}{\overline{s}^2}\Big)$$

$$
\begin{aligned}
&-\frac{1}{2}(|\sigma^*(\overline{t},\overline{x})[\alpha(\overline{x}-\overline{y})+2\varepsilon\overline{x}]|^2-|\sigma^*(\overline{s},\overline{y})[\alpha(\overline{x}-\overline{y})-2\varepsilon\overline{y}]|^2)\\
&+\langle b(\overline{t},\overline{x})-b(\overline{s},\overline{y}),\alpha(\overline{x}-\overline{y})\rangle+2\varepsilon(\langle b(\overline{t},\overline{x}),\overline{x}\rangle-\langle b(\overline{s},\overline{y}),\overline{y}\rangle)\\
&+\frac{c}{2}\mathrm{tr}\Big(\sigma\sigma^*(\overline{t},\overline{x})\Big(Q+2\varepsilon\Big(\frac{\overline{x}\otimes\overline{x}}{|\overline{x}|^2}+I\Big)\Big)\Big)-\frac{c}{2}\mathrm{tr}\Big(\sigma\sigma^*(\overline{s},\overline{y})\Big(R-2\varepsilon\Big(\frac{\overline{y}\otimes\overline{y}}{|\overline{y}|^2}+I\Big)\Big)\Big)\\
\geq& A_*(\overline{x},\alpha(\overline{x}-\overline{y})+2\varepsilon\overline{x})-A^*(\overline{y},\alpha(\overline{x}-\overline{y})-2\varepsilon\overline{y})\geq 0.
\end{aligned}
$$

Applying (6.5) here gives

$$
\begin{aligned}
&\frac{1}{2}\lambda l_0+\lambda\varepsilon(|\overline{x}|^2+|\overline{y}|^2)\leq\lambda(u(\overline{t},\overline{x})-v(\overline{s},\overline{y}))\\
\leq&\frac{1}{2}|\sigma^*(\overline{t},\overline{x})[\alpha(\overline{x}-\overline{y})+2\varepsilon\overline{x}]|^2+\frac{1}{2}|\sigma^*(\overline{s},\overline{y})[\alpha(\overline{x}-\overline{y})-2\varepsilon\overline{y}]|^2\\
&+|b(\overline{t},\overline{x})-b(\overline{s},\overline{y})||\alpha(\overline{x}-\overline{y})|+2\varepsilon|\langle b(\overline{t},\overline{x}),\overline{x}\rangle-\langle b(\overline{s},\overline{y}),\overline{y}\rangle|\\
&+\frac{c}{2}[\mathrm{tr}(\sigma\sigma^*(\overline{t},\overline{x})Q)-\mathrm{tr}(\sigma\sigma^*(\overline{s},\overline{y})R)]+4cM^2d\varepsilon\\
\leq&16M^2L^2|\overline{x}-\overline{y}|^2+36M^2\varepsilon^2(|\overline{x}|^2+|\overline{y}|^2)+4cM^2d\varepsilon+\frac{3c}{2}L^2\alpha(|\overline{t}-\overline{s}|+|\overline{x}-\overline{y}|)^2\\
&+L\alpha|\overline{x}-\overline{y}|(|\overline{t}-\overline{s}|+|\overline{x}-\overline{y}|)+2\varepsilon|\overline{x}|(1+|\overline{x}|)+2\varepsilon|\overline{y}|(1+|\overline{y}|)\\
\leq&(3c+2)L^2\alpha(|\overline{t}-\overline{s}|^2+|\overline{x}-\overline{y}|^2)+\varepsilon[|\overline{x}|^2+(1+|\overline{x}|)^2+|\overline{y}|^2+(1+|\overline{y}|)^2]\\
&+(16M^2L^2+\alpha)|\overline{x}-\overline{y}|^2+36M^2\varepsilon^2(|\overline{x}|^2+|\overline{y}|^2)+4cM^2d\varepsilon\\
\leq&(3c+2)L^2\alpha(|\overline{t}-\overline{s}|^2+|\overline{x}-\overline{y}|^2)+3\varepsilon(1+|\overline{x}|^2+|\overline{y}|^2)\\
&+(16M^2L^2+\alpha)|\overline{x}-\overline{y}|^2+36M^2\varepsilon(|\overline{x}|^2+|\overline{y}|^2)+4cM^2d\varepsilon.
\end{aligned}
$$

By taking $\lambda=(3+36M^2)$ we get

$$
\begin{aligned}
\frac{1}{2}\lambda l_0\leq&(3c+2)L^2\alpha(|\overline{t}-\overline{s}|^2+|\overline{x}-\overline{y}|^2)\\
&+(16M^2L^2+\alpha)|\overline{x}-\overline{y}|^2+(3+4cM^2d)\varepsilon,
\end{aligned}
$$

which tends to 0 by sending $\alpha\to\infty$ and then $\varepsilon\to 0$. Hence we get a contraction and $u\leq v$ on $(0,T]\times\overline{D(A)}$.

Acknowledgement The authors are very grateful to the referees for the valuable comments.

References

[1] Cépa, E., Probleme de Skorohod multivoque, *Ann. Prob.,* **26**(2), 1998, 500–532.

[2] Zălinescu, A., Second order Hamilton-Jacobi-Bellman equations with unbounded operator, *Nonlinear Analysis, TMA,* **75**, 2012, 4784–4797.

[3] Ren, J., Wu, J. and Zhang, H., General large deviations and functional iterated logarithm law for multivalued stochastic differential equations, *J. Theoret. Probab.,* **28**(2), 2015, 550–586.

[4] Ren, J., Xu, S. and Zhang, X., Large deviation for multivalued stochastic differenital equations, *J. Theoretical Prob.,* **23**(4), 2010, 1142–1156.

[5] Evans, L. C. and Ishii, H., A PDE approach to some asymptotic problems concerning random differential equations with small noise intensities, *Ann. Inst. H. Poincaré Anal. Non Linéaire,* **2**(1), 1985, 1–20.

[6] Feng, J. and Kurtz, T., Large Deviations for Stochastic Processes, Mathematical Surveys and Monographs, Vol. 131, American Mathematical Society, Providence, RI, 2006.

[7] Kobylanski, M., Large deviations principle by viscosity solutions: The case of diffusions with oblique Lipschitz reflections, *Ann. Inst. H. Poincaré, Probab. Statist.,* **49**(1), 2013, 160–181.

[8] Puhalskii, A., On functional principle of large deviations, New Trends in Probability and Statistics, **1**, V. Sazonov and T. Shervashidze (eds.), VSP/Moks'las, Vilnius, 1991, 198–219.

[9] Puhalskii, A., The method of stochastic exponentials for large deviations, *Stoch. Proc. Appl.,* **54**, 1994, 45–70.

[10] Puhlaskii, A., Large Deviations and Idempotent Probability, Chapman and Hall/CRC, New York, 2001.

[11] Święch, A., A PDE approach to large devaitons in Hilbert spaces, *Stoch. Proc. Appl.,* **119**, 2009, 1081–1123.

[12] Feng, J., Large deviations for diffusions and Hamilton-Jacobi equations in Hilbert spaces, *Ann. Probab.,* **34**(1), 2006, 321–385.

[13] Brézis, H., Opérateurs Maximaux Monotones, North-Holland, Amsterdam, London, 1973.

[14] Cardaliaguet, P., Solutions de viscosité d'équations elliptiques et paraboliques non linéaires, http://www.ann.jussieu.fr/frey/papers/levelsets/.

[15] Tataru, D., Viscosity solutions for the dynamic programming equations, *Appl. Math. Optim.,* **25**(2), 1992, 109–126.

[16] Dupuis, P. and Ellis, R. S., A Weak Convergence Approach to the Theory of Large Deviations, Wiley, New York, 1997.

[17] Friedman, A., Stochastic Diffrential Equations and Applications, Dover Publ., Mineola, New York, 2006.

[18] Dembo, A. and Zeitouni, O., Large Deviations Techniques and Applications, Springer-Verlag, New York, 1998.

[19] Revuz, D. and Yor, M., Continuous Martingales and Brownian Motion, Grundlehren der Mathematischen Wissenschaften, **293**, 3rd ed., Springer-Verlag, Berlin, 1999.

[20] Crandall, M. G., Ishii, H. and Lions, P. L., User's guide to viscosity solutions of second order partial differential equations, *Bull. Amer. Math. Soc.*, **27**, 1992, 1–67.