Stochastic Path-Dependent Hamilton-Jacobi-Bellman Equations and Controlled Stochastic Differential Equations with Random Path-Dependent Coefficients

Jinniao Qiu

June 24, 2020

Abstract

In this paper, we propose and study the stochastic path-dependent Hamilton-Jacobi-Bellman (SPHJB) equation that arises naturally from the optimal stochastic control problem of stochastic differential equations with path-dependence and measurable randomness. Both the notions of viscosity solution and classical solution are proposed, and the value function of the optimal stochastic control problem is proved to be the viscosity solution to the associated SPHJB equation. A uniqueness result about viscosity solutions is also given for certain superparabolic cases, while the uniqueness of classical solution is addressed for general cases. In addition, an Itô-Kunita-Wentzell-Krylov formula is proved for the compositions of random fields and stochastic differential equations in the path-dependent setting.

Mathematics Subject Classification (2010): 49L20, 49L25, 93E20, 35D40, 60H15

Keywords: stochastic path-dependent Hamilton-Jacobi-Bellman equation, stochastic optimal control, viscosity solution, backward stochastic partial differential equation

1 Introduction

Let \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})\) be a complete filtered probability space on which the filtration \(\{\mathcal{F}_t\}_{t \geq 0}\) satisfies the usual conditions and is generated by an \(m\)-dimensional Wiener process \(W = \{W(t) : t \in [0, \infty)\}\) together with all the \(\mathbb{P}\)-null sets in \(\mathcal{F}\). The associated predictable \(\sigma\)-algebra on \(\Omega \times [0, T]\) is denoted by \(\mathcal{P}\). Let \(C([0, T]; \mathbb{R}^d)\) be the space of \(\mathbb{R}^d\)-valued continuous functions on \([0, T]\). For each \(x \in C([0, T]; \mathbb{R}^d)\), denote by \(x_t\) its restriction to time interval \([0, t]\) for each \(t \in [0, T]\) and by \(x(t)\) its value at time \(t \in [0, T]\).

Consider the following stochastic optimal control problem

\[
\min_{\theta \in \mathcal{U}} E\left[\int_0^T f(s, X_s, \theta(s)) \, ds + G(X_T)\right]
\]  

subject to

\[
\begin{aligned}
  dX(t) &= \beta(t, X_t, \theta(t)) \, dt + \sigma(t, X_t, \theta(t)) \, dW(t), \quad t \geq 0; \\
  X_0 &= x_0 \in \mathbb{R}^d.
\end{aligned}
\]  

Here and throughout this paper, the number \(T \in (0, \infty)\) denotes a fixed deterministic terminal time, and \(\mathcal{U}\) represents the set of all the \(U\)-valued and \(\mathcal{F}_t\)-adapted processes with \(U \subset \mathbb{R}^m\) \((m \in \mathbb{N}^+)\) being a nonempty set. The state process \((X(t))_{t \in [0, T]}\), governed by the control \(\theta \in \mathcal{U}\) may be written

---

1. This work was partially supported by the National Science and Engineering Research Council of Canada (NSERC) and by the start-up funds from the University of Calgary.
2. Department of Mathematics & Statistics, University of Calgary, 2500 University Drive NW, Calgary, AB T2N 1N4, Canada. E-mail: jinniao.qiu@ucalgary.ca (J. Qiu).
as \( X^{r,x_r;\theta}(t) \) for \( 0 \leq r \leq t \leq T \) to indicate the dependence of the state process on the control \( \theta \), the initial time \( r \) and initial path \( x_r \).

In this paper, we consider the non-Markovian case where the coefficients \( \beta, \sigma, f \), and \( G \) depend not only on time and control but also explicitly on \( \omega \in \Omega \) and paths/history of the state process. Define the dynamic cost functional:

\[
J(t, x_t; \theta) = E_{\mathcal{F}_t} \left[ \int_t^T f(s, X^t_{s}; \theta(s)) \, ds + G(X^t_{T}; \theta) \right], \quad t \in [0, T],
\]

(1.3)

where \( E_{\mathcal{F}_t} \) denotes the conditional expectation with respect to \( \mathcal{F}_t \). Then, the value function is given by

\[
V(t, x_t) = \text{essinf}_{\theta \in \mathcal{U}} J(t, x_t; \theta), \quad t \in [0, T].
\]

(1.4)

Due to the randomness and path-dependence of the coefficient \( s \), the value function \( V(t, x_t) \) is generally a function of time \( t \), path \( x_t \), and \( \omega \in \Omega \), and it proves to satisfy the following stochastic path-dependent Hamilton-Jacobi-Bellman (SPHJB) equation:

\[
\begin{align*}
-\partial_t u(t, x_t) - \mathbb{H}(t, x_t, \nabla u(t, x_t), \nabla^2 u(t, x_t), \partial_\omega \nabla u(t, x_t)) &= 0, \quad (t, x) \in [0, T) \times C([0, T]; \mathbb{R}^d); \\
u(T, x) &= G(x), \quad x \in C([0, T]; \mathbb{R}^d),
\end{align*}
\]

(1.5)

with

\[
\mathbb{H}(t, x_t, p, A, B) = \text{essinf}_{v \in U} \left\{ \text{tr} \left( \frac{1}{2} \sigma \sigma'(t, x_t, v) A + \sigma(t, x_t, v) B \right) + \beta'(t, x_t, v) p + f(t, x_t, v) \right\},
\]

for \( (p, A, B) \in \mathbb{R}^d \times \mathbb{R}^{d \times d} \times \mathbb{R}^{m \times d} \). Here, \( \nabla u(t, x_t) \) and \( \nabla^2 u(t, x_t) \) represent, respectively, the first and second order vertical derivative of \( u(t, x_t) \) at the path \( x_t \) (see Definition 2.1) and the unknown adapted random field \( u \) is confined to the following form:

\[
u(t, x_t) = u(T, x_{t,T-t}) - \int_t^T \partial_\omega u(s, x_{t,s-t}) \, ds - \int_t^T \partial_\omega u(s, x_{t,s-t}) \, dW(s),
\]

(1.6)

where \( x_{t,r-t}(s) = x_t(s)1_{[0,t]}(s) + x_r(s)1_{[t,r]}(s) \) for \( 0 \leq t \leq s \leq r \leq T \). The Doob-Meyer decomposition theorem indicates the uniqueness of the pair \( (\partial_t u, \partial_\omega u) \) and thus the linear operators \( \partial_t \) and \( \partial_\omega \) are well defined in certain spaces (see Definition 2.2). The pair \( (\partial_t u, \partial_\omega u) \) may also be defined as two differential operators; see [3] Section 5.2 and [17] Theorem 4.3) for instance.

When it holds that

\[
\partial_\omega \nabla u(t, x_t) = \nabla \partial_\omega u(t, x_t), \quad \text{a.s. for all } t \in [0, T] \text{ and } x \in C([0, T]; \mathbb{R}^d),
\]

(1.7)

putting \( \psi = \partial_\omega u \) and comparing (1.5) and (1.6), we may rewrite the SPHJB equation (1.5) as:

\[
\begin{align*}
-\partial_t u(t, x_t) &= \mathbb{H}(t, x_t, \nabla u(t, x_t), \nabla^2 u(t, x_t), \nabla \psi(t, x_t)) dt - \psi(t, x_t) \, dW(t); \\
u(T, x) &= G(x), \quad x \in C([0, T]; \mathbb{R}^d),
\end{align*}
\]

which turns out to be a fully nonlinear backward stochastic partial differential equation (BSPDE), nevertheless, defined on path spaces. In fact, the relation (1.7) holds true when all the coefficients are just state-dependent, i.e., with probability one, \( (\beta, \sigma, f)(t, x_t, v) = (\beta, \sigma, f)(t, x(t), v) \), and \( G(x_T) = G(x(T)) \) for all \( (t, x, v) \in [0, T] \times C([0, T]; \mathbb{R}^d) \times U \); see [23] for instance. This sheds light on the connections between SPHJB equation (1.5) and the BSPDEs; for related research on general BSPDEs, we refer to [11][14][15][16] among many others. However, the exchangeability (1.7) generally does not hold due to the path-dependence (see Remark 2.1 for examples), which makes SPHJB equation (1.5) stand beyond the realm of BSPDEs.
When all the coefficients $\beta, \sigma, f,$ and $G$ are deterministic continuous path-dependent functions of time $t$, control $\theta$, and the paths of $(X, W)$, the SPHJB equation (1.5) falls into the range of so-called fully nonlinear path-dependent partial differential equations (PPDEs); refer to [5, 6, 11] to mention just a few. The viscosity solution theory of such PPDEs involves admirable path-dependence and in fact gives a nontrivial meaning to the proposed SPHJB equations and fully nonlinear path-dependent partial differential equations (PPDEs); refer to [5, 6, 11, 12, 13, 21, 25, 28] to mention just a few. The viscosity solution theory of such PPDEs involves admirable path-dependence and in fact gives a nontrivial meaning to the proposed SPHJB equations and fully nonlinear path-dependent partial differential equations (BSDEs); for instance, the following trivial BSDE:

\[ Y(t) = \xi - \int_t^T Z(t) \, dW(t), \quad t \in [0, T]; \quad \xi \in L^\infty(\Omega, \mathcal{F}_T; \mathbb{R}), \]

by the martingale representation theorem, admits a unique $L^2$-solution pair $(Y, Z)$ with $Y(t, \omega) = \mathbb{E}_{\mathbb{F}}[\xi](\omega)$, which does not require the continuity in $\omega$ of the solution $(Y, Z)(t, \omega)$ or the given terminal value $\xi \in L^\infty(\Omega, \mathcal{F}_T; \mathbb{R})$. This observation motivates our considerations of measurable randomness and path-dependence and in fact gives a nontrivial meaning to the proposed SPHJB equations and associated solution theory with different methods.

In this paper, we propose the SPHJB equation (1.5) for the optimal stochastic control problem (1.1). Both classical solutions and viscosity solutions are discussed. The value function $V$ is verified to be a viscosity solution. A uniqueness result about viscosity solutions is addressed for the superparabolic cases with state-dependent $\sigma$, and as a byproduct, the uniqueness of classical solution is derived for general cases. In addition, an Itô-Kunita-Wentzell-Krylov formula is proved for the compositions of random fields and stochastic differential equations in a path-dependent setting.

Due to the mixture of path-dependence and measurable randomness, a viscosity solution theory for SPHJB equation (1.5) is never a trivial task. On the one hand, due to the path-dependence, the solution $u(\omega, t, \cdot)$, for each $(\omega, t) \in \Omega \times [0, T]$, is path-wisely defined on the path space $C([0, T]; \mathbb{R}^d)$, and we have to deal with the lack of local compactness of the path space; instead of using the nonlinear expectation techniques via second order BSDEs for deterministic PPDEs (see [10, 11] for instance), we define the random test functions by taking extreme points in certain locally compact spaces (actually Hölder spaces) via conventional optimal stopping times. On the other hand, as the involved coefficients are just measurable w.r.t. $\omega$ on the sample space $(\Omega, \mathcal{F})$ without any specified topology, it is not appropriate to define the viscosity solutions in a pointwise manner w.r.t. $\omega \in (\Omega, \mathcal{F})$; instead, we use a class of random fields of form (1.6) having sufficient spatial regularity as test functions; at each point $(\tau, \xi)$ ($\tau$ may be stopping time and $\xi$ a $C([0, \tau]; \mathbb{R}^d)$-valued $\mathcal{F}_\tau$-measurable variable) the classes of test functions are also parameterized by the measurable sets $\Omega_\tau \in \mathcal{F}_\tau$ and the type of compact subspaces.

Finally, we compare the present work with the accompanying one [24]. In fact, when $\sigma(\omega, t, x, \theta(t))(\equiv \sigma(\omega, t))$ is path-invariant and uncontrolled in (1.2), we may take $\overline{X}(t) = X(t) - \xi(t)$ with $\xi(t) = \int_0^t \sigma(s) \, dW(s)$ for $t \in [0, T]$, and then the optimization (1.1)-(1.2) is equivalent to the following one:

\[
\min_{\theta \in \mathcal{U}} \mathbb{E} \left[ \int_0^T f(s, (\overline{X} + \xi)_{s}, \theta(s)) \, ds + G((\overline{X} + \xi)_T) \right], \tag{1.8}
\]

subject to

\[
\begin{align*}
\frac{d\overline{X}(t)}{dt} &= \beta(t, (\overline{X} + \xi)_t, \theta(t)), \quad t \geq 0; \\
\overline{X}_0 &= x_0 \in \mathbb{R}^d.
\end{align*} \tag{1.9}
\]

The paper [24] is devoted to the control problem (1.8)-(1.9) and the existence and uniqueness of viscosity solution is addressed for the associated stochastic path-dependent Hamilton-Jacobi equation which, we note, is first-order. In contrast, our SPHJB (1.5) is second-order, and this leads to the different methods and contents for the viscosity solution theory. For instance, to deal with the lacking of local compactness of the path space, subspaces of Lipschitz functions are used for treating viscosity solutions in [24], while we use subspaces of Hölder functions herein because of the controlled stochastic integrals in the state process (1.2). Two Lipschitz functions over two successive
time intervals with a joint point and an identical Lipschitz constant may be pieced together as a new Lipschitz function with the same Lipschitz constant, which, however, does not hold for Hölder functions. This together with the lacking of boundedness estimates of the second-order terms $\nabla^2 u$ and $\partial_x \nabla u$ for SPHJB ([L5]) gives rise to, particularly, the different discussions on the uniqueness of viscosity solutions.

The rest of this paper is organized as follows. In Section 2, we introduce some notations, show the standing assumption on the coefficients, and define both the viscosity (semi)solutions and classical (semi)solutions. In Section 3, a generalized Itô-Kunita-Wentzell-Krylov formula is proved in a path-dependent setting and then it is applied to semisolutions. Section 4 is devoted to the existence of the viscosity solution, while in Section 5, we discuss the uniqueness.

2 Preliminaries and definition of solutions

2.1 Preliminaries

For each $(k, r) \in \mathbb{N}^+ \times [0, T]$, denote by $\Lambda^0_r(\mathbb{R}^k) := C([0, r]; \mathbb{R}^k)$ the space of all $\mathbb{R}^k$-valued continuous functions on $[0, r]$ and by $\Lambda_r(\mathbb{R}^k) := D([0, r]; \mathbb{R}^k)$ the space of $\mathbb{R}^k$-valued càdlàg (right continuous with left limits) functions on $[0, r]$. Set

$$\Lambda^0(\mathbb{R}^k) = \bigcup_{r \in [0, T]} \Lambda^0_r(\mathbb{R}^k), \quad \Lambda(\mathbb{R}^k) = \bigcup_{r \in [0, T]} \Lambda_r(\mathbb{R}^k).$$

For each path $X \in \Lambda_T(\mathbb{R}^k)$ and $t \in [0, T]$, let $X_t = (X(s))_{0 \leq s \leq t}$ be its restriction to time interval $[0, t]$, and $X(t)$ its value at time $t$. When $k = d$, we write $\Lambda^0, \Lambda^0_r, \Lambda$, and $\Lambda_r$ for simplicity.

Both $\Lambda$ and $\Lambda^0$ are endowed with the following quasi-norm and metric: for each $(x_r, \bar{x}_t) \in \Lambda_r \times \Lambda_t$ or $(x_r, \bar{x}_t) \in \Lambda^0_r \times \Lambda^0_t$ with $0 \leq r \leq t \leq T$,

$$\|x_r\|_0 = \sup_{s \in [0, r]} |x_r(s)|;$$

$$d_0(x_r, \bar{x}_t) = \sqrt{|t - r|} + \sup_{s \in [0, t]} \left\{ |x_r(s) - \bar{x}_t(s)|1_{[0, r)}(s) + |x_r(r) - \bar{x}_t(s)|1_{[r, t]}(s) \right\}.$$}

Then both $(\Lambda^0_r, \| \cdot \|_0)$ and $(\Lambda_r, \| \cdot \|_0)$ are Banach spaces for each $t \in [0, T]$, while $(\Lambda^0, d_0)$ and $(\Lambda, d_0)$ are complete metric spaces. In fact, for each $t \in [0, T]$, $(\Lambda^0_r, \| \cdot \|_0)$ and $(\Lambda_r, \| \cdot \|_0)$ can be and (throughout this paper) will be thought of as the complete subspaces of $(\Lambda^0_T, \| \cdot \|_0)$ and $(\Lambda_T, \| \cdot \|_0)$, respectively; indeed, for each $x_r \in \Lambda_t$ ($x_r \in \Lambda^0_t$, respectively), we define, correspondingly, $\bar{x} \in \Lambda_T$ ($\bar{x} \in \Lambda^0_T$, respectively) with $\bar{x}(s) = x_r(t \wedge s)$ for $s \in [0, T]$. In addition, we shall use $\mathcal{B}(\Lambda^0_T)$, $\mathcal{B}(\Lambda_T)$, $\mathcal{B}(\Lambda^0_T)$ and $\mathcal{B}(\Lambda_T)$ to denote the corresponding Borel $\sigma$-algebras. By contrast, for each $\delta > 0$ and $x_r \in \Lambda$, denote by $B_\delta(x_r)$ the set of paths $y_t \in \Lambda$ satisfying $d_0(x_r, y_t) \leq \delta$.

For each $(x_t, h) \in \Lambda_t \times \mathbb{R}^d$, its vertical perturbation is given as $x^h_t \in \Lambda_t$ with $x^h_t(s) = x_t(s)1_{[0, t)}(s) + (x_t(t) + h)1_{[t, T]}(s)$ for $s \in [0, t]$.

**Definition 2.1.** Given a functional $\phi: \Lambda \to \mathbb{R}$ and a path $x_t \in \Lambda_t$, we say that $\phi$ is differentiable at $x_t$ if the function $\phi(x_t): \mathbb{R}^d \to \mathbb{R}, \ h \mapsto \phi(x^h_t)$ is differentiable at $0$. The gradient

$$\nabla \phi(x_t) := (\nabla_1 \phi(x_t), \ldots, \nabla_d \phi(x_t))' \quad \text{with} \quad \nabla_i \phi(x_t) := \lim_{\delta \to 0} \frac{\phi(x_t + \delta e_i) - \phi(x_t)}{\delta}$$

is called the vertical derivative of $\phi$ at $x_t$, where $\{e_i\}_{i=1, \ldots, d}$ is the canonical basis in $\mathbb{R}^d$.

Let $(\mathbb{B}, \| \cdot \|_\mathbb{B})$ be a Banach space. If the $\mathbb{B}$-valued functional $\phi$ is continuous and bounded at all $x_t \in \Lambda$, $\phi$ is said to be continuous on $\Lambda$ and denoted by $\phi \in C(\Lambda; \mathbb{B})$. Similarly, we define $C(\Lambda^0_0; \mathbb{B})$, $C([0, T] \times \Lambda; \mathbb{B})$, and $C([0, T] \times \Lambda^0_0; \mathbb{B})$. In particular, we define $C((a, b] \times \Lambda; \mathbb{B}) = \cap_{\delta \in (0, b - a)} C([a + \delta, b] \times \Lambda; \mathbb{B})$ as usual for $0 \leq a < b$, and so is it for $C((a, b) \times \Lambda; \mathbb{B})$. 

4
For each $t \in [0, T]$, let $L^0(\Omega \times \Lambda_t, \mathcal{F}_t \otimes \mathcal{B}(\Lambda_t); \mathbb{B})$ be the space of $\mathbb{B}$-valued $\mathcal{F}_t \otimes \mathcal{B}(\Lambda_t)$-measurable random variables. The measurable function

$$u : (\Omega \times [0, T] \times \Lambda, \mathcal{F} \otimes \mathcal{B}([0, T]) \otimes \mathcal{B}(\Lambda)) \to (\mathbb{B}, \mathcal{B}(\mathbb{B})), $$

is said to be adapted if for any time $t \in [0, T]$, $u$ is $\mathcal{F}_t \otimes \mathcal{B}(\Lambda_t)$-measurable. For $p \in [1, \infty)$, denote by $S^p(\Lambda; \mathbb{B})$ the set of all the adapted functions $u : \Omega \times [0, T] \times \Lambda \to \mathbb{B}$ such that for almost all $\omega \in \Omega$, $u$ is valued in $C([0, T] \times \Lambda; \mathbb{B})$ and

$$\|u\|_{S^p(\Lambda; \mathbb{B})} := \left( \sup_{(t, x) \in [0, T] \times \Lambda} \|u(t, x)\|_B \right)^{1/p} < \infty.$$ 

For $p \in [1, \infty)$, denote by $L^p(\Lambda; \mathbb{B})$ the set of all the adapted functions $\psi : \Omega \times [0, T] \times \Lambda \to \mathbb{B}$ such that for almost all $(\omega, t) \in \Omega \times [0, T]$, $\psi(t)$ is valued in $C(\Lambda_t; \mathbb{B})$, and

$$\|\psi\|_{L^p(\Lambda; \mathbb{B})} := \left( E \left[ \int_0^T \sup_{x \in \Lambda_t} \|\psi(t, x)\|_B \, dt \right] \right)^{1/p} < \infty.$$ 

Both $(S^p(\Lambda; \mathbb{B}), \| \cdot \|_{S^p(\Lambda; \mathbb{B})})$ and $(L^p(\Lambda; \mathbb{B}), \| \cdot \|_{L^p(\Lambda; \mathbb{B})})$ are Banach spaces. Analogously, we define $L^\infty(\Omega \times \Lambda_0, \mathcal{F}_t \otimes \mathcal{B}(\Lambda_t); \mathbb{B})$, $(S^p(\Lambda_0; \mathbb{B}), \| \cdot \|_{S^p(\Lambda_0; \mathbb{B})})$, and $(L^p(\Lambda_0; \mathbb{B}), \| \cdot \|_{L^p(\Lambda_0; \mathbb{B})})$.

As usual, we use $C$ with or without a subscript to denote a constant whose value may vary from line to line. Throughout this paper, we use the following assumption.

(A1) $G \in L^\infty(\Omega, \mathcal{F}_T; C(\Lambda_T; \mathbb{R}))$. For the coefficients $g = f, \beta^i, \sigma^{ij}, i = 1, \ldots, d, j = 1, \ldots, m$, (i) for each $v \in U$, $g(\cdot, v)$ is adapted; (ii) for almost all $(\omega, t) \in \Omega \times [0, T]$, $g(t, \cdot, \cdot)$ is continuous on $\Lambda_t \times U$; (iii) there exists $L > 0$ such that for all $x, \bar{x} \in \Lambda_T$, $t \in [0, T]$ and $\gamma_t, \bar{\gamma}_t \in \Lambda_t$, there hold

$$\text{esssup}_{\omega \in \Omega} |G(x)| + \text{esssup}_{\omega \in \Omega} \sup_{v \in U} |g(t, \gamma_t, v)| \leq L,$$

$$\text{esssup}_{\omega \in \Omega} |G(x) - G(\bar{x})| + \text{esssup}_{\omega \in \Omega} \sup_{v \in U} |g(t, \gamma_t, v) - g(t, \bar{\gamma}_t, v)| \leq L (\|x - \bar{x}\|_0 + \|\gamma_t - \bar{\gamma}_t\|_0).$$

### 2.2 Definition of the solutions

For $\delta \geq 0$ and $x_t \in \Lambda_t$, the horizontal extension $x_{t,\delta} \in \Lambda_{t+\delta}$ is given as $x_{t,\delta}(s) = x_t(s \wedge t)$ for $s \in [0, t + \delta]$.

**Definition 2.2.** For $u \in S^2(\Lambda; \mathbb{R})$ with $\nabla u \in L^2(\Lambda; \mathbb{R})$, we say $u \in \mathcal{C}^2_{\mathbb{F}}$ if there exist a constant $\alpha \in (0, 1)$ and a finite partition $0 = \underline{L}_0 < \underline{L}_1 < \ldots < \underline{L}_n = T$, for integer $n \geq 1$, such that

(i) on each subinterval $[\underline{L}_j, \underline{L}_{j+1})$, $j = 0, \ldots, n - 1$,

(a) there exists $(\partial_t u, \partial_u u)$ with

$$(\partial_t u, \partial_u u)_{|([\underline{L}_j, \underline{L}_{j+1} - \varepsilon])} \in L^2(\Lambda; \mathbb{R}^d) \times L^2(\Lambda; \mathbb{R}^m), \quad \forall \varepsilon \in (0, \underline{L}_{j+1} - \underline{L}_j),$$

satisfying for all $\underline{L}_j \leq r \leq t < \underline{L}_{j+1}$, and all $x_r \in \Lambda_r$,

$$u(t, x_r, t - r) = u(r, x_r) + \int_r^t \partial_s u(s, x_r, s - r) \, ds + \int_r^t \partial_u u(s, x_r, s - r) \, dW(s), \text{ a.s.};$$

(b) $\nabla u$ is a.s. valued in $C((\underline{L}_j, \underline{L}_{j+1}) \times \Lambda; \mathbb{R}^d)$, and there exists some adapted $C((\underline{L}_j, \underline{L}_{j+1}) \times \Lambda; \mathbb{R}^{m \times d})$-valued function denoted by $\partial_u \nabla u$ such that for each subinterval $[\underline{L}_j, \underline{L}_{j+1}] \subset$
operators in [5, Section 5.2] and [17, Theorem 4.3]. In particular, if the random function \( \partial \) is just the Malliavin derivative; if \( u(t,x) \) is a deterministic function on the time-state space \( \Omega \times [0,T] \times \mathbb{R}^d \), we may have \( \partial \omega u \equiv 0 \) and \( \partial \omega u \) coincides with the classical derivative in time. Noteworthily, the operators \( \partial_t \) and \( \partial \omega \) are different from the path derivatives \( (\partial_t, \partial \omega) \) via the functional Ito formulas (see [3] and [11] Section 2.3); if \( u(t,x,t) \) is smooth enough w.r.t. \( (\omega, t) \) in the path space, for each \( x \in \Lambda^0 \), we have the relation

\[
(\partial_t u, \partial \omega u)(t, x_s) = \lim_{s \to t^-} (\partial_t u, \partial \omega u)(t, x_{s,t-s}) = (\partial_t u, \partial \omega u)(t, \lim_{s \to t^-} x_{s,t-s}), \quad \forall x_t \in \Lambda_t.
\]

This makes sense of the two linear operators \( \partial_t \) and \( \partial \omega \), which are consistent with the differential operators in [3] Section 5.2 and [11] Theorem 4.3. In particular, if the random function \( u \) on \( \Omega \times [0,T] \times \mathbb{R}^d \) is regular enough (w.r.t. \( \omega \)), the term \( \partial \omega u \) is just the Malliavin derivative; if \( u(t,x) \) is a deterministic function on the time-state space \( [0,T] \times \mathbb{R}^d \), we may have \( \partial \omega u \equiv 0 \) and \( \partial \omega u \) coincides with the classical derivative in time. Noteworthily, the operators \( \partial_t \) and \( \partial \omega \) are different from the path derivatives \( (\partial_t, \partial \omega) \) via the functional Ito formulas (see [3] and [11] Section 2.3); if \( u(t,x,t) \) is smooth enough w.r.t. \( (\omega, t) \) in the path space, for each \( x \in \Lambda^0 \), we have the relation

\[
\partial_t u(\omega, t, x_{s,t-s}) = \left( \partial_t + \frac{1}{2} \partial^2 \omega \right) u(\omega, t, x_{s,t-s}), \quad \partial \omega u(\omega, t, x_{s,t-s}) = \partial \omega u(\omega, t, x_{s,t-s}),
\]

for \( 0 \leq s \leq t < T \) which may be seen from [11] Section 6 and [3].

By (2.1), we define \( \partial \omega \partial u \) in a way different from \( \partial \omega u \). Indeed, if there is \( \partial_t \partial u, \partial \omega \partial u \) with

\[
\{ \partial_t \partial u, \partial \omega \partial u \}^{(t, \omega, t+ \varepsilon)} \in L^2(\Lambda; \mathbb{R}^d) \times L^2(\Lambda; \mathbb{R}^d), \quad \forall \varepsilon \in (0, \theta_{j+1} - \theta_j),
\]
satisfying for all \( \theta_j \leq r \leq \tau \leq \theta_{j+1} \), and \( x_r \in \Lambda_r \),

\[
\partial \omega \partial u(\tau, x_{r,t-r}) = \partial \omega u(\tau, x_{r,t-r}) + \int_r^\tau \partial \omega \partial u(\tau, x_{r,s-r}) ds + \int_r^\tau (\partial \omega \partial u)'(\tau, x_{r,s-r}) dW(s), \quad \text{a.s.}
\]

one may easily check that \( \partial \omega \partial u = \partial \omega \partial u \) which unveils the consistency. We do not adopt the method via (2.2) to define \( \partial \omega \partial u \), basically because it requires the existence of \( \partial \omega \partial u \) that is not necessary but narrows the test function space \( \mathcal{E}_1 \).
Remark 2.1. It is worth noting that the operators $\partial_s$ and $\nabla$ are not exchangeable in many cases. For instance, taking $d = m = 1$, $u(t, x_t) = \int_0^t \sin(x_t(s)) \, dW(s)$, one has $\partial_s u(t, x_t) = \sin(x_t(t^-))$, $\nabla u(t, x_t) = 0$; however, $\nabla_\omega u(t, x_t) = \cos(x_t(t^-))$, while $\partial_\omega \nabla u(t, x_t) = 0$.

For each stopping time $t \leq T$, let $\mathcal{T}^t$ be the set of stopping times $\tau$ valued in $[t, T]$ and $\mathcal{T}^t_+$ the subset of $\mathcal{T}^t$ such that $\tau > t$ for each $\tau \in \mathcal{T}^t_+$. For each $\tau \in \mathcal{T}^0$ and $\Omega_\tau \in \mathcal{F}_\tau$, we denote by $L^0(\Omega_\tau, \mathcal{F}_\tau; \Lambda^0_\tau)$ the set of $\Lambda^0_\tau$-valued $\mathcal{F}_\tau$-measurable functions.

Given a Banach space $(\mathbb{B}, \|\cdot\|_\mathbb{B})$, for each $\alpha \in (0, 1)$ and $0 \leq t_0 < t_1 \leq T$, denote by $C^\alpha([t_0, t_1]; \mathbb{B})$ the usual $\alpha$-Hölder space of $\mathbb{B}$-valued functions equipped with the semi-norm and norm:

$$\|\gamma\|_{C^\alpha([t_0, t_1]; \mathbb{B})} = \sup_{0 \leq s < t \leq t_1} \frac{|\gamma(s) - \gamma(t)|_\mathbb{B}}{|s - t|^{1/\alpha}}, \quad \text{for } \gamma \in C^\alpha([t_0, t_1]; \mathbb{B}),$$

$$\|\gamma\|_{C^\alpha([t_0, t_1])} = \max_{t \in [t_0, t_1]} \|\gamma(t)\|_\mathbb{B} + \|\gamma\|_{C^\alpha([t_0, t_1]; \mathbb{B})}, \quad \text{for } \gamma \in C^\alpha([t_0, t_1]; \mathbb{B}).$$

For each $k \in \mathbb{N}^+$, $\alpha \in (0, 1)$, $0 \leq t \leq s \leq T$ and $\xi \in \Lambda_t$, define

$$\Lambda_{t,s}^{0,k,\alpha; \xi} = \left\{ x \in \Lambda_s : x(\tau) = 1_{[0,t]}(\tau) \xi(\tau \wedge t) + 1_{(t,s]}(\tau) g(\tau), \ \tau \in [0, s], \right.$$  

for some $g \in C^\alpha([t, s]; \mathbb{R}^d)$, with $g(t) = \xi(t)$, $\|g\|_{C^\alpha([t, s]; \mathbb{R}^d)} \leq k$,

and furthermore, we set $\Lambda_{0,t}^{0,k,\alpha; \xi} = \cup_{\xi \in \mathbb{R}^d} \Lambda_{0,t}^{0,k,\alpha; \xi}$ for each $t \in [0, T]$. Then Arzelà-Ascoli theorem indicates that each $\Lambda_{t,s}^{0,k,\alpha; \xi}$ is compact in $\Lambda_s$. Moreover, it is obvious that $\cup_{k \in \mathbb{N}^+} \Lambda_{0,t}^{0,k,\alpha; \xi}$ is dense in $\Lambda_0$ for any $\alpha \in (0, 1)$. In addition, by saying $(s, x) \rightarrow (t^+, \xi)$ for some $(t, \xi) \in [0, T] \times \Lambda_t$ we mean $(s, x) \rightarrow (t^+, \xi)$ with $s \in (t, T]$ and $x \in \cup_{k \in \mathbb{N}^+} \Lambda_{s}^{0,k,\alpha; \xi}$ for some $\alpha \in (0, 1)$.

We expect the test function space $\mathcal{C}_\mathbb{B}^2_{\mathcal{F}}$ to include the classical solutions. Nevertheless, it is typical that the classical solutions may not be differentiable in the time variable $t$ and $(\partial_t u, \partial_\omega u)$ may not be time-continuous but just measurable in $t$; see [8] [27] for the state-dependent cases, or one may even refer to the BSDEs that may be thought of as the trivial stochastic path-independent PDEs.

Definition 2.3. We say that $u \in \mathcal{C}_\mathbb{B}^2_{\mathcal{F}}$ is a classical supersolution (resp. subsolution) of SPHJB equation (1.3), if $u(T, x) \geq$ (resp. $\leq$) $G(x)$ for all $x \in \Lambda_T$ a.s. and for each $t \in (0, T)$ with $y \in \Lambda_t$,

$$\text{ess} \lim_{(s, x) \rightarrow (t^+, y)} \inf \mathcal{E}_{\mathcal{F}_t} \{ -\partial_s u(s, x) - \mathbb{H}(s, x, \nabla u(s, x), \nabla^2 u(s, x), \partial_\omega \nabla u(s, x)) \} \geq 0, \ a.s., \quad (2.3)$$

(resp. $\text{ess} \lim_{(s, x) \rightarrow (t^+, y)} \inf \mathcal{E}_{\mathcal{F}_t} \{ -\partial_s u(s, x) - \mathbb{H}(s, x, \nabla u(s, x), \nabla^2 u(s, x), \partial_\omega \nabla u(s, x)) \} \leq 0, \ a.s.) \quad (2.4)$$

The function $u$ is a classical solution of SPHJB equation (1.3) if it is both a classical subsolution and a classical supersolution.

Throughout this paper, we denote by $\mathcal{F}_t$ the set of all the classical supersolutions of SPHJB equation (1.3) and by $\mathcal{V}_t$ the set of all the classical subsolutions. Set $\overline{\phi}(t, x) = Le^{L(T-t)}$, and $\phi(t, x) = -Le^{-L(T-t)}$. Straightforward computations indicate that $\overline{\phi} \in \mathcal{F}_t$ and $\phi \in \mathcal{V}_t$ under Assumption (A1). Therefore, we have the following assertion.

Lemma 2.1. Let Assumption (A1) hold. Neither $\mathcal{F}_t$ nor $\mathcal{V}_t$ is empty.

We now introduce the notion of viscosity solutions. For each $(u, \tau) \in S^2(\Lambda^0_0; \mathbb{R}) \times \mathcal{T}_0$, $\Omega_\tau \in \mathcal{F}_\tau$, with $\mathbb{P}(\Omega_\tau) > 0$ and $\xi \in L^0(\Omega_\tau, \mathcal{F}_\tau; \Lambda^0_\tau)$, we define for each $(k, \alpha) \in \mathbb{N}^+ \times (0, 1)$,

$$\mathcal{V}u(\tau, \xi; \Omega_\tau, k, \alpha) := \left\{ \phi \in \mathcal{C}_\mathbb{B}^2_{\mathcal{F}} : \text{there exists } \tilde{\tau}_k \in \mathcal{T}^\tau_+ \text{ such that} \right.$$  

$\text{each } \xi \in L^0(\Omega_\tau, \mathcal{F}_\tau; \Lambda^0_\tau)$ is thought of as $\xi \in L^0(\Omega_\tau, \mathcal{F}_\tau; \Lambda^0_\tau)$ satisfying $\xi(\omega) \in \Lambda_\tau(\omega)$ for almost all $\omega \in \Omega_\tau$. 

$\text{7}$
where and all

In view of the assumption (1.5), if \( u(\tau, x, \Omega; t, x) \) and a viscosity supersolution.

\[ \mathcal{G}u(\tau, \xi; \Omega, k, \alpha) := \left\{ \phi \in \mathcal{C}^2_{x} : \text{there exists } \hat{\tau}_k \in \mathcal{T}_T \text{ such that} \right. \]

\[ (\phi - u)(\tau, \xi)1_{\Omega_k} = 0 = \text{essinf}_{\tau \in \mathcal{T}_T} E_{\mathcal{G}} \left[ \inf_{y \in \Lambda^{k, \alpha}, \zeta} (\phi - u)(\tau \wedge \hat{\tau}_k, y) \right] 1_{\Omega_k}, \text{ a.s.} \}

\[ \mathcal{G}u(\tau, \xi; \Omega, k, \alpha) := \left\{ \phi \in \mathcal{C}^2_{x} : \text{there exists } \hat{\tau}_k \in \mathcal{T}_T \text{ such that} \right. \]

\[ (\phi - u)(\tau, \xi)1_{\Omega_k} = 0 = \text{esssup}_{\tau \in \mathcal{T}_T} E_{\mathcal{G}} \left[ \sup_{y \in \Lambda^{k, \alpha}, \zeta} (\phi - u)(\tau \wedge \hat{\tau}_k, y) \right] 1_{\Omega_k}, \text{ a.s.} \}

Obviously, if \( \mathcal{G}u(\tau, \xi; \Omega, k, \alpha) \) or \( \mathcal{G}u(\tau, \xi; \Omega, k, \alpha) \) is nonempty, there holds \( 0 \leq \tau < T \) on \( \Omega_T \).

**Definition 2.4.** We say that \( u \in \mathcal{S}^2(\Lambda^{0}; \mathbb{R}) \) is a viscosity subsolution (resp. supersolution) of SPHJB equation (15), if \( u(T, x) \leq (\text{ resp. } \geq) G(x) \) for all \( x \in \Lambda^{0} \) a.s., and for each \((K_0, a_0) \in \mathbb{N}^{+} \times (0, 1)\), there exists \((k, \alpha) \in \mathbb{N}^{+} \times (0, 1)\) with \( k \geq K_0 \) and \( \alpha \leq a_0 \), such that for any \( \tau \in \mathcal{T}_T \), \( \Omega_T \in \mathcal{F}_T \) with \( \mathbb{P}(\Omega_T) > 0 \) and \( \xi \in L^{0}(\Omega_T, \mathcal{F}_T; \Lambda^{0}) \) and any \( \phi \in \mathcal{G}u(\tau, \xi; \Omega_T, k, \alpha) \) (resp. \( \phi \in \mathcal{G}u(\tau, \xi; \Omega_T, k, \alpha) \)) holds

\[
\text{ess lim inf}_{(s, x) \to (\tau, \xi)} E_{\mathcal{G}} \{ -\partial_s \phi(s, x) - \mathbb{H}(s, x, \nabla \phi(s, x), \nabla^2 \phi(s, x), \partial_x \nabla \phi(s, x)) \} \leq 0,
\]

for almost all \( \omega \in \Omega_T \)

\[
(\text{resp. } \text{ess lim sup}_{(s, x) \to (\tau, \xi)} E_{\mathcal{G}} \{ -\partial_s \phi(s, x) - \mathbb{H}(s, x, \nabla \phi(s, x), \nabla^2 \phi(s, x), \partial_x \nabla \phi(s, x)) \} \geq 0,
\]

for almost all \( \omega \in \Omega_T \).

The function \( u \) is a viscosity solution of SPHJB equation (15) if it is both a viscosity subsolution and a viscosity supersolution.

To make sense of the involved vertical derivatives, a classical (semi)solution is defined on path space \( \Lambda \), while the viscosity solution is just defined on \( \Lambda^{0} \). Throughout this paper, we define for each \( \phi \in \mathcal{C}^2_{x}, v \in U, t \in [0, T], \) and \( x_t \in \Lambda_t, \)

\[
\mathcal{L}^\phi(t, x_t) = \partial_t \phi(t, x_t) + \beta'(t, x_t, v) \nabla \phi(t, x_t)
\]

\[
+ \text{tr} \left\{ \frac{1}{2} \sigma(t, x_t, v) \sigma'(t, x_t, v) \nabla^2 \phi(t, x_t) + \sigma(t, x_t, v) \partial_x \nabla \phi(t, x_t) \right\}.
\]

**Remark 2.2.** In view of the assumption (A1), for each \( \phi \in \mathcal{C}^2_{x}, \) there exists a finite partition \( 0 = t_0 < t_1 < \ldots < t_n = T, \) such that for any \( 0 < \delta < \max_{0 \leq j \leq n-1} |t_{j+1} - t_j|, \) there exist an \( \mathcal{F}_t \)-adapted process \( \zeta^0 \) and a constant \( L_0^\phi \in (0, \infty) \) satisfying that a.s. for all \( t \in [0, \tau_0 \wedge \zeta^0], \)

\[
\left| -\partial_t \phi(t, x_t) - \mathbb{H}(t, x_t, \nabla \phi(t, x_t), \nabla^2 \phi(t, x_t), \partial_x \nabla \phi(t, x_t)) \right| \leq \sup_{\omega \in U} \left| \mathcal{L}^\phi(t, x_t) + f(t, x_t, v) \right| \leq \zeta^0,
\]

and

\[
\left| \left\{ -\partial_t \phi(t, x_t) - \mathbb{H}(\nabla \phi, \nabla^2 \phi, \partial_x \nabla \phi) \right\} (t, x_t) - \left\{ -\partial_t \phi - \mathbb{H}(\nabla \phi, \nabla^2 \phi, \partial_x \nabla \phi) \right\} (t, \bar{x_t}) \right|
\]

\[
\leq \sup_{\omega \in U} \left| (\mathcal{L}^\phi(t, x_t) + f(t, x_t, v)) - (\mathcal{L}^\phi(t, \bar{x_t}) + f(t, \bar{x_t}, v)) \right|
\]

\[
\leq L_0^\phi \left( \|x_t - \bar{x}_t\|_0^\alpha + \|x_t - \bar{x}_t\|_0 \right), \tag{2.7}
\]

where \( \zeta^0 \in L^2(\Omega \times [t_j, t_{j+1} - \delta]) \) for \( j = 0, \ldots, n - 1, \) and \( \alpha \) is the exponent associated to \( \phi \in \mathcal{C}^2_{x}. \)

Therefore, the conditional expectations in (2.5) and (2.6) are well-defined a.e..
3 Generalized Itô-Kunita-Wentzell-Krylov formula and its applications to semisolutions

First, under assumption (A1), the following assertions may be obtained via standard computations; refer to [7, 19, 29] for instance.

**Lemma 3.1.** Let (A1) hold. Given \( \theta \in \mathcal{U} \), for the strong solution of SDE (1.2), for all \( p > 0 \), there exists \( K > 0 \) such that, for all \( 0 \leq r \leq t \leq s \leq T \), and \( \xi \in L^0(\Omega, \mathcal{F}_r; \Lambda_r) \),

(i) the two processes \( \left( X^r_{s} \right)_{t \leq s \leq T} \) and \( \left( X^r_{t} \right)_{t \leq s \leq T} \) are indistinguishable;

(ii) \( E_{\theta_r} \left[ \max_{t \leq s \leq T} \left| X^r_{s} \right|^p \right] \leq K (1 + \left| \xi \right|^p) \) a.s.,\(^3\)

(iii) \( E_{\theta_r} \left[ d_0(X^r_{s}; X^r_{t}) \right]^p \leq K (|s - t|^p + |s - t|^{p/2}) \) a.s.;

(iv) given another \( \hat{\xi} \in L^0(\Omega, \mathcal{F}_r; \Lambda_r) \), \( E_{\theta_r} \left[ \max_{t \leq s \leq T} \left| X^r_{s} - X^r_{t} \right|^p \right] \leq K \left| \xi - \hat{\xi} \right|^p \), a.s.;

(v) the constant \( K \) depends only on \( L, p, \) and \( T \).

To investigate the Hölder continuity of the paths, we recall a general version of Kolmogorov criterion by Revuz and Yor [26, Theorem (2.1), Page 26–28].

**Lemma 3.2.** Given a Banach space \((\mathbb{B}, \| \cdot \|_\mathbb{B})\), let \((Y(t))_{t \in [0,T]}\) be a \( \mathbb{B} \)-valued stochastic process for which there are three strictly positive constants \( q, \lambda, \) and \( \delta \) such that

\[
E \left[ \| Y(t) - Y(s) \|_\mathbb{B}^q \right] \leq \lambda |t - s|^{1+\delta}, \quad \text{for all } 0 \leq t \leq s \leq T.
\]

Then, for each \( \alpha \in \left(0, \frac{\delta}{q}\right)\) the process \( Y \) admits an \( \alpha \)-Hölder continuous modification (denoted by itself) such that

\[
E \left[ \| Y \|_{0,T;\alpha}^q \right] \leq C,
\]

where the constant \( C \) depends on \( \lambda, \alpha, q, \delta, \) and \( T \).

**Remark 3.1.** The controlled SDE in Lemma 3.1 may be considered in any finite interval \([0, N]\) for \( N > 0 \) and the time \( T \) may also be general \( T > 0 \). By assertion (iii) of Lemma 3.1 the arbitrariness of \( p \) therein and Lemma 3.2 imply that for all \( \alpha \in (0, \frac{1}{2}), \tau > 0, \) and \( q > 1 \), there exists constant \( C > 0 \) such that for all \( \theta \in \mathcal{U}, \xi \in L^0(\Omega, \mathcal{F}_r; \Lambda_r) \), and \( r > 0 \), there holds \( E_{\theta_r} \left[ \left| X^r_{s} - X^r_{t} \right|^p \right] \leq K \left| \xi - \hat{\xi} \right|^p \), a.s., with \( C \) depending only on \( L, \tau, q, \) and \( \alpha \). Further, for each \( \theta \in \mathcal{U} \), recalling that for each \( 0 < \alpha < \alpha' < \frac{1}{2} \), and \( 0 \leq r < t \leq T \), there holds

\[
\left| X^r_{s} \right|_{r,t;\alpha} \leq \left| X^r_{s} \right|_{r,T;\alpha'} |t - r|^{\alpha' - \alpha}, \quad \text{a.s.},
\]

we have the stopping times

\[
\tau^\theta_{k,\alpha} := \inf \{ s > r; \left| X^r_{s} \right|_{r,s;\alpha} > k \} \wedge T, \quad \text{for } k > 0,
\]

well-defined, with \( P(r < \tau^\theta_{k,\alpha}) = 1 \) and \( \tau^\theta_{k,\alpha} \) increasingly converging to \( T \) as \( k \to \infty \).

---

\(^3\)Here, denoting by \( x_r \) a path in \( \Lambda_r \), we set \( E_{\theta_r} \left[ \max_{t \leq s \leq T} \left| X^r_{s} \right|^p \right] = E_{\theta_r} \left[ \max_{t \leq s \leq T} \left| X^r_{s} \right|^p \right]_{x_r}\); the conditional expectation in assertion (iv) is defined analogously.
We then generalize an Itô-Kunita-Wentzell-Krylov formula (see [16, Pages 118-119] for instance) for the composition of random fields and stochastic differential equations to our path-dependent setting. Recall that for each \( \phi \in \mathcal{C}^2_{\mathbb{P}}, v \in U, t \in [0, T], \) and \( x_t \in \Lambda_t, \)

\[
\mathcal{L}^v \phi(t, x_t) = \partial_v \phi(t, x_t) + \beta'(t, x_t, v) \nabla \phi(t, x_t) + \text{tr} \left\{ \frac{1}{2} \sigma(t, x_t, v) \sigma'(t, x_t, v) \nabla^2 \phi(t, x_t) + \sigma(t, x_t, v) \partial_\omega \nabla \phi(t, x_t) \right\}.
\]

**Lemma 3.3.** Let assumption (A1) hold. Suppose \( u \in \mathcal{C}^2_{\mathbb{P}} \) with the associated partition \( 0 = t_0 < t_1 < \ldots < t_n = T. \) Then, for each \( \theta \in U, \) it holds almost surely that, for each \( t_0 < t_1 < \cdots < t_N < T, \) \( x_0 \in \Lambda_0, \) it holds that

\[
u(t, X^0_{\theta, x}, \theta) = u(t_0, x_0) + \int_{t_0}^T \mathcal{L}^\theta(s) u(s, X^0_{\theta, x}, \theta) \, ds
+ \int_{t_0}^T \left( (\nabla u)'(r, X^0_{\theta, x}, \theta) \sigma(r, X^0_{\theta, x}, \theta, \theta(r)) + \partial_\omega u(r, X^0_{\theta, x}, \theta) \right) dW(r), \quad \text{a.s.} \quad (3.2)
\]

**Proof.** W.l.o.g., we only prove (3.2) for \( \tau \in (0, \frac{t}{N}), \) \( \rho = 0 \) and \( x_0 = x \in \mathbb{R}^d. \) For each \( N \in \mathbb{N}^+ \) with \( N > 2, \) letting \( t_i = \frac{r}{N} \) for \( i = 0, 1, \ldots, N, \) we get a partition of \([0, \tau]\) with \( 0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = \tau. \) For each \( \theta \in U, \) set

\[
NX(t) = \sum_{i=0}^{N-1} X^0_{x, \theta}(t_i) 1_{[t_i, t_{i+1})}(t) + X^0_{x, \theta}(\tau) 1_{(\tau]}(t), \quad \text{for } t \in [0, \tau],
\]

and

\[
NX_t- (s) = NX(s) 1_{[0, t)}(s) + \lim_{r \to t-} NX(r) 1_{(t]}(s), \quad \text{for } 0 \leq s \leq t \leq \tau. \]

Due to the time-continuity of \( X^0_{x, \theta}, \) there holds the following approximation:

\[
\lim_{N \to \infty} \|X^0_{x, \theta} - NX\|_0 + \|X^0_{x, \theta} - NX_t-\|_0 = 0 \quad \text{for all } t \in (0, \tau], \quad \text{a.s..}
\]

Then, we have

\[
u(t, X^N) - u(0, x) = \sum_{i=0}^{N-1} u(t_{i+1}, X^N t_{i+1-}) - u(t_i, X^N t_i) + \sum_{i=0}^{N-1} u(t_{i+1}, X^N t_{i+1-}) - u(t_{i+1}, X^N t_{i+1-}) = I_1^{(N)} + I_2^{(N)}.
\]

As \( u \in \mathcal{C}^2_{\mathbb{P}}, \) it holds that

\[
I_1^{(N)} = \sum_{i=0}^{N-1} u(t_{i+1}, X^N t_{i+1-}) - u(t_i, X^N t_i) = \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \partial_v u(s, X^N s-) \, ds + \int_{t_N}^{t_{N-1}} \partial_v u(r, X^N r-) \, dW(r),
\]

which, as \( N \) tends to infinity, converges in probability to

\[
\int_0^T \partial_v u(r, X^0_{r, x}, \theta) \, dr + \int_0^T \partial_\omega u(r, X^0_{r, x}, \theta) \, dW(r).
\]

On the other hand, by the definition of vertical derivatives, it holds that

\[
I_2^{(N)} = \sum_{i=0}^{N-1} (\nabla u(t_{i+1}, X^N t_{i+1-}))' (NX(t_{i+1}) - NX(t_{i+1-}))
+ \frac{1}{2} \sum_{i=0}^{N-1} (NX(t_{i+1}) - NX(t_{i+1-}))' \nabla^2 u(t_{i+1}, X^N t_{i+1-}) (NX(t_{i+1}) - NX(t_{i+1-}))
\]
\[ M_1^{(N)} = (\nabla u(t_1, X_{t_1}))' (N X(t_1) - x) + \sum_{i=1}^{N-1} (\nabla u(t_i, X_{t_i}))' (N X(t_{i+1}) - N X(t_{i+1}-)) \]
\[ + \sum_{i=1}^{N-1} (\nabla u(t_{i+1}, X_{t_{i+1}-}) - \nabla u(t_i, X_{t_i})) (N X(t_{i+1}) - N X(t_{i+1}-)), \]

for some \( h \) satisfying \( |h - N X(t_{i+1} -)| \leq |N X(t_{i+1}) - N X(t_{i+1} -)| \). Further, we have

\[ M_1^{(N)} = (\nabla u(t_1, X_{t_1}))' (N X(t_1) - x) + \sum_{i=1}^{N-1} (\nabla u(t_i, X_{t_i}))' (N X(t_{i+1}) - N X(t_{i+1}-)) \]
\[ + \sum_{i=1}^{N-1} (\nabla u(t_{i+1}, X_{t_{i+1}-}) - \nabla u(t_i, X_{t_i})) (N X(t_{i+1}) - N X(t_{i+1}-)). \]

Notice that

\[ \sum_{i=1}^{N-1} (\nabla u(t_{i+1}, X_{t_{i+1}-}) - \nabla u(t_i, X_{t_i})) (N X(t_{i+1}) - N X(t_{i+1}-)) \]
\[ = \sum_{i=1}^{N-1} (\nabla u(t_{i+1}, X_{t_{i+1}-}) - \nabla u(t_i, X_{t_i})) \int_{t_i}^{t_{i+1}} \beta(r, X_r^{0, x, \theta}, \theta(r)) dr \]
\[ + \sum_{i=1}^{N-1} (\nabla u(t_{i+1}, X_{t_{i+1}-}) - \nabla u(t_i, X_{t_i})) \int_{t_i}^{t_{i+1}} \sigma(r, X_r^{0, x, \theta}, \theta(r)) dW(r) \]
\[ := P_1^{(N)} + P_2^{(N)}, \]

where \( P_1^{(N)} \) converges to zero in probability due to the boundedness of \( \beta \) and the continuity of \( \nabla u \) and in view of (i)-(b) for \( \partial_\omega \nabla u \) in Definition 2.2, we have \( P_2^{(N)} \) converge in probability to \( \int_0^\tau \text{tr} \left\{ \sigma(t, X_t^{0, x, \theta}, \theta(t)) \partial_\omega \nabla u(t, X_t^{0, x, \theta}) \right\} dt \). This combined with some standard computations yields the convergence (in probability) of \( M_1^{(N)} \) with the limit being

\[ \int_0^\tau (\nabla u(t, X_t^{0, x, \theta}))' dX_t^{0, x, \theta}(t) + \int_0^\tau \sum_{i=1}^d \sum_{l=1}^m \sigma^{il}(t, X_t^{0, x, \theta}, \theta(t)) \partial_\omega \nabla u(t, X_t^{0, x, \theta}) dt \]
\[ = \int_0^\tau \left\{ \beta(t, X_t^{0, x, \theta}, \theta(t)) \nabla u(t, X_t^{0, x, \theta}) + \text{tr} \left\{ \sigma(t, X_t^{0, x, \theta}, \theta(t)) \partial_\omega \nabla u(t, X_t^{0, x, \theta}) \right\} \right\} dt \]
\[ + \int_0^\tau (\nabla u(t, X_t^{0, x, \theta}))' \sigma(r, X_r^{0, x, \theta}, \theta(r)) dW(r). \]

Meanwhile, straightforward standard calculations give the convergence of \( M_2^{(N)} \) to

\[ \frac{1}{2} \int_0^\tau \sum_{i,j=1}^d \nabla^2 u(t, X_t^{0, x, \theta}) d \left\{ \nabla^2 X_t^{0, x, \theta} \right\} dt \]
\[ = \int_0^\tau \frac{1}{2} \text{tr} \left\{ \sigma(t, X_t^{0, x, \theta}, \theta(t)) \sigma'(t, X_t^{0, x, \theta}, \theta(t)) \nabla^2 u(t, X_t^{0, x, \theta}) \right\} dt. \]

In the course of approaching the limits (3.4), (3.5), and (3.6), the dominated convergence and the dominated convergence theorem for stochastic integrals ([22, Chapter IV, Theorem 32]) imply that the Lebesgue integrals converge almost surely and the stochastic integrals in probability. Finally, summing up all the obtained convergences yields the desired equality.

Now, we discuss some properties of classical/viscosity semisolusions.

**Theorem 3.4.** Let Assumption (A1) hold. Each classical subsolution (resp. supersolution) is a viscosity subsolution (resp. supersolution), and thus, each classical solution is a viscosity solution.
Proof. Step 1. We first prove that each $\overline{\mu} \in \overline{\mathcal{P}}$ is a viscosity supersolution. Indeed, for each $\phi \in \mathcal{C}^2_b(\Omega, \tau, k, \alpha)$ with $k > 0$ and $\alpha \in (0, \frac{1}{2})$, $\tau \in \mathcal{T}^0$, $\Omega_r \in \mathcal{F}_r$, $\mathbb{P}(\Omega_r) > 0$, and $\xi_r \in L^0(\Omega, \mathcal{F}_r; \Lambda^0_r)$, the relation (2.1) holds for almost all $\omega \in \Omega_r$. Suppose that, to the contrary, there exist $\varepsilon, \delta > 0$ and $\Omega' \in \mathcal{F}_r$ with $\Omega' \subset \Omega_r$, $\mathbb{P}(\Omega') > 0$, such that a.e. on $\Omega'$,

$$\text{esssup}_{s \in (\tau, (\tau + 4\delta^2) \wedge T), x \in B_{2 \delta}(\xi_r) \cap \Lambda^0_{\tau, \delta} \cap \Omega} \mathbb{E}_{\xi_r} \left\{ -\partial_s \phi(s, x) - \mathbb{H}(s, x, \nabla \phi(s, x), \nabla^2 \phi(s, x), \partial_\omega \nabla \phi(s, x)) \right\} \leq -\varepsilon.$$ 

Let $\tilde{\tau}_k$ be the stopping time associated to the fact $\phi \in \mathcal{C}^2_b(\Omega, \tau, k)$, We may think of $\xi$ valued in $\Lambda^0_{\tau}$, with $\xi(t) = \xi_r(t \wedge \tau)$ for all $t \in [0, T]$. Notice that, associated to $\phi \in \mathcal{C}^2_b$ and $\overline{\mu} \in \mathcal{C}^2_b$, the two partitions may be combined into one: $0 = \tau_0 < \tau_1 < \ldots < \tau_n = T$. W.l.o.g., we assume $\delta \in (0, 1)$, and $\Omega' = \{ [\tau, \tau + 4\delta^2] \subset [\tau_j, \tau_{j+1}] \} = \Omega$ for some $j \in \{ 0, 1, \ldots, n - 1 \}$.

For each $\theta \in \mathcal{U}$, define $\tau^\theta = \inf\{ s > \tau : X^{\tau, \xi_r, \theta}_s \not\in B_{2 \delta}(\xi_r) \} \wedge T$, and set

$$\tau^\theta_{k, \alpha} = \inf\{ s > \tau; \| X^{\tau, \xi_r, \theta}_s \|_{\tau, s; \alpha} > k \} \wedge T.$$ 

Then $\mathbb{P}(\tau^\theta > \tau, \tau^\theta_{k, \alpha} > \tau) = 1$. Letting $\tilde{\tau} = \tau \wedge \tau^\theta \wedge \tau^\theta_{k, \alpha} \wedge (\tau + \delta^2) \wedge T$, we have $\mathbb{P}(\tilde{\tau} > \tau) = 1$. Then, for each $\theta \in \mathcal{U}$,

$$\text{essinf}_{s \in (\tau, \tau^\theta)} \mathbb{E}_{\tau} \left[ \mathcal{L}^{\theta(s)}(s, X^{\tau, \xi_r, \theta}_s) + f(s, X^{\tau, \xi_r, \theta}_s, \theta(s)) \right] \geq \varepsilon. \quad (3.7)$$

On the other hand, as $\overline{\mu} \in \overline{\mathcal{P}}$, it holds that for all $t \in [0, T]$ with $y \in \Lambda^0_{\tau}$,

$$\text{ess} \lim_{(s, x) \to (t^\theta, y)} \mathbb{E}_{t^\theta} \left\{ -\partial_s \overline{\mu}(s, x) - \mathbb{H}(s, x, \nabla \overline{\mu}(s, x), \nabla^2 \overline{\mu}(s, x), \partial_\omega \nabla \overline{\mu}(s, x)) \right\} \geq 0, \text{ a.s.},$$

which implies that there exists $\delta \in (0, \delta)$ such that

$$\text{essinf}_{s \in (\tau, (\tau + 4\delta^2) \wedge T)} \mathbb{E}_{\tau} \left\{ -\partial_s \overline{\mu}(s, \xi_s) - \mathbb{H}(s, \xi_s, \nabla \overline{\mu}(s, \xi_s), \nabla^2 \overline{\mu}(s, \xi_s), \partial_\omega \nabla \overline{\mu}(s, \xi_s)) \right\} \geq -\varepsilon \frac{1}{4}.$$ 

Take $\tilde{\tau} = \tau \wedge (\tau + \delta^2) \wedge T$. By the measurable selection theorem, there exists $\tilde{\theta} \in \mathcal{U}$ such that

$$\text{esssup}_{s \in (\tau, \tilde{\tau})} \mathbb{E}_{\tau} \left\{ \mathcal{L}^{\tilde{\theta}(s)}(s, \xi_s) + f(s, \xi_s, \tilde{\theta}(s)) \right\} \leq \varepsilon \frac{1}{2},$$

which together with Remark 2.2 indicates that

$$\int_\tau^{\tilde{\tau}} \mathbb{E}_{\tau} \left[ \mathcal{L}^{\tilde{\theta}(s)}(s, X^{\tau, \xi_r, \tilde{\theta}}_s) + f(s, X^{\tau, \xi_r, \tilde{\theta}}_s, \tilde{\theta}(s)) \right] ds \leq \int_\tau^{\tilde{\tau}} \mathbb{E}_{\tau} \left[ \varepsilon \frac{1}{2} + \mathcal{L}^{\tilde{\theta}(s)}(s, X^{\tau, \xi_r, \tilde{\theta}}_s) \right] ds \quad \text{a.s.,} \quad (3.8)$$

where $\tilde{\alpha}$ is the exponent associated to $\overline{\mu} \in \mathcal{C}^2_b$. Combining (3.7) and (3.8) gives

$$\int_\tau^{\tilde{\tau}} \mathbb{E}_{\tau} \left[ \mathcal{L}^{\tilde{\theta}(s)}(\phi - \overline{\mu})(s, X^{\tau, \xi_r, \tilde{\theta}}_s) \right] ds \geq \int_\tau^{\tilde{\tau}} \mathbb{E}_{\tau} \left[ \varepsilon \frac{1}{2} - \mathcal{L}^{\tilde{\theta}(s)}(s, X^{\tau, \xi_r, \tilde{\theta}}_s) \right] ds, \quad \text{a.s.}$$

Applying the Itô-Kunita-Wentzell-Krylov formula in Lemma 3.3 further yields that

$$E \left[ (\phi - \overline{\mu})(\tilde{\tau}, X^{\tau, \xi_r, \tilde{\theta}}_\tilde{\tau}) \right] = E \left[ (\phi - \overline{\mu})(\tau, \xi_\tau) + \int_\tau^{\tilde{\tau}} \mathcal{L}^{\tilde{\theta}(s)}(\phi - \overline{\mu})(s, X^{\tau, \xi_r, \tilde{\theta}}_s) ds \right]$$

$$= E \left[ \int_\tau^{\tilde{\tau}} \mathcal{L}^{\tilde{\theta}(s)}(\phi - \overline{\mu})(s, X^{\tau, \xi_r, \tilde{\theta}}_s) ds \right]$$

$$\geq E \left[ \int_\tau^{\tilde{\tau}} \left( \varepsilon \frac{1}{2} - \mathcal{L}^{\tilde{\theta}(s)}(s, X^{\tau, \xi_r, \tilde{\theta}}_s) \right) ds \right].$$
which is $> 0$ when $\tilde{\delta}$ is sufficiently small, contradicting with $\phi \in \overline{\mathcal{G}P}(\tau, \xi; \Omega, k, \alpha)$.

**Step 2.** To prove that each $\underline{\mu} \in \mathcal{L}$ is a viscosity subsolution, it is sufficient to verify that for each $\phi \in \overline{\mathcal{G}P}(\tau, \xi; \Omega, k, \alpha)$ with $k > 0$ and $\alpha \in (0, \frac{1}{2})$, $\tau \in \mathcal{T}^0$, $\Omega_\tau \in \mathcal{F}_\tau$, $\mathcal{P}(\Omega_\tau) > 0$, and $\xi_\tau \in L^0(\Omega_\tau, \mathcal{F}_\tau; \Lambda_0^0)$, the relation (2.3) holds for almost all $\omega \in \Omega_\tau$.

To the contrary, suppose that there exist $\varepsilon, \tilde{\delta} \in (0, 1)$ and $\Omega' \in \mathcal{F}_\tau$ with $\Omega' \subset \Omega_\tau$, $\mathcal{P}(\Omega') > 0$, such that a.e. on $\Omega'$,

$$
E_{\mathcal{G}_\tau} \left\{ -\partial_s \phi(s, x) - \mathbb{H}(s, x, \nabla \phi(s, x), \nabla^2 \phi(s, x), \partial_\omega \nabla \phi(s, x)) \right\} \geq 2\varepsilon.
$$

Let $\bar{\tau}_k$ be the stopping time associated to $\phi \in \overline{\mathcal{G}P}(\tau, \xi; \Omega, k, \alpha)$. Again, we think of $\xi$ as a path in $\Lambda_0^0$, with $\xi(t) = \xi_\tau(t \wedge \tau)$ for all $t \in [0, T]$, and the two partitions, associated to $\phi \in \mathcal{C}_2$ and $\underline{\mu} \in \mathcal{C}_2$, are combined into one: $0 = L_0 < L_1 < \ldots < L_n = T$. W.l.o.g., we assume $\Omega' = \{[\tau, \tau + 4\bar{\delta}^2] \subset [L_j, L_{j+1}) \} = \Omega$ for some $j \in \{0, 1, \ldots, n - 1\}$.

By the measurable selection theorem, there exists $\tilde{\theta} \in \mathcal{U}$ such that a.s.,

$$
-\mathcal{L}^{\tilde{\theta}(s)}(s, \xi, \tilde{\theta}(s)) - f(s, \xi, \tilde{\theta}(s)) \geq -\partial_s \phi(s, \xi, \tilde{\theta}(s)) - \mathbb{H}(s, \xi, \nabla \phi(s, \xi, \tilde{\theta}(s), \nabla^2 \phi(s, \xi, \tilde{\theta}(s), \partial_\omega \nabla \phi(s, \xi, \tilde{\theta}(s))) - \varepsilon,
$$

for almost all $s$ satisfying $\tau < s < (\tau + 4\bar{\delta}^2) \wedge T$. Thus, we have

$$
\text{essinf}_{\tau < s < (\tau + 4\bar{\delta}^2) \wedge T} E_{\mathcal{G}_\tau} \left\{ -\mathcal{L}^{\tilde{\theta}(s)}(s, \xi, \tilde{\theta}(s)) - f(s, \xi, \tilde{\theta}(s)) \right\} \geq \varepsilon, \quad \text{a.e.} \quad (3.9)
$$

Set $\tau^\delta = \inf\{s > \tau : X_s^{\tau, \xi, \tilde{\theta}} \notin B_\delta(\xi_\tau) \} \wedge T$, and

$$
\tau^\delta_{k, \alpha} = \inf\{s > \tau : \|X_s^{\tau, \xi, \tilde{\theta}}\|_{\tau, s, \alpha} > k \} \wedge T.
$$

Then $\mathcal{P}(\tau^\delta > \tau, \tau^\delta_{k, \alpha} > \tau) = 1$. Letting $\bar{\tau} = \bar{\tau}_k \wedge \tau^\delta \wedge \tau^\delta_{k, \alpha} \wedge (\tau + \delta^2) \wedge T$, we have $\bar{\tau} > \tau$ a.s.

Combining (3.9) and the analysis in Remark 2.2 yields that for all $\tau' \in \mathcal{T}_\tau$,

$$
\int_{\tau'}^{\tau^\delta \wedge \bar{\tau}} E_{\mathcal{G}_\tau} \left[ \mathcal{L}^{\tilde{\theta}(s)}(s, X_s^{\tau, \xi, \tilde{\theta}}) + f(s, X_s^{\tau, \xi, \tilde{\theta}}, \tilde{\theta}(s)) \right] ds \leq \int_{\tau'}^{\tau^\delta \wedge \bar{\tau}} E_{\mathcal{G}_\tau} \left[ L_{\alpha'}^\theta |\delta^{\alpha'} + \delta| - \varepsilon \right] ds, \quad (3.10)
$$

where $\alpha' \in (0, 1)$ is the exponent associated to $\phi \in \mathcal{C}_2$.

On the other hand, as $\underline{\mu} \in \mathcal{L}$, there exists $\delta \in (0, \tilde{\delta})$ such that

$$
\text{esssup}_{s \in (\tau, \tau + 4\bar{\delta}^2) \wedge T} E_{\mathcal{G}_\tau} \left\{ -\partial_s \underline{\mu}(s, \xi, \tilde{\theta}(s)) - \mathbb{H}(s, \xi, \nabla \underline{\mu}(s, \xi, \tilde{\theta}(s), \nabla^2 \underline{\mu}(s, \xi, \tilde{\theta}(s), \partial_\omega \nabla \underline{\mu}(s, \xi, \tilde{\theta}(s))) \right\} \leq 0,
$$

which particularly implies that

$$
\text{essinf}_{s \in (\tau, \tau + 4\bar{\delta}^2) \wedge T} E_{\mathcal{G}_\tau} \left\{ \mathcal{L}^{\tilde{\theta}(s)}(s, \xi, \tilde{\theta}(s)) \right\} \geq 0,
$$

Take $\bar{\tau} = \bar{\tau} \wedge (\tau + \delta^2) \wedge T$. It obviously holds that $\bar{\tau} > \tau$ a.s. Moreover, Remark 2.2 implies that

$$
\int_{\tau}^{\bar{\tau}} E_{\mathcal{G}_\tau} \left[ \mathcal{L}^{\tilde{\theta}(s)}(s, X_s^{\tau, \xi, \tilde{\theta}}) + f(s, X_s^{\tau, \xi, \tilde{\theta}}, \tilde{\theta}(s)) \right] ds \geq \int_{\tau}^{\bar{\tau}} E_{\mathcal{G}_\tau} \left[ -L_{\alpha'}^\theta |\delta^{\alpha'} + \delta| \right] ds \quad \text{a.s.}, \quad (3.11)
$$

where $\bar{\alpha}$ is the exponent associated to $\underline{\mu} \in \mathcal{C}_2$.

Combining (3.10) and (3.11) and applying Lemma 3.3 yield that

$$
E \left[ (\phi - \underline{\mu})(\bar{\tau}, X_{\bar{\tau}}^{\tau, \xi, \tilde{\theta}}) \right] = E \left[ (\phi - \underline{\mu})(\tau, \xi) + \int_{\tau}^{\tau} \mathcal{L}^{\tilde{\theta}(s)}(s, X_s^{\tau, \xi, \tilde{\theta}}) ds \right]
$$

$$
= E \left[ \int_{\tau}^{\bar{\tau}} \mathcal{L}^{\tilde{\theta}(s)}(s, X_s^{\tau, \xi, \tilde{\theta}}) ds \right]
$$

$$
\leq E \left[ \int_{\tau}^{\bar{\tau}} \left( L_{\alpha'}^\theta |\delta^{\alpha'} + \delta| + L_{\alpha'}^\phi |\delta^{\alpha'} + \delta| - \varepsilon \right) ds \right],
$$

which is $< 0$ when $\tilde{\delta}$ is sufficiently small, resulting in a contradiction with $\phi \in \overline{\mathcal{G}P}(\tau, \xi; \Omega, k, \alpha)$. \qed
4 Existence of the viscosity solution

The following properties of the value function \( V \) hold in a similar way to [23, Proposition 3.3].

**Proposition 4.1.** Let (A1) hold.

(i) For each \( t \in [0, T] \), \( \varepsilon \in (0, \infty) \), and \( \xi \in L^0(\Omega; \mathcal{F}_t; \Lambda_t) \), there exists \( \bar{\theta} \in \mathcal{U} \) such that

\[
E \left[ J(t, \xi; \bar{\theta}) - V(t, \xi) \right] < \varepsilon.
\]

(ii) For each \( (\theta, x_0) \in \mathcal{U} \times \mathbb{R}^d \), \( \left\{ J(t, X^0_{t_0}; \theta) - V(t, X^0_{t_0}; \theta) \right\}_{t \in [0, T]} \) is a supermartingale, i.e., for any \( 0 \leq t \leq \hat{t} \leq T \),

\[
V(t, X^0_{t_0}; \theta) \leq E_{\mathcal{F}_t} \left[ \hat{t}, X^0_{\hat{t}}; \theta \right] + E_{\mathcal{F}_t} \int_t^{\hat{t}} f(s, X^0_{s}; \theta(s)) \, ds, \quad \text{a.s.}. \tag{4.1}
\]

(iii) For each \( (\theta, x_0) \in \mathcal{U} \times \mathbb{R}^d \), \( \left\{ V(s, X^0_{s}; \theta) \right\}_{s \in [0, T]} \) is a continuous process.

(iv) There exists \( L_V > 0 \) such that for each \( (\theta, t) \in \mathcal{U} \times [0, T] \),

\[
|V(t, x_t) - V(t, y_t)| + |J(t, x_t; \theta) - J(t, y_t; \theta)| \leq L_V \|x_t - y_t\|_0, \quad \text{a.s.,} \quad \forall x_t, y_t \in \Lambda_t,
\]

with \( L_V \) depending only on \( T \) and \( L \).

(v) With probability 1, \( V(t, x) \) and \( J(t, x; \theta) \) for each \( \theta \in \mathcal{U} \) are continuous on \( [0, T] \times \Lambda \) and

\[
\sup_{(t, x) \in [0, T] \times \Lambda} \max \{|V(t, x_t)|, |J(t, x_t; \theta)|\} \leq L(T + 1) \quad \text{a.s.}
\]

Following is the dynamic programming principle, whose proof is the same to [23, Theorem 3.4], utilizing the separability of path spaces.

**Theorem 4.2.** Let assumption (A1) hold. For any stopping times \( \tau, \hat{\tau} \) with \( \tau \leq \hat{\tau} \leq T \), and any \( \xi \in L^0(\Omega; \mathcal{F}_\tau; \Lambda_\tau) \), we have

\[
V(\tau, \xi) = \text{essinf}_{\theta \in \mathcal{U}} E_{\mathcal{F}_\tau} \left[ \int_{\tau}^{\hat{\tau}} f \left( s, X^\tau_{s}; \theta(s) \right) \, ds + V \left( \hat{\tau}, X^\tau_{\hat{\tau}}; \theta \right) \right] \quad \text{a.s.}
\]

Then we are ready to give the existence of the viscosity solution.

**Theorem 4.3.** Let (A1) hold. The value function \( V \) defined by (1.4) is a viscosity solution of the SPHJB equation (1.3).

**Proof.** First, we have \( V \in \mathcal{S}^\infty(\Lambda; \mathbb{R}) \) by Proposition 4.1. The proof is divided into two steps.

**Step 1.** To the contrary, suppose that for each \( (k, \alpha) \in \mathbb{N}^+ \times (0, \frac{1}{2}) \) with \( k \geq K_0 \) and \( \alpha \leq \alpha_0 \) for some existing \( (K_0, \alpha_0) \in \mathbb{N}^+ \times (0, 1) \), there exists \( \phi \in \mathcal{L}^0(\mathcal{F}_\tau; \Omega, \mathcal{F}_\tau; \Lambda_\tau) \), \( \tau \in \mathcal{T}_0 \), \( \Omega_\tau \in \mathcal{F}_\tau \), \( \mathbb{P}(\Omega_\tau) > 0 \), and \( \xi_\tau \in L^0(\Omega_\tau; \mathcal{F}_\tau; \Lambda^0_\tau) \), such that there exist \( \varepsilon, \delta \in (0, 1) \), and \( \Omega' \in \mathcal{F}_\tau \) with \( \Omega' \subset \Omega_\tau \), \( \mathbb{P}(\Omega') > 0 \), satisfying a.e. on \( \Omega' \),

\[
\text{essinf}_{s \in (\tau, (\tau + \delta \tau)^2) \times T \cap B_{\delta \tau}(\xi_{\tau}) \cap \Lambda^0_\tau \times T} E_{\mathcal{F}_s} \left\{ -\Delta s \phi(s, x) - \mathbb{H}(s, x, \nabla \phi(s, x), \nabla^2 \phi(s, x), \delta \omega \nabla \phi(s, x)) \right\} \geq 2 \varepsilon. \tag{4.2}
\]

Denote by \( \hat{\tau}_j \) the stopping time associated to \( \phi \in \mathcal{L}^0(\mathcal{F}_\tau; \Omega, \mathcal{F}_\tau; \Lambda_\tau) \). Note that we may think of \( \xi \) valued in \( \Lambda^0_\tau \) with \( \xi(t) = \xi_\tau (t \wedge \tau) \) for all \( t \in [0, T] \). Moreover, associated to \( \phi \in \mathcal{L}^2_\mathcal{F} \), there is a partition: \( 0 = \frac{\tau_j}{n} < \frac{\tau_1}{n} < \cdots < \frac{\tau_n}{n} = T \). W.l.o.g., we assume \( \Omega' = \{ [\tau, \tau + 4\delta \tau^2] \subset \frac{\tau_j}{n}, \frac{\tau_{j+1}}{n} \} = \Omega \) for some \( j \in \{0, \ldots, n - 1 \} \).
By assumption (ii) of (A1) and the measurable selection theorem, there exists $\tilde{\theta} \in \mathcal{U}$ such that a.s.,
\[-\mathcal{L}^{\tilde{\theta}}(s) \phi(s, \xi_s) - f(s, \xi_s, \tilde{\theta}(s)) \geq -\partial_s \phi(s, \xi_s) - \nabla \phi(s, \xi_s, \nabla^2 \phi(s, \xi_s), \partial_{xx} \nabla \phi(s, \xi_s)) - \varepsilon,
\]
for almost all $s$ satisfying $\tau \leq s < (\tau + 4\tilde{\delta}^2) \land T$. This together with (4.2) implies
\[
\essinf_{\tau < s < (\tau + 4\tilde{\delta}^2) \land T} E_{\tilde{\theta}, \tau} \{-\mathcal{L}^{\tilde{\theta}}(s) \phi(s, \xi_s) - f(s, \xi_s, \tilde{\theta}(s))\} \geq \varepsilon, \quad \text{a.s.}
\]
(4.3)

Define $\tau^{\theta}_{k,\alpha} = \inf\{s > \tau; \|X^{\tau,\xi,\theta}\|_{\tau,s,\alpha} > k\} \land T$, and for each $\delta \in (0, \tilde{\delta})$, set $\bar{\tau}_{k,\alpha} = \inf\{s > \tau; X^{\tau,\xi,\theta} \notin B_{\delta}(\xi_s)\} \land T$.

Then $\mathbb{P}(\bar{\tau}_{k,\alpha} > \tau) = 1$. Putting $\bar{\tau} = \bar{\tau}_{k} \land \tau^{\theta}_{k,\alpha} \land (\tau + \tilde{\delta}^2) \land T$, we have $\mathbb{P}(\bar{\tau} > \tau) = 1$. Combining (4.3) and the analysis in Remark 2.2 yields that
\[
\int_{\tau}^{\bar{\tau}} E_{\tilde{\theta}, \tau} \left[ \mathcal{L}^{\tilde{\theta}}(s) \phi(s, X^{\tau,\xi,\theta}_{s}) + f(s, X^{\tau,\xi,\theta}_{s}, \tilde{\theta}(s)) \right] ds \leq \int_{\tau}^{\bar{\tau}} E_{\tilde{\theta}, \tau} \left[ L^{\phi}_{\alpha', \delta} \alpha' + \delta \right] ds,
\]
where $\alpha' \in (0, 1)$ is the exponent associated to $\phi \in C^{2}_{\mathcal{F}}$.

Choose a small $\delta \in (0, \tilde{\delta})$ such that $2L^{\phi}_{\alpha'} \delta^\alpha + \delta < \varepsilon$. Using the dynamic programming principle of Theorem 1.2 and the Itô-Kunita-Wentzell-Krylov formula of Lemma 3.3, we have
\[
0 \leq E_{\tilde{\theta}, \tau} \left[ (\phi - V)(\tau, X^{\tau,\xi,\theta}_{\tau}) - (\phi - V)(\tau, \xi_{\tau}) \right]
\leq E_{\tilde{\theta}, \tau} \left[ \phi(\tau, X^{\tau,\xi,\theta}_{\tau}) - \phi(\tau, \xi_{\tau}) + \int_{\tau}^{\bar{\tau}} f(s, X^{\tau,\xi,\theta}_{s}, \tilde{\theta}(s)) ds \right]
\leq E_{\tilde{\theta}, \tau} \left[ \int_{\tau}^{\bar{\tau}} \left( L^{\phi}_{\alpha'} \alpha' + \delta \right) ds \right]
\leq E_{\tilde{\theta}, \tau} \left[ \int_{\tau}^{\bar{\tau}} \left( L^{\phi}_{\alpha'} \alpha' + \delta \right) ds \right]
< -\frac{\varepsilon}{2} E_{\tilde{\theta}, \tau} [\bar{\tau} - \tau],
\]
(4.4)

which gives rise to a contradiction. Hence, $V$ is a viscosity subsolution of SPHJB equation (1.5).

**Step 2.** We prove that $V$ is a viscosity supersolution of (1.5). To the contrary, assume that for each $(k, \alpha) \in \mathbb{N}^+ \times (0, \frac{1}{3})$ with $k \geq K_0$ and $\alpha \leq \alpha_0$ for some existing $(K_0, \alpha_0) \in \mathbb{N}^+ \times (0, 1)$, there exists $\phi \in C^0(T, \Omega, \mathbb{R}^n, \mathbb{R})$, $\xi(t) \in \Omega_{\tau}$, $\mathbb{P}(\Omega) > 0$, and $\xi_{\tau} \in L^0(\Omega_{\tau}, \mathcal{F}_{\tau}; \mathbb{A}^k_{\mathcal{F}})$ such that there exist $\varepsilon > 0$, $\delta \in (0, 1)$ and $\Omega' \in \mathcal{F}_{\tau}$ with $\Omega' \subset \Omega_{\tau}$, $\mathbb{P}(\Omega') > 0$, satisfying a.e. on $\Omega'$,
\[
\esssup_{s \in (\tau, (\tau + 4\tilde{\delta}^2) \land T), x \in B_{\delta}(\xi_{\tau})} E_{\tilde{\theta}, \tau} \left\{-\partial_s \phi(s, x) - \nabla \phi(s, x) \right\} \leq -\varepsilon.
\]

Denote by $\tilde{\tau}_k$ the stopping time associated to $\phi \in C^0(T, \Omega, \mathbb{R}^n, \mathbb{R})$. Again, we think of $\xi$ valued in $\Lambda_{\tau}$ with $\xi(t) = \xi_{\tau}(t \land \tau)$ for all $t \in [0, T]$, and associated to $\phi \in C^{2}_{\mathcal{F}}$, there is a partition: $0 = t_0 < t_1 < \ldots < t_n = T$. W.l.o.g., we assume $\Omega' = \{\tau, (\tau + 4\tilde{\delta}^2) \subseteq [t_j, t_{j+1})\} = \Omega$ for some $j \in \{0, \ldots, n - 1\}$.

For each $\theta \in \mathcal{U}$, define $\tau^\theta = \inf\{s > \tau; X^{\tau,\xi,\theta}_{s} \notin B_{\delta}(\xi_{\tau})\}$, and set $\tau_{k,\alpha}^\theta = \inf\{s > \tau; \|X^{\tau,\xi,\theta}_{s}\|_{\tau,s,\alpha} > k\}$, and $\bar{\tau}_{k,\alpha}^\theta = \tau^\theta \land \tau_{k,\alpha}^\theta$. 


Then $\mathbb{P}(\tilde{\tau}^\theta > \tau) = 1$. Recalling $\alpha \in (0, \frac{1}{4})$, we have for each $h \in (0, \frac{\delta^2}{4})$,

\[
E_{\tilde{\mathcal{F}}_\tau}\left[1_{\{\tilde{\tau}^\theta < \tau + h\}}\right] \leq E_{\tilde{\mathcal{F}}_\tau}\left[1_{\{\tilde{\tau}^\theta < \tau + h\}} + E_{\tilde{\mathcal{F}}_\tau}\left[1_{\{\tilde{\tau}^\theta,\alpha < \tau + h\}}\right]\right]
= E_{\tilde{\mathcal{F}}_\tau}\left[1_{\{\max_{\tau \leq \tau + h} |X^{\tau,\xi,\theta}(s) - \xi(\tau)| > \sqrt{\delta}\}}\right] + E_{\tilde{\mathcal{F}}_\tau}\left[1_{\{|X^{\tau,\xi,\theta}|_{\tau + h, \alpha} > h\}}\right]
\leq \frac{1}{(\delta - \sqrt{h})^8} E_{\tilde{\mathcal{F}}_\tau}\left[\max_{\tau \leq \tau + h} |X^{\tau,\xi,\theta}(s) - \xi(\tau)|^8 \right] + \frac{1}{k_16} E_{\tilde{\mathcal{F}}_\tau}\left[|X^{\tau,\xi,\theta}|_{\tau, \tau + h, \alpha}^{16}\right]
\leq \frac{256 \cdot K^8}{\delta^8} (h + \sqrt{h})^8 + \frac{h^4}{k_16} E_{\tilde{\mathcal{F}}_\tau}\left[|X^{\tau,\xi,\theta}|_{\tau, \tau + \hat{\delta} + h}^{16}\right]
\leq \hat{C} h^4 \quad \text{a.s.}
\] (4.5)

Here, the constants $K$ and $C$, independent of $(\theta, h, k, \tau)$, are from Lemma 3.1 and Remark 3.1 respectively, and thus, the positive constant $\hat{C}$ is independent from $\theta, h,$ and $\tau$.

In view of Remark 2.2, Lemma 3.3, Theorem 4.2, and the estimate (4.5), we have for each $h \in (0, \frac{\delta^2}{4})$,

\[
0 = \frac{V(\tau, \xi, \tau) - \phi(\tau, \xi, \tau)}{h}
= \frac{1}{h} \text{essinf}_{\theta \in U} E_{\tilde{\mathcal{F}}_\tau}\left[\int_{\tau}^{\tilde{\tau}_k \wedge (\tau + h)} f(s, X_s^{\tau, \xi, \theta}, \theta(s)) \, ds + V\left(\hat{\tau}_k \wedge (\tau + h), X_{\hat{\tau}_k \wedge (\tau + h)}^{\tau, \xi, \theta}\right) - \phi(\tau, \xi, \tau)\right]
\geq \frac{1}{h} \text{essinf}_{\theta \in U} E_{\tilde{\mathcal{F}}_\tau}\left[\int_{\tau}^{\tilde{\tau}_k \wedge (\tau + h)} f(s, X_s^{\tau, \xi, \theta}, \theta(s)) \, ds + \phi\left(\hat{\tau}_k \wedge (\tau + h), X_{\hat{\tau}_k \wedge (\tau + h)}^{\tau, \xi, \theta}\right) - \phi(\tau, \xi, \tau)\right]
\geq \frac{1}{h} \text{essinf}_{\theta \in U} E_{\tilde{\mathcal{F}}_\tau}\left[\int_{\tau}^{\tilde{\tau}_k \wedge (\tau + h)} \left(\mathcal{L}^{\theta}(s) \phi\left(s, X_s^{\tau, \xi, \theta}\right) + f(s, X_s^{\tau, \xi, \theta}, \theta(s))\right) \, ds\right]
\geq \frac{1}{h} \text{essinf}_{\theta \in U} E_{\tilde{\mathcal{F}}_\tau}\left[\int_{\tau}^{\tilde{\tau}_k \wedge (\tau + h)} \left|\mathcal{L}^{\theta}(s) \phi\left(s, X_s^{\tau, \xi, \theta}\right) + f(s, X_s^{\tau, \xi, \theta}, \theta(s))\right| \, ds\right]
\geq \frac{\varepsilon}{h} \text{essinf}_{\theta \in U} E_{\tilde{\mathcal{F}}_\tau}\left[\frac{\left(\hat{\tau}_k \wedge (\tau + h)\right) - \tau}{h}\right]
\rightarrow \varepsilon, \quad \text{as } h \rightarrow 0^+.
\]
5 Uniqueness

Recall that $\overline{\mathcal{V}}$ is the set of all the classical supersolutions of SPHJB equation (1.5) and $\underline{\mathcal{V}}$ the set of all the classical subsolutions, and by Lemma 2.1, both $\overline{\mathcal{V}}$ and $\underline{\mathcal{V}}$ are nonempty. Set

$$\underline{\mathcal{V}} = \text{essinf} \phi, \quad \overline{\mathcal{V}} = \text{esssup} \phi.$$  

Letting $V$ be the value function defined in (1.4), we shall prove the uniqueness results: (i) a classical solution must be $V$ if it exists; (ii) $V$ is the unique viscosity solution approximated by classical supersolutions from above and by classical subsolutions from below.

5.1 Comparison relation $\underline{\mathcal{V}} \geq V \geq \overline{\mathcal{V}}$ and uniqueness of classical solution under Assumption (A1)

**Theorem 5.1.** Letting Assumption (A1) hold, we have $\underline{\mathcal{V}} \geq V \geq \overline{\mathcal{V}},$ i.e., for all $t \in [0, T]$ and $\xi_t \in \Lambda_t,$ there holds $\underline{\mathcal{V}}(t, \xi_t) \geq V(t, \xi_t) \geq \overline{\mathcal{V}}(t, \xi_t)$ a.s.

**Proof.** Step 1. To prove $V \geq \overline{\mathcal{V}},$ we need only verify the relation $\phi \leq V$ for each $\phi \in \underline{\mathcal{V}}.$ Recall that, associated to $\phi \in \mathcal{C}_2^\mathcal{A},$ there is a partition: $0 = t_0 < t_1 < \ldots < t_n = T.$ Also, we have $\phi(T, x) \leq G(x)$ for all $x \in \Lambda_T$ a.s. and for each $s \in [0, T)$ with $y_s \in \Lambda_s,$ there holds

$$\text{ess lim sup}_{(\tau, x) \to (s^+, y_s)} E_{\mathfrak{F}_s} \left\{ -\partial_s \phi(\tau, x) - \mathbb{H}(\tau, x, \nabla \phi(\tau, x), \nabla^2 \phi(\tau, x), \partial_\omega \nabla \phi(\tau, x)) \right\} \leq 0, \text{ a.s.},$$

i.e.,

$$\text{ess lim sup}_{(\tau, x) \to (s^+, y_s)} E_{\mathfrak{F}_s} \left[ \sup_{v \in \mathcal{U}} \{-\mathcal{L}^v \phi(\tau, x) - f(\tau, x, v)\} \right] \leq 0, \text{ a.s.}.$$  

Then, for each $\theta \in \mathcal{U},$ and $\xi_t \in \Lambda_t$ with $t_{n-1} \leq t \leq T,$ we have by Lemma 3.3

$$\phi(t, \xi_t) = E_{\mathfrak{F}_t} \left[ \phi(\tau, X^{t, \xi_t}_\tau; \theta) + \int_t^T \left( -\mathcal{L}^0(s) \phi(s, X^{t, \xi_t}_s; \theta) \right) ds \right]$$

$$\leq E_{\mathfrak{F}_t} \left[ \phi(\tau, X^{t, \xi_t}_\tau; \theta) + \int_t^T f(s, X^{t, \xi_t}_s; \theta(s)) ds \right]$$

$$\to E_{\mathfrak{F}_t} \left[ \phi(T, X_T^{t, \xi_t}; \theta) + \int_t^T f(s, X^{t, \xi_t}_s; \theta(s)) ds \right], \quad \text{as } \tau \to T.$$  

Thus, $\phi(t, \xi_t) \leq J(t, \xi_t; \theta)$ as $\phi(T, x_T) \leq G(x_T) \text{ a.s.}$ for all $x_T \in \Lambda_T$ and this together with the arbitrariness of $\theta$ implies that $\phi(t, \xi_t) \leq V(t, \xi_t) \text{ a.s.}$ for $t \in [l_{n-1}, T).$ Similarly, we may verify $\phi \leq V$ recursively over the time intervals $[l_{n-2}, l_{n-1}), \ldots, [0, l_1).$

**Step 2.** We prove $\overline{\mathcal{V}} \geq V.$ For each $\phi \in \overline{\mathcal{V}},$ recall that $\phi(T, x) \geq G(x)$ for all $x \in \Lambda_T$ a.s. and for each $s \in [0, T)$ with $y_s \in \Lambda_s,$ there holds

$$\text{ess lim inf}_{(\tau, x) \to (s^+, y_s)} E_{\mathfrak{F}_s} \left\{ -\partial_s \phi(\tau, x) - \mathbb{H}(\tau, x, \nabla \phi(\tau, x), \nabla^2 \phi(\tau, x), \partial_\omega \nabla \phi(\tau, x)) \right\} \geq 0, \text{ a.s.},$$

i.e.,

$$\text{ess lim inf}_{(\tau, x) \to (s^+, y_s)} E_{\mathfrak{F}_s} \left[ \sup_{v \in \mathcal{U}} \{-\mathcal{L}^v \phi(\tau, x) - f(\tau, x, v)\} \right] \geq 0, \text{ a.s.} \quad (5.1)$$

Also, associated to $\phi \in \mathcal{C}_2^\mathcal{A},$ there is a partition: $0 = t_0 < t_1 < \ldots < t_n = T.$

As in Step 1, we first prove the comparison on the interval $[l_{n-1}, T).$ For each $t \in [l_{n-1}, T)$ and $\xi_t \in L^0(\Omega, \mathcal{F}_t; \Lambda_t),$ we may extend $\xi_t$ to be valued in $\Lambda_T$ with $\xi(s) = \xi_t(t \wedge s)$ for all $s \in [0, T].$ Take
\[ \delta \in (0, (T - t) \wedge 1). \] Then by (5.1) and the measurable selection theorem, for each \( \varepsilon \in (0, 1) \), there exists \( \bar{\theta} \in \mathcal{U} \) such that for all \( N \in \mathbb{N}^+ \setminus \{1, 2\} \), it holds that

\[
E_{\mathcal{F}_t} \left[ \int_t^{t + \frac{\delta}{N}} \left\{ -L^\theta(s) \phi(s, \xi_s) - f(s, \xi_s, \bar{\theta}(s)) \right\} ds \right] \geq \int_t^{t + \frac{\delta}{N}} E_{\mathcal{F}_t} \left[ \sup_{v \in \mathcal{U}} \left\{ -L^v(s) \phi(s, \xi_s) - f(s, \xi_s, v) \right\} - \varepsilon \right] ds \geq \frac{-\delta \varepsilon}{N}, \text{ a.s.} \tag{5.2}
\]

In view of Remark 2.2 and (iii) of Lemma 3.1, we have

\[
E_{\mathcal{F}_t} \left[ \int_t^{t + \frac{\delta}{N}} \left\{ \left( L^\theta(s) \phi(s, \xi_s) + f(s, \xi_s, \bar{\theta}(s)) \right) - \left( L^\bar{\theta}(s) \phi(s, X_s^{t, \bar{\theta}}) + f(s, X_s^{t, \bar{\theta}}, \bar{\theta}(s)) \right) \right\} ds \right] \leq E_{\mathcal{F}_t} \left[ \int_t^{t + \frac{\delta}{N}} \left| L^\phi \left( \| X_s^{t, \bar{\theta}} - \xi_s \|_0^\alpha + \| X_s^{t, \bar{\theta}} - \xi_s \|_0 \right) \right| ds \right] \leq C_1 \left( \frac{\delta}{N} \right)^{1+\frac{\alpha}{2}}, \tag{5.3}
\]

where \( \alpha \in (0, 1) \) is the exponent associated to \( \phi \in C^2_{\mathcal{F}} \), the constant \( L^\phi_\alpha \), depending on \( \delta \), is associated to the interval \([t-1, t+\delta] \), and \( C_1 \) depends only on \( L^\phi_\alpha \) and \( K \).

Starting with the obtained \( \left( t + \frac{\delta}{N}, X_s^{t, \bar{\theta}} \right) \), we may conduct the same discussions as in (5.2) and (5.3) recursively over the time interval \([t + \frac{\delta}{N}, t + \frac{2\delta}{N}], \ldots, [t + \frac{(N-1)\delta}{N}, t + \delta] \) for \( N - 1 \) steps. Step by step, the controls \( \bar{\theta} \)s may be patched together, and there exists some \( \bar{\theta} \in \mathcal{U} \) such that

\[ \phi(t, \xi_t) = \int_t^{t+\delta} \left( -L^\theta(s) \phi(s, X_s^{t, \bar{\theta}}) \right) ds \]

Choosing a big \( N \) so that \( \varepsilon \cdot N^{\frac{\alpha}{2}} > C_1 \), we have

\[ \phi(t, \xi_t) \geq J(t, \xi_t, \bar{\theta}) + E_{\mathcal{F}_t} \left[ \phi(t + \delta, X_s^{t, \bar{\theta}} - \phi(T, X_T^{t, \bar{\theta}})) \right] - L(T - t - \delta) - \varepsilon \left\{ \delta^{1+\frac{\alpha}{2}} + \delta \right\} \]

The arbitrariness of \((t, \xi_t, \varepsilon)\) further implies that \( V \leq \phi \) on \([L_{n-1}, T] \), and recursively, the comparison may be verifed over the time intervals \([L_{n-2}, L_{n-1}], \ldots, [0, t] \). Finally, we obtain \( V \leq \overline{u} \) over the whole time interval \([0, T] \). \( \square \)

A straightforward application of Theorem 5.1 gives the uniqueness of classical solution.

**Corollary 5.2.** Let Assumption (A1) hold. If \( u \) is a classical solution of SPHJB equation (1.3), then \( u(t, \xi_t) = V(t, \xi_t) \), a.s. for all \( t \in [0, T] \) and \( \xi_t \in \Lambda_t \).
5.2 Uniqueness: \( \pi = V = u \) for superparabolic cases with state-dependent \( \sigma \)

First, we write the Wiener process \( W = (\tilde{W}, \hat{W}) \), where \( \tilde{W} \) and \( \hat{W} \) are two mutually independent and respectively, \( m_0 \) and \( m_1 (= m - m_0) \)-dimensional Wiener processes. In what follows, we adopt the decomposition \( \sigma = (\tilde{\sigma}, \hat{\sigma}) \) with \( \tilde{\sigma} \) and \( \hat{\sigma} \) valued in \( \mathbb{R}^{d \times m_0} \) and \( \mathbb{R}^{d \times m_1} \) respectively associated to \( \tilde{W} \) and \( \hat{W} \). Denote by \( \{\mathcal{F}_t\}_{t \geq 0} \) the natural filtration generated by \( \tilde{W} \) and augmented by all the \( \mathbb{P} \)-null sets.

(Superparabolicity) \[
\sum_{i,j=1}^d \sum_{k=1}^{m_1} \bar{\sigma}^{ik} \bar{\sigma}^{jk}(t, x, v) \xi_i \xi_j \geq \lambda |\xi|^2 \quad \forall (t, x, v, \xi) \in [0, T) \times \mathbb{R}^d \times U \times \mathbb{R}^d.
\]

Moreover, we assume the following continuity properties of the coefficients \( \beta, f, \) and \( G \).

(A3) For each \( \varepsilon > 0 \), there exist partition \( 0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T \) for some \( N > 3 \) and functions \( G^N \in C^3([0, T] \times \mathbb{R}^{d + (m_0 + m_1)}; \mathbb{R}) \), such that \( G^N \) is \( \mathcal{F}_t \)-adapted with

\[
\|G^N\|_{L^2(\Omega, \mathcal{F}_T; \mathbb{R})} + \|f^N\|_{L^2([0, T]; \mathbb{R})} + \|\beta^N\|_{L^2([0, T]; \mathbb{R})} < \varepsilon,
\]

and \( G^N, f^N \) and \( \beta^N \) are uniformly Lipschitz-continuous in the space variable \( x \) with an identical Lipschitz-constant \( L_\varepsilon \) independent of \( N \) and \( \varepsilon \).

Remark 5.1. In Assumption (A3), the coefficients \( \beta, f, \) and \( G \) are approximated via regular functions. Indeed, such approximations may be proved in a similar way to [24] Lemma 4.2 if we assume the uniform time-continuity: for \( g = \beta_i, i = 1, \ldots, d \), there is a continuously increasing function \( \rho : [0, \infty) \to [0, \infty) \) with \( \rho(0) = 0 \) such that for all \( r, s \in [0, T] \) with \( s \leq t \leq r \), and \( x \in \Lambda_T \),

\[
|g(r, (x_{L_N})_r) - g(r, (x_{L_N})_s)| + |g(t, x_t) - g(s, x_s)| + |G((x_{L_N})_r) - G((x_{L_N})_s)| \leq \rho(|t - s|), \text{ a.s.} \quad (5.4)
\]

For example, the relation \( (5.4) \) is obviously satisfied if we take \( g(t, x_t) = \int_0^t \phi(x_s) \, ds \) for \( t \in [0, T] \), and \( G(x_T) = \zeta \cdot \int_0^T \psi(x_T(s)) \, ds \), where \( \zeta \in L^\infty(\Omega, \mathcal{F}_T; \mathbb{R}) \) and \( \phi \) and \( \psi \) are bounded and uniformly continuous functions on \( \mathbb{R}^d \).

Theorem 5.3. Letting (A1) – (A3) hold and \( V \) be the value function in (1.4), we have \( \overline{u} = V = u \), i.e., for each \( t \in [0, T] \) and \( x_t \in \Lambda_T^0 \), there holds \( \overline{u}(t, x_t) = V(t, x_t) = u(t, x_t) \) a.s.
Proof. By Theorem 5.1, we have \( \underline{\Omega} \leq V \leq \overline{\Omega} \). Therefore, it is sufficient to construct functions from \( \overline{\Omega} \) and \( \underline{\Omega} \) to approximate the value function \( V \) from above and from below respectively.

For each \( \varepsilon \in (0, 1) \), we take \( (G^\varepsilon, f^\varepsilon, \beta^\varepsilon) \) and \( (G^N, f^N, \beta^N) \) as in Assumption (A3). By the theory of backward SDEs (see [2] for instance), let the pair \( (Y^\varepsilon, Z^\varepsilon) \) be the unique adapted solution to backward SDE

\[
Y^\varepsilon(s) = G^\varepsilon + \int_s^T \left( f^\varepsilon(t) + \tilde{L}^\varepsilon(t) \right) dt - \int_s^T Z^\varepsilon(t) d\tilde{W}(t),
\]

where the constant \( \tilde{L} \) is to be determined later. For each \( s \in [0, T) \) and \( x_s \in \Lambda_s \), let

\[
V^\varepsilon(s, x_s) = \text{essinf}_{\theta \in \mathcal{U}} \left[ \int_s^T f^N \left( \tilde{W}(t_1 \wedge t), \ldots, \tilde{W}(t_N \wedge t), t, X^s, \theta, N(0), X^s, \theta, N(t_1 \wedge t), \ldots, X^s, \theta, N(t \wedge t), \theta(t) \right) dt \\
+ G^N \left( \tilde{W}(t_1), \ldots, \tilde{W}(t_N), X^s, \theta, N(0), X^s, \theta, N(t_1), \ldots, X^s, \theta, N(t) \right) \right],
\]

where \( X^s, \theta, N \) satisfies the SDE

\[
\begin{aligned}
\begin{cases}
\d X(t) = \beta^N \left( \tilde{W}(t_1 \wedge t), \ldots, \tilde{W}(t_N \wedge t), t, X(t), X(t_1 \wedge t), \ldots, X(t_N \wedge t), \theta(t) \right) dt \\
\quad + \sigma(t, X(t), \theta(t)) d\tilde{W}(t), & t \in [s, T]; \\
X(t) = x_s(t), & t \in [0, s].
\end{cases}
\end{aligned}
\]

For each \( s \in [t_{N-1}, T) \), we have the representation

\[
V^\varepsilon(s, x_s) = \tilde{V}^\varepsilon(\tilde{W}(t_1), \ldots, \tilde{W}(t_{N-1}), \tilde{W}(s), s, x(0), \ldots, x(t_{N-1}), x(s))
\]

with

\[
\tilde{V}^\varepsilon(\tilde{W}(t_1), \ldots, \tilde{W}(t_{N-1}), \tilde{y}, s, x(0), \ldots, x(t_{N-1}), \tilde{x})
\]

\[
= \text{essinf}_{\theta \in \mathcal{U}} \left[ \int_s^T f^N \left( \tilde{W}(t_1), \ldots, \tilde{W}(t_{N-1}), \tilde{W}(t), t, \ldots, x(t_{N-1}), X^s, \theta, N(t), \theta(t) \right) dt \\
+ G^N \left( \tilde{W}(t_1), \ldots, \tilde{W}(t_N), x(0), \ldots, x(t_{N-1}), X^s, \theta, N(T) \right) \right].
\]

By the viscosity solution theory of fully nonlinear parabolic PDEs (see [18] Theorems I.1 and II.1) for instance), the function \( \tilde{V}^\varepsilon(\tilde{W}(t_1), \ldots, \tilde{W}(t_{N-1}), \tilde{y}, s, x(0), \ldots, x(t_{N-1}), \tilde{x}) \) satisfies the following HJB equation over time interval \([t_{N-1}, t_N]\):

\[
\begin{aligned}
\begin{cases}
-\partial_t u(y, t, \tilde{x}) = \frac{1}{2} \text{tr} (D^2 u(y, t, \tilde{x})) + \text{essinf}_{v \in \mathcal{U}} \left\{ \text{tr} \left( \frac{1}{2} \sigma'(t, \tilde{x}, v) D_{\tilde{x}}^2 u(y, t, \tilde{x}) + \sigma(t, \tilde{x}, v) D_{\tilde{x}} u(y, t, \tilde{x}) \right) \\
\quad + \beta^N'(\tilde{W}(t_1), \ldots, \tilde{W}(t_{N-1}), \tilde{y}, t, x(0), \ldots, x(t_{N-1}), \tilde{x}, v) D_{\tilde{x}} u(y, t, \tilde{x}) \\
\quad + f^N(\tilde{W}(t_1), \ldots, \tilde{W}(t_{N-1}), \tilde{y}, t, x(0), \ldots, x(t_{N-1}), \tilde{x}, v) \right\}; \\
u(y, T, \tilde{x}) = G^N(\tilde{W}(t_1), \ldots, \tilde{W}(t_{N-1}), \tilde{y}, x(0), \ldots, x(t_{N-1}), \tilde{x}).
\end{cases}
\end{aligned}
\]

Here, we just write \( x(t_j) = x_s(t_j) \) for \( j = 0, \ldots, N - 1 \), as they are deemed to be fixed for \( s \in (t_{N-1}, T) \); the classical derivatives are denoted by \( D_{\tilde{x}}, D_{\tilde{x}}^2, D_{\tilde{y}, \tilde{x}}, D_t \), and \( D_{\tilde{y}, \tilde{x}} \). The regularity theory of viscosity solutions then implies that for each \( (x(0), \ldots, x(t_{N-1})) \in \mathbb{R}^{N \times d} \),

\[
\tilde{V}^\varepsilon(\tilde{W}(t_1), \ldots, \tilde{W}(t_{N-1}), x(0), \ldots, x(t_{N-1})), \\
\in \cap_{t \in (t_{N-1}, T)} L^\infty \left( \Omega, \tilde{\mathcal{F}}_{t_{N-1}}; G^{1+\frac{N}{2}+2+\alpha}([t_{N-1}, T] \times \mathbb{R}^{m_0+d}) \right),
\]

(5.5)
for some $\alpha \in (0, 1)$, where the time-space H"older space $C^{1+\frac{\alpha}{2},2+\alpha}([t_{N-1},t] \times \mathbb{R}^d)$ is defined as usual. At time $t_{N-1}$, we check that $V^\varepsilon$ is still uniformly Lipschitz-continuous w.r.t. $(\tilde{W}(t_1), \ldots, \tilde{W}(t_{N-1}))$ and $(x(0), \ldots, x(t_{N-1}))$. Then, we may conduct the same discussions on time interval $[t_{N-2},t_{N-1}]$ with the previously obtained $V^\varepsilon(t_{N-1},x)$ as the terminal value, and recursively on intervals $[t_{N-3},t_{N-2}], \ldots, [0, t_1]$.

Meanwhile, applying the Itô-Kunita formula of [16, Pages 118-119] to $\tilde{V}^\varepsilon$ on $[t_{N-1}, T]$ yields

$$-dV^\varepsilon(t,x_t) = \text{essinf}_{v \in U} \left\{ \frac{1}{2} \sigma^2 \sigma'^{-1}_D \tilde{V}^\varepsilon + \tilde{d} \tilde{G} \tilde{V}^\varepsilon \right\} (\tilde{W}(t_1), \ldots, \tilde{W}(t), t, x(0), \ldots, x(t_{N-1}), x(t), v)$$

$$+ \beta(t_1, \frac{N}{\varepsilon}) (\tilde{W}(t_1), \ldots, \tilde{W}(t_{N-1}), \tilde{W}(t), t, x(0), \ldots, x(t_{N-1}), x(t), v) \nabla V^\varepsilon(t,x_t)$$

$$+ f^N(\tilde{W}(t_1), \ldots, \tilde{W}(t_{N-1}), \tilde{W}(t), t, x(0), \ldots, x(t_{N-1}), x(t), v) \right\} dt$$

$$- D_g \tilde{V}^\varepsilon(\tilde{W}(t_1), \ldots, \tilde{W}(t), t, x(0), \ldots, x(t_{N-1}), x(t)) d\tilde{W}(t), \ t \in [t_{N-1}, T) \ and \ x \in \Lambda_t;$$

$$V^\varepsilon(t,x_T) = G^N(\tilde{W}(t_1), \ldots, \tilde{W}(T), x(0), \ldots, x(t_{N-1}), x(T)), \ x_T \in \Lambda_T.$$
where we have used the Hölder estimate (5.5). For $F_2$, we notice that

$$
\frac{\tilde{V}^\varepsilon\left(\tilde{W}(\tau_k), x(\tau_k), e_i \Delta \tilde{W}(\tau_k)\right) - \tilde{V}^\varepsilon\left(\tilde{W}(\tau_{k+1}), x(\tau_k)\right)}{|\Delta \tilde{W}(\tau_k)|} - D_{\tilde{z}} \tilde{V}^\varepsilon(\tilde{W}(\tau_k), \tau_k, x(\tau_k))
$$

\small
\begin{align*}
&= \int_0^1 \left(D_{\tilde{z}} \tilde{V}^\varepsilon(\tilde{W}(\tau_k), \tau_k, x(\tau_k) + \lambda e_i |\Delta \tilde{W}(\tau_k)|) - D_{\tilde{z}} \tilde{V}^\varepsilon(\tilde{W}(\tau_k), \tau_k, x(\tau_k))\right) d\lambda \\
&= \int_0^1 \int_0^1 D_{\tilde{z}, \tilde{z}} \tilde{V}^\varepsilon(\tilde{W}(\tau_k), \tau_k, x(\tau_k) + \lambda_1 \lambda e_i |\Delta \tilde{W}(\tau_k)|) \lambda d\lambda_1 d\lambda \cdot |\Delta \tilde{W}(\tau_k)|, \text{ a.s.,} \\
&\text{(5.9)}
\end{align*}

and similarly,

\small
\begin{align*}
\frac{\tilde{V}^\varepsilon\left(\tilde{W}(\tau_k), x(\tau_k), e_i \Delta \tilde{W}(\tau_k)\right) - \tilde{V}^\varepsilon\left(\tilde{W}(\tau_{k+1}), x(\tau_k)\right)}{|\Delta \tilde{W}(\tau_k)|} - D_{\tilde{z}, \tilde{z}} \tilde{V}^\varepsilon(\tilde{W}(\tau_k), \tau_k, x(\tau_k))
\end{align*}

\small
\begin{align*}
&= \int_0^1 \int_0^1 D_{\tilde{z}, \tilde{z}} \tilde{V}^\varepsilon(\tilde{W}(\tau_k), \tau_k, x(\tau_k) + \lambda_1 \lambda e_i |\Delta \tilde{W}(\tau_k)|) \lambda d\lambda_1 d\lambda \cdot |\Delta \tilde{W}(\tau_k)|, \text{ a.s.} \\
&\text{(5.10)}
\end{align*}

Subtracting (5.9) from (5.10) and applying similarly the integration-by-parts formula yield

\small
\begin{align*}
&\leq C_2 \left(|\tau_{k+1} - \tau_k|^{\frac{3}{2}} \cdot |\Delta \tilde{W}(\tau_k)| + |\tau_{k+1} - \tau_k| \cdot |\Delta \tilde{W}(\tau_k)|^{\bar{a} - 1}\right), \text{ a.s.,}
\end{align*}

which combined with (5.8) yields that

\small
\begin{align*}
&\sum_{k=0}^{N-1} \left(\nabla_i V^\varepsilon(\tau_{k+1}, x_{\tau_k}, \tau_{k+1} - \tau_k) - \nabla_i V^\varepsilon(\tau_k, x_{\tau_k})\right) \int_{\tau_k}^{\tau_{k+1}} g(s) dW^i(s) \\
&= \sum_{k=0}^{N-1} \sum_{j=1}^{m_0} \left(D_{\tilde{y}, \tilde{z}} \tilde{V}^\varepsilon(\tilde{W}(\tau_k), \tau_{k+1}, x(\tau_k)) \cdot \left(\Delta \tilde{W}(\tau_k)\right)^j + \varepsilon_k^2 \right) \int_{\tau_k}^{\tau_{k+1}} g(s) dW^i(s) \\
&= \sum_{k=0}^{N-1} \int_{\tau_k}^{\tau_{k+1}} \sum_{j=1}^{m_0} D_{\tilde{y}, \tilde{z}} \tilde{V}^\varepsilon(\tilde{W}(\tau), \tau, x(\tau_k)) d\tilde{W}^j(\tau) \int_{\tau_k}^{\tau_{k+1}} g(s) dW^i(s) \\
&+ \sum_{k=0}^{N-1} \int_{\tau_k}^{\tau_{k+1}} \left(D_{\tilde{y}, \tilde{z}} \tilde{V}^\varepsilon(\tilde{W}(\tau), \tau, x(\tau_k)) - D_{\tilde{y}, \tilde{z}} \tilde{V}^\varepsilon(\tilde{W}(\tau_k), \tau, x(\tau_k))\right) \cdot d\tilde{W}(\tau) \\
&\quad \cdot \int_{\tau_k}^{\tau_{k+1}} g(s) dW^i(s) + \sum_{k=0}^{N-1} \varepsilon_k^2 \int_{\tau_k}^{\tau_{k+1}} g(s) dW^i(s),
\end{align*}

\small
\begin{align*}
&= \sum_{k=0}^{N-1} \int_{\tau_k}^{\tau_{k+1}} g(s) dW^i(s) + \sum_{k=0}^{N-1} \varepsilon_k^2 \int_{\tau_k}^{\tau_{k+1}} g(s) dW^i(s),
\end{align*}

22
with
\[ |\varepsilon_k^2| \leq C_3 \left( |\Delta \tilde{W}(\tau_k)|^{1+\alpha} + |\tau_{k+1} - \tau_k|^{\frac{p}{2}} \cdot |\Delta \tilde{W}(\tau_k)| + |\tau_{k+1} - \tau_k| \cdot |\Delta \tilde{W}(\tau_k)|^{\alpha-1} \right), \text{ a.s.,} \]
where the constant $C_3$ is independent of the partition $\Pi$. As
\[
E \sum_{k=0}^{N-1} \left( |\Delta \tilde{W}(\tau_k)|^{1+\alpha} + |\tau_{k+1} - \tau_k|^{\frac{p}{2}} \cdot |\Delta \tilde{W}(\tau_k)| + |\tau_{k+1} - \tau_k| \cdot |\Delta \tilde{W}(\tau_k)|^{\alpha-1} \right) \int_{\tau_k}^{\tau_{k+1}} g(s) \, dW^i(s)
\]
\[ \leq C \sum_{k=0}^{N-1} \left( E \left[ |\Delta \tilde{W}(\tau_k)|^{p+p\alpha} + |\tau_{k+1} - \tau_k|^{\frac{p}{2}} \cdot |\Delta \tilde{W}(\tau_k)|^p + |\tau_{k+1} - \tau_k| \cdot |\Delta \tilde{W}(\tau_k)|^{p\alpha-p} \right] \right)^{\frac{1}{p}}
\]
\[ \leq C \sum_{k=0}^{N-1} |\tau_{k+1} - \tau_k|^{1+\frac{p}{2}} \to 0, \text{ as } |\Pi| \to 0,
\]
with $1 < p < 1/(1-\alpha)$ and
\[
E \sum_{k=0}^{N-1} \int_{\tau_k}^{\tau_{k+1}} \left( D_{\tilde{g}^x} \tilde{V}^\varepsilon(\tilde{W}(\tau), \tau, x(\tau_k)) - D_{\tilde{g}^x} \tilde{V}^\varepsilon(\tilde{W}(\tau_k), \tau_{k+1}, x(\tau_k)) \right) \, d\tilde{W}(\tau) \int_{\tau_k}^{\tau_{k+1}} g(s) \, dW^j(s)
\]
\[ \leq C \sum_{k=0}^{N-1} \left( E \int_{\tau_k}^{\tau_{k+1}} \left| D_{\tilde{g}^x} \tilde{V}^\varepsilon(\tilde{W}(\tau), \tau, x(\tau_k)) - D_{\tilde{g}^x} \tilde{V}^\varepsilon(\tilde{W}(\tau_k), \tau_{k+1}, x(\tau_k)) \right|^2 \, d\tau \cdot E \int_{\tau_k}^{\tau_{k+1}} |g(s)|^2 \, ds \right)^{1/2}
\]
\[ \leq C \sum_{k=0}^{N-1} \left( E \int_{\tau_k}^{\tau_{k+1}} \left( |\tau_{k+1} - \tau|^\alpha + |W(\tau) - W(\tau_k)|^{2\alpha} \right) \, d\tau \cdot |\tau_{k+1} - \tau_k| \right)^{1/2}
\]
\[ \leq C \sum_{k=0}^{N-1} |\tau_{k+1} - \tau_k|^{1+\frac{p}{2}} \to 0, \text{ as } |\Pi| \to 0,
\]
using standard computations for covariation (see [22] Section 6 of Chapter II) for instance gives
\[
\lim_{|\Pi| \to 0^+} \sum_{k=0}^{N-1} \left( \nabla_i V^\varepsilon(\tau_{k+1}, x_{\tau_k}, x_{\tau_{k+1}}) - \nabla_i V^\varepsilon(\tau_k, x_{\tau_k}) \right) \int_{\tau_k}^{\tau_{k+1}} g(s) \, dW^i(s)
\]
\[ = \lim_{|\Pi| \to 0^+} \sum_{k=0}^{N-1} \int_{\tau_k}^{\tau_{k+1}} \sum_{j=1}^{m_0} D_{g^x} \tilde{V}^\varepsilon(\tilde{W}(\tau), \tau, x(\tau_k)) \, dW^j(\tau) \int_{\tau_k}^{\tau_{k+1}} g(s) \, dW^i(s)
\]
\[ = \int_{t_{N-1}}^{T} (\partial_x \nabla \tilde{V}^\varepsilon)^{L^1}(t,x) g(t) \, dt, \text{ in probability.}
\]
It follows similarly for subintervals $(t_{j-1}, t_j)$ for $j = 1, \ldots, N-1$ and this yields [5.7] and $V^\varepsilon \in C_2$. In view of the approximations in Assumption (A3) and with an analogy to (iv) in Proposition 4.4, we may select the constant $\tilde{L} > 0$ such that for all $t \in [0, T] \setminus \{t_0, \ldots, t_{N-1}\}$ with $x_t \in \Lambda_t$, $|\nabla V^\varepsilon(t,x)| \leq \tilde{L}$, a.s., with $\tilde{L}$ being independent of $\varepsilon$ and $N$. Put
\[ V^\varepsilon(s,x) = V^\varepsilon(s,x_s) + Y^\varepsilon(s), \quad \overline{V}^\varepsilon(s,x) = V^\varepsilon(s,x_s) - Y^\varepsilon(s).
\]
It remains to verify $\nabla \overline{V}^\varepsilon \in \overline{F}$ and $V^\varepsilon \in \overline{F}$ and find a constant $C_4$ independent of $\varepsilon$ and $N$ s.t.
\[ E \left[ |\nabla \overline{V}^\varepsilon(s,x)| - V(s,x) \right] + E \left[ |V^\varepsilon(s,x) - V(s,x)| \right] \leq C_4 \cdot \varepsilon, \quad \forall s \in [0,T] \text{ with } x_s \in \Lambda_0^0,
\]
which together with the relation $\nabla \overline{V}^\varepsilon \geq V \geq \overline{V}^\varepsilon$ finally yields $u = V = \overline{W}$. As the remaining part of the proof is analogous to that of [23] Theorem 5.6, it is omitted. \qed
In the above proof, we construct the approximations of $V$ with $V^\varepsilon$ and $\tilde{V}^\varepsilon$ that are just $\tilde{\mathcal{F}}_t$-adapted. Thus, for each $t \in [0, T]$ and $x_t \in \Lambda^0_t$, $V(t, x_t)$ is just $\tilde{\mathcal{F}}_t$-measurable, which indicates that $(\partial_\omega \nabla V^\varepsilon)^{ij}$ is actually vanishing for $l = m_0 + 1, \ldots, m$, $j = 1, \ldots, d$. Hence, under assumptions (A1) – (A3), the SPHJB equation may be equivalently written as

$$\begin{cases}
-\partial_t u(t, x_t) = \essinf_{v \in U} \left\{ \text{tr} \left( \frac{1}{2} \sigma'(t, x_t, v) \nabla^2 u(t, x_t) + \tilde{\sigma}(t, x_t, v) \partial_\omega \nabla u(t, x_t) \right) 
+ \beta'(t, x_t, v) \nabla^2 u(t, x_t) + f(t, x_t, v) \right\}, (t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d); \\
u(T, x) = G(x), \quad x \in C([0, T]; \mathbb{R}^d),
\end{cases}$$

where we use the notation $\partial_\omega \nabla u = \left( (\partial_\omega \nabla u)^{kj} \right)_{1 \leq k \leq m_0, 1 \leq j \leq d}$.

References

[1] E. Bayraktar and J. Qiu, Controlled reflected SDEs and Neumann problem for backward SPDEs, Ann. Appl. Probab., 29 (2019), pp. 2819–2848.

[2] P. Briand, B. Delyon, Y. Hu, E. Pardoux, and L. Stoica, $L^p$ solutions of backward stochastic differential equations, Stoch. Process. Appl., 108 (2003), pp. 604–618.

[3] R. Buckdahn, C. Keller, J. Ma, and J. Zhang, Pathwise viscosity solutions of stochastic PDEs and forward path-dependent PDEs—a rough path view, arXiv:1501.06978 (2015).

[4] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions, The Master Equation and the Convergence Problem in Mean Field Games: (AMS-201), vol. 201, Princeton University Press, 2019.

[5] R. Cont and D.-A. Fournié, Functional Itô calculus and stochastic integral representation of martingales, Ann. Probab., 41 (2013), pp. 109–133.

[6] A. Cosso, S. Federico, F. Gozzi, M. Rosestolato, and N. Touzi, Path-dependent equations and viscosity solutions in infinite dimension, Ann. Probab., 46 (2018), pp. 126–174.

[7] G. Da Prato and J. Zabczyk, Stochastic equations in infinite dimensions, Cambridge university press, 2014.

[8] K. Du, J. Qiu, and S. Tang, $L^p$ theory for super-parabolic backward stochastic partial differential equations in the whole space, Appl. Math. Optim., 65 (2011), pp. 175–219.

[9] K. Du, S. Tang, and Q. Zhang, $W^{m,p}$-solution ($p \geq 2$) of linear degenerate backward stochastic partial differential equations in the whole space, J. Differ. Eq., 254 (2013), pp. 2877–2904.

[10] I. Ekren, C. Keller, N. Touzi, and J. Zhang, On viscosity solutions of path dependent PDEs, Ann. Probab., 42 (2014), pp. 204–236.

[11] I. Ekren, N. Touzi, and J. Zhang, Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part I, Ann. Probab., 44 (2016), pp. 1212–1253.

[12] ———, Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part II, Ann. Probab., 44 (2016), pp. 2507–2553.

[13] I. Ekren and J. Zhang, Pseudo-Markovian viscosity solutions of fully nonlinear degenerate PPDEs, Probability, Uncertainty and Quantitative Risk, 1 (2016), pp. 1–34.
[14] N. Englezos and I. Karatzas, Utility maximization with habit formation: Dynamic programming and stochastic PDEs, SIAM J. Control Optim., 48 (2009), pp. 481–520.

[15] Y. Hu, J. Ma, and J. Yong, On semi-linear degenerate backward stochastic partial differential equations, Probab. Theory Relat. Fields, 123 (2002), pp. 381–411.

[16] H. Kunita, Some extensions of Itô’s formula, in Séminaire de Probabilités XV 1979/80, Springer, 1981, pp. 118–141.

[17] D. Leão, A. Ohashi, and A. B. Simas, A weak version of path-dependent functional Itô calculus, Ann. Probab., 46 (2018), pp. 3399–3441.

[18] P. L. Lions, Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations, Part II, Commun. Partial Differ. Equ., 8 (1983), pp. 1229–1276.

[19] S.-E. A. Mohammed, Stochastic functional differential equations, vol. 99, Pitman Advanced Publishing Program, 1984.

[20] S. Peng, Stochastic Hamilton-Jacobi-Bellman equations, SIAM J. Control Optim., 30 (1992), pp. 284–304.

[21] S. Peng and F. Wang, BSDE, path-dependent PDE and nonlinear Feynman-Kac formula, Science China Mathematics, 59 (2016), pp. 19–36.

[22] P. E. Protter, Stochastic Integration and Differential Equations, Springer, Berlin Heidelberg, 2005.

[23] J. Qiu, Viscosity solutions of stochastic Hamilton–Jacobi–Bellman equations, SIAM J. Control Optim., 56 (2018), pp. 3708–3730.

[24] ——, Controlled ordinary differential equations with random path-dependent coefficients and stochastic path-dependent Hamilton-Jacobi equations, [arXiv:2005.01232] (2020).

[25] Z. Ren, N. Touzi, and J. Zhang, Comparison of viscosity solutions of fully nonlinear degenerate parabolic path-dependent PDEs, SIAM J. Math. Anal., 49 (2017), pp. 4093–4116.

[26] D. Revuz and M. Yor, Continuous martingales and Brownian motion, vol. 293, Springer, third ed., 1999.

[27] S. Tang and W. Wei, On the cauchy problem for backward stochastic partial differential equations in Hölder spaces, Ann. Probab., 44 (2016), pp. 360–398.

[28] S. Tang and F. Zhang, Path-dependent optimal stochastic control and viscosity solution of associated Bellman equations, Discrete and Continuous Dynamical Systems - Series A, 35 (2015), pp. 5521–5553.

[29] J. Yong and X. Y. Zhou, Stochastic controls: Hamiltonian systems and HJB equations, vol. 43, Springer Science & Business Media, 1999.