Acta Mathematica Scientia, 2019, **39B**(3): 857–873
https://doi.org/10.1007/s10473-019-0314-3
©Wuhan Institute Physics and Mathematics,
Chinese Academy of Sciences, 2019
Mathematica Acta Scientia 数学物理学报
http://actams.wipm.ac.cn

UNIQUENESS OF VISCOSITY SOLUTIONS OF STOCHASTIC HAMILTON-JACOBI EQUATIONS*

Jinniao QIU (仇金鸟) *Wenning WEI* (魏文宁)
Department of Mathematics & Statistics, University of Calgary, 2500 University Drive NW, Calgary, AB T2N 1N4, Canada
E-mail: *jinniao.qiu@ucalgary.ca; wenning.wei@ucalgary.ca*

Abstract This article is devoted to the study of fully nonlinear stochastic Hamilton-Jacobi (HJ) equations for the optimal stochastic control problem of ordinary differential equations with random coefficients. Under the standard Lipschitz continuity assumptions on the coefficients, the value function is proved to be the unique viscosity solution of the associated stochastic HJ equation.

Key words Stochastic Hamilton-Jacobi equation; optimal stochastic control; backward stochastic partial differential equation; viscosity solution

2010 MR Subject Classification 49L20; 49L25; 93E20; 35D40; 60H15

1 Introduction

Let $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$ be a complete filtered probability space with the filtration satisfying the usual conditions and generated by an m-dimensional Wiener process $W = \{W_t : t \in [0,\infty)\}$ together with all the $\mathbb{P}$-null sets in $\mathscr{F}$. The predictable σ-algebra on $\Omega \times [0,T]$ associated with $\{\mathscr{F}_t\}_{t\geq 0}$ is denoted by $\mathscr{P}$.

This article is devoted to the uniqueness of viscosity solution to the following stochastic Hamilton-Jacobi (HJ) equation:

$$\begin{cases} -\mathrm{d}u(t,x) = \mathbb{H}(t,x,Du)\mathrm{d}t - \psi(t,x)\,\mathrm{d}W_t, & (t,x) \in Q := [0,T) \times \mathbb{R}^d; \\ u(T,x) = G(x), & x \in \mathbb{R}^d, \end{cases} \tag{1.1}$$

with

$$\mathbb{H}(t,x,p) = \operatorname*{essinf}_{v\in U} \Big\{ \beta'(t,x,v)p + f(t,x,v) \Big\}, \quad \text{for } p \in \mathbb{R}^d,$$

where $T \in (0,\infty)$ is a fixed deterministic terminal time, $U \subset \mathbb{R}^n$ is a nonempty compact set, and both the random fields $u(t,x)$ and $\psi(t,x)$ are unknown.

*Received July 18, 2018; revised January 5, 2019. Jinniao Qiu is partially supported by the National Science and Engineering Research Council of Canada (NSERC) and by the start-up funds from the University of Calgary. The support of the NSERC grant of Professor Robert Elliott for Wenning Wei is gratefully acknowledged.

Stochastic HJ equations like (1.1) arise naturally from optimal stochastic control problems of the following form:

$$\inf_{\theta\in\mathcal{U}} E\left[\int_0^T f(s,X_s,\theta_s)\mathrm{d}s + G(X_T)\right] \tag{1.2}$$

subject to

$$\begin{cases} \mathrm{d}X_t = \beta(t,X_t,\theta_t)\mathrm{d}t, \ \ t\in[0,T]; \\ X_0 = x, \end{cases} \tag{1.3}$$

where $\mathcal{U}$ is the set of all the U-valued and $\mathscr{F}_t$-adapted processes and the coefficients β, f, and G depend not only on time, space. and control but also explicitly on $\omega\in\Omega$ (see assumption $(\mathcal{A}1)$). The state process $(X_t)_{t\in[0,T]}$ is governed by the control $\theta\in\mathcal{U}$, and to indicate the dependence of the state process on the control θ, the initial time r, and initial state $x\in\mathbb{R}^d$, we also write $X_t^{r,x;\theta}$ for $0\le r\le t\le T$. Following the dynamic programming method, we may define the dynamic cost functional by

$$J(t,x;\theta) = E_{\mathscr{F}_t}\left[\int_t^T f(s,X_s^{t,x;\theta},\theta_s)\mathrm{d}s + G(X_T^{t,x;\theta})\right], \ \ t\in[0,T]. \tag{1.4}$$

Here and throughout this work, we use $E_{\mathscr{F}_t}[\,\cdot\,]$ to denote the conditional expectation given σ-algebra $\mathscr{F}_t$ for each $t\ge 0$. Then, it is proved that the value function

$$V(t,x) = \operatorname*{essinf}_{\theta\in\mathcal{U}} J(t,x;\theta), \quad t\in[0,T], \tag{1.5}$$

is a viscosity solution of the stochastic HJ equation (1.1) (see [28, Theorem 4.2]).

The stochastic HJ equation (1.1), because of the vanishing diffusion coefficients in the controlled differential equation (1.3), may be regarded as a degenerate case of fully nonlinear stochastic Hamilton-Jacobi-Bellman (HJB) equations that were first introduced by Peng [24]. Peng proved the existence and uniqueness of weak solutions in Sobolev spaces for the superparabolic semilinear stochastic HJB equations in [24], while the wellposedness of general cases was claimed as an open problem, referring to Peng's plenary lecture of ICM 2010 [25]. In fact, the stochastic HJ equations are a class of backward stochastic partial differential equations (BSPDEs) which have been studied since about forty years ago (see[23]). The linear and semilinear BSPDEs have been extensively studied; we refer to [8, 13, 21, 31] among many others. For the weak solutions and associated local behavior analysis for general quasi-linear BSPDEs, see [29], and we refer to [12] for BSPDEs with singular terminal conditions. In the recent work [27], the first author studied the weak solution in Sobolev spaces for a special class of the fully nonlinear stochastic HJB equations (with $\beta\equiv 0$ and $\sigma(t,x,v)\equiv v$).

More recently, a notion of viscosity solution was proposed in [28] for general fully nonlinear stochastic HJB equations. In [28], the value function V was verified to be the maximal viscosity solution under certain assumptions on the regularity of coefficients (see $(\mathcal{A}^*)$ in Remark 2.2), and further for the superparabolic cases when the diffusion coefficients σ do not depend explicitly on $\omega\in\Omega$, the uniqueness is proved. In this article, we shall drop the strong assumptions on regularity of coefficients (see Remark 2.2) and prove the uniqueness of viscosity solution to stochastic HJ equation (1.1) corresponding to a degenerate fully nonlinear case of [28].

Recalling heuristically the notion of viscosity solution proposed in [28], we may think of the concerned random fields like the first unknown variable u and the value function V as stochastic

differential equations (SDEs) of the following form:

$$u(t,x) = u(T,x) - \int_t^T \mathfrak{d}_s u(s,x)\mathrm{d}s - \int_t^T \mathfrak{d}_w u(s,x)\,\mathrm{d}W_s, \quad (t,x) \in [0,T] \times \mathbb{R}^d. \tag{1.6}$$

The Doob-Meyer decomposition theorem implies the uniqueness of the pair $(\mathfrak{d}_t u,\, \mathfrak{d}_\omega u)$ and thus makes sense of the linear operators $\mathfrak{d}_t$ and $\mathfrak{d}_\omega$ which actually coincide with the two differential operators introduced by Leão, Ohashi, and Simas in [17, Theorem 4.3]. In fact, an earlier discussion on operator $\mathfrak{d}_\omega u$ may be found in [4, Section 5.2]. Through comparison, we have $\psi = \mathfrak{d}_\omega u$ and solving (1.1) with a pair (u,ψ) is equivalent to seeking u (of form (1.6)) satisfying

$$\begin{cases} -\mathfrak{d}_t u(t,x) - \mathbb{H}(t,x,Du(t,x)) = 0, & (t,x) \in Q; \\ u(T,x) = G(x), & x \in \mathbb{R}^d. \end{cases} \tag{1.7}$$

The equivalence relation between (1.1) and (1.7) provides the key to defining the viscosity solutions for stochastic HJ equations. As a standard assumption in the general stochastic control theory, all the involved coefficients herein are only measurable w.r.t. ω on the sample space $(\Omega, \mathscr{F})$ and this challenge prevents us from defining the viscosity solutions in a point-wise manner, while motivating us to use a class of random fields of form (1.6) with sufficient spacial regularity as test functions. At each point (τ,ξ) (τ may be stopping time and ξ may be an $\mathbb{R}^d$-valued $\mathscr{F}_\tau$-measurable variable) the classes of test functions are also parameterized by $\Omega_\tau \in \mathscr{F}_\tau$. Another challenge is from the nonanticipativity constraints on the unknown variables, which makes the classical variable-doubling techniques for deterministic HJ equations inapplicable in the proof of uniqueness for stochastic equations like (1.1). In this work, we first prove that the value function is the maximal viscosity (sub)solution which in fact reveals a weak version of comparison principle, and then through approximations, the value function is verified to be the unique one on the basis of the established comparison results.

We refer to [5, 6, 14, 32] among many others for the theory of (deterministic) viscosity solutions and [3, 18] for the stochastic viscosity solutions of (forward) SPDEs. Note that the (backward) stochastic HJB equations like (1.1) and the (forward) ones studied in [3, 18] are essentially different, that is, the noise term in the latter is exogenous, while in the former it is governed by the coefficients through the martingale representation and thus endogenous.

When the coefficients β, f, and G are deterministic functions of time t, control θ, and the paths of X and W, the optimal stochastic control problem is beyond the classical Markovian framework and the value function can be characterized by a path-dependent PDE. We refer to [9, 10, 20, 26] for the theory of viscosity solutions of such nonlinear path-dependent PDEs. In particular, in [9, 10], the authors applied the path-dependent viscosity solution theory to some classes of stochastic HJB equations which, however, required all the coefficients to be continuous in $\omega \in \Omega$ because of the involved pathwise analysis. We would stress that, in this article, all the involved coefficients are only measurable w.r.t. $\omega \in \Omega$ and we even do not need to specify any topology on Ω, which allows the general random variables to appear in the coefficients.

The rest of this article is organized as follows. In Section 2, we introduce in the first subsection some notations and the standing assumptions on the coefficients, and in the second subsection, the main result is exhibited. Two auxiliary results are presented in Section 3. Finally, Section 4 is devoted to the proof of our main result; we verify in the first subsection

that the value function is the maximal viscosity solution and then the uniqueness of viscosity solution is derived in the second subsection.

2 Preliminaries and Main Result

2.1 Preliminaries

Throughout this article, we write $(s,y)\to(t^+,x)$, meaning that $s\downarrow t$ and $y\to x$.

Let $\mathbb{B}$ be a Banach space equipped with norm $\|\cdot\|_{\mathbb{B}}$. For each $t\in[0,T]$, denote by $L^0(\Omega,\mathscr{F}_t;\mathbb{B})$ the space of $\mathbb{B}$-valued $\mathscr{F}_t$-measurable random variables. For $p\in[1,\infty]$, $\mathcal{S}^p(\mathbb{B})$ is the set of all the $\mathbb{B}$-valued, $\mathscr{P}$-measurable continuous processes $\{\mathcal{X}_t\}_{t\in[0,T]}$ such that

$$\|\mathcal{X}\|_{\mathcal{S}^p(\mathbb{B})}:=\left\|\sup_{t\in[0,T]}\|\mathcal{X}_t\|_{\mathbb{B}}\right\|_{L^p(\Omega,\mathscr{F},\mathbb{P})}<\infty.$$

Denote by $\mathcal{L}^p(\mathbb{B})$ the totality of all the $\mathbb{B}$-valued, $\mathscr{P}$-measurable processes $\{\mathcal{X}_t\}_{t\in[0,T]}$ such that

$$\|\mathcal{X}\|_{\mathcal{L}^p(\mathbb{B})}:=\left\|\left(\int_0^T\|\mathcal{X}_t\|_{\mathbb{B}}^p\mathrm{d}t\right)^{1/p}\right\|_{L^p(\Omega,\mathscr{F},\mathbb{P})}<\infty.$$

Obviously, $(\mathcal{S}^p(\mathbb{B}),\|\cdot\|_{\mathcal{S}^p(\mathbb{B})})$ and $(\mathcal{L}^p(\mathbb{B}),\|\cdot\|_{\mathcal{L}^p(\mathbb{B})})$ are Banach spaces. For each $(k,q)\in\mathbb{N}_0\times[1,\infty]$, we define the k-th Sobolev space $(H^{k,q},\|\cdot\|_{k,q})$ as usual, and for each domain $\mathcal{O}\subset\mathbb{R}^d$, denote by $C^k(\mathcal{O};\mathbb{B})$ the space of $\mathbb{B}$-valued functions with the up to k-th order derivatives being bounded and continuous on $\mathcal{O}$, $C_0^k(\mathcal{O};\mathbb{B})$ being the subspace of $C^k(\mathcal{O};\mathbb{B})$ vanishing on the boundary $\partial\mathcal{O}$. If there is no confusion about $\mathbb{B}$, we will omit $\mathbb{B}$ and just write $C^k(\mathcal{O})$ and $C_0^k(\mathcal{O})$. When $k=0$, write $C_0(\mathcal{O})$ and $C(\mathcal{O})$ simply. Through this article, we define $C^\infty(\mathcal{O};\mathbb{B})=\bigcap\limits_{k\in\mathbb{N}^+}C^k(\mathcal{O};\mathbb{B})$ and

$$C_0^\infty(\mathcal{O};\mathbb{B})=\bigcap_{k\in\mathbb{N}^+}C_0^k(\mathcal{O};\mathbb{B}),\quad \mathcal{S}^p(C_{loc}(\mathbb{R}^d))=\bigcap_{N>0}\mathcal{S}^p(C(B_N(0))),\qquad \text{for } p\in[1,\infty].$$

Throughout this work, we use the following assumption.

$(\mathcal{A}1)$ $G\in L^\infty(\Omega,\mathscr{F}_T;H^{1,\infty})$. For the coefficients $g=f,\beta^i$ $(1\le i\le d)$,

(i) $g:\ \Omega\times[0,T]\times\mathbb{R}^d\times U\to\mathbb{R}$ is $\mathscr{P}\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{B}(U)$-measurable;

(ii) for almost all $(\omega,t)\in\Omega\times[0,T]$, $g(t,x,v)$ is uniformly continuous on $\mathbb{R}^d\times U$;

(iii) there exists $L>0$ such that

$$\|G\|_{L^\infty(\Omega,\mathscr{F}_T;H^{1,\infty})}+\sup_{v\in\mathcal{U}}\|g(\cdot,\cdot,v)\|_{\mathcal{S}^\infty(H^{1,\infty})}\le L.$$

2.2 Main result

We first introduce the test function space for viscosity solutions.

Definition 2.1 For $u\in\mathcal{S}^2(C_{loc}(\mathbb{R}^d))$ with $Du\in\mathcal{L}^2(C(\mathbb{R}^d))$, we say $u\in\mathscr{C}^1_{\mathscr{F}}$ if there exists $(\mathfrak{d}_tu,\,\mathfrak{d}_\omega u)\in\mathcal{L}^2(C(\mathbb{R}^d))\times\mathcal{L}^2(C(\mathbb{R}^d))$ such that with probability 1,

$$u(r,x)=u(T,x)-\int_r^T\mathfrak{d}_su(s,x)\mathrm{d}s-\int_r^T\mathfrak{d}_\omega u(s,x)\,\mathrm{d}W_s,\quad\forall\,(r,x)\in[0,T]\times\mathbb{R}^d.$$

Remark 2.1 Instead of $\mathscr{C}^2_{\mathscr{F}}$ defined in [28] which requires D^2u and $D\mathfrak{d}_\omega u$ to be lying in $\mathcal{L}^2(C(\mathbb{R}^d))$, we use $\mathscr{C}^1_{\mathscr{F}}$ which imposes no requirement on D^2u or $D\mathfrak{d}_\omega u$. This is basically because the two terms D^2u and $D\mathfrak{d}_\omega u$ are not involved in the first-order BSPDE (1.1). Analogous to

the space $\mathscr{C}^2_{\mathscr{F}}$ in [28], by Definition 2.1, we have in fact characterized the two linear operators $\mathfrak{d}_t$ and $\mathfrak{d}_\omega$ which is consistent with the two differential operators w.r.t. the paths of Wiener process W in the sense of [17], defined via a finite-dimensional approximation procedure based on controlled inter-arrival times and approximating martingales; in particular, for the operator $\mathfrak{d}_\omega u$, an earlier discussion may be found in [4, Section 5.2]. We would also note that the operators $\mathfrak{d}_t$ and $\mathfrak{d}_\omega$ here are different from the path derivatives $(\partial_t, \partial_\omega)$ via the functional Itô formulas (see [2] and [10, Section 2.3]). If $u(\omega, t, x)$ is smooth enough w.r.t. (ω, t) in the path space, for each x, we have the relation

$$\mathfrak{d}_t u(\omega,t,x) = \left(\partial_t + \frac{1}{2}\partial^2_{\omega\omega}\right) u(\omega,t,x), \quad \mathfrak{d}_\omega u(\omega,t,x) = \partial_\omega u(\omega,t,x),$$

which can be seen either from the applications in [10, Section 6] to BSPDEs or from a rough view on the pathwise viscosity solution of (forward) SPDEs in [2].

For each stopping time $t \le T$, denote by $\mathcal{T}^t$ the set of stopping times τ satisfying $t \le \tau \le T$ and by $\mathcal{T}^t_+$ the subset of $\mathcal{T}^t$ such that $\tau > t$ for any $\tau \in \mathcal{T}^t_+$. For each $\tau \in \mathcal{T}^0$ and $\Omega_\tau \in \mathscr{F}_\tau$, we denote by $L^0(\Omega_\tau, \mathscr{F}_\tau; \mathbb{R}^d)$ the set of $\mathbb{R}^d$-valued $\mathscr{F}_\tau$-measurable functions.

We now introduce the notion of viscosity solutions. For each $(u,\tau) \in \mathcal{S}^2(C_{loc}(\mathbb{R}^d)) \times \mathcal{T}^0$, $\Omega_\tau \in \mathscr{F}_\tau$ with $\mathbb{P}(\Omega_\tau) > 0$, and $\xi \in L^0(\Omega_\tau, \mathscr{F}_\tau; \mathbb{R}^d)$, we define

$$\begin{aligned}
\underline{\mathcal{G}}u(\tau,\xi;\Omega_\tau) := \Big\{ & \phi \in \mathscr{C}^1_{\mathscr{F}} : (\phi - u)(\tau,\xi) 1_{\Omega_\tau} = 0 \\
& = \operatorname*{essinf}_{\bar\tau \in \mathcal{T}^\tau} E_{\mathscr{F}_\tau}\left[\inf_{y \in B_\delta(\xi)} (\phi - u)(\bar\tau \wedge \hat\tau, y)\right] 1_{\Omega_\tau} \text{ a.s.} \\
& \text{for some } (\delta, \hat\tau) \in (0,\infty) \times \mathcal{T}^\tau_+ \Big\}, \\
\overline{\mathcal{G}}u(\tau,\xi;\Omega_\tau) := \Big\{ & \phi \in \mathscr{C}^1_{\mathscr{F}} : (\phi - u)(\tau,\xi) 1_{\Omega_\tau} = 0 \\
& = \operatorname*{esssup}_{\bar\tau \in \mathcal{T}^\tau} E_{\mathscr{F}_\tau}\left[\sup_{y \in B_\delta(\xi)} (\phi - u)(\bar\tau \wedge \hat\tau, y)\right] 1_{\Omega_\tau} \text{ a.s.} \\
& \text{for some } (\delta, \hat\tau) \in (0,\infty) \times \mathcal{T}^\tau_+ \Big\}.
\end{aligned}$$

It is obvious that if $\underline{\mathcal{G}}u(\tau,\xi;\Omega_\tau)$ or $\overline{\mathcal{G}}u(\tau,\xi;\Omega_\tau)$ is nonempty, we must have $0 \le \tau < T$ on Ω_τ.

Now, it is at the stage to introduce the definition of viscosity solutions.

Definition 2.2 We say $u \in \mathcal{S}^2(C_{loc}(\mathbb{R}^d))$ is a viscosity subsolution (resp. supersolution) of BSPDE (1.1), if $u(T,x) \le$ (resp. $\ge$)$G(x)$ for all $x \in \mathbb{R}^d$ a.s., and for any $\tau \in \mathcal{T}^0$, $\Omega_\tau \in \mathscr{F}_\tau$ with $\mathbb{P}(\Omega_\tau) > 0$ and $\xi \in L^0(\Omega_\tau, \mathscr{F}_\tau; \mathbb{R}^d)$ and any $\phi \in \underline{\mathcal{G}}u(\tau,\xi;\Omega_\tau)$ (resp. $\phi \in \overline{\mathcal{G}}u(\tau,\xi;\Omega_\tau)$), there holds

$$\operatorname{ess} \liminf_{(s,x)\to(\tau^+,\xi)} E_{\mathscr{F}_\tau}\left\{-\mathfrak{d}_s\phi(s,x) - \mathbb{H}(s,x,D\phi(s,x))\right\} \le \ 0, \text{ for almost all } \omega \in \Omega_\tau \tag{2.1}$$

$$\text{(resp. } \operatorname{ess} \limsup_{(s,x)\to(\tau^+,\xi)} E_{\mathscr{F}_\tau}\left\{-\mathfrak{d}_s\phi(s,x) - \mathbb{H}(s,x,D\phi(s,x))\right\} \ge \ 0, \text{ for almost all } \omega \in \Omega_\tau). \tag{2.2}$$

The function u is a viscosity solution of BSPDE (1.1) if it is both a viscosity subsolution and a viscosity supersolution of (1.1).

The stochastic HJ equation (1.1) is a particular case of [28, Theorem 4.2] with vanishing diffusion coefficients. Therefore, as a straightforward consequence, we have the following existence of viscosity solution of BSPDE (1.1).

Theorem 2.1 (see [28, Theorem 4.2]) Let $(\mathcal{A}1)$ hold. The value function V defined by (1.5) is a viscosity solution of the stochastic Hamilton-Jacobi equation (1.1) in $\mathcal{S}^2(C(\mathbb{R}^d))$.

We note that even though the test function space used in [28] is $\mathscr{C}^2_{\mathscr{F}}$ instead of $\mathscr{C}^1_{\mathscr{F}}$, Proof of Theorem 2.1 follows exactly the same as that of [28, Theorem 4.2] as there would be no term involving D^2u or $D\mathfrak{d}_\omega u$ in the proof.

Our main result is focused on the uniqueness.

Theorem 2.2 Let $(\mathcal{A}1)$ hold. The viscosity solution to stochastic HJ equation (1.1) is unique in $\mathcal{S}^2(C(\mathbb{R}^d))$.

Remark 2.2 The uniqueness is twofold, consisting of the maximality and minimality of the value function V defined by (1.5). In [28, Theorem 5.2], it was concerned with the controlled stochastic differential equation:

$$\begin{cases} \mathrm{d}X_t = \beta(t, X_t, \theta_t)\mathrm{d}t + \sigma(t, X_t, \theta_t)\,\mathrm{d}W_t, \ \ t \in [0, T]; \\ X_0 = x, \end{cases} \tag{2.3}$$

instead of the controlled ordinary differential equation (1.3) with random coefficients, and the value function was just proved to be the maximal viscosity (sub)solution which, however, relies on the following additional strong assumption on the coefficients:

$(\mathcal{A}^*)$ There exists $q > 2 + \frac{d}{2}$ such that $(G(\cdot), f(\cdot,\cdot,\theta)) \in L^2(\Omega, \mathscr{F}_T; H^{q,2}) \times \mathcal{L}^2(H^{q,2})$ for any $\theta \in \mathcal{U}$, and $g(\cdot,\cdot,\theta) \in \mathcal{L}^\infty(H^{q,\infty})$ for $g = \beta^i, \sigma^{ij}$ $(1 \le i \le d,\ 1 \le j \le m)$.

In fact, the author in [28] only gave a complete uniqueness for superparabolic stochastic HJB equations with the diffusion coefficients depending only on time, state, and control (see [28, Theorem 5.6]), while stochastic HJ equation (1.1) has vanishing diffusion coefficients ($\sigma \equiv 0$) and thus is degenerate.

3 Auxiliary Results

In view of assumption $(\mathcal{A}1)$ and the vanishing diffusion coefficients of stochastic differential equation (1.3), we may conclude the following assertions straightforwardly from [28, Lemma 3.1].

Lemma 3.1 Let $(\mathcal{A}1)$ hold. Given $\theta \in \mathcal{U}$, for the strong solution of SDE (1.3), there exists $K > 0$ such that, for any $0 \le r \le t \le s \le T$ and $\xi \in L^0(\Omega, \mathscr{F}_r; \mathbb{R}^d)$,

(i) the two processes $\left(X_s^{r,\xi;\theta}\right)_{t\le s\le T}$ and $\left(X_s^{t,X_t^{r,\xi;\theta};\theta}\right)_{t\le s\le T}$ are indistinguishable;

(ii) $\max\limits_{r\le l\le T}\left|X_l^{r,\xi;\theta}\right| \le K\,(1+|\xi|)$ a.s.;

(iii) $\left|X_s^{r,\xi;\theta} - X_t^{r,\xi;\theta}\right| \le K\,(1+|\xi|)\,(s-t)$ a.s.;

(iv) given another $\hat{\xi} \in L^0(\Omega, \mathscr{F}_r; \mathbb{R}^d)$,

$$\max_{r\le l\le T}\left|X_l^{r,\xi;\theta} - X_l^{r,\hat{\xi};\theta}\right| \le K|\xi - \hat{\xi}| \quad \text{a.s.};$$

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

The following regular properties of the value function V are from [28, Proposition 3.3].

Proposition 3.2 Let $(\mathcal{A}1)$ hold.

(i) For each $t \in [0,T]$ and $\xi \in L^0(\Omega, \mathscr{F}_t; \mathbb{R}^d)$, there exists $\bar{\theta} \in \mathcal{U}$ such that

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

(ii) For each $(\bar{\theta}, x) \in \mathcal{U} \times \mathbb{R}^d$, $\left\{J(t, X_t^{0,x;\bar{\theta}}; \bar{\theta}) - V(t, X_t^{0,x;\bar{\theta}})\right\}_{t\in[0,T]}$ is a supermartingale, that is, for any $0 \le t \le \tilde{t} \le T$,

$$V(t, X_t^{0,x;\bar{\theta}}) \le E_{\mathscr{F}_t} V(\tilde{t}, X_{\tilde{t}}^{0,x;\bar{\theta}}) + E_{\mathscr{F}_t} \int_t^{\tilde{t}} f(s, X_s^{0,x;\bar{\theta}}, \bar{\theta}_s)\mathrm{d}s, \ \text{ a.s.} \tag{3.1}$$

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

(iv) There exists $L_V > 0$ such that for any $\theta \in \mathcal{U}$,

$$|V(t,x) - V(t,y)| + |J(t,x;\theta) - J(t,y;\theta)| \le L_V|x-y|, \ \text{ a.s.}, \quad \forall\, x, y \in \mathbb{R}^d,$$

with L_V depending only on T and the uniform Lipschitz constants of the coefficients β, σ, f and G w.r.t. the spatial variable x.

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

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \max\{|V(t,x)|, |J(t,x;\theta)|\} \le L(T+1) \quad \text{a.s.}$$

4 Proof of Theorem 2.2

The proof consists of two steps. In the first subsection, we prove that the value function is the maximal viscosity (sub)solution of the stochastic HJ equation (1.1), which essentially yields a weak version of comparison principle. In the second subsection, the uniqueness is addressed on the basis of the established comparison results through approximations.

Throughout this section, we define for any $\phi \in \mathscr{C}^1_{\mathscr{F}}$ and $v \in U$,

$$\mathscr{L}^v\phi(t,x) = \mathfrak{d}_t\phi(t,x) + \beta'(t,x,v)D\phi(t,x).$$

4.1 Maximal viscosity subsolution

We first prove that the value function is the maximal viscosity (sub)solution of BSPDE (1.1). Such maximality is parallel to that of [28, Theorem 5.2], but, as we want to achieve this without the additional strong regularity assumption (see $(\mathcal{A}^*)$ in Remark 2.2) required in [28], some new techniques are needed. The first one is based on smooth approximations.

Let

$$\rho(x) = \begin{cases} \tilde{c}\, e^{\frac{1}{|x|^2-1}} & \text{if } |x| < 1; \\ 0 & \text{otherwise;} \end{cases} \quad \text{with} \quad \tilde{c} := \left(\int_{|x|<1} e^{\frac{1}{x^2-1}}\mathrm{d}x\right)^{-1}, \tag{4.1}$$

and we define mollifier $\rho_l(x) = l^d\rho(lx)$, $x \in \mathbb{R}^d$ for each $l \in \mathbb{N}^+$. For $g = \beta^i(t,\cdot,v)$, $f(t,\cdot,v)$, $G(\cdot)$ $(1 \le i \le d,\ 1 \le j \le m)$, take convolutions

$$g_l(x) = \int_{\mathbb{R}^d} \rho_l(x-y)g(y)\mathrm{d}y, \quad x \in \mathbb{R}^d.$$

Then, the coefficients β_l, f_l, and G_l satisfy assumption $(\mathcal{A}1)$ for each $l \in \mathbb{N}^+$ and

$$\lim_{l\to\infty} \|G - G_l\|_{L^\infty(\Omega,\mathscr{F}_T;L^\infty(\mathbb{R}^d))} + \sup_{v\in\mathcal{U}} \big(\|f_l - f\|_{\mathcal{S}^\infty(L^\infty(\mathbb{R}^d))} + \|\beta_l - \beta\|_{\mathcal{S}^\infty(L^\infty(\mathbb{R}^d))}\big) = 0. \quad (4.2)$$

Moreover, as $G_l \in L^\infty(\Omega, \mathscr{F}_T; C^k(\mathbb{R}^d))$, $\beta_l^i(\cdot,\cdot,\theta), f_l(\cdot,\cdot,\theta) \in \mathcal{S}^\infty(C^k(\mathbb{R}^d))$, $i = 1, \cdots, d$ for any $(k,\theta) \in \mathbb{N}^+ \times \mathcal{U}$, by the classical solution theory for BSPDEs in [30, Lemma 5.1, Theorems 4.6, 5.1&5.2], we have

Proposition 4.1 For each $(l,\theta) \in \mathbb{N}^+ \times \mathcal{U}$, there exists a unique solution u_l in $\mathscr{C}^1_{\mathscr{F}} \cap \mathcal{S}^\infty(C^2(\mathbb{R}^d))$ to the following BSPDE:

$$\begin{cases} -\mathrm{d}u_l(t,x) = [\beta_l'(t,x,\theta_t)Du_l(t,x) + f_l(t,x,\theta_t)]\,\mathrm{d}t - \psi_l(t,x)\,\mathrm{d}W_t, & (t,x) \in Q; \\ u_l(T,x) = G_l(x), & x \in \mathbb{R}^d, \end{cases} \quad (4.3)$$

with $\psi_l = \mathfrak{d}_\omega u_l$, and for each $x \in \mathbb{R}^d$ and $0 \le t \le s \le T$, the random processes

$$X_s^{t,x;\theta,l}, \quad Y_s^{t,x;\theta,l} := u_l(s, X_s^{t,x;\theta,l}), \quad \text{and} \quad Z_s^{t,x;\theta,l} := \psi_l(s, X_s^{t,x;\theta,l})$$

satisfy the following forward-backward SDEs:

$$\begin{cases} X_s^{t,x;\theta} = x + \displaystyle\int_t^s \beta_l(r, X_r^{t,x;\theta}, \theta_r)\mathrm{d}r, & 0 \le t \le s \le T; \\ Y_s^{t,x;\theta} = G_l(X_T^{t,x;\theta}) + \displaystyle\int_s^T f_l(r, X_r^{t,x;\theta}, \theta_r)\mathrm{d}r - \int_s^T Z_r^{t,x;\theta}\,\mathrm{d}W_r, & 0 \le t \le s \le T. \end{cases} \quad (4.4)$$

We now introduce another two space-invariant stochastic processes. Put

$$\begin{aligned} \delta G_l &:= \operatorname*{esssup}_{x\in\mathbb{R}^d} |G_l(x) - G(x)|, \\ \delta f_t^{l,\theta} &:= \sup_{x\in\mathbb{R}^d} |f_l(t,x,\theta_t) - f(t,x,\theta_t)|, \quad \text{for } t \in [0,T], \\ \delta \beta_t^{l,\theta} &:= \sup_{x\in\mathbb{R}^d} |\beta_l(t,x,\theta_t) - \beta(t,x,\theta_t)|, \quad \text{for } t \in [0,T]. \end{aligned}$$

Let $(Y^l, Z^l) \in \mathcal{S}^2(\mathbb{R}) \times \mathcal{L}^2(\mathbb{R}^m)$ be the solution to BSDE:

$$Y_t^l = \delta G_l + \int_t^T (\delta f_s^{l,\theta} + K\delta\beta_s^{l,\theta})\mathrm{d}s - \int_t^T Z_s^l\,\mathrm{d}W_s,$$

where $K = L_V$ is the constant from (iv) of Proposition 3.2. Recalling relation (4.2), we have by the theory of BSDE,

$$\|Y^l\|_{\mathcal{S}^2(\mathbb{R})} \le C\delta_l, \quad \text{with } \delta_l = \|\delta G_l\|_{L^2(\Omega,\mathscr{F}_T)} + \|\delta f^{l,\theta} + K\delta\beta^{l,\theta}\|_{\mathcal{L}^2(\mathbb{R})} \to 0, \quad (4.5)$$

where the constant C is independent of l.

Lemma 4.2 For each $\theta \in \mathcal{U}$, setting $\hat{J}_l(t,x) = u_l(t,x) + Y_t^l$, we have $\hat{J}_l(t,x) \in \mathscr{C}^1_{\mathscr{F}}$ with

$$\lim_{l\to\infty} \|\hat{J}_l(t,x) - J(t,x;\theta)\|_{\mathcal{S}^2(C(\mathbb{R}^d))} = 0$$

and

$$\operatorname{ess}\liminf_{(s,x)\to(t^+,y)} E_{\mathscr{F}_t}\Big[-\mathfrak{d}_s\hat{J}_l(s,x) - \mathbb{H}(s,x,D\hat{J}_l(s,x))\Big] \ge 0, \text{ a.s.} \quad \forall (t,y) \in [0,T) \times \mathbb{R}^d. \quad (4.6)$$

Proof It follows obviously that $\hat{J}_l(t,x) \in \mathscr{C}^1_{\mathscr{F}}$ from Proposition 4.1.

Using Gronwall's inequality through standard computations gives for any $(s,x)\in[0,T]\times\mathbb{R}^d$,

$$\begin{aligned}\sup_{s\le t\le T}\left|X_t^{s,x;\theta,l}-X_t^{s,x;\theta}\right| &\le C\int_s^T\left|\beta_l\left(X_t^{s,x;\theta,l},\theta_t\right)-\beta\left(t,X_t^{s,x;\theta,l},\theta_t\right)\right|\mathrm{d}t\\ &\le C\delta_l,\quad \text{a.s.}\end{aligned}$$

with constant C depending only on L and T. Then

$$\begin{aligned}|u_l(s,x)-J(s,x;\theta)| &\le E_{\mathscr{F}_s}\Bigg[\int_s^T\left(\delta f_t^{l,\theta}+\left|f\left(t,X_t^{s,x;\theta,l},\theta_t\right)-f\left(t,X_t^{s,x;\theta},\theta_t\right)\right|\right)\mathrm{d}t\\ &\quad+\delta G_l+\left|G\left(X_T^{s,x;\theta,l}\right)-G\left(X_T^{s,x;\theta}\right)\right|\Bigg]\\ &\le |Y_s^l|+2L(T+1)E_{\mathscr{F}_s}\left[\sup_{s\le t\le T}\left|X_t^{s,x;\theta,l}-X_t^{s,x;\theta}\right|\right]\\ &\le C\delta_l,\quad \text{a.s.}\end{aligned}$$

with the constant C independent of l,s, and x, which together with (4.5) implies

$$\lim_{l\to\infty}\|\hat{J}_l(t,x)-J(t,x;\theta)\|_{\mathcal{S}^2(C(\mathbb{R}^d))}=0.$$

Notice that the coefficients β_l, f_l, and G_l satisfy assumption $(\mathcal{A}1)$ with the identical constant L. In view of Proposition (4.5) and the BSDE for Y^l, the random field u_l satisfies (iv) of Proposition 3.2 with the same Lipschitz constant L_V, and we have

$$\begin{aligned}-\mathfrak{d}_t u_l &= \beta_l' Du_l+f_l,\\ -\mathfrak{d}_t Y^l &= \delta f^{l,\theta}+L_V\delta\beta^{l,\theta}\end{aligned}$$

and thus

$$\begin{aligned}-\mathscr{L}^{\theta_s}\hat{J}_l-f &= -\mathfrak{d}_t\hat{J}_l-\beta_l'D\hat{J}_l-f_l-(\beta-\beta_l)'D\hat{J}_l-f+f_l\\ &= -\mathfrak{d}_t u_l+\delta f^{l,\theta}+L_V\delta\beta^{l,\theta}-\beta_l'Du_l-f_l-(\beta-\beta_l)'Du_l-f+f_l\\ &= \delta f^{l,\theta}+L_V\delta\beta^{l,\theta}-(\beta-\beta_l)'Du_l-f+f_l\\ &\ge 0,\end{aligned}$$

where we omitted the inputs for each involved function for the sake of convenience. Therefore, it holds that

$$\operatorname{ess}\liminf_{(s,x)\to(t^+,y)}E_{\mathscr{F}_t}\left[-\mathfrak{d}_s\hat{J}_l(s,x)-\mathbb{H}(s,x,D\hat{J}_l(s,x))\right]\ge 0,\ \text{a.s.}\quad\forall(t,y)\in[0,T)\times\mathbb{R}^d.$$

□

Recalling the compactly-supported smooth (bump) function $\rho(x)$ defined in (4.1), set

$$h(x)=\int_{\mathbb{R}^d}1_{\{|y|>1\}}(|y|-1)\,\rho(x-y)\mathrm{d}y,\quad x\in\mathbb{R}^d.$$

Then, the function $h(x)$ is convex and continuously differentiable with $h(0)=0$, $h(x)>0$ whenever $|x|>0$, and

$$h(x)>|x|-2,\quad |Dh(x)|\le 1 \text{ for any } x\in\mathbb{R}^d. \tag{4.7}$$

Theorem 4.3 Let $(\mathcal{A}1)$ hold. Let $u\in\mathcal{S}^2(C(\mathbb{R}^d))$ be a viscosity subsolution of the stochastic HJB equation (1.1). It holds a.s. that $u(t,x)\le V(t,x)$ for any $(t,x)\in[0,T]\times\mathbb{R}^d$, where V is the value function defined by (1.5).

Proof We argue by contradiction. Suppose that with a positive probability, $u(t,\bar{x}) > V(t,\bar{x})$ at some point $(t,\bar{x}) \in [0,T) \times \mathbb{R}^d$. In view of the approximating relations between $V(t,x)$, $J(t,x;\theta)$ and $\hat{J}_l(t,x)$ in Proposition 3.2 and Lemma 4.2, we have some $(l,\theta) \in \mathbb{N}^+ \times \mathcal{U}$ such that $u(t,\bar{x}) > \hat{J}_l(t,\bar{x})$ with a positive probability; more precisely, there exists $\kappa > 0$ such that $\mathbb{P}(\overline{\Omega}_t) > 0$ with $\overline{\Omega}_t := \{u(t,\bar{x}) - \hat{J}_l(t,\bar{x}) > \kappa\}$. Furthermore, for any $\varepsilon \in (0,1)$, there exists $\xi_t \in L^0(\overline{\Omega}_t, \mathscr{F}_t; \mathbb{R}^d)$ such that

$$\begin{aligned}\alpha :&= u(t,\xi_t) - \hat{J}_l(t,\xi_t) - \varepsilon h(\xi_t - \bar{x}) \\ &= \max_{x\in\mathbb{R}^d}\{u(t,x) - \hat{J}_l(t,x) - \varepsilon h(x-\bar{x})\} \geq \kappa \text{ for almost all } \omega \in \overline{\Omega}_t,\end{aligned}$$

where the existence and the measurablity of ξ_t follow from the measurable selection, the linear growth of function $h(x)$ (see (4.7)), and the fact that $u, \hat{J}_l \in \mathcal{S}^2(C(\mathbb{R}^d))$. Note that κ and $\overline{\Omega}_t$ are independent of ε. W.l.o.g, we take $\overline{\Omega}_t = \Omega$ in what follows.

For each $s \in (t,T]$, choose an $\mathscr{F}_s$-measurable variable ξ_s such that

$$\Big(u(s,\xi_s) - \hat{J}_l(s,\xi_s) - \varepsilon h(\xi_s - \bar{x})\Big)^+ = \max_{x\in\mathbb{R}^d}\Big(u(s,x) - \hat{J}_l(s,x) - \varepsilon h(x-\bar{x})\Big)^+. \tag{4.8}$$

Set

$$\begin{aligned}Y_s &= (u(s,\xi_s) - \hat{J}_l(s,\xi_s) - \varepsilon h(\xi_s - \bar{x}))^+ + \frac{\alpha(s-t)}{2(T-t)};\\ Z_s &= \operatorname*{esssup}_{\tau\in\mathcal{T}^s} E_{\mathscr{F}_s}[Y_\tau],\end{aligned}$$

where we recall that $\mathcal{T}^s$ denotes the set of stopping times valued in $[s,T]$. As $u, \hat{J}_l \in \mathcal{S}^2(C(\mathbb{R}^d))$, it follows obviously the time-continuity of

$$\max_{x\in\mathbb{R}^d}\Big(u(s,x) - \hat{J}_l(s,x) - \varepsilon h(x-\bar{x})\Big)^+$$

and thus that of $\Big(u(s,\xi_s) - \hat{J}_l(s,\xi_s) - \varepsilon h(\xi_s - \bar{x})\Big)^+$. Therefore, the process $(Y_s)_{t\leq s\leq T}$ has continuous trajectories. Define $\tau = \inf\{s \geq t : Y_s = Z_s\}$. In view of the optimal stopping theory, observe that

$$E_{\mathscr{F}_t} Y_T = \frac{\alpha}{2} < \alpha = Y_t \leq Z_t = E_{\mathscr{F}_t} Y_\tau = E_{\mathscr{F}_t} Z_\tau.$$

It follows that $\mathbb{P}(\tau < T) > 0$. As

$$(u(\tau,\xi_\tau) - \hat{J}_l(\tau,\xi_\tau) - \varepsilon h(\xi_\tau - \bar{x}))^+ + \frac{\alpha(\tau-t)}{2(T-t)} = Z_\tau \geq E_{\mathscr{F}_\tau}[Y_T] = \frac{\alpha}{2},$$

we have

$$\mathbb{P}((u(\tau,\xi_\tau) - \hat{J}_l(\tau,\xi_\tau) - \varepsilon h(\xi_\tau - \bar{x}))^+ > 0) > 0. \tag{4.9}$$

Define

$$\hat{\tau} = \inf\{s \geq \tau : (u(s,\xi_s) - \hat{J}_l(s,\xi_s) - \varepsilon h(\xi_s - \bar{x}))^+ \leq 0\}.$$

Obviously, $\tau \leq \hat{\tau} \leq T$. Put $\Omega_\tau = \{\tau < \hat{\tau}\}$. Then $\Omega_\tau \in \mathscr{F}_\tau$ and in view of relation (4.9), and the definition of $\hat{\tau}$, we have $\mathbb{P}(\Omega_\tau) > 0$.

Set

$$\phi(s,x) = \hat{J}_l(s,x) + \varepsilon h(x-\bar{x}) - \frac{\alpha(s-t)}{2(T-t)} + E_{\mathscr{F}_s} Y_\tau.$$

Then $\phi \in \mathscr{C}^1_{\mathscr{F}}$ because $\hat{J}_l \in \mathscr{C}^1_{\mathscr{F}}$. For each $\bar{\tau} \in \mathcal{T}^\tau$, [1] we have for almost all $\omega \in \Omega_\tau$,

$$(\phi - u)(\tau, \xi_\tau) = 0 = Y_\tau - Z_\tau \le Y_\tau - E_{\mathscr{F}_\tau}[Y_{\bar{\tau}\wedge\hat{\tau}}] = E_{\mathscr{F}_\tau}\left[\inf_{y\in\mathbb{R}^d}(\phi - u)(\bar{\tau}\wedge\hat{\tau}, y)\right],$$

which together with the arbitrariness of $\bar{\tau}$ implies that $\phi \in \underline{\mathcal{G}}u(\tau, \xi_\tau; \Omega_\tau)$. As u is a viscosity subsolution, by Lemma 4.2 it holds that for almost all $\omega \in \Omega_\tau$,

$$\begin{aligned}
0 &\ge \operatorname{ess}\liminf_{(s,x)\to(\tau^+,\xi_\tau)} E_{\mathscr{F}_\tau}\left\{-\mathfrak{d}_s\phi(s,x) - \mathbb{H}(s,x,D\phi(s,x))\right\}\\
&= \frac{\alpha}{2(T-t)} + \operatorname{ess}\liminf_{(s,x)\to(\tau^+,\xi_\tau)} E_{\mathscr{F}_\tau}\left\{-\mathfrak{d}_s\hat{J}_l(s,x) - \mathbb{H}(s,x,D\hat{J}_l(s,x) + \varepsilon Dh(x-\bar{x}))\right\}\\
&\ge \frac{\kappa}{2(T-t)} + \operatorname{ess}\liminf_{(s,x)\to(\tau^+,\xi_\tau)} E_{\mathscr{F}_\tau}\left\{-\mathfrak{d}_s\hat{J}_l(s,x) - \mathbb{H}(s,x,D\hat{J}_l(s,x))\right\}\\
&\quad - \varepsilon E_{\mathscr{F}_\tau}\left[\sup_{(s,x,v)\in[0,T]\times\mathbb{R}^d\times v}|\beta_l(s,x,v)|\cdot|Dh(x-\bar{x})|\right]\\
&\ge \frac{\kappa}{2(T-t)} - \varepsilon L.
\end{aligned}$$

This is an obvious contradiction as ε is sufficiently small. □

Remark 4.1 Compared with the proof of [28, Theorem 5.2], we added two new techniques in the above proof: (i) because of the lack of regularity of coefficients, we construct sequences $\{\hat{J}_l\}$ in $\mathscr{C}^1_{\mathscr{F}}$ to approximate $J(\cdot,\cdot;\theta)$; (ii) lack of spatial integrability of V (or V possibly being nonzero at infinity) motivates us to introduce a penalty function h in the proof to ensure the existence of maximums (for instance, in (4.8)).

Throughout the proof of Theorem 4.3, we see that only the viscosity subsolution property of u and the property (4.6) of $\hat{J}_l \in \mathscr{C}^1_{\mathscr{F}}$ are used. Hence, omitting the proofs, we have the following weak version of comparison principle.

Corollary 4.4 Let $(\mathcal{A}1)$ hold and u be a viscosity subsolution (resp. supersolution) of BSPDE (1.1) and $\phi \in \mathscr{C}^1_{\mathscr{F}}$, $\phi(T,x) \ge$ (resp. $\le$)$G(x)$ for all $x \in \mathbb{R}^d$ a.s. and with probability 1

$$\operatorname{ess}\liminf_{(s,x)\to(t^+,y)} E_{\mathscr{F}_t}\left\{-\mathfrak{d}_s\phi(s,x) - \mathbb{H}(s,x,D\phi(s,x))\right\} \ge 0$$

$$(\text{resp. } \operatorname{ess}\limsup_{(s,x)\to(t^+,y)} E_{\mathscr{F}_t}\left\{-\mathfrak{d}_s\phi(s,x) - \mathbb{H}(s,x,D\phi(s,x))\right\} \le 0)$$

for all $(t,y) \in [0,T)\times\mathbb{R}^d$. It holds a.s. that $u(t,x) \le$ (resp., $\ge$) $\phi(t,x)$, $\forall\,(t,x) \in [0,T]\times\mathbb{R}^d$.

4.2 Uniqueness of viscosity solution

We shall prove the uniqueness on the basis of the established comparison results. First, we approximate the coefficients β, f, and G via regular functions.

Lemma 4.5 Let $(\mathcal{A}1)$ hold. 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

$$\begin{aligned}
(G^N, f^N, \beta^N) \in C^3(\mathbb{R}^{m_0\times N}\times\mathbb{R}^d) &\times C(U; C^3([0,T]\times\mathbb{R}^{m_0\times N}\times\mathbb{R}^d))\\
&\times C(U; C^3([0,T]\times\mathbb{R}^{m_0\times N}\times\mathbb{R}^d))
\end{aligned}$$

such that

$$G^\varepsilon := \operatorname{esssup}_{x\in\mathbb{R}^d}\left|G^N(W_{t_1},\cdots,W_{t_N},x) - G(x)\right|,$$

[1] Recall that $\mathcal{T}^\tau$ denotes the set of stopping times ζ satifying $\tau \le \zeta \le T$ as defined in Section 2.2.

$$f_t^\varepsilon := \operatorname*{esssup}_{(x,v)\in\mathbb{R}^d\times U} \left|f^N(W_{t_1\wedge t},\cdots,W_{t_N\wedge t},t,x,v)-f(t,x,v)\right|, \quad \text{for } t\in[0,T],$$

$$\beta_t^\varepsilon := \operatorname*{esssup}_{(x,v)\in\mathbb{R}^d\times U} \left|\beta^N(W_{t_1\wedge t},\cdots,W_{t_N\wedge t},t,x,v)-\beta(t,x,v)\right|, \quad \text{for } t\in[0,T],$$

are $\mathscr{F}_t$-adapted with

$$\|G^\varepsilon\|_{L^2(\Omega,\mathscr{F}_T;\mathbb{R})}+\|f^\varepsilon\|_{\mathcal{L}^2(\mathbb{R})}+\|\beta^\varepsilon\|_{\mathcal{L}^2(\mathbb{R}^d)}<\varepsilon,$$

and G^N, f^N, and β^N are uniformly Lipschitz-continuous in the space variable x with an identical Lipschitz-constant L_c independent of N and ε.

Although the proof of Lemma 4.5 is an application of standard density arguments, we would sketch the proof for the readers who are interested.

Sketched Proof of Lemma 4.5 Consider the approximations of f. First, in a similar way to [11, case (c) in the proof of Proposition 2.2, Page 29], the dominated convergence theorem indicates that f may be approximated in $\mathcal{L}^2(C(U\times\mathbb{R}^d))$ by random fields of the form:

$$\bar{f}^l(\omega,t,x,v)=\phi_1(\omega,x,v)1_{[0,t_1]}(t)+\sum_{j=2}^{l}\phi_j(\omega,x,v)1_{(t_{j-1},t_j]}(t),$$

where $0=t_0<t_1<\cdots<t_l<T$, and for $j=1,\cdots,l$, $\phi_j=f(t_{j-1},\cdot,\cdot,\cdot)\in L^2(\Omega,\mathscr{F}_{t_{j-1}};C(U\times\mathbb{R}^d))$. In fact, with the identity approximations as in (4.2), we may take instead

$$\phi_j\in L^2(\Omega,\mathscr{F}_{t_{j-1}};C(U,C^\infty(\mathbb{R}^d)),\quad j=1,\cdots,l.$$

Furthermore, for each $j\geq 2$, ϕ_j may be approximated monotonically (see [7, Lemma 1.2, Page 16] for instance) by simple random variables of the following form:

$$\sum_{i=1}^{l_j}1_{A_i^j}(\omega)h_i^j(x,v),\quad \text{with } h_i^j\in C(U;C^\infty(\mathbb{R}^d)),\quad A_i^j\in\mathscr{F}_{t_{j-1}},\quad i=1,\cdots,l_j,$$

and [22, Lemma 4.3.1., page 50] implies that each $1_{A_i^j}$ may be approximated in $L^2(\Omega,\mathscr{F}_{t_{j-1}})$ by functions in the following set

$$\{g(W_{\tilde{t}_1},\cdots,W_{\tilde{t}_{l_i^j}}):\ \tilde{t}_r\in[0,t_{j-1}],\ g\in C_0^\infty(\mathbb{R}^{l_i^j})\}.$$

In addition, each $1_{(t_{j-1},t_j]}$ may be increasingly approximated by compactly-supported nonnegative functions $\varphi_j\in C^\infty((t_{j-1},T];\mathbb{R})$. To sum up, f may be approximated in $\mathcal{L}^2(C(U\times\mathbb{R}^d))$ by the following random fields:

$$f^N(W_{\bar{t}_1\wedge t},\cdots,W_{\bar{t}_N\wedge t},t,x,v)=\sum_{j=1}^{l}\sum_{i=1}^{l_j}g_i^j(W_{\bar{t}_1\wedge t_{j-1}},\cdots,W_{\bar{t}_N\wedge t_{j-1}})h_i^j(x,v)\varphi_j(t),$$

where $0=\bar{t}_0<\bar{t}_1<\cdots<\bar{t}_{N-1}<\bar{t}_N=T$, and g_i^j, h_i^j, and φ_j are smooth functions. The required approximations for G and β follow similarly. □

We are now ready to present the proof for the uniqueness of viscosity solution.

Proof of Theorem 2.2

Define

$$\overline{\mathscr{V}}=\Big\{\phi\in\mathscr{C}_{\mathscr{F}}^1:\ \phi(T,x)\geq G(x)\ \forall x\in\mathbb{R}^d,\ \text{a.s., and with probability 1},$$

$$\operatorname{ess}\liminf_{(s,x)\to(t^+,y)} E_{\mathscr{F}_t}\left[-\mathfrak{d}_s\phi(s,x)-\mathbb{H}(s,x,D\phi(s,x))\right]\geq 0,\quad \forall (t,y)\in[0,T)\times\mathbb{R}^d\Big\},$$

$$\underline{\mathscr{V}}=\Big\{\phi\in\mathscr{C}^1_{\mathscr{F}}:\phi(T,x)\leq G(x)\ \ \forall x\in\mathbb{R}^d,\ \text{a.s., and with probability 1,}$$

$$\operatorname{ess}\limsup_{(s,x)\to(t^+,y)} E_{\mathscr{F}_t}\left[-\mathfrak{d}_s\phi(s,x)-\mathbb{H}(s,x,D\phi(s,x))\right]\leq 0,\quad \forall (t,y)\in[0,T)\times\mathbb{R}^d\Big\},$$

and set

$$\overline{u}=\operatorname*{essinf}_{\phi\in\overline{\mathscr{V}}}\phi,\quad \underline{u}=\operatorname*{esssup}_{\phi\in\underline{\mathscr{V}}}\phi.$$

In view of Corollary 4.4, for any viscosity solution $u\in\mathcal{S}^2(C(\mathbb{R}^d))$, we have $\underline{u}\leq u\leq\overline{u}$. Therefore, for the uniqueness of viscosity solution, it is sufficient to check $\underline{u}=V=\overline{u}$.

Let $(\Omega',\mathscr{F}',\{\mathscr{F}'_t\}_{t\geq 0},\mathbb{P}')$ be another complete filtered probability space which carries a d-dimensional standard Brownian motion $B=\{B_t:t\geq 0\}$ with $\{\mathscr{F}'_t\}_{t\geq 0}$ generated by B and augmented by all the $\mathbb{P}'$-null sets in $\mathscr{F}'$. Set

$$(\bar{\Omega},\bar{\mathscr{F}},\{\bar{\mathscr{F}}_t\}_{t\geq 0},\bar{\mathbb{P}})=(\Omega\times\Omega',\mathscr{F}\otimes\mathscr{F}',\{\mathscr{F}_t\otimes\mathscr{F}'_t\}_{t\geq 0},\mathbb{P}\otimes\mathbb{P}').$$

Then B and W are independent on $(\bar{\Omega},\bar{\mathscr{F}},\{\bar{\mathscr{F}}_t\}_{t\geq 0},\bar{\mathbb{P}})$ and it is easy to see that all the theory established in previous sections still hold on the enlarged probability space.

For each fixed $\varepsilon\in(0,1)$, choose $(G^\varepsilon,\ f^\varepsilon,\ \beta^\varepsilon)$ and (G^N,f^N,β^N) as in Lemma 4.5. Recalling the standard theory of backward SDEs (see [1] for instance), let the pairs $(Y^\varepsilon,Z^\varepsilon)\in\mathcal{S}^2_{\mathscr{F}}(\mathbb{R})\times\mathcal{L}^2_{\mathscr{F}}(\mathbb{R}^m)$ and $(y_t,z_t)\in\mathcal{S}^2_{\mathscr{F}'}(\mathbb{R})\times\mathcal{L}^2_{\mathscr{F}'}(\mathbb{R}^d)$ be the solutions of backward SDEs

$$Y^\varepsilon_s=G^\varepsilon+\int_s^T(f^\varepsilon_t+K\beta^\varepsilon_t)\,\mathrm{d}t-\int_s^T Z^\varepsilon_s\,\mathrm{d}W_s,$$

and

$$y_s=|B_T|+\int_s^T|B_r|\mathrm{d}r-\int_s^T z_r\,\mathrm{d}B_r,$$

respectively, and for each $(s,x)\in[0,T)\times\mathbb{R}^d$, set

$$\begin{aligned}V^\varepsilon(s,x)=\operatorname*{essinf}_{\theta\in\mathcal{U}}E_{\bar{\mathscr{F}}_s}\Bigg[&\int_s^T f^N\left(W_{t_1\wedge t},\cdots,W_{t_N\wedge t},t,X^{s,x;\theta,N}_t,\theta_t\right)\mathrm{d}t\\&+G^N\left(W_{t_1},\cdots,W_{t_N},X^{s,x;\theta,N}_T\right)\Bigg],\end{aligned}$$

where the constant $K\geq 0$ is to be determined later and $X^{s,x;\theta,N}_t$ satisfies the SDE

$$\begin{cases}\mathrm{d}X_t=\beta^N(t,X_t,\theta_t)\mathrm{d}t+\delta_N\mathrm{d}B_t,\ \ t\in[s,T];\\ X_s=x\end{cases}$$

with $\delta_N>0$ being a constant.

By the viscosity solution theory of fully nonlinear parabolic PDEs (see [19, Theorems I.1 and II.1] for instance), when $s\in[t_{N-1},T)$, we have

$$V^\varepsilon(s,x)=\tilde{V}^\varepsilon(s,x,W_{t_1},\cdots,W_{t_{N-1}},W_s)$$

with

$$\tilde{V}^\varepsilon(s,x,W_{t_1},\cdots,W_{t_{N-1}},y)$$

$$= \operatorname*{essinf}_{\theta\in\mathcal{U}} E_{\bar{\mathscr{F}}_s, W_s=y}\Bigg[\int_s^T f^N\left(W_{t_1},\cdots,W_{t_{N-1}},W_{t_N\wedge t},t,X_t^{s,x;\theta,N},\theta_t\right)\mathrm{d}t$$
$$+ G^N\left(W_{t_1},\cdots,W_{t_N},X_T^{s,x;\theta,N}\right)\Bigg]$$

satisfying the HJB equation of the following form

$$\begin{cases} -D_t u(t,x,y) = \dfrac{1}{2}\mathrm{tr}\left(D_{yy}u(t,x,y)\right) + \dfrac{\delta_N^2}{2}\mathrm{tr}\left(D_{xx}u(t,x,y)\right) \\ \qquad\qquad + \operatorname*{essinf}_{v\in U}\Big\{(\beta^N)'(W_{t_1},\cdots,W_{t_{N-1}},y,t,x,v)D_x u(t,x,y) \\ \qquad\qquad + f^N(W_{t_1},\cdots,W_{t_{N-1}},y,t,x,v)\Big\},\ (t,x,y)\in[t_{N-1},T)\times\mathbb{R}^d\times\mathbb{R}^{m_0}; \\ u(T,x,y) = G^N(W_{t_1},\cdots,W_{t_{N-1}},y,x), \quad (x,y)\in\mathbb{R}^d\times\mathbb{R}^m, \end{cases} \tag{4.10}$$

and thus the regularity theory of viscosity solutions (see [15, Theorem 1.1] or [16, Chapter 6] for instance[2]) gives

$$\tilde{V}^\varepsilon(\cdot,\cdot,W_{t_1},\cdots,W_{t_{N-1}},\cdot)\in L^\infty\left(\Omega,\mathscr{F}_{t_{N-1}};C^{1+\frac{\bar{\alpha}}{2},2+\bar{\alpha}}([t_{N-1},T]\times\mathbb{R}^d)\right),$$

for some $\bar{\alpha}\in(0,1)$, where the time-space Hölder space $C^{1+\frac{\bar{\alpha}}{2},2+\bar{\alpha}}([t_{N-1},T]\times\mathbb{R}^d)$ is defined as usual. We can make similar arguments on time interval $[t_{N-2},t_{N-1})$ taking the obtained $V^\varepsilon(t_{N-1},x)$ as the terminal value, and recursively on intervals $[t_{N-3},t_{N-2})$, $\cdots$, $[0,t_1)$. Furthermore, applying the Itô-Kunita formula to $\tilde{V}^\varepsilon(s,x,\tilde{W}_{t_1},\cdots,\tilde{W}_{t_{N-1}},y)$ on $[t_{N-1},T]$ yields that

$$\begin{cases} \quad -\mathrm{d}V^\varepsilon(t,x-\delta_N B_t) \\ = \operatorname*{essinf}_{v\in U}\Big\{(\beta^N)'(W_{t_1},\cdots,W_{t_{N-1}},W_t,t,x-\delta_N B_t,v)D_x V^\varepsilon(t,x-\delta_N B_t) \\ \quad + f^N(W_{t_1},\cdots,W_{t_{N-1}},W_t,t,x-\delta_N B_t,v)\Big\}\mathrm{d}t \\ \quad - D_y\tilde{V}^\varepsilon(s,x,\tilde{W}_{t_1},\cdots,\tilde{W}_{t_{N-1}},W_t)\,\mathrm{d}W_t + \delta_N D_x V^\varepsilon(t,x-\delta_N B_t)\,\mathrm{d}B_t, \\ \quad (t,x)\in[t_{N-1},T)\times\mathbb{R}^d; \\ V^\varepsilon(T,x) = G^N(W_{t_1},\cdots,W_{t_{N-1}},W_T,x-\delta_N B_T), \quad x\in\mathbb{R}^d. \end{cases} \tag{4.11}$$

It follows similarly on intervals $[t_{N-2},t_{N-1})$, $\cdots$, $[0,t_1)$, and finally we have $V^\varepsilon(\cdot,\cdot-\delta_N B_\cdot)\in\mathscr{C}^1_{\bar{\mathscr{F}}}$.

In view of the approximation in Lemma 4.5 and with an analogy to the proof of (iv) in Proposition 3.2, there exists $\tilde{L}>0$ such that

$$\max_{(t,x)\in[0,T]\times\mathbb{R}^d}\{|DV^\varepsilon(t,x)|\}\le\tilde{L},\ \text{ a.s.}$$

with $\tilde{L}$ being independent of ε and N. Set $K=\tilde{L}$ and

$$\overline{V}^\varepsilon(s,x) = V^\varepsilon(s,x-\delta_N B_s) + Y_s^\varepsilon + \delta_N\bar{K}y_t,$$

[2]As $U\subset\mathbb{R}^n$ is a nonempty compact set, it has a denumerable subset $\mathcal{K}\subset U$ that is dense in U, and by the continuity of the coefficients, the essential infimum may be taken over $\mathcal{K}$. This together with some basic properties of viscosity solutions (see [32, Proposition 3.7] for instance) allows [15, Theorem 1.1] to be applied straightforwardly.

$$\underline{V}^\varepsilon(s,x) = V^\varepsilon(s, x-\delta_N B_s) - Y_s^\varepsilon - \delta_N \bar{K} y_t,$$

with $\bar{K} = 4L(\tilde{L}+1)$ and L the constant in $(\mathcal{A}1)$.

Notice that

$$|\beta(t,x,v) - \beta(t, x-\delta_N B_t, v)| + |f(t,x,v) - f(t, x-\delta_N B_t, v)| \le 2\delta_N L|B_t|,$$
$$|G(x) - G(x-\delta_N B_T)| \le \delta_N L|B_T|.$$

Then for $\overline{V}^\varepsilon$ on $[t_{N-1}, T)$, omitting the inputs for some involved functions, we have

$$\begin{aligned}
-\mathfrak{d}_t\overline{V}^\varepsilon - \mathbb{H}(D\overline{V}^\varepsilon) &= -\mathfrak{d}_t\overline{V}^\varepsilon - \operatorname*{essinf}_{v\in U}\Big\{(\beta^N)'D\overline{V}^\varepsilon + f^N + f^\varepsilon + \tilde{L}\beta^\varepsilon + \delta_N\bar{K}|B_t| \\
&\qquad + \left(\beta - \beta^N\right)' D\overline{V}^\varepsilon - \beta^\varepsilon\tilde{L} + f - f^N - f^\varepsilon - \delta_N\bar{K}|B_t|\Big\} \\
&\ge -\mathfrak{d}_t\overline{V}^\varepsilon - \operatorname*{essinf}_{v\in U}\Big\{(\beta^N)'D\overline{V}^\varepsilon + f^N + f^\varepsilon + \beta^\varepsilon\tilde{L} + \delta_N\bar{K}|B_t|\Big\} \\
&= 0,
\end{aligned} \tag{4.12}$$

and it follows similarly on intervals $[t_{N-2}, t_{N-1})$, $\cdots$, $[0, t_1)$ that

$$-\mathfrak{d}_t\overline{V}^\varepsilon - \mathbb{H}(D\overline{V}^\varepsilon) \ge 0,$$

which together with the obvious relation $\overline{V}^\varepsilon(T) = G^\varepsilon + G^N + \delta\bar{K}|B_T| \ge G$ indicates that $\overline{V}^\varepsilon \in \overline{\mathscr{V}}$. Analogously, $\underline{V}^\varepsilon \in \underline{\mathscr{V}}$.

Now, let us measure the distance between $\underline{V}^\varepsilon$, $\overline{V}^\varepsilon$, and V. By the estimates for solutions of backward SDEs (see [1, Proposition 3.2] for instance), we first have

$$\begin{aligned}
\|Y^\varepsilon\|_{\mathcal{S}^2(\mathbb{R})} + \|Z^\varepsilon\|_{\mathcal{L}^2(\mathbb{R}^{m_0})} &\le C\left(\|G^\varepsilon\|_{L^2(\Omega,\mathscr{F}_T;\mathbb{R})} + \|f^\varepsilon + \tilde{L}\beta^\varepsilon\|_{\mathcal{L}^2(\mathbb{R})}\right) \\
&\le C(1+\tilde{L})\varepsilon
\end{aligned}$$

with the constant C independent of N and ε. Fix some $(s,x) \in [0,T)\times\mathbb{R}^d$. In view of the approximation in Lemma 4.5, using Itô's formula, Burkholder-Davis-Gundy's inequality, and Gronwall's inequality, we have through standard computations, for any $\theta \in \mathcal{U}$,

$$\begin{aligned}
&E_{\mathscr{F}_s}\left[\sup_{s\le t\le T}\left|X_t^{s,x;\theta,N} - X_t^{s,x;\theta}\right|^2\right] \\
&\le \tilde{K}\left(\delta_N^2 + E_{\mathscr{F}_s}\int_s^T\left|\beta^N\left(\tilde{W}_{t_1\wedge t},\cdots,\tilde{W}_{t_N\wedge t}, t, X_t^{s,x;\theta,N},\theta_t\right) - \beta\left(t, X_t^{s,x;\theta,N},\theta_t\right)\right|^2 \mathrm{d}t\right) \\
&\le \tilde{K}\left(\delta_N^2 + E_{\mathscr{F}_s}\int_s^T|\beta_t^\varepsilon|^2\,\mathrm{d}t\right),
\end{aligned}$$

with $\tilde{K}$ being independent of s, x, N, ε, and θ. Then,

$$\begin{aligned}
&E\left|V^\varepsilon(s,x) - V(s,x)\right| \\
&\le E\operatorname*{esssup}_{\theta\in\mathcal{U}} E_{\mathscr{F}_s}\Bigg[\int_s^T\left(f_t^\varepsilon + \left|f\left(t, X_t^{s,x;\theta,N},\theta_t\right) - f\left(t, X_t^{s,x;\theta},\theta_t\right)\right|\right)\mathrm{d}t \\
&\quad + G^\varepsilon + \left|G\left(X_T^{s,x;\theta,N}\right) - G\left(X_T^{s,x;\theta}\right)\right|\Bigg] \\
&\le E|Y_s^\varepsilon| + 2L(T^{1/2}+1)(\tilde{K}+1)\left(\delta_N + E\operatorname*{esssup}_{\theta\in\mathcal{U}}\left(E_{\mathscr{F}_s}\int_s^T|\beta_t^\varepsilon|^2\,\mathrm{d}t\right)^{1/2}\right)
\end{aligned}$$

$$\leq \|Y^\varepsilon\|_{\mathcal{S}^2(\mathbb{R})} + 2L(\tilde{K}+1)(T^{1/2}+1)\left(\delta_N + \|\beta^\varepsilon\|_{\mathcal{L}^2(\mathbb{R}^d)}\right)$$
$$\leq K_0(\varepsilon + \delta_N),$$

with the constant K_0 being independent of N, ε, and (s,x). Furthermore, in view of the definitions of $\overline{V}^\varepsilon$ and $\underline{V}^\varepsilon$, there exists some constant K_1 independent of ε and N such that

$$E\left|\overline{V}^\varepsilon(s,x) - V(s,x)\right| + E\left|\underline{V}^\varepsilon(s,x) - V(s,x)\right| \leq K_1(\varepsilon+\delta_N), \quad \forall\,(s,x)\in[0,T]\times\mathbb{R}^d.$$

The arbitrariness of (ε, δ_N) together with the relation $\overline{V}^\varepsilon \geq V \geq \underline{V}^\varepsilon$ finally implies that $\underline{u} = V = \overline{u}$. □

Remark 4.2 In the above proof, by enlarging the original filtered probability space with an independent Brownian motion B, we have actually constructed the regular approximations of V with a regular perturbation induced by εB, which does not necessitate the superparabolicity assumed in [28]. This method may help to address the uniqueness of viscosity solution for certain classes of fully nonlinear degenerate stochastic HJB equations. Nevertheless, we would not do such a generalization in order to focus on the study of stochastic HJ equations in this work.

References

[1] Briand P, Delyon B, Hu Y, Pardoux E, Stoica L. L^p solutions of backward stochastic differential equations. Stoch Process Appl, 2003, **108**(4): 604–618

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

[3] Buckdahn R, Ma J. Pathwise stochastic control problems and stochastic HJB equations. SIAM J Control Optim, 2007, **45**(6): 2224–2256

[4] Cont R, Fournié D -A, et al. Functional itô calculus and stochastic integral representation of martingales. The Annals of Probability, 2013, **41**(1): 109–133

[5] Crandall M G, Ishii H, Lions P -L. Users guide to viscosity solutions of second order partial differential equations. Bull Amer Math Soc, 1992, **27**(1): 1–67

[6] Crandall M G, Kocan M, Świech A. L^p-theory for fully nonlinear uniformly parabolic equations. Commun Partial Differ Equ, 2000, **25**(11/12): 1997–2053

[7] Da Prato G, Zabczyk J. Stochastic equations in infinite dimensions. Cambridge University Press, 2014

[8] Du K, Qiu J, Tang S. L^p theory for super-parabolic backward stochastic partial differential equations in the whole space. Appl Math Optim, 2011, **65**(2): 175–219

[9] Ekren I, Keller C, Touzi N, Zhang J. On viscosity solutions of path dependent PDEs. Ann Probab, 2014, **42**(1): 204–236

[10] Ekren I, Touzi N, Zhang J. Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part I. Ann Probab, 2016, **44**(2): 1212–1253

[11] Gawarecki L, Mandrekar V. Stochastic differential equations in infinite dimensions: with applications to stochastic partial differential equations. Springer Science & Business Media, 2010

[12] Horst U, Qiu J, Zhang Q. A constrained control problem with degenerate coefficients and degenerate backward SPDEs with singular terminal condition. SIAM J Control Optim, 2016, **54**(2): 946–963

[13] Hu Y, Ma J, Yong J. On semi-linear degenerate backward stochastic partial differential equations. Probab Theory Relat Fields, 2002, **123**: 381–411

[14] Juutinen P. On the definition of viscosity solutions for parabolic equations. Proceedings of the American Mathematical Society, 2001, **129**(10): 2907–2911

[15] Krylov N V. Boundedly nonhomogeneous elliptic and parabolic equations. Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya, 1982, **46**(3): 487–523

[16] Krylov N V. Nonlinear Elliptic and Parabolic Equations of the Second Order. Dordrecht: D Reidel, 1987

[17] Leão D, Ohashi A, Simas A. A weak version of path-dependent functional Itô calculus. Ann Probab, 2018, **46**(6): 3399–3441

[18] Lions P, Souganidis P. Fully nonlinear stochastic partial differential equations: Non-smooth equations and applications. C R Acad Sci paris, 1998, **327**(1): 735–741

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

[20] Lukoyanov N Y. On viscosity solution of functional hamilton-jacobi type equations for hereditary systems. Proceedings of the Steklov Institute of Mathematics, 2007, **259**(2): S190–S200

[21] Ma J, Yin H, Zhang J. On non-Markovian forward–backward SDEs and backward stochastic PDEs. Stoch Process Appl, 2012, **122**(12): 3980–4004

[22] Øksendal B. Stochastic differential equations. Springer, 2003

[23] Pardoux E. Stochastic partial differential equations and filtering of diffusion processes. Stoch, 1979: 127–167

[24] Peng S. Stochastic Hamilton-Jacobi-Bellman equations. SIAM J Control Optim, 1992, **30**: 284–304

[25] Peng S. Backward stochastic differential equation, nonlinear expectation and their applications. Proceedings of the International Congress of Mathematicians, 2010: 393–432

[26] Peng S. Note on viscosity solution of path-dependent PDE and G-martingales. arXiv:1106.1144, 2011

[27] Qiu J. Weak solution for a class of fully nonlinear stochastic hamilton–jacobi–bellman equations. Stoch Process Appl, 2017, **127**(6): 1926–1959

[28] Qiu J. Viscosity solutions of stochastic Hamilton–Jacobi–Bellman equations. SIAM J Control Optim, 2018, **56**(5): 3708–3730

[29] Qiu J, Tang S. Maximum principles for backward stochastic partial differential equations. J Funct Anal, 2012, **262**: 2436–2480

[30] Tang S. Semi-linear systems of backward stochastic partial differential equations in $\mathbb{R}^n$. Chin Ann Math, 2005, **26B**(3): 437–456

[31] Tang S, Wei W. On the cauchy problem for backward stochastic partial differential equations in Hölder spaces. Ann Probab, 2016, **44**(1): 360–398

[32] Wang L. On the regularity theory of fully nonlinear parabolic equations: I. Commun Pure Appl Math, 1992, **45**(1): 27–76