# REFLECTED FORWARD-BACKWARD SDEs AND OBSTACLE PROBLEMS WITH BOUNDARY CONDITIONS

JIN MA[1]
*Purdue University, Department of Mathematics*
*West Lafayette, IN 47907-1395 USA*
*E-mail: majin@math.purdue.edu*

JAKŠA CVITANIĆ[2]
*Columbia University, Department of Statistics*
*New York, NY 10027 USA*
*E-mail: cj@stat.columbia.edu*

In this paper we study a class of forward-backward stochastic differential equations with reflecting boundary conditions (FBSDER for short). More precisely, we consider the case in which the forward component of the FBSDER is restricted to a fixed, convex region, and the backward component will stay, at each fixed time, in a convex region that may depend on time and is possibly random. The solvability of such FBSDER is studied in a fairly general way. We also prove that if the coefficients are all deterministic and the backward equation is one-dimensional, then the adapted solution of such FBSDER will give the viscosity solution of a quasilinear variational inequality (obstacle problem) with a Neumann boundary condition. As an application, we study how the solvability of FBSDERs is related to the solvability of an *American game option*.

**Key words:** Forward-Backward Stochastic Differential Equations, Quasilinear Variational Inequalities, Viscosity Solutions, American Game Option.

**AMS subject classifications:** 60H30, 93E05, 49J40.

## 1. Introduction

In this paper we are interested in the following system of stochastic differential equations:

$$
\begin{cases}
dX_t = b(t, X_t, Y_t, Z_t)dt + \sigma(t, X_t, Y_t, Z_t)dW_t + d\eta_t, \\
dY_t = h(s, X_t, Y_t Z_t)dt - Z_t dW_t + d\zeta_t, \quad t \in [0, T], \\
\qquad\qquad X_0 = x, \quad Y_T = g(X_T),
\end{cases}
\tag{1.1}
$$

where $\eta$ and $\zeta$ are two adapted processes with bounded variations, which are the *reflecting processes* that keep $X$ and $Y$ from leaving certain prescribed regions, respectively. Since the second equation is determined by its *terminal* value at $t = T$, it is known as a *backward stochastic differential equation*, for which the process $Z$ has to come into play. For this reason, we call system (1.1) a *forward-backward stochastic differential equation with reflections* (FBSDER for short) throughout this paper.

The amount of literature on the backward and forward-backward stochastic differential equations (BSDE and FBSDE) has exploded since the initial works of Pardoux and Peng [15] and Antonelli [1]. We refer the readers to Ma, Protter, and Yong [14], Hu and Peng [12], Yong [18] and Pardoux and Tang [17] for different techniques developed thus far for solving FBSDEs. Recently, efforts have also been made to investigate BSDEs with *state constraints*, among others are BSDEs with reflections. For references in this regard we refer to El Karoui et al. [8], Cvitanić and Karatzas [5] for one-dimensional BSDERs; and to Gegout-Petit and Pardoux [10] and Pardoux and Rascanu [16] for higher dimensional BSDERs.

The FBSDERs considered in this paper have the following features: the forward component of the solution will live in a fixed, convex domain; while the backward component lives in a convex domain that is allowed to move, even randomly. We note that although in this paper we contend ourselves only to the FBSDEs in which the backward equation is one-dimensional, our method is general enough to cover many higher dimensional cases, as long as both forward SDER and backward SDER are separately solvable. It should be pointed out here that in the reflected FBSDE case, the frequently used methods in FBSDE literature (e.g., the Four Step Scheme by Ma, Protter and Yong [14] and the method of continuation by Hu and Peng [12] and Yong [18] do not seem to apply. Therefore we will return to the most primitive technique: the method of *contraction mapping*, in the spirit of Pardoux and Tang [17] with some modifications. We should note, however, that in case the backward SDE is higher dimensional, the characterization of moving domains becomes more subtle, and extra conditions might be needed. We plan to address this issue in future work.

Another topic of this paper is to relate FBSDERs to a class of quasilinear variational inequalities. More precisely, in the case when all the coefficients are of Markovian type, we show that the adapted solution of an FBSDER gives a probabilistic representation of the *unique* viscosity solution to a variational inequality with Neumann boundary conditions. An interesting but not surprising application is the connection between the solvability of such FBSDERs and optimal stopping problems. In particular, combining the results of Cvitanić and Karatzas [5] on Dynkin Games and the solvability of FBSDERs derived in this paper we can study a class of option pricing problems in which the underlying market involves large investors as proposed by Cvitanić and Ma [6] and later by Buckdahn and Hu [3]. The option under study is of

American type, but, unlike the standard American option, both parties are allowed to choose the exercise time. We call it an *American game option*, and the minimum hedging price of such an option is found in terms of the adapted solution to a suitable FBSDER.

The paper is organized as follows. In Section 2 we formulate the problem and introduce all the notations. In Section 3 we state the main results. The American game option is presented in Section 4. Finally, Sections 5, 6 and 7 are devoted to the proofs of the results stated in Section 3.

## 2. Formulation of the Problem

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ be a complete filtered probability space on which is defined a $d$-dimensional standard Brownian motion $W = \{W_t : t \geq 0\}$. We assume that $\{\mathcal{F}_t\}_{t \geq 0}$ is the natural filtration generated by $W$, augmented by all the $P$-null sets. Thus, $\{\mathcal{F}_t\}_{t \geq 0}$ satisfies the *usual hypotheses*.

Throughout this paper we will make use of the following notations: Let $\mathbb{E}$ be any Euclidean space; denote

- for any sub-$\sigma$-field $\mathcal{G}$ of $\mathcal{F}$, $L^2_{\mathcal{G}}(\Omega; \mathbb{E})$ to be the set of all $\mathcal{G}$-measurable, $\mathbb{E}$-valued square-integrable random variables;
- $L^2_{\mathcal{F}}(0, T; \mathbb{E})$ to be the set of all $\mathbb{E}$-valued, $\{\mathcal{F}_t\}$-progressively measurable processes $X$, such that $E \int_0^T |X_t|^2 dt < \infty$;
- $L^2_{\mathcal{F}}(\Omega; C(0, T; \mathbb{E})) \subset L^2_{\mathcal{F}}(0, T; \mathbb{E})$ to be the set of all $\mathbb{E}$-valued, continuous, $\{\mathcal{F}_t\}$-progressively measurable processes $X$, such that $E \sup_{t \in [0, T]} |X_t|^2 < \infty$;
- $BV_{\mathcal{F}}(0, T; \mathbb{E})$ to be the set of all $\mathbb{E}$-valued, continuous, $\{\mathcal{F}_t\}$-progressively measurable processes $\xi$ whose paths are of bounded variation, and $\xi_0 = 0$.

Let $\mathcal{O}_1$ be a convex, closed set in $\mathbb{R}^n$, with boundary $\partial \mathcal{O}_1$. Define for any $x \in \partial \mathcal{O}_1$, the set of inward normals to $\mathcal{O}_1$ at $x$ by

$$\mathcal{N}_x(\mathcal{O}_1) \triangleq \{\gamma : |\gamma| = 1, \text{ and } \langle \gamma, x - y \rangle \leq 0, \ \forall y \in \mathcal{O}_1\}. \tag{2.1}$$

Next, consider a *moving interval* $\mathcal{O}_2(\omega, t, x) = [L(\omega, t, x), U(\omega, t, x)]$, where both $L$ and $U$ are continuous random fields such that $L(\omega, t, x) \leq U(\omega, t, x)$, $\forall (\omega, t, x) \in \Omega \times [0, T] \times \mathbb{R}^n$. We say that $L$ (or $U$) is an *Itô-process with parameter $x$* if

$$L(t, x) \text{ (or } U(t, x)) = \int_0^t J(s, x) ds + \int_0^t \langle K(s, x), dW_s \rangle,$$

where $J$ and $K$ are measurable random fields, such that for fixed $x$, $J(\cdot, \cdot, x) \in L^2_{\mathcal{F}}(0, T; \mathbb{R})$, and $K(\cdot, \cdot, x) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^d)$, respectively. Clearly, if $L$ and $U$ are $C^2$ in the $x$ variable, then for any Itô process $\xi = \{\xi_t : t \in [0, T]\}$, the composition $L_t \triangleq L(t, \xi_t)$, $U_t \triangleq U(t, \xi_t)$, $t \in [0, T]$ are also Itô processes, thanks to the Itô-Ventzell formula. We shall call such $L$ (or $U$) an *Itô-process with parameter $x$ of class $C^2$*.

Now let $\mathcal{O}_1$ and $\mathcal{O}_2(\omega, t, x) = [L(\omega, t, x), U(\omega, t, x)]$, $(\omega, t, x) \in \Omega \times [0, T] \times \mathbb{R}^n$ be given. Consider the following FBSDER: for $t \in [0, T]$,

$$\begin{cases} X_t = x + \int\limits_0^t b(s,X_s,Y_s,Z_s)ds + \int\limits_0^t \sigma(s,X_s,Y_s,Z_s)dW_s + \eta_t; \\ Y_t = g(X_T) + \int\limits_t^T h(s,X_s,Y_s,Z_s)ds - \int\limits_t^T \langle Z_s, dW_s \rangle + \zeta_T - \zeta_t. \end{cases} \qquad (2.2)$$

In light of the structure of FBSDEs and SDEs with reflections, we now give the definition of the *adapted solution* to an FBSDER.

**Definition 2.1:** A quintuple of processes $(X, Y, Z, \eta, \zeta)$ is called *an adapted solution of FBSDER* (2.2) if

(1)  $(X, Y) \in L^2_{\mathcal{F}}(\Omega, C(0, T; \mathbb{R}^n \times \mathbb{R}))$, $\quad Z \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^d)$, $\quad (\eta, \zeta) \in BV_{\mathcal{F}}(0, T; \mathbb{R}^n \times \mathbb{R})$;

(2)  $X_t \in \mathcal{O}_1$, $Y_t \in \mathcal{O}_2(\cdot, t, X_t)$, $\forall t \in [0, T]$, $P$-a.s.;

(3)  $|\eta|_t = \int_0^t 1_{\partial \mathcal{O}_1}(X_s) d|\eta|_s$; $\eta_t = \int_0^t \gamma_s d|\eta|_s$, $\quad \forall t \in [0, T]$,-a.s., for some progressively measurable process $\gamma$ such that $\gamma_s \in \mathcal{N}_{X_s}(\mathcal{O}_1)$, $d|\eta|$-a.e.;

(4)  for any continuous and progressively measurable process $\rho$ such that $\rho_t \in \mathcal{O}_2(t, X_t, \cdot)$, $\forall t \in [0, T]$, $P$-a.e., one has $\langle Y_t - \rho_t, d\zeta_t \rangle \le 0$, $\forall t \in [0, T]$, $P$-a.e.;

(5)  $(X, Y, Z, \eta, \zeta)$ satisfies SDE (2.2) almost surely.

We remark that Definition 2.1 (3), combined with (2.1), implies that for any process $K = \{K_t : t \ge 0\}$ with RCLL paths such that $K_t \in \mathcal{O}_1$, $\forall t$, $P$-a.s.,

$$\int\limits_0^t \langle X_s - K_s, d\eta_s \rangle = \int\limits_0^t \langle X_s - K_s, \gamma(s) \rangle 1_{\partial \mathcal{O}}(X_s) d|\eta_s| \le 0, \qquad (2.3)$$

$$\forall t \in [0, T], \ P\text{-a.s.}$$

Throughout this paper we shall make use of the following *standing assumptions* on the domains $\mathcal{O}_1$, $\mathcal{O}_2$ and the coefficients of the FBSDER (2.2). we note that some of the conditions are made stronger than necessary for the sake of simpler presentation.

(A1)  (i)  $\mathcal{O}_1$ has a smooth boundary, and $0 \in \mathcal{O}_1$.

(ii)  $L$ and $U$ are both Itô processes with parameter $x$ of class $C^2$ such that for $P$-a.e. $\omega \in \Omega$, we have $0 \in (L(\omega, t, x), U(\omega, t, x))$, $\forall(t, x) \in [0, T) \times \mathbb{R}^n$, and $L(\omega, T, x) \le g(\omega, x) \le U(\omega, T, x)$, $\forall x$.

(iii)  The coefficients $b, h, \sigma$, and $g$ are random fields defined on $[0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^d$ such that for fixed $(x, y, z)$, the processes $b(\cdot, \cdot, x, y, z)$, $h(\cdot, \cdot, x, y, z)$, and $\sigma(\cdot, \cdot, x, y, z)$ are $\{\mathcal{F}_t\}$-progressively measurable, and $g(\cdot, x)$ is $\mathcal{F}_T$-measurable.

(iv)  For fixed $(t, x, z)$ and $P$-a.e. $\omega$, $h(\omega, t, x, \cdot, z)$ is continuous, and there exists a constant $K > 0$ such that $|h(\omega, t, x, y, z)| \le K(1 + |x| + |y|)$, for all $(\omega, t, x, y, z)$. Moreover,

$$E \int\limits_0^T |b(t, 0, 0, 0)|^2 dt + E \int\limits_0^T |\sigma(t, 0, 0, 0)|^2 dt + E|g(0)|^2 < \infty.$$

(v)  There exists constants $k_i \ge 0$, $i = 1, 2$ and $\gamma \in \mathbb{R}$ such that for all

$$(t, \omega) \in [0, T] \times \Omega \quad \text{and} \quad \mathbf{x} \stackrel{\Delta}{=} (x, y, z), \mathbf{x}_i \stackrel{\Delta}{=} (x_i, y_i, z_i) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^d,$$
$$i = 1, 2, \text{ and } \mathbf{x}^0 \stackrel{\Delta}{=} (x, y) \text{ for } \mathbf{x} = (x, y, z).$$

$$\begin{cases} |b(\omega, t, \mathbf{x}_1) - b(\omega, t, \mathbf{x}_2)| \leq K |\mathbf{x}_1 - \mathbf{x}_2|; \\ \langle h(\omega, t, x, y_1, z) - h(\omega, t, x, y_2, z), y_1 - y_2 \rangle \leq \gamma |y_1 - y_2|^2; \\ |h(\omega, t, x_1, y, z_1) - h(\omega, t, x_2, y, z_2)| < K(|x_1 - x_2| + |z_1 - z_2|); \qquad (2.4) \\ |\sigma(\omega, t, \mathbf{x}_1) - \sigma(\omega, t, \mathbf{x}_2)|^2 \leq K^2 |\mathbf{x}_1^0 - \mathbf{x}_2^0|^2 + k_1^2 |z_1 - z_2|^2; \\ |g(\omega, x_1) - g(\omega, x_2)| \leq k_2 |x_1 - x_2|. \end{cases}$$

To end this section, let us introduce some spaces that will be used in the future. We shall denote, for $t \in [0, T)$, $\mathbf{H}(t, T) = L_{\mathcal{F}}^2(t, T; \mathbb{R})$, and let $\mathbf{H}^{\mathbf{C}}(t, T)$ be the subset of $\mathbf{H}(t, T)$ consisting of all continuous processes. For any $\lambda \in \mathbb{R}$, define an equivalent norm on $\mathbf{H}(t, T)$ by: $\|\xi\|_{t, \lambda} \stackrel{\Delta}{=} \{E \int_t^T e^{-\lambda s} |\xi(s)|^2 ds\}^{\frac{1}{2}}$. Then $\mathbf{H}_\lambda(t, T) \stackrel{\Delta}{=} \{\xi \in \mathbf{H}(t, T): \|\xi\|_{t, \lambda} < \infty\} = \mathbf{H}(t, T)$. We shall also use the following norm on $\mathbf{H}^{\mathbf{C}}(t, T)$:

$$|\xi|_{t, \lambda, \beta} \stackrel{\Delta}{=} e^{-\lambda T} E |\xi_T|^2 + \beta \|\xi\|_{t, \lambda}^2, \quad \xi \in \mathbf{H}^{\mathbf{C}}(t, T), \ \lambda \in \mathbb{R}, \ \beta > 0,$$

and denote by $\mathbf{H}_{\lambda, \beta}(t, T)$ the completion of $\mathbf{H}^{\mathbf{C}}(t, T)$ under norm $|\cdot|_{t, \lambda, \beta}$. Then for any $\lambda$ and $\beta$, $\mathbf{H}_{\lambda, \beta}(t, T)$ is a Banach space. Further, if $t = 0$, we simply denote $\|\cdot\|_\lambda \stackrel{\Delta}{=} \|\cdot\|_{0, \lambda}$; $|\cdot|_{\lambda, \beta}^2 \stackrel{\Delta}{=} |\cdot|_{0, \lambda, \beta}^2$; $\mathbf{H} = \mathbf{H}(0, T)$; $\mathbf{H}^{\mathbf{C}} = \mathbf{H}^{\mathbf{C}}(0, T)$; $\mathbf{H}_\lambda = \mathbf{H}_\lambda(t, T)$, and $\mathbf{H}_{\lambda, \beta} = \mathbf{H}_{\lambda, \beta}(t, T)$.

## 3. Statements of Main Results

In this section, we state the main results of this paper. The proof of these results will be carried out in the remaining sections. Let us first introduce some notations: for any $\lambda \in \mathbb{R}$, $\alpha \in (0, 1)$, and constants $C_1, \ldots, C_4 > 0$,

$$\begin{cases} A(\lambda, t) = e^{-(\lambda \wedge 0)t}; \ B(\lambda, t) = \dfrac{1 - e^{-\lambda t}}{\lambda} = t \int_0^1 e^{-\lambda t \theta} d\theta, \\ \bar{\lambda}_1 = \lambda - K(2 + C_1^{-1} + C_2^{-1}) - K^2; \ \bar{\lambda}_2 = -\lambda - 2\gamma - K(C_3^{-1} + C_4^{-1}), \end{cases} \qquad (3.1)$$

and

$$\mu(\alpha, T) \stackrel{\Delta}{=} K(C_1 + K)B(\bar{\lambda}_2, T) + \frac{A(\bar{\lambda}_2, T)}{\alpha}(KC_2 + k_1^2). \qquad (3.2)$$

Clearly, for all $\lambda \in \mathbb{R}^n$, $B(\lambda, \cdot)$ is a nonnegative, increasing function, $B(\lambda, 0) = 0$; $A(\lambda, t) \geq 1$ and $A(\lambda, 0) = 1$. Further, the function $\mu(\cdot, \cdot)$ depends on the constants $K$, $k_1$, $k_2$, $\gamma$, $\lambda$, $\alpha$, the duration $T > 0$, and the choice of $C_1, \ldots, C_4$.

One of the main ingredients of our discussion involves the following relations among the parameters introduced above which we will call *compatibility conditions* in the sequel.

(C1) $0 \leq k_1, k_2 < 1$;

(C2) $k_2 = 0$; $\exists \alpha \in (0, 1)$ such that $\mu(\alpha, T)KC_3 < \bar{\lambda}_1$;

**(C3)** $k_2 > 0$; $\exists \alpha_0 \in (k_1, k_2, 1)$, such that $\mu(\alpha_0^2, T)k_2^2 < 1$ and $\bar{\lambda}_1 = \dfrac{KC_3}{k_2^2}$.

**Theorem 3.1:** *Assume* (A1). *Assume also that the compatibility conditions* (C1), *and either* (C2) *or* (C3) *hold for some choices of constants* $\lambda$, $\alpha$, $C_1$, $C_2$, $C_3$, *and* $C_4 = \dfrac{1 - \alpha_0^2}{K}$. *Then the FBSDER* (2.2) *has a unique adapted solution over* $[0, T]$. *Furthermore, the process* $\zeta$ *is absolutely continuous with respect to the Lebesgue measure, and for any RCLL process* $V$ *such that* $V_t \in \mathcal{O}_2(\cdot, t, X_t)$, $\forall t \in [0, T]$, *P-a.s., we have*

$$\langle \frac{d\zeta_t}{dt}, Y_t - V_t \rangle \le 0, \quad \forall t \in [0, T], \quad P\text{-}a.s. \tag{3.3}$$

*Moreover, if the compatibility condition* (C1) *holds, then there exists* $T_0 > 0$ *such that for all* $T \in (0, T_0]$, *the FBSDER* (2.2) *has a unique adapted solution. In particular, if either* $k_1 = 0$ *or* $k_2 = 0$, *then the FBSDER* (2.2) *is always uniquely solvable on* $[0, T]$ *for* $T$ *small.*

**Remark 3.2:** The compatibility conditions (C1)-(C3) do not contain the case when the duration is arbitrary. However, by combining the methods of this paper (Theorem 3.1) and that of Pardoux and Tang [17], it is not too hard to prove that under the compatibility condition (C1), there exist a constant $\Lambda > 0$, depending only on the constants $K$, $k_1$, $k_2$, such that whenever $\gamma < -\Lambda$, the FBSDER (3.2) has a unique adapted solution for all $T > 0$. We omit the details.

**Remark 3.3:** Let us denote $\zeta = \zeta^+ - \zeta^-$, $\zeta_0^+ = \zeta_0^- = 0$, the minimal decomposition of $\zeta$ as a difference of two non-decreasing processes. By replacing $V$ in (3.3) by

$$V_t^L = L_t 1_{\{\frac{d\zeta_t}{dt} \ge 0\}} + Y_t 1_{\{\frac{d\zeta_t}{dt} < 0\}}, \quad V_t^U = U_t 1_{\{\frac{d\zeta_t}{dt} \le 0\}} + Y_t 1_{\{\frac{d\zeta_t}{dt} > 0\}}, \quad t \in [0, T],$$

respectively, we obtain

$$\langle Y_t - L_t, d\zeta_t^+ \rangle = 0, \quad \langle Y_t - U_t, d\zeta_t^- \rangle = 0, \forall t \in [0, T], \text{ a.s.} \tag{3.4}$$

Let us now consider the FBSDER (2.2) over a subinterval $[t, T] \subseteq [0, T]$. In other words, for $t \le s \le T$,

$$\begin{cases} X_s = x + \displaystyle\int_t^s b(r, X_r, Y_r, Z_r)dr + \int_s^t \sigma(r, X_r, Y_r, Z_r)dW_r - \int_t^s n(X_s)d\,|\,\eta\,|_s; \\[2mm] Y_s = g(X_T) + \displaystyle\int_s^T h(r, X_r, Y_r, Z_r)dr - \int_s^T \langle Z_r, dW_r \rangle + \zeta_T - \zeta_s. \end{cases} \tag{3.5}$$

Clearly, Theorem 3.1 remains true if we replace $T$ there by $T - t$. We denote its solution by $(X^{t,x}, Y^{t,x}, Z^{t,x}, \eta^{t,x}, \zeta^{t,x})$. Moreover, we shall need a further assumption:

**(A2)** The coefficients $b, h, \sigma, g$, and the moving boundaries $L$ and $U$ are all deterministic. The function $\sigma$ does not depend on $z$; and the functions $b$ and $h$ are differentiable in $z$.

Since under (A2) the coefficients of (3.5) are "Markovian", the standard technique of "time shifting" can thus be used to show that the process $\{Y_s^{t,x}\}_{s \ge t}$ is $\mathcal{F}_s^t$-adapted, where $\mathcal{F}_s^t = \sigma\{W_r, t \le r < s\}$. Consequently, by Blumenthal 0-1 law, the function $u(t, x) = Y_t^{t,x}$ is deterministic. The following "continuous dependence" result is not

surprising.

**Theorem 3.4:** *Assume* (A1) *and* (A2). *Assume also that the compatibility conditions* (C1) *and either* (C2) *or* (C3) *hold. Then the function* $u(t,x) = Y_t^{t,x}$, $t \in [0,T] \times \mathcal{O}_1$ *is continuous and enjoys the following property:*

$$| u(t_1, x_1) - u(t_2, x_2) |^2 \leq C\Big( | x_1 - x_2 |^2 + (1 + | x_1 |^2 \vee | x_2 |^2) | t_1 - t_2 | \Big), \qquad (3.6)$$

*where* $C > 0$ *is some constant depending only on* $b$, $h$, $\sigma$, $g$ *and* $T$.

The next theorem concerns the connection between the adapted solution of an FBSDER and the viscosity solution to a variational inequality, for which we shall assume in the sequel the solvability of (3.5) whenever it is mentioned. Further, let us denote $\mathcal{G}(n)$ to be the set of $n \times n$ symmetric nonnegative matrices, and for $p \in \mathbb{R}^n$, $Q \in \mathcal{G}(n)$, define

$$F(t,x,u,p,Q) \triangleq \tfrac{1}{2}\mathrm{tr}\{\sigma\sigma^T(t,x,u)Q + \langle b(t,x,u,\sigma(t,x,u)p), p\rangle \\ + h(t,x,u,\sigma(t,x,u)p). \qquad (3.7)$$

Consider the following two *variational inequalities* (or *obstacle problems*) with boundary conditions: denoting $Du = \nabla u = (\partial_{x_1} u, ..., \partial_{x_n} u)^T$, $D^2u = (\partial^2_{x_i x_j} u)_{i,j}$ (the Hessian of $u$), $u_t = \partial_t u$; and $a \wedge b \triangleq \min\{a, b\}$, $a \vee b \triangleq \max\{a, b\}$,

$$\begin{cases} (u - L) \wedge \{(u - U) \wedge [-u_t - F(t,x,u,Du,D^2u)]\} = 0, \\ u(T,x) = g(x), \ x \in \mathcal{O}_1; \ \langle n(x), Du \rangle = 0, \ x \in \partial\mathcal{O}_1 \end{cases} \qquad (3.8)$$

and

$$\begin{cases} (u - U) \vee \{(u - L) \wedge [-u_t - F(t,x,u,Du,D^2u)]\} = 0, \\ u(T,x) = g(x), \ x \in \mathcal{O}_1; \ \langle n(x), Du \rangle = 0, \ x \in \partial\mathcal{O}_1. \end{cases} \qquad (3.9)$$

Observe that if $a, b \in \mathbb{R}$ such that $a \geq b$, then for any $c \in \mathbb{R}$, $a \wedge (b \vee c) = b \vee (a \wedge c)$. Therefore, (3.8) and (3.9) are equivalent, because $L(t,x) \leq U(t,x)$ for all $(t,x)$. This observation will facilitate our discussion quite a bit in Section 6 in the proof of Theorem 3.8 below.

**Definition 3.5:** Let $u \in C([0,T] \times \mathcal{O}_1)$ and $(t,x) \in [0,T] \times \mathcal{O}_1$. The *parabolic superjet* (resp. *parabolic subjet*) of $u$ at $(t,x)$, denoted by $\mathcal{P}^{2,+} u(t,x)$, (resp $\mathcal{P}^{2,-} u(t,x)$), is the set of triples $(p, q, X) \in \mathbb{R} \times \mathbb{R}^n \times S(d)$ such that for $(s,y) \in [0,T] \times \mathcal{O}_1$,

$$u(s,y) \leq \ (\mathrm{resp.} \ \geq )u(t,x) + p(s-t) + \langle q, y-x \rangle$$

$$+ \tfrac{1}{2}\langle X(y-x), y-x \rangle + o(|s-t| + |y-x|^2).$$

**Definition 3.6:** A function $u \in C([0,T] \times \mathcal{O}_1)$ is called a *viscosity sub-* (resp. *super-*) *solution* of (3.8) or (3.9), if $u(T,x) \leq g(x)$ (resp. $u(T,x) \geq g(x)$), $\forall x \in \mathcal{O}_1$; and for any $(p, q, X) \in \mathcal{P}^{2,-} u(t,x)$ (resp. $\mathcal{P}^{2,+} u(t,x)$), we have

$$\left\{ \begin{array}{l} (u - L) \wedge \{(u - U) \vee [\, - p - F(t,x,u,q,X)]\}\} \leq 0, (t,x) \in [0,T] \times \mathcal{O}_1; \\ \langle n(x), q \rangle \wedge \{(u - L) \wedge \{(u - U) \vee [\, - p - F(t,x,u,q,X)]\}\}\} \leq 0, \qquad (3.10) \\ \hfill (t,x) \in [0,T] \times \partial\mathcal{O}_1 \end{array} \right.$$

(resp.

$$\left\{ \begin{array}{l} (u - U) \wedge \{(u - L) \vee [\, - p - F(t,x,u,q,X)]\}\} \leq 0, (t,x) \in [0,T] \times \mathcal{O}_1; \\ \langle n(x), q \rangle \vee \{(u - U) \vee \{(u - U) \wedge [\, - p - F(t,x,u,q,X)]\}\}\} \geq 0, \qquad (3.11) \\ \hfill (t,x) \in [0,T] \times \partial\mathcal{O}_1.) \end{array} \right.$$

Finally, $u$ is a *viscosity solution* to (3.8) (or (3.9)) if it is both a viscosity subsolution and a viscosity supersolution.

We shall use the following equivalent definition of a viscosity solution often.

**Definition 3.7:** A function $u \in C([0,T] \times \mathcal{O}_1)$ is called a *viscosity sub-* (resp. *super-*) *solution* of (3.8) (or (3.9)), if $u(T,x) \leq g(x)$ (resp. $u(T,x) \geq g(x)$), $\forall x \in \mathcal{O}_1$; and for any $\varphi \in C^{1,2}([0,T] \times \mathcal{O}_1)$ such that whenever $(t,x) \in [0,T) \times \mathcal{O}_1$ is a global maximum (resp. minimum) of $u - \varphi$, we have, suppressing dependence on $(t,x)$,

$$\left\{ \begin{array}{l} (u - L) \wedge \{(u - U) \vee [\, - \varphi_t - F(t,x,u,D\varphi,D^2\varphi)]\}\} \leq 0, \\ \langle n(x), D\varphi \rangle \wedge \{(u - L) \wedge \{(u - U) \vee [\, - \varphi_t - F(t,x,u,D\varphi,D^2\varphi)]\}\}\} \leq 0, \\ \hfill (t,x) \in [0,T] \times \partial\mathcal{O}_1 \end{array} \right.$$

(resp.

$$\left\{ \begin{array}{l} (u - U) \vee \{(u - L) \wedge [\, - \varphi_t - F(t,x,u,D\varphi,D^2\varphi)]\}\} \leq 0, (t,x) \in [0,T] \times \mathcal{O}_1; \\ \langle n(x), D\varphi \rangle \vee \{(u - U) \vee \{(u - L) \wedge [\, - \varphi_t - F(t,x,u,D\varphi,D^2\varphi)]\}\}\} \geq 0, \\ \hfill (t,x) \in [0,T] \times \partial\mathcal{O}_1.) \end{array} \right.$$

Finally, $u$ is a *viscosity solution* to (3.8) (or (3.9)) if it is both a viscosity subsolution and a viscosity supersolution.

We then have the following theorem.

**Theorem 3.8:** *Assume* (A2) *and the conditions of Theorem 3.1. Suppose that for a given time duration* $[t,T]$, *the FBSDER* (3.5) *has an adapted solution* $(X^{t,x}, Y^{t,x}, Z^{t,x}, \eta^{t,x}, \zeta^{t,x})$. *Then the function* $u(t,x) \triangleq Y_t^{t,x}$, $(t,x) \in [0,T] \times \mathcal{O}_1$ *is a viscosity solution of the variational inequality with boundary condition* (3.8) *(or equivalently* (3.9))*. Furthermore, if the coefficient* $\sigma$ *is independent of* $y$ *as well (i.e.,* $\sigma = \sigma(t,x)$)*, then the viscosity solution of* (3.8) *is unique among those that satisfy* (3.6)*.*

## 4. An American Game Option

Before we go into the technical proofs of Theorems 3.1 and 3.8, let us look at a possible application of FBSDERs in option pricing theory.

Let us assume that $\mathcal{O}_1 = \mathbb{R}^n$ (i.e., no reflection of $X$). Denote by $\mathcal{M}_{t,T}$ the set of $\{\mathcal{F}_t\}$-stopping times taking values in $[t,T]$. In light of Cvitanić and Ma [6], we interpret $X$ in (2.2) as a price process of financial assets; $Y$ and $Z$ are the wealth process and the trading strategy, respectively, of a "large" investor in the market in the sense that both $Y$ and $Z$ might affect the price $X$. We note that such an interpretation should only be used up until the first time we have $d\zeta < 0$, when the external funds are brought in, contradicting the conventional "self-financing" assumption.

The *American game option* can be described as follows. Unlike the usual American option where only the buyer has the right to choose the exercise time, in game option the seller (the large investor) has the same right as well. However, in order to get a nontrivial option (i.e., to avoid immediate exercise being optimal), it is required that the payoff be higher if the seller opts to force the exercise. Of course, the seller may choose not to do anything, then the game option becomes the usual American option. To be more precise, we let $\tau < T$ be the (random) time the buyer chooses to exercise the option; and $\sigma < T$ be that of the seller. If $\tau \le \sigma$, then the seller pays $L(\tau, X_\tau)$; if $\sigma < \tau$, then the seller pays $U(\sigma, X_\sigma)$. If neither exercises the option by the maturity date $T$, then the seller pays $B = g(X_T)$. We define the *minimal hedging prices* of this contract to be the infimum of initial wealth amounts $Y_0$, such that the seller can deliver the payoff, a.s., without having to use outside funds. In other words, the wealth process has to follow the dynamics of $Y$ (with $d\zeta \ge 0$), up to the exercise time $\sigma \wedge \tau \wedge T$, and at that tie

$$Y_{\sigma \wedge \tau \wedge T} \ge g(X_T)\mathbf{1}_{\{\sigma \wedge \tau = T\}} + L(\tau, X_\tau)\mathbf{1}_{\{\tau < T, \tau \le \sigma\}} + U(\sigma, X_\sigma)\mathbf{1}_{\{\sigma < \tau\}} \quad (4.1)$$

must hold. Our purpose is to determine the minimal hedging price, as well as the corresponding minimal hedging process.

In order to solve this problem, we first reformulate it as the following stochastic (Dynkin) game: for any fixed $0 \le t \le T$, two players both have the right to choose a time, say, $\sigma, \tau \in \mathcal{M}_{t,T}$, respectively, to stop the game. If player $I$ stops the game first, i.e., $\sigma < \tau$, he/she pays $U(\sigma)$ to player $II$. If player $II$ stops first, $\tau \le \sigma < T$, player $I$ pays $L(\tau)$. If no one stops by time $T$, player $I$ pays $B$. There is also a *running cost $h(t)$*. In other words, the payoff player $I$ has to pay is given by

$$R_t^B(\sigma, \tau) \triangleq \int_t^{\sigma \wedge \tau} h(u)du + B\mathbf{1}_{\{\sigma \wedge \tau = T\}} + L(\tau)\mathbf{1}_{\{\tau < T, \tau \le \sigma\}} + U(\sigma)\mathbf{1}_{\{\sigma < \tau\}}, \quad (4.2)$$

where $B \in L^2(\Omega)$ is a given $\mathcal{F}_T$-measurable random variable satisfying $L(T) \le B \le U(T)$. Suppose that player $II$ is trying to maximize the payoff, while player $I$ attempts to minimize it. Define the *upper* and *lower* values of the game by

$$\overline{V}(t) \triangleq \underset{\sigma \in \mathcal{M}_{t,T}}{\operatorname{essinf}} \quad \underset{\tau \in \mathcal{M}_{t,T}}{\operatorname{esssup}} \; E\{R_t^B(\sigma, \tau) \mid \mathcal{F}_t\},$$

$$\underline{V}(t) \triangleq \underset{\tau \in \mathcal{M}_{t,T}}{\operatorname{esssup}} \quad \underset{\sigma \in \mathcal{M}_{t,T}}{\operatorname{essinf}} \; E\{R_t^B(\sigma, \tau) \mid \mathcal{F}_t\}, \quad (4.3)$$

respectively; and we say that the game has a value if $\overline{V}(t) = \underline{V}(t) \triangleq V(t)$.

The following theorem can be proved exactly as Theorem 4.1 in Cvitanić and Karatzas [5], after recalling (3.4). We give only the statement.

**Theorem 4.1:** *Suppose there exists a solution* $(X, Y, Z, \eta, \zeta)$ *to FBSDER* (2.2). *Then the game* (4.2) *with* $B = g(X_T)$, $h(t) = h(t, X_t, Y_t, Z_t)$, *and* $L(t, \omega) = L(t, X_t(\omega))$, $U(t, \omega) = U(t, X_t(\omega))$ *has value* $V(t)$, *given by the backward component* $Y$ *of the solution to the FBSDER, i.e.,* $V(t) = \bar{V}(t) = \underline{V}(t) = Y_t$, *-as., for all* $0 \le t \le T$. *Moreover, there exists a saddle-point* $(\hat{\sigma}_t, \hat{\tau}_t) \in \mathcal{M}_{t,T} \times \mathcal{M}_{t,T}$, *given by*

$$\hat{\sigma}_t \triangleq \inf\{s \in [t, T] : Y_s = U(s, X_s)\} \wedge T,$$

$$\hat{\tau}_t := \inf\{x \in [t, T] : Y_s = L(s, X_s)\} \wedge T, \tag{4.4}$$

*namely, we have*

$$E\{R_t^{g(X_T)}(\hat{\sigma}_t, \tau) \mid \mathcal{F}_t\} \le E\{R_t^{g(X_T)}(\hat{\sigma}_t, \hat{\tau}_t) \mid \mathcal{F}_t\}$$

$$Y_t \le E\{R_t^{g(X_T)}(\sigma, \hat{\tau}_t) \mid \mathcal{F}_t\}, \quad a.s. \tag{4.5}$$

*for every* $(\sigma, \tau) \in \mathcal{M}_{t,T} \times \mathcal{M}_{t,T}$.

Our main result of this section is the following.

**Theorem 4.2:** *The minimal hedging price of the American Game Option is greater or equal to* $\bar{V}(0)$, *the upper value of the game (at* $t = 0$) *of Theorem 4.1. If the corresponding FBSDER has a solution* $(\widetilde{X}, \widetilde{Y}, \widetilde{Z}, \zeta)$, *then the minimal hedging prices is equal to* $\widetilde{Y}_0$.

**Proof:** Fix the exercise times $\sigma, \tau$ of the seller and the buyer, respectively. If $Y$ is the seller's hedging process, it satisfies the following dynamics for $t \le \tau \wedge \sigma \wedge T$:

$$Y_t + \int_0^t h(s, X_s, Y_s, Z_s) ds = \int_0^t Z_s dW_s - \zeta_t,$$

with $\zeta$ non-decreasing. Hence, the left-hand side is a supermartingale From this and the requirement (4.1) that $Y$ be a hedging process, we get $Y_t \ge E\{R_t^{\gamma(X_T)}(\sigma, \tau) \mid \mathcal{F}_t\}$, $\forall t$, -a.s. in the notation of Theorem 4.1. Since the buyer is trying to maximize the payoff, and the seller to minimize it, we get $Y_t \ge \bar{V}_t$, $\forall t$, -a.s. Consequently, the minimal hedging prices is no less than $\bar{V}(0)$.

Conversely, if the FBSDER has a solution with $\widetilde{Y}$ as the backward component, then by Theorem 4.1, process $\widetilde{Y}$ is equal to the value process of the game. By (4.4) (with $t = 0$) and (3.4), $\zeta_t$ is nondecreasing for $t \le \hat{\sigma}_0$. Therefore, for $t$ less than the optimal exercise time for the seller $\hat{\sigma} = \hat{\sigma}_0$, $\widetilde{Y}_t$ obeys the dynamics of a wealth process. So, the seller can start with $\widetilde{Y}_0$, follow the dynamics of $\widetilde{Y}$ until $t = \hat{\sigma}$ and then exercise, if the buyer has not exercised first. In general, from the saddle-point property we know that for any $\tau \in \mathcal{M}_{0,T}$,

$$\widetilde{Y}_{\hat{\sigma} \wedge \tau} \ge g(X_T) 1_{\{\hat{\sigma} \wedge \tau = T\}} + L(\tau, X_\tau) 1_{\{\tau < T, \tau \le \hat{\sigma}\}} + U(\hat{\sigma}, X_{\hat{\sigma}}) 1_{\{\hat{\sigma} < \tau\}}.$$

This implies that the seller can deliver th required payoff if he uses $\hat{\sigma}$ as his exercise time, no matter what the buyer's exercise time $\tau$ is. Consequently, $\widetilde{Y}_0 = V(0)$ is no less than the minimal hedging price. □

## 5. Proofs of Theorem 3.1 and 3.4

We start by establishing some *a priori* estimates. The following lemma is a variation of a similar one in Pardoux and Tang [17]. Since no proof was given there and in the present case, the reflection processes have to be considered, we shall provide a short proof for completeness.

**Lemma 5.1:** *Assume* (A1). *Let* $(X, Y, Z, \eta, \zeta)$ *and* $(X', Y', Z', \eta', \zeta')$ *be two solutions to the FBSDE* (2.2), *and let* $\widehat{\xi} \triangleq \xi - \xi'$, *where* $\xi = X, Y, Z, \eta, \zeta$, *respectively.*

(i)  *Let* $\lambda \in \mathbb{R}$, $C_1, C_2 > 0$, *and let* $\bar{\lambda}_1 = \lambda - K(2 + C_1^{-1} + C_2^{-1}) - K^2$. *Then, for all* $\lambda' \in \mathbb{R}$,

$$e^{-\lambda t} E \mid \widehat{X}_t \mid^2 + (\bar{\lambda}_1 - \lambda') \int_0^t e^{-\lambda\tau} e^{-\lambda'(t-\tau)} E \mid \widehat{X}_\tau \mid^2 d\tau$$

$$\leq \int_0^t e^{-\lambda\tau} e^{-\lambda'(t-\tau)} \{ K(C_1 + K) E \mid \widehat{Y}_\tau \mid^2 + (KC_2 + k_1^2) E \mid \widehat{Z}_\tau \mid^2 \} d\tau. \tag{5.1}$$

(ii)  *Let* $\lambda \in \mathbb{R}$ *and* $C_3, C_4 > 0$, *and let* $\bar{\lambda}_2 = -\lambda - 2\gamma - K(C_3^{-1} + C_4^{-1})$. *Then, for all* $\lambda' \in \mathbb{R}$,

$$e^{-\lambda t} E \mid \widehat{Y}_t \mid^2 + (\bar{\lambda}_2 - \lambda') \int_t^T e^{-\lambda\tau} e^{-\lambda'(\tau-t)} E \mid \widehat{Y}_\tau \mid^2 d\tau$$

$$+ (1 - KC_4) \int_t^T e^{-\lambda\tau} e^{-\lambda'(\tau-t)} E \mid \widehat{Z}_\tau \mid^2 d\tau \tag{5.2}$$

$$\leq k_2^2 e^{-\lambda T} e^{-\lambda'(T-t)} E \mid \widehat{X}_T \mid^2 + KC_3 \int_t^T e^{-\lambda\tau} e^{-\lambda'(\tau-t)} \mid \widehat{X}_\tau \mid^2 d\tau.$$

*Consequently, if* $KC_4 = 1 - \alpha$ *for some* $\alpha \in (0, 1)$, *then*

$$e^{-\lambda T} E \mid \widehat{X}_T \mid^2 + \bar{\lambda}_1 \parallel \widehat{X} \parallel_\lambda^2 \leq K(C_1 + K) \parallel \widehat{Y} \parallel_\lambda^2 + (KC_2 + k_1^2) \parallel \widehat{Z} \parallel_\lambda^2, \tag{5.3}$$

$$\parallel \widehat{X} \parallel_\lambda^2 \leq B(\bar{\Lambda}_1, T)[K(C_1 + K) \parallel \widehat{Y} \parallel_\lambda^2 + (KC_2 + k_1^2) \parallel \widehat{Z} \parallel_\lambda^2], \tag{5.4}$$

$$\parallel \widehat{Y} \parallel_\lambda^2 \leq B(\bar{\lambda}_2, T)[k_2^2 e^{-\lambda T} E \mid \widehat{X}_T \mid^2 + KC_3 \parallel \widehat{X} \parallel_\lambda^2], \tag{5.5}$$

$$\parallel \widehat{Z} \parallel_\lambda^2 \leq \frac{A(\bar{\lambda}_2, T)}{\alpha}[k_2^2 e^{-\lambda T} E \mid \widehat{X}_T \mid^2 + KC_3 \parallel \widehat{X} \parallel_\lambda^2]. \tag{5.6}$$

**Proof:** We first show (5.1). Let $(X, Y, Z, \eta, \zeta)$ and $(X', Y', Z', \eta', \zeta')$ be two solutions to the FBSDE (2.2), and let $\widehat{\xi} \triangleq \xi - \xi'$, where $\xi = X, Y, Z, \eta, \zeta$, respectively. Let $t \in (0, T]$, $\lambda, \lambda'$ be arbitrarily given, and consider the function $F_t(s, x) \triangleq e^{-\lambda s} e^{-\lambda'(t-s)} \mid x \mid^2$, for $(s, x) \in [0, t] \times \mathbb{R}^n$. Applying Itô's formula to $F_t(s, \widehat{X}_s)$ from 0 to $t$, and then taking expectations, we have

$$e^{-\lambda t} E \mid \widehat{X}_t \mid^2 + (\lambda - \lambda') E \int_0^t e^{-\lambda\tau} e^{-\lambda'(t-\tau)} \mid \widehat{X}_\tau \mid^2 d\tau$$

$$= \int_0^t e^{-\lambda\tau} e^{-\lambda'(t-\tau)} \{2\langle \widehat{X}_\tau, b(\tau, X_\tau, Y_\tau, Z_\tau) - b(\tau, X'_\tau, Y'_\tau, Z'_\tau)\rangle \qquad (5.7)$$

$$+ \mid \sigma(\tau, X_\tau, Y_\tau, Z_\tau) = \sigma(\tau, X'_\tau, Y'_\tau, Z'_\tau)\mid^2\}d\tau + 2E\int_0^t e^{-\lambda\tau} e^{-\lambda'(t-\tau)}\langle \widehat{X}_\tau, d\widehat{\eta}_\tau\rangle.$$

Since $X_t, X'_t \in \mathcal{O}_1$, $\forall t \in [0,T]$, -a.s., by (2.3) we have $\langle \widehat{X}_t, d\widehat{\eta}_\tau\rangle \le 0$ (as a signed measure), $\forall s \in [0,T]$, -a.s. Therefore, repeatedly applying the Schwartz inequality and the fact that $2ab \le ca^2 + c^{-1}b^2$, $\forall c > 0$ to the right side of (5.7), using the definition of $\bar{\lambda}_1$, together with some elementary computation with the help of (A1), we derive (5.1).

The estimate (5.2) can be proved similarly by applying Itô's formula to $\widetilde{F}_t(s, Y_s)$ from $t$ to $T$, where $\widetilde{F}_t(s,x) = e^{-\lambda s} e^{-\lambda'(s-t)}\mid x\mid^2$, using Definition 2.14, and some by now standard estimates for backward SDEs. We omit the details.

Finally, setting $\lambda' = 0$ and $t = T$ in (5.1) we have (5.3); while setting $\lambda' = \bar{\lambda}_1$ in (5.1) then integrating both sides from 0 to $T$ (noting that $B(\lambda_1, \cdot)$ is increasing), we obtain (5.4). Similarly, setting $\lambda' = \bar{\lambda}_2$ in (5.2) and integrating from 0 to $T$ we get (5.5). In order to prove (5.6), we note that if $\bar{\lambda}_2 \le 0$, by setting $\lambda' = \bar{\lambda}_2$ and $t = 0$ in (5.2) we get (remember $k_4 C_4 = \alpha$)

$$\parallel \widehat{Z} \parallel_\lambda^2 \le \int_0^T e^{-\lambda\tau} e^{-\bar{\lambda}_2(\tau - t)}\parallel \widehat{Z}_\tau \parallel^2 d\tau$$

$$\le \frac{e^{\mid\bar{\lambda}_2\mid T}}{\alpha}\{k_2^2 e^{-\lambda T} E\mid \widehat{X}_T\mid^2 + KC_3\parallel \widehat{X}\parallel_\lambda^2\}.$$

If $\bar{\lambda}_2 > 0$, then we let $\lambda' = 0$ in (5.2) to get $\parallel \widehat{Z} \parallel_\lambda^2 \le \frac{1}{\alpha}\{k_2^2 e^{-\lambda T} E\mid \widehat{X}_T\mid^2 + KC_3\parallel \widehat{X}\parallel_\lambda^2\}$, and (5.6) thus follows.  $\square$

Similarly to Lemma 5.1, we can derive an a priori estimate for an adapted solution to FBSDER (2.2). Denote $\sigma^0(\omega, t) = \sigma(\omega, t, 0, 0, 0)$, $f^0(\omega, t) = f(\omega, t, 0, 0, 0)$, $h^0(\omega, t) = h(\omega, t, 0, 0, 0)$, and $g^0(\omega) = g(0, \omega)$. Then we have

**Lemma 5.2:**  *Assume* (A1). *Let* $(X, Y, Z, \eta, \zeta)$ *be an adapted solution to the FBSDER* (2.2). *For any* $\lambda, \lambda' \in \mathbb{R}$, $\varepsilon > 0$, $C_1, ..., C_4 > 0$, *we define* $\bar{\lambda}_1^\varepsilon = \bar{\lambda}_1 - (1 + K^2)\varepsilon$ *and* $\bar{\lambda}_2^\varepsilon = \bar{\lambda}_2 - \varepsilon$, *where* $\bar{\lambda}_1$ *and* $\bar{\lambda}_2$ *are those defined in Lemma 3.2. Then*

$$e^{-\lambda t} E\mid X_t\mid^2 + (\bar{\lambda}_1^\varepsilon - \lambda')\int_0^t e^{-\lambda'(t-\tau)} e^{-\lambda s} E\mid X_\tau\mid^2 d\tau \le e^{-\lambda't}\mid x\mid^2$$

$$+ \int_0^t e^{-\lambda'(t-\tau)} e^{-\lambda\tau}\frac{1}{\varepsilon}E\mid f^0(\tau)\mid^2 + \left(1 + \frac{1}{\varepsilon}\right)\mid\sigma^0(\tau)\mid^2\right) \qquad (5.8)$$

$$+ K(C_1 + K(1 + \varepsilon))E\mid Y_\tau\mid^2 + (KC_2 + k_1^2(1 + \varepsilon))E\mid Z_\tau\mid^2\}d\tau$$

*and*

$$e^{-\lambda t}E\,|\,Y_t\,|^2 + (\bar{\lambda}_2^\varepsilon - \lambda')\int_t^T e^{-\lambda'(\tau - t)}e^{-\lambda\tau}E\,|\,Y_\tau\,|^2 d\tau$$

$$+ (1 - k_4 C_4)\int_t^T e^{-\lambda'(\tau - t)}e^{-\lambda\tau}E\,|\,Z_\tau\,|^2 d\tau \qquad (5.9)$$

$$\leq k_2^2(1 + \varepsilon)e^{-\lambda'(T - t)}e^{-\lambda T}E\,|\,X_T\,|^2 + \left(1 + \frac{1}{\varepsilon}\right)e^{-\lambda'(T - t)}e^{-\lambda T}E\,|\,g^0\,|^2$$

$$+ \int_t^T e^{-\lambda'(\tau - t)}e^{-\lambda\tau}\Big\{KC_3 E\,|\,X_\tau\,|^2 + \frac{1}{\varepsilon}E\,|\,h^0(\tau)\,|^2\Big\}d\tau.$$

*Consequently, if* $C_4 = \frac{1 - \alpha}{K}$, *for some* $\alpha \in (0,1)$, *we have*

$$e^{-\lambda T}E\,|\,X_T\,|^2 + \bar{\lambda}_1^\varepsilon\,\|\,X\,\|_\lambda^2 \leq [\,|\,x\,|^2 + K(C_1 + K(1 + \varepsilon))\,\|\,Y\,\|_\lambda^2$$

$$+ (KC_2 + k_1^2(1 + \varepsilon))\,\|\,Z\,\|_\lambda^2 + \frac{1}{\varepsilon}\,\|\,f^0\,\|_\lambda^2 + \left(1 + \frac{1}{\varepsilon}\right)\|\,\sigma^0\,\|_\lambda^2], \qquad (5.10)$$

$$\|\,X\,\|_\lambda^2 \leq B(\bar{\lambda}_1^\varepsilon, T)[\,|\,x\,|^2 + K(C_1 + K(1 + \varepsilon))\,\|\,Y\,\|_\lambda^2$$

$$+ (KC_2 + k_1^2(1 + \varepsilon))\,\|\,Z\,\|_\lambda^2 + \frac{1}{\varepsilon}\,\|\,f^0\,\|_\lambda^2 + \left(1 + \frac{1}{\varepsilon}\right)\|\,\sigma^0\,\|_\lambda^2], \qquad (5.11)$$

$$\|\,Y\,\|_\lambda^2 \leq B(\bar{\lambda}_2^\varepsilon, T)[k_2^2(1 + \varepsilon)e^{-\lambda T}E\,|\,X_T\,|^2 + KC_3\,\|\,X\,\|_\lambda^2$$

$$+ \left(1 + \frac{1}{\varepsilon}\right)e^{-\lambda T}E\,|\,g^0\,|^2 + \frac{1}{\varepsilon}\,\|\,h^0\,\|_\lambda^2], \qquad (5.12)$$

$$\|\,Z\,\|_\lambda^2 \leq \frac{A(\bar{\lambda}_2^\varepsilon, T)}{\alpha}[k_2^2(1 + \varepsilon)e^{-\lambda T}E\,|\,X_T\,|^2 + KC_3\,\|\,X\,\|_\lambda^2$$

$$+ \left(1 + \frac{1}{\varepsilon}\right)e^{-\lambda T}E\,|\,g^0\,|^2 + \frac{1}{\varepsilon}\,\|\,h^0\,\|^2]. \qquad (5.13)$$

**Proof of Theorem 3.1:** We introduce a mapping $\Gamma: \mathbf{H^C} \mapsto \mathbf{H^C}$ defined as follows: for fixed $x \in \mathbb{R}^n$, let $\bar{X} \triangleq \Gamma(X)$ be the solution to the FSDER:

$$\bar{X}_t = x + \int_0^t b(s, \bar{X}_s, Y_s, Z_s)ds + \int_0^t \sigma(s, \bar{X}_s, y_s, Z_s)dW_s + \bar{\eta}_t, \qquad (5.14)$$

where the processes $Y$ and $Z$ are the solution to the following BSDER:

$$Y_t = g(X_T) + \int_t^T h(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s + \zeta_T - \zeta_t. \qquad (5.15)$$

We note that the assumption (A1) on the random fields $L$ and $U$ and the coefficients enables us to apply the result of Cvitanić and Karatzas [5] to conclude that the

BSDER (5.15) has a unique solution $(Y, Z, \zeta)$ such that (3.3) hold. This in turn guarantees the existence and uniqueness of the adapted solution $\bar{X}$ to the FSDER (5.14) (cf. e.g. Lions and Sznitman [13]). Furthermore, by definition of $\lambda_1^\varepsilon$ (Lemma 5.2) we see that if $\lambda$ is chosen so that $\bar{\lambda}_1 > 0$, then it is always possible to choose $\varepsilon > 0$ small enough so that $\bar{\lambda}_1^\varepsilon > 0$ as well and (5.10) will lead to $\bar{X} \in \mathbf{H}_{\lambda, \bar{\lambda}_1}$ since $\bar{\lambda}_1 > 0$ and $\bar{\lambda}_1^\varepsilon > 0$. Let us try to find a suitable $\bar{\lambda}_1 > 0$ so that $\Gamma$ is a contraction on $\mathbf{H}_{\lambda, \bar{\lambda}_1}$, which will lead to the existence and uniqueness of the adapted solution to the FBSDER (2.2) immediately.

To this end, let $X^1, X^2 \in \mathbf{H^c}$ and let $(Y^i, Z^i, \zeta^i)$ and $(\bar{X}^i, \bar{\eta}^i)$, $i = 1, 2$, be the corresponding solutions to (5.15) and (5.14), respectively. Denote $\Delta\xi = \xi^1 - \xi^2$, for $\xi = X, Y, Z, \bar{X}$. Applying (5.3)-(5.6) (with $C_4 = \frac{1-\alpha}{K}$) we easily deduce that

$$e^{-\lambda T} E \, | \, \Delta\bar{X}_T \, |^2 + \bar{\lambda}_1 \, \| \, \Delta\bar{X} \, \|^2_\lambda$$

$$\leq \mu(\alpha, T)\{k_2^2 e^{-\lambda T} E \, | \, \Delta X_T \, |^2 + K C_3 \, \| \, \Delta X \, \|^2_\lambda\} \tag{5.16}$$

where $\mu(\alpha, T)$ is defined by (3.2).

Now fix $C_4 = \frac{1 - \alpha_0^2}{K}$. First assume that (C1) and (C2) hold. Since $k_2 = 0$, (5.16) leads to

$$\| \, \Delta\bar{X} \, \|^2_\lambda \leq \frac{\mu(\alpha, T) K C_3)}{\bar{\lambda}_1} \, \| \, \Delta X \, \|^2_\lambda.$$

Since we can find $C_1, C_2, C_3$ and $\alpha \in (0, 1)$ so that $\mu(\alpha, T) K C_3 < 1$, $\Gamma$ is a contraction mapping on $(H, \| \cdot \|_\lambda)$. Thus the first part of the theorem follows in this case.

Similarly, if (C1) and (C3) hold, then we can choose $\lambda$ in (3.1) so that $\bar{\lambda}_1 = K C_3 / k_2^2$, and then derive from (5.16) that

$$| \, \Delta\bar{X} \, |^2_{\lambda^0, \bar{\lambda}_1} \leq \mu(\alpha_0^2, T) k_2^2 \, | \, \Delta X \, |^2_{\lambda^0, \bar{\lambda}_1}.$$

Let $C_i$, $i = 1, 2, 3$ and $\alpha_0 \in (k_1 k_2, 1)$ be such that $\mu(\alpha_0^2, T) k_2^2 < 1$. The mapping $\Gamma$ is again a contraction, but on the space $\mathbf{H}_{\lambda, \bar{\lambda}_1}$, proving the first part of the theorem again.

To prove the last assertion of the theorem, we again assume $k_2 = 0$ first. For fixed $C_1, C_2, C_3, \lambda$ and $\alpha \in (0, 1)$ we have from (3.2)

$$\mu(\alpha, 0) K C_3 = \frac{(K C_2 + k_1^2) K C_3}{\alpha}.$$

Therefore, with $C_1, C_2, C_3$ and $\alpha$ fixed, we can choose $\lambda$ large enough so that $\mu(\alpha, 0) K C_3 < \bar{\lambda}_1$. Then, by the continuity of the functions $A(\alpha, \cdot)$ and $B(\alpha, \cdot)$, for this fixed $\lambda$ we can find $T_0 > 0$ such that $\mu(\alpha, T) K C_3 < \bar{\lambda}_1$ for all $T \in (0, T_0]$. thus (C2) holds for all $T \in (0, T_0]$ and the conclusion follows by the first part of the theorem.

Now assuming $k_2 > 0$. In this case we pick $\alpha_0 \in (k_1 k_2, 1)$, and define

$$\delta \triangleq \frac{1}{k_2^2} - \frac{k_1^2}{\alpha_0^2} > 0. \tag{5.17}$$

Let $C_2 = \frac{\alpha_0^2 \delta}{2K}$, $C_4 = \frac{1-\alpha_0^2}{K}$, and choose $\lambda$ so that $\bar{\lambda}_1 = (k_3 C_3)/k_2^2 > 0$. In this case we have

$$\mu(\alpha_0^2, 0) = \frac{KC_2 + k_1^2}{\alpha_0^2} = \frac{1}{2k_2^2} + \frac{k_1^2}{2\alpha_0^2} < \frac{1}{k_2^2},$$

thanks to (5.17). Using the continuity of $\mu(\alpha_0^2, \cdot)$ again, for any $C_1, C_3 > 0$ we can find $T_0(C_1, C_3) > 0$ such that $\mu(\alpha_0^2, T)k_2^2 < 1$ for all $T \in (0, T_0]$. In other words, the compatibility condition (C3) holds for all $T \in (0, T_0]$, proving our assertion again.

Finally, if $k_1 = 0$ then (C1) becomes trivial, and the conclusion follows. □

**Proof of Theorem 3.4:** Let $(t_1, x_1)$ and $(t_2, x_2)$ be given. For $\xi = X, Y, Z$, let $\xi^i = \xi^{t_i, x_i}$, $i = 1, 2$ and $\widehat{\xi} = \xi^1 - \xi^2$. Assume first $t_1 \geq t_2$, and recall the norms $\| \cdot \|_{t, \lambda}$ and $| \cdot |_{t, \lambda, \beta}$ at the end of Section 2. Repeating the arguments of Theorem 3.1 over the interval $[t_2, T]$, we see that (5.5) and (5.6) will look the same, and $\| \cdot \|_{\lambda}$ being replaced by $\| \cdot \|_{t_2, \lambda}$; but (5.3) and (5.4) become

$$e^{-\lambda T} E | \widehat{X}_T |^2 + \bar{\lambda}_1 \| \widehat{X} \|_{t_1, \lambda}^2$$
$$\leq K(c_1 + K) \| \widehat{Y} \|_{t_2, \lambda}^2 + (KC_2 + k_1^2) \| \widehat{Z} \|_{t_2, \lambda}^2 + E | \widehat{X}_{t_2} |^2, \tag{5.3}'$$

$$\| \widehat{X} \|_{t_2, \lambda}^2 \leq \widetilde{B}(\bar{\lambda}_1, T)[K(C_1 + K) \| \widehat{Y} \|_{t_2, \lambda}^2$$
$$+ (KC_2 + k_1^2) \| \widehat{Z} \|_{t_2, \lambda}^2 + E | \widehat{X}_{t_2} |^2], \tag{5.4}'$$

where $\widetilde{B}(\lambda, T) \triangleq \frac{e^{-\lambda t_2} - e^{-\lambda T}}{\lambda}$. Now similarly to (5.16), one shows that

$$e^{-\lambda T} E | \widehat{X}_T |^2 + \bar{\lambda}_1 \| \widehat{X} \|_{t_2, \lambda}^2$$
$$\leq \mu(\alpha, T) \{ k_2^2 e^{-\lambda T} E | \widehat{X}_T |^2 + KC_3 \| \widehat{X} \|_{t_2, \lambda}^2 \} + E | \widehat{X}_{t_2} |^2. \tag{5.16}'$$

Arguing as in the proof of Theorem 3.1 and using compatibility conditions (C1)-(C3), we can find a constant $C > 0$ depending only on $T > 0$ and $K$, $k_1, k_2$ such that

$$| \widehat{X} |_{t_2, \lambda, \beta}^2 \leq CE | \widehat{X}_{t_2} |^2 = CE | x_2 - X_{t_2}^1 |^2, \tag{5.18}$$

where $\beta = \bar{\lambda}_1 - \mu(\alpha, T)KC_3$ if $k_2 = 0$ and $\beta = \mu(\alpha, T)k_2^2$ if $k_2 > 0$.

From now on by slight abuse of notation, we let $C > 0$ be a generic constant depending only on $T, K, k_1$ and $k_2$ and it is allowed to vary from line to line. Applying standard arguments using Burkholder-Davis-Gundy inequality, we obtain

$$E \sup_{t_2 \leq s \leq T} | X_s^1 |^2 + E \sup_{t_2 \leq s \leq T} | Y_s^1 |^2 \leq CE | \widehat{X}_{t_2} |^2. \tag{5.19}$$

To estimate $E | \widehat{X}_{t_2} |^2$ let us recall the parameters $\lambda_1^\varepsilon$ and $\lambda_2^\varepsilon$ defined in Lemma 5.2. For each $\varepsilon > 0$, define

$$\mu^\varepsilon(\alpha, T) \triangleq K(C_1 K(1+\varepsilon))B(\lambda_2^\varepsilon, T) + \frac{A(\lambda_2^\varepsilon, T)}{1 - KC_4} KC_2.$$

Since $\lambda_1^\varepsilon \to \lambda_1$, $\lambda_2^\varepsilon \to \lambda_2$, and $\mu^\varepsilon(\alpha, T) \to \mu(\alpha, T)$, as $\varepsilon \to 0$, if the compatibility condition (C1) and either (C2) or (C3) hold, then we can choose $\varepsilon > 0$ such that $\mu^\varepsilon(\alpha, T)k_2^2(1 + \varepsilon) < 1$ when $k_2 = 0$ and $\mu^\varepsilon(\alpha, T)KC_3 < \lambda_1^\varepsilon$ when $k_2 \neq 0$. For this fixed $\varepsilon > 0$ we can then repeat the argument of Theorem 3.1 by using (5.10)-(5.13) to derive

$$\left(1 - \frac{\mu^\varepsilon(\alpha, T)KC_3}{\overline{\lambda}_1^\varepsilon}\right) \| X^1 \|_\lambda^2 \leq C(\varepsilon)\left[|x_1|^2 + \left(1 + \frac{1}{\varepsilon}\right)\right], \quad k_2 = 0$$

or

$$\left(1 - \mu^\varepsilon(\alpha, T)k_2^2\right)|X^1|_{\lambda,\beta}^2 \leq C(\varepsilon)\left[|x_1|^2 + \left(1 + \frac{1}{\varepsilon}\right)\right], \quad k_2 \neq 0,$$

where $C(\varepsilon)$ is some constant depending on $T, K, k_1, k_2$ and $\varepsilon$. Since $\varepsilon > 0$ is now fixed, in either case we have, for a generic constant $C > 0$, $\| X^1 \|_\lambda^2 \leq C(1 + |x_1|^2)$ which in turn shows that, in light of (5.10)-(5.13), $\| Y^1 \|_\lambda^2 \leq C(1 + |x_1|^2)$, and $\| Z \|_\lambda^2 \leq C(1 + |x_1|^2)$. Again, applying the Burkholder and Hölder inequalities we can then derive

$$E\left\{\sup_{t_1 \leq s \leq T} |X_t^1|^2\right\} + E\left\{\sup_{t_1 \leq s \leq T} |Y_t^1|^2\right\} \leq C(1 + |x_1|^2). \quad (5.20)$$

Next, note that on the interval $[t_1, t_2]$, the process $(\widehat{X}, \widehat{Y}, \widehat{Z})$ satisfies the following SDE:

$$\begin{cases} \widehat{X}_s = (x_1 - x_2) + \displaystyle\int_{t_1}^s b_r^1 dr + \int_{t_1}^s \sigma_r^1 dW(r), \\[4mm] \widehat{Y}_s = \widehat{Y}_{t_2} + \displaystyle\int_s^{t_2} h_r^1 dr + \int_s^{t_2} Z_r^1 dW(r), \end{cases} \quad s \in [t_1, t_2], \quad (5.21)$$

where $b_r^1 = b(r, X_r^1, Y_r^1, Z_r^1)$, $\sigma_r^1 = \sigma(r, X_r^1, Y_r^1, Z_r^1)$ and $h_r^1 = h(r, X_r^1, Y_r^1, Z_r^1)$. Now from the first equation of (5.21) we easily derive

$$E\left\{\sup_{t_1 \leq s \leq t_2} |\widehat{X}_s|^2\right\} \leq C\{|x_1 - x_2|^2 + (1 + |x_1|^2)|t_1 - t_2|\}.$$

Combining this with (5.19), (5.20), as well as the assumption (A1) (iii), we derive from the second equation of (5.21)

$$E|\widehat{Y}_{t_1}|^2 \leq E|\widehat{Y}_{t_2}|^2 + C(1 + |x_1|^2 \vee |x_2|^2)|t_1 - t_2|$$

$$\leq C\{|x_1 - x_2|^2 + (1 + |x_1|^2 \vee |x_2|^2)|t_1 - t_2|\}.$$

Since $\widehat{Y}(t_1) = u(t_1, x_1) - u(t_2, x_2)$ is deterministic, (3.6) follows. The case when $t_1 \leq t_2$ can be proved by symmetry, and the proof is complete.                    □

## 6. Proof of Theorem 3.8 (I. - Viscosity Property)

Let us assume that FBSDER (3.5) has a unique adapted solution on $[t,T]$, denoted by $(X^{t,x}, Y^{t,x}, Z^{t,x}, \eta^{t,x}, \zeta^{t,x})$. By Proposition 3.4, we see that $u(t,x) = Y_t^{t,x}$ is continuous on $[0,T] \times \bar{\mathcal{O}}_1$: it is Lipschitz-continuous in $x$, and Hölder-$\frac{1}{2}$ in $t$, respectively. We first verify that $u$ is a viscosity solution to (3.8). Since the proof for the "supersolution" is virtually the same as that of the "subsolution", thanks to the equivalency of (3.8) and (3.9), we shall carry out only the proof for the latter.

To this end, let $(t,x) \in [0,T) \times \bar{\mathcal{O}}_1$ be given and let $\varphi \in C^{1,2}([0,T] \times \bar{\Omega}_1)$ be such that $(t,x)$ is a global maximum point of $u - \varphi$. By slightly modifying $\varphi$ we can assume without loss of generality that $u(t,x) = \varphi(t,x)$ and $D\varphi$ is uniformly bounded, thanks to the uniform Lipschitz property of $u$ in $x$.

Note that the pathwise uniqueness implies that $u(s, X_s^{t,x}) = Y_s^{t,x}$, $P$-a.s., for each $s \in [t,T]$. Using the continuity of the function $u$ and processes $X^{t,x}$ and $Y^{t,x}$, one can easily check that for all $\tau \in \mathcal{M}_{t,T}$, the set of all $\mathcal{F}_t$-stopping time $\tau'$ such that $t \leq \tau' \leq T$, $P$-a.s., we have $u(\tau, X_\tau^{t,x}) = Y_\tau^{t,x}$, $P$-a.s. In what follows, we shall suppress the superscript "$t,x$" for the adapted solution, and denote $u_s = u(s, X_s)$, $(D\varphi)_s = D\varphi(s, X_s)$, $(D^2\varphi)_s = D^2\varphi(s, X_s)$, to simplify notations.

For any $\tau \in \mathcal{M}_{t,T}$, we have

$$
\begin{aligned}
u(t,x) = Y_t &= Y_\tau + \int_t^\tau h(s, X_s, Y_s, Z_s) ds - \int_t^\tau Z_s dW_s + \zeta_\tau - \zeta_s \\
&= u(\tau, X_\tau) + \int_t^\tau h(s, X_s, Y_s, Z_s) ds - \int_t^\tau Z_s dW_s + \zeta_\tau - \zeta_t.
\end{aligned}
\tag{6.1}
$$

On the other hand, applying Itô's formula to $\varphi(\cdot, X)$ from $t$ to $\tau$ we have

$$
\varphi(\tau, X_\tau) = \varphi(t,x) + \int_t^\tau \varphi_t(s, X_s) ds + \int_t^\tau \langle (D\varphi)_s, b(s, X_s, u_s, Z_s) \rangle ds
$$

$$
+ \int_t^\tau \frac{1}{2} \mathrm{tr}\{\sigma\sigma^T(s, X_s, u_s)(D^2\varphi)_s\} ds + \int_t^\tau \langle (D\varphi)_s, \sigma(s, X_s, u_s) dW_s \rangle
\tag{6.2}
$$

$$
- \int_t^\tau \langle (D\varphi)_s, n(X_s) \rangle 1_{\partial\mathcal{O}_1}(X_s) d \, |\, \eta \, |_s.
$$

Now let us write

$$
h(s, X_s, Y_s, Z_s) = h(s, X_s, u_s, (D\varphi)_s \sigma(s, X_s, u_s)) + \langle \alpha_s, Z_s - \sigma^T(s, X_s, u_s)(D\varphi)_s \rangle;
$$

$$
b(s, X_s, Y_s, Z_s) = b(s, X_s, u_s, (D\varphi)_s \sigma(s, X_s, u_s)) + \beta_s(Z_s - \sigma^T(s, X_s, u_s)(D\varphi)_s),
\tag{6.3}
$$

where

$$
\begin{cases}
\alpha_s = \int_0^1 \frac{\partial h}{\partial z}(s, X_s, u_s, \mu Z_s + (1-\mu)\sigma^T(s, X_s, u_s)(D\varphi)_s)d\mu; \\[2em]
\beta_s = \int_0^1 \frac{\partial b}{\partial z}(s, X_s, u_s, \mu Z_s + (1-\mu)\sigma^T(s, X_s, u_s)(D\varphi)_s)d\mu.
\end{cases}
\tag{6.4}
$$

By assumption (A1), we see that $\alpha$ and $\beta$ are both bounded. Subtracting (6.2) from (6.1), using (6.3) and (6.4), and noting the facts that $u(t,x) = \varphi(t,x)$ and $u(\tau, X_\tau) \leq \varphi(\tau, X_\tau)$, we obtain

$$
0 \geq u(\tau, X_\tau) - \varphi(\tau, X_\tau) = \int_0^\tau \{ -\frac{\partial \varphi}{\partial t}(s, X_s) - F(s, X_s, u_s, (D\varphi)_s \sigma(s, X_s, u_s))
$$

$$
- \langle Z_s - (D\varphi)_s \sigma(s, X_s, u_s), \alpha_s - (D\varphi)_s \beta_s \rangle \} ds - (\zeta_\tau - \zeta_t)
\tag{6.5}
$$

$$
+ \int_t^\tau \langle Z_s - (D\varphi)_s \sigma(s, X_s, u_s), dW_s \rangle + \int_t^\tau \langle (D\varphi)_s, n(x_s) \rangle 1_{\partial \mathcal{O}_1}(X_s) d \mid \eta \mid_s.
$$

Now define $\theta_s \triangleq \alpha_s + (D\varphi)_s \beta_s$, $s \in [t, T]$. The uniform boundedness of $\alpha, \beta$, and $D\varphi$ renders the process $\Theta_r^t \triangleq \exp\{ -\int_t^s \langle \theta_s, dW_s \rangle - \frac{1}{2} \int_t^r \mid \theta_s \mid^2 ds \}$, $r \in [t, T]$, a $P$-martingale on $[t, T]$. Thus, by Girsanov's Theorem we can define a new probability measure $\widetilde{P}$ via $\frac{d\widetilde{P}}{dP} = \Theta_T^t$, so that $\widetilde{W}_r^t = W_r - W_t - \int_t^r (\alpha_s - D\varphi(s, X_s)\beta_s)ds$ is a $\widetilde{P}$-Brownian motion on $[t, T]$. Furthermore, since $(X, Y, Z) \in L_{\mathcal{F}}^2(\Omega; C([0,T] \times \mathbb{R}^{n+1})) \times L_{\mathcal{F}}^2(0, T; \mathbb{R}^d)$ by definition, the boundedness of $D\varphi$ and the uniform Lipschitz property of $\sigma$ imply that, for some constant $C > 0$ depending only on $K$ and $k_1$

$$
\widetilde{E} \left\{ \int_t^T \mid Z_s - (D\varphi)_s \sigma(s, X_s, u_s) \mid^2 \right\}^{\frac{1}{2}}
$$

$$
\leq C\widetilde{E} \left\{ \int_t^T [\mid Z_s \mid^2 + \mid X_s \mid^2 + \mid Y_s \mid^2] ds \right\}^{\frac{1}{2}}
$$

$$
\leq C\{E(\Theta_T^t)^2\}^{\frac{1}{2}} \left\{ E \int_t^T [\mid Z_s \mid^2 + \mid X_s \mid^2 + \mid Y_s \mid^2] ds \right\}^{\frac{1}{2}} < \infty.
$$

Consequentially, the integral

$$
M_r^t \triangleq \int_t^r \langle Z_s - (D\varphi)_s \sigma(s, X_s, u_s), \theta_s \rangle ds + \int_t^r \langle Z_s - (D\varphi)_s \sigma(s, X_s, u_s), dW_s \rangle
$$

$$
= \int_t^r \langle Z_s - (D\varphi)_s \sigma(s, X_s, u_s), d\widetilde{W}_s \rangle, \quad r \in [t, T]
$$

is a $\widetilde{P}$ martingale on $[t, T]$. Hence, by taking expectation $\widetilde{E}$ on both sides of (6.5), we obtain

$$
0 \geq \widetilde{E} \int_t^\tau \left\{ -\frac{\partial \varphi}{\partial t}(s, X_s) - F(s, X_s, u_s, (D\varphi)_s \sigma(s, X_s, u_s) \right\} ds
$$

$$
- \widetilde{E}(\zeta_\tau - \zeta_t) + \widetilde{E} \int_t^\tau \langle (D\varphi)_s, n(x_s) \rangle 1_{\partial O_1}(X_s) d \mid \eta \mid_s.
$$

(6.6)

We are now ready to prove that $u(t, x)$ is a viscosity subsolution. We shall use Definition 3.5 for this purpose. Let us consider the following two cases:

**Case 1:** $(t, x) \in [0, T) \times \text{int} O_1$. Note that $L(t, x) \leq u(t, x) = Y_t^{t, x} \leq U(t, x)$ by definition of FBSDER, and that $u(t, x) = L(t, x)$ would imply that the first inequality of (3.12) holds, so we assume without loss of generality that $u(t, x) > L(t, x)$. Now it is easy to see that it suffices to show that

$$
H(t, x; u, \varphi) \triangleq -\varphi_t(t, x) - F(t, x, u(t, x), D\varphi(t, x), D^2\varphi(t, x)) \leq 0, \qquad (6.7)
$$

since $u(t, x) - U(t, x) \leq 0$ is always true.

Suppose (6.7) is not true, namely, $H(t, x; u, \varphi) > 0$. Then by the continuity of all the functions involved, we can find a constant $\varepsilon_0 > 0$ such that

$$
\Theta(t, x) \triangleq (u(t, x) - L(t, x)) \vee H(t, x; u, \varphi) \vee d(x, \partial O_1) \geq \varepsilon_0.
$$

Define $\vartheta_s^t \triangleq \Theta(s, X_s), s \in [t, T]$ and define a stopping time $\tau_1$ by

$$
\tau_1 \triangleq \inf\{s > t \mid \vartheta_s^t \leq \frac{\varepsilon_0}{2}\} \wedge T \in \mathcal{M}_{t, T}.
$$

Replacing $\tau$ in (6.6) by $\tau_1$ we see that the first integral on the right side of (6.6) is no less than $\frac{\varepsilon_0}{2}\widetilde{E}(\tau_1 - t)$ and the second integral is zero because $X_s \notin \partial O_1$ for $s \in [t, \tau_1)$. To analyze the second term there, we define a process $V_s \triangleq 1_{[t, \tau_1)}(s)[u(s, X_s) - \varepsilon_0/2]$. It is easy to check that $V$ is an RCLL, $\{\mathcal{F}_t\}$-adapted process such that

$$
L(s, X_s(\omega)) \leq V_s(\omega) = u(s, X_s(\omega)) - \frac{\varepsilon_0}{2} < u(s, X_s(\omega)) \leq U(s, X_s(\omega)),
$$

for all $s \in [t, \tau_1(\omega))$, $P$-a.e. $\omega \in \Omega$. Moreover, for $s \notin [t, \tau_1(\omega))$, one has $V_s(\omega) = 0 \in [L(s, L_s(\omega)), U(s, X_s(\omega))]$ by assumption (A1). Therefore, by the definition of the reflecting process $\zeta$ (Definition 2.1 *(iv)*), we have

$$
0 \geq \int_t^{\tau_1} (Y_s - V_s) d\zeta_s = \int_t^{\tau_1} (u(s, X_s) - V_s) d\zeta_s = \frac{\varepsilon_0}{2}(\zeta_{\tau_1} - \zeta_t),
$$

which implies that $\widetilde{E}(\zeta_{\tau_1} - \zeta_t) \leq 0$, i.e. the second term on the right side of (6.6) is non-negative. Consequently, we have from (6.6), with $\tau = \tau_1$, that

$$
0 \geq \frac{\varepsilon_0}{2}\widetilde{E}(\tau_1 - t). \qquad (6.8)
$$

But since $\vartheta^t$ is continuous and $\vartheta^t_t > \frac{\varepsilon_0}{2}$, we have $\tau_1 > t$, $P$-a.s. (whence $\widetilde{P}$-a.s.), contradicting (6.8). Therefore (6.7) must hold.

**Case 2:** $(t,x) \in [0,T) \times \partial \mathcal{O}_1$. We are to show that the second inequality in (3.12) holds in this case. As argued in the previous case, we assume that $u(t,x) > L(t,x)$ and claim that either $\langle n(x), D\varphi(t,x) \rangle \le 0$ or $H(t,x;u,\varphi) \le 0$ must hold.

Again, suppose not. Then we can find a constant $\varepsilon_0 > 0$ such that

$$\widehat{\Theta}(t,x) \triangleq (u(t,x) - L(t,x)) \vee H(t,x;u,\varphi) \vee \langle n(x), D\varphi(t,x) \rangle \ge \varepsilon_0.$$

Define $\widehat{\vartheta}^t_s \triangleq \widehat{\Theta}(s, X_s)$, for $s \in [t,T]$ and $\tau_2 \triangleq \inf\{s > t \mid \widehat{\vartheta}^t_s \le \varepsilon_0/2\} \wedge T \in \mathcal{M}_{t,T}$. setting $\tau = \tau_2$ in (6.6) analyzing the first and second term on the right side as before, we see that (6.6) becomes

$$0 \ge \frac{\varepsilon_0}{2}\widetilde{E}(\tau_2 - t) + \widetilde{E}\int_t^{\tau_2} \langle n(X_s),(D\varphi)_s \rangle \mathbf{1}_{\partial \mathcal{O}_1}(X_s) d\,|\eta|_s. \tag{6.9}$$

But since $|\eta|$ is non-decreasing, we have

$$\widetilde{E}\int_t^{\tau_2} \langle n(X_s),(D\varphi)_s \rangle \mathbf{1}_{\partial \mathcal{O}_1}(X_s) d\,|\eta|_s$$

$$\ge \frac{\varepsilon_0}{2}\widetilde{E}\int_t^{\tau_2} \mathbf{1}_{\partial \mathcal{O}_1}(X_s) d\,|\eta|_s = \frac{\varepsilon_0}{2}\widetilde{E}(|\eta|_{\tau_2} - |\eta|_t) \ge 0.$$

Hence, from (6.9), with $\tau = \tau_2$, we obtain $\widetilde{E}(\tau_2 - t) \le 0$, a contradiction again. Thus the first part of Theorem 3.8 is proved.

# 7. Proof of Theorem 3.8 (II. Uniqueness of the Viscosity Solution)

In this section we prove the uniqueness of Theorem 3.8, thus we shall assume that $\sigma(t,s,y) = \sigma(t,x)$. Our proof borrows some ideas of Barles, Buckdahn and Pardoux [2], with necessary modifications. First we note that the uniqueness result will be a direct consequence of the following comparison theorem which is interesting in its own right.

**Theorem 7.1:** *Assume that all assumptions in Theorem 3.8 are in force. Let $u$ be a subsolution and $v$ a supersolution to the variational inequality (3.8) (or (3.9)), respectively, such that $u(t,x) \ge L(t,x)$ and $v(t,x) \le U(t,x)$ for all $(t,x) \in [0,T] \times \mathcal{O}_1$. Suppose further that both $u$ and $v$ are uniformly Lipschitz continuous in $x$ and continuous in $t$. Then $u(t,x) \le v(t,x)$ for all $(t,x) \in [0,T] \times \mathcal{O}_1$.*

In order to prove Theorem 7.1, we note that since both $u$ and $v$ are continuous, we only need to show $u \le v$ on $(0,T) \times \text{int}\mathcal{O}_1$. Let us define for each $\alpha > 0$ a subset of $\mathcal{O}_1$:

$$\mathcal{O}_1^\alpha \triangleq \{x \in \mathcal{O}_1 \mid d(x, \partial \mathcal{O}_1) \ge \alpha\}, \tag{7.1}$$

and choose $\alpha_0 > 0$ such that $\mathcal{O}_1^\alpha \ne \emptyset$ for all $0 \le \alpha \le \alpha_0$. Then, it suffices to prove that for every $0 \le \alpha \le \alpha_0$, $u(t,x) \le v(t,x)$ for all $(t,x) \in (0,T) \times \mathcal{O}_1^\alpha$. In what follows,

we denote $w = u - v$, and define

$$
\begin{aligned}
F_{u,v}(t,x,r,p,q,X) &\triangleq -p - \tfrac{1}{2}\mathrm{tr}\{\sigma\sigma^T(t,x)X\} - \langle b(t,x,u(t,x),0),q\rangle \\
&\quad - \widetilde{K}[\,|\,r\,| + |\,q\,| \cdot |\,\sigma(t,x)\,|\,],
\end{aligned}
\tag{7.2}
$$

where $\widetilde{K} > 0$ is some constant depending only on the Lipschitz constants of the functions $b, h, u, v$ in $(x, y, z)$. The following key lemma is adapted from Barles, Buckdahn, and Pardoux [2]. Since our situation is more involved, we shall provide a sketch of proof for completeness.

**Lemma 7.2:** *Suppose that the assumptions of Theorem 3.8 hold, and $u, v$ are those in Theorem 7.1. Then for any $0 \le \alpha \le \alpha_0$, the function $w = u - v$ is a viscosity subsolution to the following variational inequality:*

$$
\begin{cases}
0 = \min\{w, F_{u,v}(t,x,w,w_t,Dw,D^2 w)\} & \forall (t,x) \in [0,T) \times \mathcal{O}_1^\alpha, \\
w(T,x) = 0, & \forall x \in \mathcal{O}_1^\alpha.
\end{cases}
\tag{7.3}
$$

**Proof:** Fix $\alpha \in [0, \alpha_0]$ and $(t_0, x_0) \in [0, T) \times \mathcal{O}_1^\alpha$. Let $\varphi \in C^2([0,T] \times \bar{\mathcal{O}}_1)$ be such that $(t_0, x_0)$ is a maximum point of $w - \varphi$ and that $w(t_0, x_0) = \varphi(t_0, x_0)$. Modifying $\varphi$ if necessary, we may assume without loss of generality that $(t_0, x_0)$ is a strict, global maximum point of $w - \varphi$, and that $D\varphi$ is uniformly bounded, thanks to the Lipschitz property of $u$ and $v$ (whence $w$). Our purpose is to show the following inequality:

$$
\min\{w(t_0,x_0), F_{u,v}(t_0,x_0,w,\varphi_t,D\varphi,D^2\varphi)\} \le 0,
\tag{7.4}
$$

It is easy to see that (by Definition 3.6 or 3.7), if $u$ is a subsolution and $v$ is a supersolution, then $u(t,x) \le U(t,x)$, and $L(t,x) \le v(t,x)$. Combining with the assumption of the lemma, we have $l(t,x) \le u(t,x)$, $v(t,x) \le U(t,x)$, $\forall (t,x) \in [0,T] \times \mathcal{O}_1^\alpha$. Thus if $(t_0, x_0)$ is such that either $u(t_0, x_0) = L(t_0, x_0)$ or $v(t_0, x_0) = U(t_0, x_0)$, then $w(t_0, x_0) \le 0$ and (7.4) holds. Thus we assume without loss of generality that

$$
[u(t_0,x_0) - L(t_0,x_0)] \cdot [U(t_0,x_0) - v(t_0,x_0)] \ne 0.
\tag{7.5}
$$

Let us define for a given $\varepsilon > 0$, a function $\psi_\varepsilon(t,x,y) = u(t,x) - v(t,y) - \dfrac{|\,x-y\,|^2}{\varepsilon^2} - \varphi(t,x)$. Choose $R > 0$ large enough and define $\mathcal{O}_1^{\alpha,R} = B_R \cap \mathcal{O}_1^\alpha$ so that $(t_0, x_0) \in \mathcal{O}_1^{\alpha,R}$, where $B_R$ is the open ball in $\mathbb{R}^n$ centered at origin with radius $R$. Let

$$
(t^\varepsilon, x^\varepsilon, y^\varepsilon) \in \mathrm{argmax}_{[0,T] \times \mathcal{O}_1^{\alpha,R}} \psi_\varepsilon(t,x,y).
$$

Applying Proposition 3.7 of Crandall, Ishii and Lions [4], we can easily conclude that

(i)    $(t_\varepsilon, x_\varepsilon, y_\varepsilon) \to (t_0, x_0, x_0)$, as $\varepsilon \to 0$;

(ii)    $\dfrac{|\,x_\varepsilon - y_\varepsilon\,|^2}{\varepsilon^2}$ is bounded and tends to zero as $\varepsilon \to 0$.

Now, fixing $\varepsilon > 0$, and then applying Theorem 8.3 in Crandall, Ishii and Lions [4], we obtain that, for any $\delta > 0$ there exists $(X^\delta, Y^\delta) \in \mathcal{S}(n) \times \mathcal{S}(n)$, $c^\delta \in \mathbb{R}^n$ such that

$$(c^\delta + \varphi_t(t^\varepsilon, x^\varepsilon), P_\varepsilon + D\varphi(t^\varepsilon, x^\varepsilon), X^\delta) \in \bar{\mathscr{P}}^{2,+} u(t^\varepsilon, x^\varepsilon);$$

$$(c^\delta, P_\varepsilon, Y^\delta) \in \bar{\mathscr{P}}^{2,-} v(t^\varepsilon, y^\varepsilon),$$

(7.6)

and

$$\begin{pmatrix} X^\delta & 0 \\ 0 & -Y^\delta \end{pmatrix} \le A + \delta A^2,$$

(7.7)

where $P_\varepsilon = \dfrac{2(x^\varepsilon - y^\varepsilon)}{\varepsilon^2}$, and $A = \begin{pmatrix} D^2\varphi(t^\varepsilon, x^\varepsilon) + 2/\varepsilon^2 & -2/\varepsilon^2 \\ -2/\varepsilon^2 & 2/\varepsilon^2 \end{pmatrix}$. An easy

computation shows, that for $\gamma \triangleq 1 + \dfrac{4\delta}{\varepsilon^2}$,

$$A + \delta A^2 = \begin{pmatrix} \delta(D_x^2\varphi)^2 + \gamma D_x^2\varphi & -\dfrac{2\delta}{\varepsilon^2} D_x^2\varphi \\ -\dfrac{2\delta}{\varepsilon^2} D_x^2\varphi & 0 \end{pmatrix} + \dfrac{2}{\varepsilon^2}\left(1 + \dfrac{4\delta}{\varepsilon^2}\right)\begin{pmatrix} I & -I \\ -I & I \end{pmatrix}$$

$$= \begin{pmatrix} D_x^2\varphi & 0 \\ 0 & 0 \end{pmatrix} + \dfrac{2}{\varepsilon^2}\begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + \delta M(\varepsilon),$$

(7.8)

where

$$M(\varepsilon) = \begin{pmatrix} (D_x^2\varphi)^2 + \dfrac{4}{\varepsilon^2} D_x^2\varphi & -\dfrac{2\delta}{\varepsilon^2} D_x^2\varphi \\ -\dfrac{2\delta}{\varepsilon^2} D_x^2\varphi & 0 \end{pmatrix} + \dfrac{8}{\varepsilon^4}\begin{pmatrix} I & -I \\ -I & I \end{pmatrix}.$$

(7.9)

Let us now make use of Definition 3.6 of viscosity solution. Thus for the subjet and superjet shown in (7.6) we have

$$\begin{cases} (u - L) \wedge \{(u - U) \vee [-(c^\delta + \varphi_t) - F(t, x, u, P_\varepsilon + D\varphi, X^\delta)]\} \mid_{(t,x) = (t^\varepsilon, x^\varepsilon)} \le 0, \\ (v - U) \vee \{(v - L) \wedge [-c^\delta - F(t, y, v, P_\varepsilon, Y^\delta)]\} \mid_{(t,y) = (t^\varepsilon, y^\varepsilon)} \ge 0. \end{cases}$$

(7.10)

Since $(t_0, x_0)$ satisfies (7.5) and $(t^\varepsilon, x^\varepsilon, y^\varepsilon) \to (t_0, x_0, x_0)$, we may assume that $(t^\varepsilon, x^\varepsilon)$ and $(t^\varepsilon, y^\varepsilon)$ also satisfy (7.50, thanks to the continuity of the functions $u, v, L, U$. Thus we must have $u(t^\varepsilon, x^\varepsilon) > L(t^\varepsilon, x^\varepsilon)$ and $v(t^\varepsilon, y^\varepsilon) < U(t^\varepsilon, y^\varepsilon)$ and (710) leads to

$$\begin{cases} -(c^\delta + \varphi_t(t^\varepsilon, x^\varepsilon)) - F(t^\varepsilon, x^\varepsilon, u, P_\varepsilon + D\varphi, X^\delta) \le 0, \\ -c^\delta - F(t^\varepsilon, y^\varepsilon, v, P_\varepsilon, Y^\delta) \ge 0. \end{cases}$$

(7.11)

Subtracting the second inequality from the first one in (7.11) and recalling Definition 3.7 of $F$, we obtain

$$0 \ge -\varphi_t(t^\varepsilon, x^\varepsilon) - \{\tfrac{1}{2}\mathrm{tr}[\sigma\sigma^T(t^\varepsilon, x^\varepsilon)X^\delta] - \tfrac{1}{2}\mathrm{tr}[\sigma\sigma^T(t^\varepsilon, x^\varepsilon)Y^\delta]\}$$

$$- \{\langle b(t^\varepsilon, x^\varepsilon, u(t^\varepsilon, x^\varepsilon), P_\varepsilon + D\varphi(t^\varepsilon, x^\varepsilon))\sigma(t^\varepsilon, x^\varepsilon)), P_\varepsilon + D\varphi(t^\varepsilon, x^\varepsilon)\rangle$$

$$- \langle b(t^\varepsilon, y^\varepsilon, v(t^\varepsilon, y^\varepsilon), P_\varepsilon \sigma(t^\varepsilon, y^\varepsilon)), P_\varepsilon \rangle \} \tag{7.12}$$

$$- \{ h(t^\varepsilon, x^\varepsilon, u(t^\varepsilon, x^\varepsilon), (P_\varepsilon + D\varphi(t^\varepsilon, x^\varepsilon))\sigma(t^\varepsilon, x^\varepsilon))$$

$$- h(t^\varepsilon, y^\varepsilon, v(t^\varepsilon, y^\varepsilon), P_\varepsilon \sigma(t^\varepsilon, y^\varepsilon)) \}$$

$$\triangleq - \varphi_t(t^\varepsilon, x^\varepsilon) - I_1^\varepsilon(\delta) - I_2^\varepsilon(\delta) - I_3^\varepsilon(\delta),$$

where $I_1, I_2, I_3$ are three $\{\dots\}$'s on the right side of (7.12) respectively. We now estimate $I_i$, $i = 1, 2, 3$ separately. To simplify notations, we hereafter denote $f^{\varepsilon, x} = f(t^\varepsilon, x^\varepsilon)$ and $f^{\varepsilon, y} = f(t^\varepsilon, y^\varepsilon)$ for any given function $f(t, x)$; and by $\Theta_\varepsilon$ any quantity that satisfies

$$|\Theta_\varepsilon| \leq \widetilde{K} \left\{ |x^\varepsilon - y^\varepsilon| + \frac{|x^\varepsilon - y^\varepsilon|}{\varepsilon^2} \right\},$$

for some generic constant $\widetilde{K}$ that depends only on $b, h, \sigma g, u$ and $v$, which may vary from line to line. Let $\sigma_i$ be the $i$th column vector of the matrix $\sigma$. From (7.7), (7.8) and (7.9), we have

$$I_1^\varepsilon(\delta) = \tfrac{1}{2} \{ \mathrm{tr}[\sigma^{\varepsilon, x}(\sigma^{\varepsilon, x})^T X^\delta] - \mathrm{tr}[\sigma^{\varepsilon, y}(\sigma^{\varepsilon, y})^T Y^\delta]$$

$$= \tfrac{1}{2} \sum_{i=1}^{d} |\langle \sigma_i^{\varepsilon, x}, X^\delta \sigma_i^{\varepsilon, x} \rangle - \langle \sigma_i^{\varepsilon, y}, Y^\delta \sigma_i^{\varepsilon, y} \rangle]$$

$$\leq \tfrac{1}{2} \mathrm{tr}[\sigma^{\varepsilon, x}(\sigma^{\varepsilon, x})^T D_x^2 \varphi^{\varepsilon, x}] + \tfrac{1}{\varepsilon^2} |\sigma^{\varepsilon, x} - \sigma^{\varepsilon, y}|^2 + \tfrac{\delta}{2} R^\varepsilon(\sigma^{\varepsilon, x}, \sigma^{\varepsilon, y}) \tag{7.13}$$

$$\leq \tfrac{1}{2} \mathrm{tr}[\sigma^{\varepsilon, x}(\sigma^{\varepsilon, x})^T D_x^2 \varphi_x^2 \varphi^{\varepsilon, x}] + \tfrac{\delta}{2} R^\varepsilon(\sigma^{\varepsilon, x}, \sigma^{\varepsilon, y}) + \Theta_\varepsilon,$$

where $R^\varepsilon(\sigma^{\varepsilon, x}, \sigma^{\varepsilon, y}) \triangleq \langle \begin{pmatrix} \sigma^{\varepsilon, x} \\ \sigma^{\varepsilon, y} \end{pmatrix}, M(\varepsilon) \begin{pmatrix} \sigma^{\varepsilon, x} \\ \sigma^{\varepsilon, y} \end{pmatrix} \rangle$ and $M(\varepsilon)$ is given by (7.9).

Next it is fairly easy to check

$$I_2^\varepsilon(\delta) = \langle b(t^\varepsilon, x^\varepsilon, u^{\varepsilon, x}, (P_\varepsilon + D\varphi^{\varepsilon, x})\sigma^{\varepsilon, x}), P_\varepsilon + D\varphi^{\varepsilon, x} \rangle$$

$$- \langle b(t^\varepsilon, y^\varepsilon, v^{\varepsilon, y}, P_\varepsilon \sigma^{\varepsilon, y}), P_\varepsilon \rangle \tag{7.14}$$

$$= \langle b(t^\varepsilon, x^\varepsilon, u^{\varepsilon, x}, 0), D\varphi^{\varepsilon, x} \rangle + \langle \Delta b_1, D\varphi^{\varepsilon, x} \rangle + \langle \Delta b_2, P_\varepsilon \rangle$$

where

$$\Delta b_1 \triangleq b(t^\varepsilon, x^\varepsilon, u^{\varepsilon, x}, (P_\varepsilon + D\varphi^{\varepsilon, x})\sigma^{\varepsilon, x}) - b(t^\varepsilon, x^\varepsilon, u^{\varepsilon, x}, 0) \leq K_b |P_\varepsilon + D\varphi^{\varepsilon, x}| |\sigma^{\varepsilon, x}|,$$

$$\Delta b_2 \triangleq b(t^\varepsilon, x^\varepsilon, u^{\varepsilon, x}, (P_\varepsilon + D\varphi^{\varepsilon, x})\sigma^{\varepsilon, x}) - b(t^\varepsilon, y^\varepsilon, v^{\varepsilon, y}, P_\varepsilon \sigma^{\varepsilon, y})$$

$$\leq K_b \{ |u^{\varepsilon, x} - v^{\varepsilon, y}| + |D\varphi^{\varepsilon, x}| |\sigma^{\varepsilon, x}| \} + \Theta_\varepsilon,$$

where $K_b$ denotes some generic constant depending only on $b$. Consequently, from (7.14) we get

$$I_2^\varepsilon(\delta) \leq \langle b(t^\varepsilon, x^\varepsilon, u^{\varepsilon, x}, 0), D\varphi^{\varepsilon, x}\rangle + K_b\{\,|\,P_\varepsilon + D\varphi^{\varepsilon, x}\,|\,|\,\sigma^{\varepsilon, x}\,|\,|\,D\varphi^{\varepsilon, x}\,|$$
$$+ |\,P_\varepsilon\,|\,\{\,|\,u^{\varepsilon, x} - v^{\varepsilon, y}\,| + |\,D\varphi^{\varepsilon, x}\,|\,|\,\sigma^{\varepsilon, x}\,|\,\} + \Theta_\varepsilon\}. \tag{7.15}$$

Similarly, we have, for some constant $K_h$ depending on $h$,

$$I_3^\varepsilon(\delta) \leq |\,h(t^\varepsilon, x^\varepsilon, u^{\varepsilon, x}, (P_\varepsilon + D\varphi^{\varepsilon, x})\sigma^{\varepsilon, x}) - h(t^\varepsilon, y^\varepsilon, v^{\varepsilon, y}, P_\varepsilon \sigma^{\varepsilon, y})\,|$$
$$\leq K_h\{\,|\,u^{\varepsilon, x} - v^{\varepsilon, y}\,| + |\,D\varphi^{\varepsilon, x}\,|\,|\,\sigma^{\varepsilon, x}\,|\,\}\} + \Theta_\varepsilon. \tag{7.16}$$

Now by definition of $(t^\varepsilon, x^\varepsilon)$, and the fact that $w^{\varepsilon, x} - \varphi^{\varepsilon, x} \leq w(t_0, x_0) - \varphi(t_0, x_0) = 0$, we have

$$0 = w(t_0, x_0) - \varphi(t_0, x_0) = \psi_\varepsilon(t_0, x_0, x_0) \leq \psi(t^\varepsilon, x^\varepsilon, y^\varepsilon)$$

$$= (w^{\varepsilon, x} - \varphi^{\varepsilon, x}) + (v^{\varepsilon, x} - v^{\varepsilon, y}) - \frac{|\,x^\varepsilon - y^\varepsilon\,|^2}{\varepsilon^2} \leq (v^{\varepsilon, x} - v^{\varepsilon, y}) - \frac{|\,x^\varepsilon - y^\varepsilon\,|^2}{\varepsilon^2},$$

therefore, $|\,P_\varepsilon\,| = 2\dfrac{|\,x^\varepsilon - y^\varepsilon\,|}{\varepsilon^2} \leq \dfrac{2[v^{\varepsilon, x} - v^{\varepsilon, y}]}{|\,x^\varepsilon - y^\varepsilon\,|} \leq 2K_v < \infty$. Furthermore, by the uni form boundedness of $D\varphi$ we also have $|\,P_\varepsilon + D\varphi^{\varepsilon, x}\,| \leq K_{u, v}$. Therefore, we derive from (7.15) and (7.16) that

$$I_2^\varepsilon(\delta) + I_3^\varepsilon(\delta) \leq \langle b(t^\varepsilon, x^\varepsilon, u^{\varepsilon, x}, 0), D\varphi^{\varepsilon, x}\rangle$$

$$+ \widetilde{K}\{\,|\,\sigma^{\varepsilon, x}\,|\,|\,D\varphi^{\varepsilon, x}\,| + |\,u^{\varepsilon, x} - v^{\varepsilon, y}\,|\,\} + \Theta_\varepsilon.$$

Combining the above with (7.13), we obtain

$$\lim_{\varepsilon \to 0} \lim_{\delta \to 0} \sum_{i=1}^3 I_i^\varepsilon(\delta) \leq \lim_{\varepsilon \to 0} \left\{\frac{1}{2}\mathrm{tr}[\sigma^{\varepsilon, x}(\sigma^{\varepsilon, x})^T D_x^2 \varphi^{\varepsilon, x}]\,|\,\langle b(t^\varepsilon, x^\varepsilon, u^{\varepsilon, x}, 0), D\varphi^{\varepsilon, x}\rangle\right.$$

$$+ \widetilde{K}\{\,|\,\sigma^{\varepsilon x}\,|\,|\,D\varphi^{\varepsilon, x}\,| + |\,u^{\varepsilon, x} - v^{\varepsilon, y}\,|\,\} + \Theta_\varepsilon \Big\}$$

$$= \frac{1}{2}\mathrm{tr}[\sigma\sigma^T(t_0, x_0)D_x^2\varphi(t_0, x_0)] + \langle b(t_0, x_0, u(t_0, x_0), 0), D\varphi(t_0, x_0)\rangle$$

$$+ \widetilde{K}\{\,|\,\sigma(t_0, x_0)\,|\,|\,D\varphi(t_0, x_0)\,| + |\,w(t_0, x_0)\,|\,\}.$$

Therefore, first letting $\delta \to 0$ and then letting $\varepsilon \to 0$ in (7.12) we obtain

$$0 \geq -\varphi_t(t_0, x_0) - F_{u, v}(t_0, x_0, w, D\varphi, D^2\varphi),$$

and hence (7.4) holds. Finally, noting that $w(T, w) \leq 0$ by the definitions of subsolution and supersolution, the lemma is proved. $\quad\square$

**Proof of Theorem 7.1:** We need only show that $w \leq 0$ on $[0, T] \times \mathbb{O}_1^\alpha$ for any $\alpha \in [0, \alpha_0]$. Consider the function $\Psi(t, x) = \exp\{(C_1(T - t) + a)\langle x\rangle\}$, where $\langle x\rangle = (1 + |\,x\,|^2)^{1/2}$, and $A, C_1 > 0$ are to be determined. Let $\bar{t} = (T - A/C_1)^+$. It is easy to check that

$$
\left\{
\begin{aligned}
&D\langle x\rangle = \frac{x}{\langle x\rangle}, \quad D^2\langle x\rangle = \frac{1}{\langle x\rangle} - \frac{x\otimes x}{\langle x\rangle^3}, \text{ where } x\otimes x = xx^T; \\[4pt]
&|D\langle x\rangle| \le 1, \quad \|D^2\langle x\rangle\| \le \frac{1}{\langle x\rangle} + \frac{|x|^2}{\langle x\rangle^3} \le 2; \\[4pt]
&D\Psi = \Psi[C_1(T-t)+A]D\langle x\rangle; \\[4pt]
&D^2\Psi = \Psi\{(C_1(T-t)+A)^2(Dx\otimes Dx) + (C_1(T-t)+A)D^2x\}, \\[4pt]
&|D\Psi| \le 2A\Psi, \quad \|D^2\Psi\| \le 4(A^2+A)\Psi, \ \forall t \in [\bar t, T]\times \mathbb{O}_1^\alpha.
\end{aligned}
\right.
\tag{7.17}
$$

We claim that for any fixed $A$, there exists $C_1 > 0$ sufficiently large such that

$$
\min\{\Psi, F_{u,v}(T,x,\Psi,\Psi_t,D\Psi,D^2\Psi)\} > 0, \quad \forall(t,x)\in[\bar t, T]\times\mathbb{O}_1^\alpha.
\tag{7.18}
$$

In fact, some direct computations using the facts listed in (7.17) imply

$$
F_{u,v}(t,x,\Psi,\Psi_t,D\Psi,D^2\Psi)
$$

$$
\ge C_1\Psi\langle x\rangle - \tfrac12 K_\sigma \|D^2\Psi\| - K_b |D\Psi| - \widetilde K \Psi - \widetilde K K_\sigma |D\Psi|
$$

$$
\ge \Psi\Big\{C_1 - \frac{K_\sigma}{2}\cdot 4(A^2+A) - 2K_b A - \widetilde K - \widetilde K K_\sigma\cdot 2A\Big\} > 0,
$$

provided $C_1$ is large enough and $t\in[\bar t_1, T]$, for some constant $K_\sigma > 0$, depending only on $\sigma$. Thus (7.18) holds because $\Psi > 0$ is always true.

Now let us choose $A$ and $C_1$ such that $\lim_{|x|\to+\infty}|w(t,x)|e^{-A|x|} = 0$ and $\Psi$ satisfies (7.18). Define for $\alpha,\varepsilon > 0$ and $0\le s < t\le T$,

$$
M^\alpha_{s,t}(\varepsilon) = \max_{[s,t]\times\mathbb{O}_1^\alpha}\{w(t,x) - \varepsilon\Psi(t,x)\}e^{-\widetilde K(T-t)},
$$

with $\widetilde K$ as in (7.2). We claim that $M^\alpha_{\bar t, T}(\varepsilon)\le 0$, $\forall\alpha,\varepsilon > 0$. Suppose not. Then $M^\alpha_{\bar t, T}(\varepsilon) > 0$, for some $\alpha > 0$ and $\varepsilon > 0$. Let $(t_0,x_0)\in\operatorname{argmax}\{w(t,x) - \varepsilon\Phi(t,x)\}e^{-wtK(T-t_0)}$. (Note that this is possible even if $\mathbb{O}_1$ is unbounded, since for $|x|$ sufficiently large we have $w(t,x) - \varepsilon\Psi(t,x) < 0$.) Define $\varphi(t,x) = \varepsilon\Psi(t,x) + M^\alpha(\varepsilon)e^{\widetilde K(T-t)}$. Then $\varphi\in C^{1,2}$, $\varphi(t_0,x_0) = w(t_0,x_0)$ and $w(t,x) - \varphi(t,x)\le 0$ for all $(t,x)\in[\bar t, T]\times\mathbb{O}_1^\alpha$.

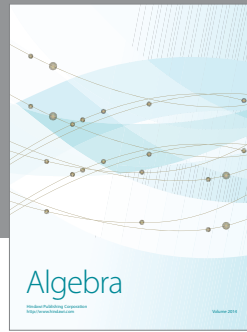Note that $w$ is a viscosity subsolution to (7.3) by Lemma 7.2, and that $\omega(t_0,x_0) > \varepsilon\Psi(t_0,x_0)e^{\widetilde K(T_0-t)} > 0$ by definition, we must have $F_{u,v}(t_0,x_0,w,D\varphi,D^2\varphi)\le 0$ by the definition of a subsolution. But on the other hand, one can easily check that

$$
F_{u,v}(t_0,x_0,w,\varphi_t,D\varphi,D^2\varphi) = \varepsilon F_{u,v}(t_0,x_0,\Psi,\Psi_t,D\Psi,D^2\Psi) > 0,
$$

thanks to (7.17). This contradiction proves the claim. Consequently, $w(t,x) - \varepsilon\Psi(t,x)\le 0$ for all $(t,x)\in[\bar t, T]\times\mathbb{O}_1^\alpha$. Letting $\varepsilon\to 0$, we have $w(t,x)\le 0$ for all $(t,x)\in[\bar t, T]\times\bar{\mathbb{O}}_1^\alpha$. Repeating this argument for $[\bar t_i, \bar t_{i-1}]$, for $i = 1,2,\dots$ if necessary, where $\bar t_{i+1} = (\bar t_{i-1} - A/C_1)^+$, $\bar t_0 = T$, $\bar t_1 = \bar t$, one shows that $w(t,x)\le 0$ for all $(t,x)\in[0,T]\times\mathbb{O}_1^\alpha$, $\alpha\le\alpha_0$ proving the theorem. $\qquad\square$

## References

[1]   Antonelli, F., Backward-forward stochastic differential equations, *The Annals of Appl. Probab.* **3** (1993), 773-793.

[2]   Barles, G., Buckdahn, R. and Pardoux, E., BSDEs and integral-partial differential equations, *Stoch. and Stoch. Reports* **60** (1997), 57-83.

[3]   Buckdahn, R. and Hu, Y., Hedging contingent claims for a large investor in an incomplete market, *Adv. in Appl. Probab.* **30** (1998), 239-255.

[4]   Crandall, M., Ishii, H. and Lions, P.L., User's guide to the viscosity solutions of second order partial differential equations, *Bull. AMS* **27** (1992), 1-67.

[5]   Cvitanić, J. and Karatzas, I., Backward stochastic differential equations with reflection and Dynkin games, *The Annals of Probab.* **24** (1996), 2024-2056.

[6]   Cvitanić, J. and Ma, J., Hedging options for a large investor and forward-backward SDEs, *The Annals of Appl. Probab.* **6** (1996), 370-398.

[7]   Duffie, D. and Epstein, L.G., Stochastic differential utility, *Econometrica* **60** (1992), 353-394 (Appendix with C. Skiadas).

[8]   El Karoui, N., Kapoudjian, C., Pardoux, E., Peng, S. and Quenez, M.C., Reflected solutions of backward SDEs and related obstacle problems for PDEs, *Ann. Probab.* **25**:2 (1997), 702-737.

[9]   Fleming, W. and Soner, H.M., *Controlled Markov Processes and Viscosity Solutions*, Springer-Verlag, Berlin 1993.

[10]  Gegout-Petit, A. and Pardoux, E., Equations différentielles stochastiques rétrogrades réfléchies dans un convexe, *Stoch. and Stoch. Reports* **57** (1996), 111-128.

[11]  Hamadene, S., Lepeltier, J.P. and Matoussi, A., Double barriers reflected backward SDEs with continuous coefficient, (1996), preprint.

[12]  Hu, Y. and Peng, S., Solution of a forward-backward stochastic differential equation, *Probab. Theory Relat. Fields* **103** (1995), 273-283.

[13]  Lions, P.L. and Sznitman, A.S., Stochastic differential equations with reflecting boundary conditions, *Comm. Pure Appl. Math.* **37** (1984), 511-537.

[14]  Ma, J., Protter, P. and Yong, J., Solving forward-backward stochastic differential equations explicitly - A four step scheme, *Probab. Theory Relat. Fields* **98** (1994), 339-359.

[15]  Pardoux, E. and Peng, S., Adapted solution of a backward stochastic differential equation, *Systems Control Lett.* **14** (1990), 55-61.

[16]  Pardoux, E. and Rascanu, A., Backward stochastic differential equations with subdifferential operator and related variational inequalities, *Stoch. Process. Appl.* **76**:2 (1998), 191-215.

[17]  Pardoux, E. and Tang, S., The study of forward-backward stochastic differential equations and its applications to quasilinear PDEs, (1996), preprint.

[18]  Yong, J., Finding adapted solutions of forward-backward stochastic differential equations-Method of continuation, *Probab. Theory Relat. Fields* **107** (1997), 537-572.