

Approximate Solvability of Forward–Backward Stochastic Differential Equations*

J. Ma¹ and J. Yong²

¹Department of Mathematics, Purdue University,
West Lafayette, IN 47907-1395, USA
majin@math.purdue.edu

²Laboratory of Mathematics for Nonlinear Sciences, Department of Mathematics,
and Institute of Mathematical Finance, Fudan University,
Shanghai 200433, People's Republic of China

Abstract. The solvability of forward–backward stochastic differential equations (FBSDEs for short) has been studied extensively in recent years. To guarantee the existence and uniqueness of adapted solutions, many different conditions, some quite restrictive, have been imposed. In this paper we propose a new notion: the *approximate solvability* of FBSDEs, based on the method of optimal control introduced in our primary work [15]. The approximate solvability of a class of FBSDEs is shown under mild conditions; and a general scheme for constructing approximate adapted solutions is proposed.

Key Words. Forward–backward stochastic differential equation, Approximate solvability, Approximate adapted solution, Nodal set.

AMS Classification. 60H15, 35R60, 34F05, 93E20.

* The first author was supported in part by the Office of Naval Research Grant #N00014-96-1-0262 and the U.S. National Science Foundation Grant #9971720. The second author's work was supported in part by the NSFC, under Grant #79790130, the National Distinguished Youth Science Foundation of China under Grant #19725106, the Chinese Education Ministry Science Foundation under Grant #2000024605, and the Cheung Kong Scholars Programme. Part of this work was completed when this author was visiting the Department of Mathematics, Purdue University.

1. Introduction

In recent years the following type of stochastic differential equations (SDEs) has been brought into attention:

$$\begin{cases} dX(t) = b(t, X(t), Y(t), Z(t)) dt + \sigma(t, X(t), Y(t), Z(t)) dW(t), \\ dY(t) = h(t, X(t), Y(t), Z(t)) dt + Z(t) dW(t), & t \in [0, T], \\ X(0) = x, & Y(T) = g(X(T)). \end{cases} \quad (1.1)$$

Here, processes X , Y and Z are unknown, taking values in \mathbb{R}^n , \mathbb{R}^m and $\mathbb{R}^{m \times d}$, respectively, functions b , σ , h and g are all given and deterministic with appropriate dimensions, and W is a d -dimensional standard Brownian motion. Note that an initial value is specified for the component X , whereas a terminal value is specified for the component Y . It is conceivable that the second equation will have to “evolve” in the opposite time direction as opposed to that of the first one. Therefore (1.1) has been called a *forward-backward stochastic differential equation* (FBSDE for short). Our purpose is to find a solution (X, Y, Z) that is *adapted* to the (forward) filtration $\{\mathcal{F}_t\}$ generated by the Brownian motion W . We should note that it is the extra process Z that makes finding such an adapted solution possible.

The study of FBSDEs of this kind can be traced back to early 1970s, mainly in the context of stochastic control (see, e.g., [3] and [2]). The recent development was ignited by the seminal work of Pardoux and Peng [16] on *backward stochastic differential equations* (BSDE for short). Since the first paper on FBSDEs by Antonelli [1], many methods have been proposed for various types of FBSDEs over an arbitrary time duration $[0, T]$ (see [15], [14], [11], [18], [20], [21], and [17]); and applications, mainly in mathematical finance, have also been found in recent years (see [8], [6], [4], and also [7]). However, compared with the pure forward or pure backward SDEs, all the existing works require conditions on the coefficients that are restrictive and non-standard in different ways, at times they are quite painful to verify. In fact, as a (stochastic) two-point boundary-value problem, the theory seems to welcome methods that offer simpler ways of accessing the adapted solution.

In this paper we revisit a method proposed in our earlier work on the subject [15]: the method of optimal control; and introduce a new notion of *approximate solvability* for FBSDEs, by which we mean that for any $\varepsilon > 0$, there exist $\{\mathcal{F}_t\}$ -adapted processes X , Y and Z , such that (1.1) is satisfied except that the last condition is replaced by the following:

$$E|Y(T) - g(X(T))| < \varepsilon. \quad (1.2)$$

Such a notion is particularly useful in applications where the terminal condition in (1.1) is the main concern and some other approximations (e.g., numerical simulations) are involved, because then (1.2) becomes a practical way of saying “ $Y(T) = g(X(T))$ ” if ε is small enough. Our purpose is to use the equivalent relation between the solvability of an FBSDE and a stochastic control problem to construct approximate solutions of FBSDE (1.1) by choosing appropriate initial states and control processes. In the case when the coefficients b , σ and h are independent of Z , we do this by attacking again the “nodal set” of the value function to the control problem as we did in [15], but in a general higher-dimensional setting. We prove the non-emptiness of the nodal set under

standard conditions assuming neither the non-degeneracy of the forward diffusion nor the monotonicity conditions on the coefficients, hence we considerably extend the result of [15].

A topic related to approximate solvability is numerical feasibility. We note that besides the techniques of constructing *nearly optimal controls* that will be naturally involved in a numerical scheme, the difficulty also lies in finding the correct initial state. In an ideal situation, one can find a manifold (or “nodal surface”) on which the value function vanishes, and choose initial values from there (see [15] and [14]). However, in the general case when σ is allowed to be degenerate, finding such a nodal set explicitly seems to be quite remote. In this paper we show how to overcome this difficulty. In Section 5 we essentially construct an *approximate nodal surface*; and in Section 6 we give a more general scheme.

This paper is organized as follows. In Section 2 we introduce the optimal control problem associated with our FBSDE (1.1). In Section 3 *approximate solvability* is carefully discussed. An equivalence relation between the approximate solvability and the non-emptiness of the nodal set of the value function is established. In Section 4 we introduce an approximation for the value function. Section 5 is devoted to the synthetic analysis of a special class of FBSDEs. Finally in Section 6 we propose a general scheme of constructing approximate solutions with the help of approximate value functions.

2. Formulation of the Optimal Control Problem

Throughout this paper we assume that $(\Omega, \mathcal{F}, P; \{\mathcal{F}_t\}_{t \geq 0})$ is a complete filtered probability space on which is defined a d -dimensional standard Brownian motion $W = \{W(t) : t \in [0, T]\}$. We further assume that the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ is generated by W , augmented by all the P -null sets in \mathcal{F} , and thus satisfies the “usual hypotheses” [19]. We denote by $|\cdot|$ and $\langle \cdot, \cdot \rangle$ the usual Euclidean norm and inner product in \mathbb{R}^n , respectively. Also, let $\mathbb{R}^{m \times d}$ be the Euclidean space consisting of all $(m \times d)$ -matrices with the inner product

$$\langle A, B \rangle \triangleq \text{tr}\{AB^T\}, \quad \forall A, B \in \mathbb{R}^{m \times d},$$

whose induced norm is denoted by $|\cdot|$. Next, we introduce the following spaces:

- For $s \in [0, T)$ and Euclidean space \mathbb{E} , $L^2_{\mathcal{F}}(s, T; \mathbb{E})$ is the set of all $\{\mathcal{F}_t\}$ -progressively measurable, square-integrable, \mathbb{E} -valued processes ξ defined on $[s, T]$, endowed with the norm

$$\|\xi\|_{L^2_{\mathcal{F}}(s, T; \mathbb{E})} \triangleq \left\{ E \int_s^T |\xi(t, \cdot)|^2 dt \right\}^{1/2}.$$

- For Euclidean space \mathbb{E} , $L^2_{\mathcal{F}}(\Omega; C([0, T]; \mathbb{E}))$ is the set of all $\{\mathcal{F}_t\}$ -progressively measurable, continuous, \mathbb{E} -valued processes ξ , endowed with the norm

$$\|\xi\|_{L^2_{\mathcal{F}}(\Omega; C([0, T]; \mathbb{E}))} \triangleq \left\{ E \left\{ \sup_{0 \leq t \leq T} |\varphi(t, \cdot)|^2 \right\} \right\}^{1/2}.$$

- $\mathcal{M}[0, T] \triangleq L^2_{\mathcal{F}}(\Omega; C([0, T]; \mathbb{R}^n \times \mathbb{R}^m)) \times L^2_{\mathcal{F}}(0, T; \mathbb{R}^{m \times d})$.

We consider the following controlled stochastic system: for $s \in [0, T]$,

$$\begin{cases} dX(t) = b(t, X(t), Y(t), Z(t)) dt + \sigma(t, X(t), Y(t), Z(t)) dW(t), \\ dY(t) = h(t, X(t), Y(t), Z(t)) dt + Z(t) dW(t), & t \in [s, T], \\ X(s) = x, & Y(s) = y. \end{cases} \quad (2.1)$$

Here (X, Y) is the *state process*, $Z \in \mathcal{Z}[0, T] \triangleq L^2_{\mathcal{F}}(0, T; \mathbb{R}^{m \times d})$ is the *control process*, $s \in [0, T]$ is the *initial time* and $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ is the *initial state*.

We make use of the following assumption throughout the paper.

(H1) Functions $b(t, x, y, z)$, $\sigma(t, x, y, z)$, $h(t, x, y, z)$ and $g(x)$ are continuous and there exists a constant $L > 0$, such that for $\varphi = b, \sigma, h, g$, it holds that

$$\begin{cases} |\varphi(t, x, y, z) - \varphi(t, \bar{x}, \bar{y}, \bar{z})| \leq L(|x - \bar{x}| + |y - \bar{y}| + |z - \bar{z}|), \\ |\varphi(t, 0, 0, 0)|, |\sigma(t, x, y, 0)| \leq L, \\ \forall t \in [0, T], \quad x, \bar{x} \in \mathbb{R}^n, \quad y, \bar{y} \in \mathbb{R}^m, \quad z, \bar{z} \in \mathbb{R}^{m \times d}. \end{cases} \quad (2.2)$$

Clearly, under (H1), for any initial datum $(s, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m$ and control $Z \in \mathcal{Z}[s, T]$, SDE (2.1) admits a unique strong solution $(X, Y) \equiv (X^{s,x,y,Z}, Y^{s,x,y,Z})$. We define the *cost functional*:

$$J(s, x, y; Z) \triangleq Ef(X^{s,x,y,Z}(T), Y^{s,x,y,Z}(T)), \quad (2.3)$$

where

$$f(x, y) = \sqrt{1 + |y - g(x)|^2} - 1, \quad \forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^m. \quad (2.4)$$

It is clear that $f(x, y) \geq 0$, for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$; $f(x, y) = 0$ if and only if $y = g(x)$; and the following elementary inequality holds:

$$|f(x, y) - f(\bar{x}, \bar{y})| \leq |y - \bar{y}| + |g(x) - g(\bar{x})|, \quad \forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^m. \quad (2.5)$$

Therefore, assumption (H1) would imply that f is uniformly Lipschitz in (x, y) as well.

The following is the optimal control problem associated with FBSDE (1.1).

Problem (OC)_s. For any given $(s, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m$, minimize (2.3) subject to (2.1) over $Z(\cdot) \in \mathcal{Z}[s, T]$.

We define the *value function* of Problem (OC)_s as follows:

$$\begin{cases} V(s, x, y) \triangleq \inf_{Z(\cdot) \in \mathcal{Z}[s, T]} J(s, x, y; Z(\cdot)), & s \in [0, T], \\ V(T, x, y) = f(x, y). \end{cases} \quad (2.6)$$

We are particularly interested in the case when $s = 0$, which is directly related to the (approximate) solvability of FBSDE (1.1). Thus, we set

$$\bar{V}(x, y) \triangleq V(0, x, y), \quad (x, y) \in \mathbb{R}^n \times \mathbb{R}^m. \quad (2.7)$$

Finally, we denote the *nodal set* of function $\bar{V}: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ by

$$\mathcal{N}(\bar{V}) \triangleq \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid \bar{V}(x, y) = 0\}. \quad (2.8)$$

In [15] we proved that FBSDE (1.1) admits an adapted solution if and only if the following two conditions are satisfied:

Condition (N). $\mathcal{N}(\bar{V}) \cap [\{x\} \times \mathbb{R}^m] \neq \emptyset$.

Condition (E). For some $(x, y) \in \mathcal{N}(\bar{V})$, there exists an *optimal control* $Z(\cdot) \in \mathcal{Z}[0, T]$, such that $J(0, x, y; Z(\cdot)) = \bar{V}(x, y) = 0$.

Here, Condition (N) amounts to saying that the nodal set of \bar{V} is non-empty, while Condition (E) is a problem of the existence of optimal controls. Since this paper deals only with approximate solvability, the problem of the existence of optimal control will be naturally relaxed to finding the nearly optimal controls; and the main technical difficulty will therefore lie in the treatment of Condition (N). We note that in [15] only a very special one-dimensional case was discussed, this paper can thus be regarded as a continuation of [15] as well.

3. Approximate Solvability

The precise definition of the *approximate solvability* of an FBSDE is the following.

Definition 3.1. For given $x \in \mathbb{R}^n$, (1.1) is said to be *approximately solvable* if for any $\varepsilon > 0$, there exists a triple $(X_\varepsilon, Y_\varepsilon, Z_\varepsilon) \in \mathcal{M}[0, T]$, such that (1.1) is satisfied except for the last condition, which is replaced by the following:

$$E|Y_\varepsilon(T) - g(X_\varepsilon(T))| < \varepsilon. \quad (3.1)$$

We call $(X_\varepsilon, Y_\varepsilon, Z_\varepsilon)$ an *approximate adapted solution* of (1.1) with accuracy ε .

It is clear that for a given $x \in \mathbb{R}^n$, if (1.1) is solvable, then it is approximately solvable. On the other hand, for an FBSDE, even if all the coefficients are all uniformly Lipschitz continuous, we still cannot guarantee its approximate solvability. Here is a simple example, adopted from [20].

Example 3.2. Consider the following simple FBSDE:

$$\begin{cases} dX(t) = Y(t) dt + dW(t), \\ dY(t) = -X(t) dt + Z(t) dW(t), \\ X(0) = x, \quad Y(T) = -X(T), \end{cases} \quad (3.2)$$

with $T = 3\pi/4$ and $x \neq 0$. It is obvious that the coefficients of this FBSDE are all uniformly Lipschitz. However, we claim that (3.2) is not approximately solvable. To see

this, note that by the variation of constants formula with $y = Y(0)$, we have

$$\begin{aligned} \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} &= \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &+ \int_0^t \begin{pmatrix} \cos(t-s) & \sin(t-s) \\ -\sin(t-s) & \cos(t-s) \end{pmatrix} \begin{pmatrix} 1 \\ Z(s) \end{pmatrix} dW(s). \end{aligned} \quad (3.3)$$

Plugging $t = T = 3\pi/4$ into (3.3), we obtain that

$$X(T) + Y(T) = -\sqrt{2}x + \int_0^T \eta(s) dW(s),$$

where η is some process in $L^2_{\mathcal{F}}(0, T; \mathbb{R})$. Consequently, by Jensen's inequality we have

$$E|Y(T) - g(X(T))| = E|X(T) + Y(T)| \geq |E[X(T) + Y(T)]| = \sqrt{2}|x| > 0,$$

for all $(y, Z) \in \mathbb{R}^m \times \mathcal{Z}[0, T]$. Thus, by Definition 3.1, FBSDE (3.2) is not approximately solvable (whence not solvable, as was concluded in [20]).

The following result establishes the relationship between the approximate solvability of FBSDE (1.1) and the value function of the associated control problem.

Proposition 3.3. *Let (H1) hold. For a given $x \in \mathbb{R}^n$, the FBSDE (1.1) is approximately solvable if and only if the following holds:*

$$\inf_{y \in \mathbb{R}^m} \bar{V}(x, y) = 0. \quad (3.4)$$

Proof. We first claim that inequality (3.1) in Definition 3.1 can be replaced by

$$Ef(X_\varepsilon(T), Y_\varepsilon(T)) < \varepsilon. \quad (3.5)$$

Indeed, by the following elementary inequalities,

$$\frac{r \wedge r^2}{3} \leq \sqrt{1+r^2} - 1 \leq r, \quad \forall r \in [0, \infty), \quad (3.6)$$

we see that if (3.1) holds, so does (3.5). Conversely, (3.6) implies

$$\begin{aligned} Ef(X_\varepsilon(T), Y_\varepsilon(T)) &\geq \frac{1}{3}E(|Y_\varepsilon(T) - g(X_\varepsilon(T))|^2 I_{(|Y_\varepsilon(T) - g(X_\varepsilon(T))| \leq 1)}) \\ &+ \frac{1}{3}E(|Y_\varepsilon(T) - g(X_\varepsilon(T))| I_{(|Y_\varepsilon(T) - g(X_\varepsilon(T))| > 1)}). \end{aligned}$$

Here I_A is the indicator function of A . Consequently, we have

$$E|Y_\varepsilon(T) - g(X_\varepsilon(T))| \leq 3Ef(X_\varepsilon(T), Y_\varepsilon(T)) + \sqrt{3Ef(X_\varepsilon(T), Y_\varepsilon(T))}. \quad (3.7)$$

Thus (3.5) implies (3.1) with ε being replaced by $\varepsilon' = 3\varepsilon + \sqrt{3\varepsilon}$. Namely, (3.4) is equivalent to the approximate solvability, by Definition 3.1 and the definition of \bar{V} .

Using Proposition 3.3 we can now claim the non-approximate solvability of the FBSDE (3.2) in a different way. By a direct computation using (3.7), one shows that

$$\begin{aligned} J(x, y; Z(\cdot)) &= Ef(X(T), Y(T)) \geq \frac{1}{3} \left[\sqrt{\sqrt{2}|x| + \frac{1}{4}} - \frac{1}{2} \right]^2 \\ &> 0, \quad \forall Z(\cdot) \in \mathcal{Z}[0, T]. \end{aligned}$$

Thus,

$$\bar{V}(x, y) \geq \frac{1}{3} \left[\sqrt{\sqrt{2}|x| + \frac{1}{4}} - \frac{1}{3} \right]^2 > 0,$$

violating (3.4), whence not approximately solvable.

Next, we relate the approximate solvability to Condition (N). To this end, we introduce the following supplementary assumption.

(H2) There exists a constant $L > 0$, such that for all $(t, x, y, z) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d}$, one of the following holds:

$$\begin{cases} |b(t, x, y, z)| + |\sigma(t, x, y, z)| \leq L(1 + |x|), \\ \langle h(t, x, y, z), y \rangle \geq -L(1 + |x||y| + |y|^2), \end{cases} \quad (3.8)$$

$$\begin{cases} \langle h(t, x, y, z), y \rangle \geq -L(1 + |y|^2), \\ |g(x)| \leq L. \end{cases} \quad (3.9)$$

Proposition 3.4. *Let (H1) hold. Then Condition (N) implies (3.4); conversely, if $\bar{V}(x, \cdot)$ is continuous, and (H2) holds, then (3.4) implies Condition (N).*

Proof. That Condition (N) implies (3.4) is obvious. We need only prove the converse. We first assume that \bar{V} is continuous and (3.8) holds.

Since (3.4) implies the approximate solvability of (1.1), for every $\varepsilon \in (0, 1]$, we may let $(X_\varepsilon, Y_\varepsilon, Z_\varepsilon) \in \mathcal{M}[0, T]$ be the approximate adapted solution of (1.1) with accuracy ε . Some standard arguments using Itô's formula, Gronwall's inequality and condition (3.8) yield the following estimate:

$$E|X_\varepsilon(t)|^2 \leq C(1 + |x|^2), \quad \forall t \in [0, T], \quad \varepsilon \in (0, 1]. \quad (3.10)$$

Here and in what follows, the constant $C > 0$ is a generic one, depending only on L and T , and may change from line to line. By (3.10) and (3.1) we obtain

$$\begin{aligned} E|Y_\varepsilon(T)| &\leq E|g(X_\varepsilon(T))| + E|Y_\varepsilon(T) - g(X_\varepsilon(T))| \\ &\leq C(1 + |x|) + \varepsilon \leq C(1 + |x|). \end{aligned} \quad (3.11)$$

Next, let $\langle x \rangle \triangleq \sqrt{1 + |x|^2}$. It is not hard to check that both $D\langle x \rangle$ and $D^2\langle x \rangle$ are uniformly bounded, thus applying Itô's formula to $\langle Y_\varepsilon(t) \rangle$, and noting (3.8) and (3.10), we have

$$\begin{aligned}
& E\langle Y_\varepsilon(T) \rangle - E\langle Y_\varepsilon(t) \rangle \\
&= E \int_t^T \frac{1}{\langle Y_\varepsilon(s) \rangle} \left\{ \langle Y_\varepsilon(s), h(s, X_\varepsilon(s), Y_\varepsilon(s), Z_\varepsilon(s)) \rangle \right. \\
&\quad \left. + \frac{1}{2} \left[|Z_\varepsilon(s)|^2 - \left| Z_\varepsilon(s)^T \frac{Y_\varepsilon(s)}{\langle Y_\varepsilon(s) \rangle} \right|^2 \right] \right\} ds \\
&\geq -LE \int_t^T (1 + |X_\varepsilon(s)| + \langle Y_\varepsilon(s) \rangle) ds \\
&\geq -C(1 + |x|) - LE \int_t^T \langle Y_\varepsilon(s) \rangle ds, \quad \forall t \in [0, T]. \tag{3.12}
\end{aligned}$$

Now note that $|y| \leq \langle y \rangle \leq 1 + |y|$, we have by Gronwall's inequality and (3.11) that

$$E\langle Y_\varepsilon(t) \rangle \leq C(1 + |x|), \quad \forall t \in [0, T], \quad \varepsilon \in (0, 1]. \tag{3.13}$$

In particular, (3.13) leads to the boundedness of the set $\{|Y_\varepsilon(0)|\}_{\varepsilon > 0}$, since $Y_\varepsilon(0)$'s are deterministic, thanks to the Blumenthal 0–1 law. Thus, along a sequence we have $Y_{\varepsilon_k}(0) \rightarrow y$, as $k \rightarrow \infty$. Condition (N) will now follow easily from the continuity of $\bar{V}(x, \cdot)$ and the following equalities:

$$0 \leq \bar{V}(x, Y_{\varepsilon_k}(0)) \leq Ef(X_{\varepsilon_k}(T), Y_{\varepsilon_k}(T)) < \varepsilon_k. \tag{3.14}$$

Finally, if (3.9) holds, then redoing (3.11) and (3.12), we see that (3.13) can be replaced by $E\langle Y_\varepsilon(t) \rangle \leq C, \forall t \in [0, T], \varepsilon \in (0, 1]$. Thus the same conclusion holds. \square

4. Approximation of the Value Function

In this section we introduce an approximation of the value function $V(s, x, y)$. This approximation plays an important role in our future discussion: it removes the degeneracy of the forward diffusion, and restricts the controls to a compact set.

To begin, we present some basic properties of the value function V of Problem $(OC)_s$. We note that these properties are slightly weaker than the standard ones because of the non-compactness of the control domain, but are sufficient for our purpose.

Proposition 4.1. *Let (H1) hold. Then $V(s, x, y)$ is right-continuous in $s \in [0, T)$ and there exists a constant $C > 0$, such that for all $(s, x, y), (s, \bar{x}, \bar{y}) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m$,*

$$0 \leq V(s, x, y) \leq C(1 + |x| + |y|), \tag{4.1}$$

$$|V(s, x, y) - V(s, \bar{x}, \bar{y})| \leq C(|x - \bar{x}| + |y - \bar{y}|). \tag{4.2}$$

In particular, $\bar{V}(\cdot, \cdot)$ (defined by (2.7)) is continuous in $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$.

Proof. The proof of (4.1) follows directly from the following inequalities:

$$\begin{aligned} 0 \leq V(s, x, y) &\leq J(s, x, y; 0) \\ &\leq C(1 + |x| + |y|), \quad \forall (s, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m. \end{aligned}$$

The proof of (4.2) is standard, by using (H1) and (2.5) (directly estimate the cost functional with the initial states (x, y) and (\bar{x}, \bar{y}) , and with the same $s \in [0, T)$ and $Z \in \mathcal{Z}[s, T]$). It remains to prove the right-continuity of $V(s, x, y)$ in $s \in [0, T)$. To this end, we note that for any $Z \in \mathcal{Z}[s, T]$, and $\bar{s} > s$, $Z^{\bar{s}} \triangleq Z|_{[\bar{s}, T]} \in \mathcal{Z}[\bar{s}, T]$. Therefore, denoting $(X, Y) \triangleq (X^{s,x,y,Z}, Y^{s,x,y,Z})$ and $(\bar{X}, \bar{Y}) \triangleq (X^{\bar{s},x,y,Z^{\bar{s}}}, Y^{\bar{s},x,y,Z^{\bar{s}}})$, it is standard to show that

$$E|X(T) - \bar{X}(T)| + E|Y(T) - \bar{Y}(T)| \leq C_Z |s - \bar{s}|^{1/2}, \quad (4.3)$$

where C_Z is some constant depending on L, T , as well as $Z|_{[s, \bar{s}]}$. Now for any $\bar{Z} \in \mathcal{Z}[\bar{s}, T]$, let $Z \in \mathcal{Z}[s, T]$ be such that $Z = 0$ on $[s, \bar{s}]$ and $Z^{\bar{s}} = \bar{Z}$, then we have from (2.5), (H1) and (4.3) that

$$\begin{aligned} V(s, x, y) &\leq J(s, x, y; Z) \\ &= J(\bar{s}, x, y; \bar{Z}) + [Ef(X(T), Y(T)) - Ef(\bar{X}(T), \bar{Y}(T))] \\ &\leq J(\bar{s}, x, y; \bar{Z}) + C_0 |s - \bar{s}|^{1/2}, \end{aligned}$$

where C_0 is some constant depending only on L and T , since $Z|_{[s, \bar{s}]} = 0$. Thus, taking infimum over $\bar{Z} \in \mathcal{Z}[\bar{s}, T]$, we obtain

$$V(s, x, y) \leq V(\bar{s}, x, y) + C_0 |s - \bar{s}|^{1/2}. \quad (4.4)$$

Next, note that the mapping $(s, x, y) \mapsto J(s, x, y; Z|_{[s, T]}(\cdot))$ is continuous for each fixed Z . Thus, as the infimum of continuous functions, $V(s, x, y)$ is upper semi-continuous. Now from (4.4) we easily deduce that $\lim_{\bar{s} \downarrow s} V(\bar{s}, x, y) = V(s, x, y)$, proving the right-continuity of V in $s \in [0, T)$. \square

We now introduce the approximation of the value function. First, let $\tilde{W}(t) \equiv (\tilde{W}_1(t), \tilde{W}_2(t))$ be an $(n + m)$ -dimensional Brownian motion which is independent of $W(t)$ (embedded into an enlarged probability space, if necessary) and let $\{\tilde{\mathcal{F}}_t\}_{t \geq 0}$ be the filtration generated by $W(t)$ and $\tilde{W}(t)$, augmented by all the P -null sets in \mathcal{F} . Define

$$\left\{ \begin{array}{l} \mathcal{Z}_0[s, T] \triangleq \mathcal{Z}[s, T], \\ \tilde{\mathcal{Z}}_0[s, T] \triangleq \left\{ Z: [s, T] \times \Omega \rightarrow \mathbb{R}^{m \times d} \mid Z \text{ is } \{\tilde{\mathcal{F}}_t\}\text{-adapted,} \right. \\ \left. \int_0^T E|Z(t)|^2 dt < \infty \right\}. \end{array} \right. \quad (4.5)$$

Next, for any $\delta > 0$, we define

$$\left\{ \begin{array}{l} \mathcal{Z}_\delta[s, T] \triangleq \left\{ Z \in \mathcal{Z}[s, T] \mid |Z(t)| \leq \frac{1}{\delta}, \text{ a.e. } t \in [s, T], \text{ a.s.} \right\}, \\ \tilde{\mathcal{Z}}_\delta[s, T] \triangleq \left\{ Z \in \tilde{\mathcal{Z}}_0[s, T] \mid |Z(t)| \leq \frac{1}{\delta}, \text{ a.e. } t \in [s, T], \text{ a.s.} \right\}. \end{array} \right. \quad (4.6)$$

The following inclusions are obvious:

$$\begin{aligned} \mathcal{Z}_0[s, T] \supset \mathcal{Z}_{\delta_1}[s, T] \supset \mathcal{Z}_{\delta_2}[s, T] \\ \cap \quad \quad \quad \cap \quad \quad \quad \cap \quad \quad \quad \forall \delta_2 \geq \delta_1 \geq 0. \end{aligned} \quad (4.7)$$

$$\tilde{\mathcal{Z}}_0[s, T] \supset \tilde{\mathcal{Z}}_{\delta_1}[s, T] \supset \tilde{\mathcal{Z}}_{\delta_2}[s, T]$$

In what follows, for any $Z \in \mathcal{Z}_0[s, T]$ (resp. $\tilde{\mathcal{Z}}_0[s, T]$) and $\delta > 0$, we define the $(1/\delta)$ -truncation of Z as follows:

$$Z_\delta(t, \omega) = \begin{cases} Z(t, \omega), & \text{if } |Z(t, \omega)| \leq \frac{1}{\delta}, \\ \frac{Z(t, \omega)}{\delta|Z(t, \omega)|}, & \text{if } |Z(t, \omega)| > \frac{1}{\delta}. \end{cases}$$

Clearly, $Z_\delta \in \mathcal{Z}_\delta[s, T]$ (resp. $\tilde{\mathcal{Z}}_\delta[s, T]$).

We now consider, for any $\varepsilon > 0$, the following *regularized state equation* (compare with (2.1)):

$$\begin{cases} dX(t) = b(t, X(t), Y(t), Z(t)) dt + \sigma(t, X(t), Y(t), Z(t)) dW(t) \\ \quad + \sqrt{2\varepsilon} d\tilde{W}_1(t), \\ dY(t) = h(t, X(t), Y(t), Z(t)) dt + Z(t) dW(t) \\ \quad + \sqrt{2\varepsilon} d\tilde{W}_2(t), \quad t \in [s, T], \\ X(s) = x, \quad Y(s) = y. \end{cases} \quad (4.8)$$

Define the cost functional by $J^{\delta, \varepsilon}(s, x, y; Z(\cdot))$ (resp. $\tilde{J}^{\delta, \varepsilon}(s, x, y; Z(\cdot))$) which has the same form as (2.3) with the control being taken in $\mathcal{Z}_\delta[s, T]$ (resp. $\tilde{\mathcal{Z}}_\delta[s, T]$) and the state satisfying (4.8), indicating the dependence on $\delta \geq 0$ and $\varepsilon \geq 0$. The corresponding optimal control problem is called Problem $(OC)_s^{\delta, \varepsilon}$ (resp. Problem $(\widetilde{OC})_s^{\delta, \varepsilon}$). The corresponding (approximate) value functions are then defined as, respectively,

$$\begin{cases} \tilde{V}^{\delta, \varepsilon}(s, x, y) = \inf_{Z(\cdot) \in \tilde{\mathcal{Z}}_\delta[s, T]} \tilde{J}^{\delta, \varepsilon}(s, x, y; Z(\cdot)), \\ V^{\delta, \varepsilon}(s, x, y) = \inf_{Z(\cdot) \in \mathcal{Z}_\delta[s, T]} J^{\delta, \varepsilon}(s, x, y; Z(\cdot)). \end{cases} \quad (4.9)$$

Due to the inclusions in (4.6), we see that for any $(s, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m$,

$$\begin{cases} V^{\delta, \varepsilon}(s, x, y) \geq \tilde{V}^{\delta, \varepsilon}(s, x, y) \geq 0, \quad \forall \delta, \varepsilon \geq 0, \\ \tilde{V}^{\delta_2, \varepsilon}(s, x, y) \geq \tilde{V}^{\delta_1, \varepsilon}(s, x, y), \quad \forall \delta_2 \geq \delta_1 \geq 0, \quad \varepsilon \geq 0, \\ V^{\delta_2, \varepsilon}(s, x, y) \geq V^{\delta_1, \varepsilon}(s, x, y), \quad \forall \delta_2 \geq \delta_1 \geq 0, \quad \varepsilon \geq 0. \end{cases} \quad (4.10)$$

Also, it is an easy observation that $V^{0,0}(s, x, y) = V(s, x, y)$, $\forall (s, x, y)$. Note that for $\delta > 0$ and $\varepsilon \geq 0$, the corresponding HJB equation for the value function $\tilde{V}^{\delta, \varepsilon}(s, x, y)$ takes the following form:

$$\begin{cases} \tilde{V}_s^{\delta, \varepsilon} + \varepsilon \Delta \tilde{V}^{\delta, \varepsilon} + H^\delta(s, x, y, D\tilde{V}^{\delta, \varepsilon}, D^2\tilde{V}^{\delta, \varepsilon}) = 0, \\ \quad (s, x, y) \in (0, T) \times \mathbb{R}^n \times \mathbb{R}^m; \\ \tilde{V}^{\delta, \varepsilon}(T, x, y) = f(x, y), \\ \quad (x, y) \in \mathbb{R}^n \times \mathbb{R}^m, \end{cases} \quad (4.11)$$

where Δ is the Laplacian operator in \mathbb{R}^{n+m} , and H^δ is defined by the following:

$$H^\delta(s, x, y, q, Q) \triangleq \inf_{z \in \mathbb{R}^{m \times d}, |z| \leq 1/\delta} \left\{ \left\langle q, \begin{pmatrix} b(s, x, y, z) \\ h(s, x, y, z) \end{pmatrix} \right\rangle + \frac{1}{2} \text{tr} \left[Q \begin{pmatrix} \sigma(s, x, y, z) \\ z \end{pmatrix} \begin{pmatrix} \sigma(s, x, y, z) \\ z \end{pmatrix}^T \right] \right\},$$

for $(s, x, y, q, Q) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{n+m} \times S^{n+m}$, where S^{n+m} is the set of all $(n+m) \times (n+m)$ symmetric matrices. We observe that for $\varepsilon > 0$, (4.11) is a *non-degenerate* non-linear parabolic PDE. The following proposition collects some basic properties of the approximate value functions, whose proof can be obtained by slightly modifying relevant arguments in [10].

Proposition 4.2. *Let (H1) hold. Then:*

- (i) $\tilde{V}^{\delta, \varepsilon}(s, x, y)$ and $V^{\delta, \varepsilon}(s, x, y)$ are continuous in $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$, uniformly in $s \in [0, T]$ and $\delta, \varepsilon \geq 0$. For fixed $\delta > 0$ and $\varepsilon \geq 0$, $\tilde{V}^{\delta, \varepsilon}(s, x, y)$ and $V^{\delta, \varepsilon}(s, x, y)$ are continuous in $(s, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m$.
- (ii) For $\delta > 0$ and $\varepsilon \geq 0$, $\tilde{V}^{\delta, \varepsilon}(s, x, y)$ is the unique viscosity solution of (4.11), and for $\delta, \varepsilon > 0$, $\tilde{V}^{\delta, \varepsilon}(s, x, y)$ is the unique strong solution of (4.11).
- (iii) For $\delta > 0$ and $\varepsilon \geq 0$, $V^{\delta, \varepsilon}(s, x, y)$ is a viscosity super solution of (4.11), $V^{\delta, 0}(s, x, y)$ is the unique viscosity solution of (4.11) (with $\varepsilon = 0$).

The following result gives the continuous dependence of the approximate value functions on the parameters δ and ε .

Theorem 4.3. *Let (H1) hold. Then, for any $s \in [0, T]$, there exists a continuous function $\eta_s: [0, \infty) \times [0, \infty) \rightarrow [0, \infty)$, with $\eta_s(0, r) = 0$ for all $r \geq 0$, such that*

$$\begin{aligned} |\tilde{V}^{\delta, \varepsilon}(s, x, y) - \tilde{V}^{\hat{\delta}, \hat{\varepsilon}}(s, x, y)| &\leq \eta(|\delta - \hat{\delta}| + |\varepsilon - \hat{\varepsilon}|, |x| + |y|), \\ |V^{\delta, \varepsilon}(s, x, y) - V^{\hat{\delta}, \hat{\varepsilon}}(s, x, y)| &\leq \eta(|\delta - \hat{\delta}| + |\varepsilon - \hat{\varepsilon}|, |x| + |y|), \\ \forall (s, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m, \quad \delta, \hat{\delta}, \varepsilon, \hat{\varepsilon} &\in [0, 1]. \end{aligned} \quad (4.12)$$

Proof. Fix $(s, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m$, $\delta, \hat{\delta}, \varepsilon, \hat{\varepsilon} \geq 0$ and $Z \in \mathcal{Z}[s, T]$. Let Z_δ (resp. $Z_{\hat{\delta}}$) be the $(1/\delta)$ -truncation (resp. $(1/\hat{\delta})$ -truncation) of Z , and let (X, Y) (resp. (\hat{X}, \hat{Y})) be the solution of (4.8) corresponding to (ε, Z_δ) (resp. $(\hat{\varepsilon}, Z_{\hat{\delta}})$). By Itô's formula and Gronwall's inequality,

$$\begin{aligned} E\{|X(T) - \hat{X}(T)|^2 + |Y(T) - \hat{Y}(T)|^2\} \\ \leq C \left\{ E \int_s^T |Z_\delta(t) - Z_{\hat{\delta}}(t)|^2 dt + |\sqrt{\varepsilon} - \sqrt{\hat{\varepsilon}}|^2 \right\}, \end{aligned} \quad (4.13)$$

where $C > 0$ depends only on L and T . Thus, we obtain

$$|V^{\delta, \varepsilon}(s, x, y) - V^{\hat{\delta}, \hat{\varepsilon}}(s, x, y)| \leq C|\sqrt{\varepsilon} - \sqrt{\hat{\varepsilon}}|, \quad \forall (s, x, y), \delta, \varepsilon, \hat{\delta}, \hat{\varepsilon} \geq 0. \quad (4.14)$$

Combining with Proposition 4.2, we see that $V^{\delta,\varepsilon}(s, x, y)$ is continuous in $(\varepsilon, x, y) \in [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^m$ uniformly in $\delta \geq 0$ and $s \in [0, T]$.

Next, for fixed $(s, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m$, $\varepsilon \geq 0$ and $\hat{\delta} \geq \delta \geq 0$, by (4.10), we have

$$0 \leq V^{\hat{\delta},\varepsilon}(s, x, y) - V^{\delta,\varepsilon}(s, x, y). \quad (4.15)$$

On the other hand, for any $\delta > 0$ and $\varepsilon_0 > 0$, we can choose $Z^{\varepsilon_0} \in \mathcal{Z}_\delta[s, T]$ so that

$$V^{\delta,\varepsilon}(s, x, y) + \varepsilon_0 > J^{\delta,\varepsilon}(s, x, y; Z^{\varepsilon_0}). \quad (4.16)$$

Let $Z_\delta^{\varepsilon_0}$ be the $(1/\hat{\delta})$ -truncation of Z^{ε_0} , and denote the corresponding solution of (4.8) with Z^{ε_0} (resp. $Z_\delta^{\varepsilon_0}$) by $(X^{\varepsilon_0}, Y^{\varepsilon_0})$ (resp. $(\hat{X}^{\varepsilon_0}, \hat{Y}^{\varepsilon_0})$). Setting $(X, Y) = (X^{\varepsilon_0}, Y^{\varepsilon_0})$, $(\hat{X}, \hat{Y}) = (\hat{X}^{\varepsilon_0}, \hat{Y}^{\varepsilon_0})$, $\varepsilon = \hat{\varepsilon}$, $Z_\delta = Z^{\varepsilon_0}$ and $Z_\delta^{\varepsilon_0} = Z_\delta^{\varepsilon_0}$ in (4.13), we obtain

$$\begin{aligned} & E\{|X^{\varepsilon_0}(T) - \hat{X}^{\varepsilon_0}(T)|^2 + |Y^{\varepsilon_0}(T) - \hat{Y}^{\varepsilon_0}(T)|^2\} \\ & \leq CE \int_s^T |Z^{\varepsilon_0}(t) - Z_\delta^{\varepsilon_0}(t)|^2 dt. \end{aligned} \quad (4.17)$$

We consider the following two cases:

Case 1: $\delta > 0$. In this case note that $|Z^{\varepsilon_0}(t) - Z_\delta^{\varepsilon_0}(t)| \leq |1/\delta - 1/\hat{\delta}|$, a.e. $t \in [s, T]$, a.s. By (2.5) and (H1), one easily checks that

$$\begin{aligned} J^{\delta,\varepsilon}(s, x, y; Z^{\varepsilon_0}) & \geq J^{\hat{\delta},\varepsilon}(s, x, y; Z_\delta^{\varepsilon_0}) - C \left| \frac{1}{\delta} - \frac{1}{\hat{\delta}} \right| \\ & \geq V^{\hat{\delta},\varepsilon}(s, x, y) - C \left| \frac{1}{\delta} - \frac{1}{\hat{\delta}} \right|. \end{aligned} \quad (4.18)$$

Combining (4.15), (4.16) and (4.18), we obtain (note $\varepsilon_0 > 0$ is arbitrary)

$$\begin{aligned} 0 & \leq V^{\hat{\delta},\varepsilon}(s, x, y) - V^{\delta,\varepsilon}(s, x, y) \\ & \leq C \left| \frac{1}{\delta} - \frac{1}{\hat{\delta}} \right|, \quad \forall (s, x, y), \delta, \hat{\delta} > 0, \quad \varepsilon \geq 0, \end{aligned} \quad (4.19)$$

where C is again an absolute constant.

Case 2: $\delta = 0$. Now let $\hat{\delta} > 0$ be small enough so that the right side of (4.17) is no greater than ε_0^2 . Then, similar to (4.18), we have

$$J^{0,\varepsilon}(s, x, y; Z^{\varepsilon_0}) \geq V^{\hat{\delta},\varepsilon}(s, x, y) - \varepsilon_0. \quad (4.20)$$

Combing (4.15), (4.16) and (4.20), one has $0 \leq V^{\hat{\delta},\varepsilon}(s, x, y) - V^{0,\varepsilon}(s, x, y) \leq 2\varepsilon_0$, which shows that

$$V^{\hat{\delta},\varepsilon}(s, x, y) \downarrow V^{0,\varepsilon}(s, x, y), \quad \hat{\delta} \downarrow 0. \quad (4.21)$$

Since $V^{0,\varepsilon}(s, x, y)$ is continuous in (ε, x, y) (see (4.14) and Proposition 4.2(i)), by Dini's theorem, we obtain that the convergence in (4.21) is uniform in (ε, x, y) on compact sets.

Thus, for some continuous function $\eta_s: [0, \infty) \times [0, \infty) \rightarrow [0, \infty)$ with $\eta_s(0, r) = 0$ for all $r \geq 0$, one has

$$\begin{aligned} 0 &\leq V^{\hat{\delta}, \varepsilon}(s, x, y) - V^{0, \varepsilon}(s, x, y) \\ &\leq \eta_s(\hat{\delta}, |x| + |y|), \quad \forall (s, x, y), \quad \varepsilon \in [0, 1], \quad \hat{\delta} \geq 0. \end{aligned} \quad (4.22)$$

Combining (4.14), (4.19) and (4.22), we have that $V^{\delta, \varepsilon}(s, x, y)$ is continuous in $(\delta, \varepsilon, x, y) \in [0, \infty) \times [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^m$. The proof for $\tilde{V}^{\delta, \varepsilon}$ is exactly the same. \square

Corollary 4.4. *Let (H1) hold. Then*

$$\tilde{V}^{\delta, 0}(s, x, y) = V^{\delta, 0}(s, x, y), \quad \forall (s, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m, \quad \delta \geq 0. \quad (4.23)$$

Proof. If $\delta > 0$, then both $\tilde{V}^{\delta, 0}$ and $V^{\delta, 0}$ are the viscosity solutions of the HJB equation (4.11). Thus, (4.23) follows from the uniqueness. By the continuity of $\tilde{V}^{\delta, 0}$ and $V^{\delta, 0}$ in $\delta \geq 0$, we obtain (4.23) for $\delta = 0$. \square

As a matter of fact, for $\delta > 0$, we can regard Problem $(\widetilde{\text{OC}})_s^{\delta, 0}$ and Problem $(\text{OC})_s^{\delta, 0}$ as the same control problem under two different reference probability systems [10]. Thus, (4.23) holds according to Theorem 7.1 on p. 185 of [10] (see also Remark 6.1 on p. 178 of [10]).

Corollary 4.5. *Let $V(0, x, y) = 0$. Then, for any $\hat{\varepsilon} > 0$, there exist $\delta, \varepsilon > 0$ and $Z^{\delta, \varepsilon}(\cdot) \in \mathcal{Z}_\delta[0, T]$ satisfying*

$$J^{\delta, \varepsilon}(0, x, y; Z^{\delta, \varepsilon}(\cdot)) < \hat{\varepsilon}, \quad (4.24)$$

such that if $(X^{\delta, \varepsilon}(\cdot), Y^{\delta, \varepsilon}(\cdot))$ is the solution of (2.1) with $Z(\cdot) = Z^{\delta, \varepsilon}(\cdot)$, then the triplet $(X^{\delta, \varepsilon}, Y^{\delta, \varepsilon}, Z^{\delta, \varepsilon})$ is an approximate solution of (1.1) with accuracy $3\hat{\varepsilon} + \sqrt{3\hat{\varepsilon}}$.

Proof. Let $V(0, x, y) = 0$. Since $V = V^{0, 0}$, by Theorem 4.3, there exist $\delta, \varepsilon > 0$, such that $V^{\delta, \varepsilon}(0, x, y) < \hat{\varepsilon}$. Now by (4.9) we can find a $Z^{\varepsilon, \delta} \in \mathcal{Z}_\delta[0, T]$ such that (4.24) is satisfied. Let $(X^{\delta, \varepsilon}, Y^{\delta, \varepsilon})$ be the solutions of (2.1) with $s = 0$ and $Z = Z^{\delta, \varepsilon}$. Then we have (see (3.7))

$$\begin{aligned} E|Y^{\delta, \varepsilon}(T) - g(X^{\delta, \varepsilon}(T))| &\leq 3Ef(X^{\delta, \varepsilon}(T), Y^{\delta, \varepsilon}(T)) + \sqrt{3Ef(X^{\delta, \varepsilon}(T), Y^{\delta, \varepsilon}(T))} \\ &= 3J^{\delta, \varepsilon}(0, x, y; Z^{\delta, \varepsilon}(\cdot)) + \sqrt{3J^{\delta, \varepsilon}(0, x, y; Z^{\delta, \varepsilon}(\cdot))} \\ &\leq 3\hat{\varepsilon} + \sqrt{3\hat{\varepsilon}}. \end{aligned} \quad \square$$

To conclude this section we present another property of the approximate value function $\tilde{V}^{\delta, \varepsilon}$, which will be useful in the next section. The following definition is standard (see, e.g., [12] or [10]).

Definition 4.6. A function $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be *semi-concave* if there exists a constant $C > 0$, such that the function $\Phi(x) \equiv \varphi(x) - C|x|^2$ is concave on \mathbb{R}^n , i.e.,

$$\Phi(\lambda x + (1 - \lambda)\bar{x}) \geq \lambda\Phi(x) + (1 - \lambda)\Phi(\bar{x}), \quad \forall \lambda \in [0, 1], \quad x, \bar{x} \in \mathbb{R}^n.$$

A family of functions $\varphi_\varepsilon: \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be semi-concave *uniformly* in ε if there exists a constant $C > 0$, independent of ε , such that $\varphi_\varepsilon(x) - C|x|^2$ is concave for all ε .

We will need a further assumption.

(H3) Functions b, σ, h and g are differentiable in (x, y) with the derivatives being uniformly Lipschitz continuous in $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$, uniformly in $(t, z) \in [0, T] \times \mathbb{R}^{m \times d}$.

The following property of the value functions is a simple modification of those in [10] or [12], we omit the proof here.

Proposition 4.7. *Let (H1) and (H3) hold. Then $\tilde{V}^{\delta, \varepsilon}(s, x, y)$ is semi-concave uniformly in $s \in [0, T]$, $\delta \in (0, 1]$ and $\varepsilon \in [0, 1]$. In particular, there exists a constant $C > 0$, such that*

$$\Delta_y \tilde{V}^{\delta, \varepsilon}(s, x, y) \leq C, \quad \forall (s, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m, \quad \delta, \varepsilon \in (0, 1], \quad (4.25)$$

where $\Delta_y = \sum_{j=1}^m \partial_{y_j}^2$.

5. A Class of Approximately Solvable FBSDEs

In this section we consider the following FBSDE:

$$\begin{cases} dX(t) = b(t, X(t), Y(t)) dt + \sigma(t, X(t), Y(t)) dW(t), \\ dY(t) = h(t, X(t), Y(t)) dt + Z(t) dW(t), & t \in [0, T], \\ X(0) = x, & Y(T) = g(X(T)). \end{cases} \quad (5.1)$$

It is worth noting that the solvability of this seemingly simple class of FBSDEs has not yet been completely understood. Our example in Section 3 shows that an FBSDE of this kind might not even be approximately solvable. However, on the other hand, in our previous works [15] and [14] we did prove the solvability of (5.1), under the condition that the coefficient σ is non-degenerate (i.e., the matrix $\sigma\sigma^T$ is uniformly positive definite). Unfortunately, the methods we used there seem to fail when σ is allowed to be degenerate.

The main result of this section is the following theorem.

Theorem 5.1. *Let b, σ, h be continuous, independent of Z , C^2 in variables x and y with bounded first- and second-order partial derivatives; and*

$$|b(t, x, 0)| + |\sigma(t, x, 0)| + |h(t, x, 0)| \leq L, \quad \forall (t, x) \in [0, T] \times \mathbb{R}^n. \quad (5.2)$$

Further, let g be bounded in $C^{2+\alpha}(\mathbb{R}^n)$ for some $\alpha \in (0, 1)$. Then (1.1) is approximately solvable.

We note that under the assumption of Theorem 5.1, (H1) and (H3) hold. To prove Theorem 5.1, we need the following result.

Lemma 5.2. *Let the assumption of Theorem 5.1 hold. Then, for any $\varepsilon > 0$, there exists a unique classical solution $\theta^\varepsilon: [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ of the following (non-degenerate) parabolic system:*

$$\begin{cases} \theta_s^\varepsilon + \varepsilon \Delta \theta^\varepsilon + \frac{1}{2} \text{tr}[\sigma \sigma^T \theta_{xx}^\varepsilon] + \theta_x^\varepsilon b - h = 0, & (s, x) \in [0, T] \times \mathbb{R}^n, \\ \theta^\varepsilon|_{s=T} = g, \end{cases} \quad (5.3)$$

with θ^ε , $\theta_{x_i}^\varepsilon$ and $\theta_{x_i x_j}^\varepsilon$ all being bounded (with the bounds depending on $\varepsilon > 0$, in general). Moreover, there exists a constant $C > 0$, independent of $\varepsilon \in (0, 1]$, such that

$$|\theta^\varepsilon(s, x)| \leq C, \quad \forall (s, x) \in [0, T] \times \mathbb{R}^n, \quad \varepsilon \in (0, 1]. \quad (5.4)$$

Proof. We note that under our conditions, the following hold:

$$\begin{cases} 0 \leq (\sigma \sigma^T)(s, x, y) \leq C(1 + |y|^2)I, \\ |(\sigma_{x_i} \sigma^T)(s, x, y)| + |(\sigma_{y_k} \sigma^T)(s, x, y)| \leq C(1 + |y|), \\ \quad 1 \leq i \leq n, \quad 1 \leq k \leq m, \\ |b(s, x, y)| \leq L(1 + |y|), \\ -\langle h(s, x, y), y \rangle \leq L(1 + |y|^2). \end{cases} \quad (5.5)$$

Thus, it is standard (see, e.g., [13]) that for any $\varepsilon > 0$, there exists a unique classical solution θ^ε to (5.3) with θ^ε , $\theta_{x_i}^\varepsilon$ and $\theta_{x_i x_j}^\varepsilon$ all being bounded (with the bounds depending on $\varepsilon > 0$). Next, we prove (5.4). To this end, we fix an $\varepsilon \in (0, 1]$ and denote

$$\begin{aligned} \mathcal{A}_\varepsilon w &\triangleq \varepsilon \Delta w + \frac{1}{2} \text{tr}[\sigma \sigma^T(s, x, \theta^\varepsilon(s, x)) w_{xx}] + \langle b(s, x, \theta^\varepsilon(s, x)), w_x \rangle \\ &\equiv \sum_{i,j=1}^n a_{ij}^\varepsilon w_{x_i x_j} + \sum_{i=1}^n b_i^\varepsilon w_{x_i}. \end{aligned} \quad (5.6)$$

Set

$$\bar{w}(s, x) \triangleq \frac{1}{2} |\theta^\varepsilon(s, x)|^2 \equiv \frac{1}{2} \sum_{k=1}^m \theta^{\varepsilon,k}(s, x)^2. \quad (5.7)$$

Then it holds that (note (5.5)–(5.6))

$$\begin{aligned} \bar{w}_s &= \sum_{k=1}^m \theta^{\varepsilon,k} \theta_s^{\varepsilon,k} = \sum_{k=1}^m \theta^{\varepsilon,k} [-\mathcal{A}_\varepsilon \theta^{\varepsilon,k} + h^k(s, x, \theta^\varepsilon)] \\ &= \sum_{k=1}^m \theta^{\varepsilon,k} \left[-\sum_{i,j=1}^n a_{ij}^\varepsilon \theta_{x_i x_j}^{\varepsilon,k} - \sum_{i=1}^n b_i^\varepsilon \theta_{x_i}^{\varepsilon,k} + h^k(s, x, \theta^\varepsilon) \right] \\ &= -\sum_{k=1}^m \sum_{i,j=1}^n a_{ij}^\varepsilon \{ [(\frac{1}{2} \theta^{\varepsilon,k})^2]_{x_i x_j} - \theta_{x_i}^{\varepsilon,k} \theta_{x_j}^{\varepsilon,k} \} \\ &\quad - \sum_{k=1}^m \sum_{i=1}^n b_i^\varepsilon [(\frac{1}{2} \theta^{\varepsilon,k})^2]_{x_i} + \sum_{k=1}^m \theta^{\varepsilon,k} h^k(s, x, \theta^\varepsilon) \\ &\geq -\mathcal{A}_\varepsilon \bar{w} - 2L\bar{w} - L. \end{aligned}$$

Thus, \bar{w} is a bounded (with the bound depending on $\varepsilon > 0$) solution of the following:

$$\begin{cases} \bar{w}_s + \mathcal{A}_\varepsilon \bar{w} + 2L\bar{w} \geq -L, & (s, x) \in [0, T) \times \mathbb{R}^n, \\ \bar{w}|_{s=T} \leq \frac{1}{2} \|g\|_\infty. \end{cases} \quad (5.8)$$

By Lemma 5.3 below, we obtain

$$\bar{w}(s, x) \leq C, \quad \forall (s, x) \in [0, T] \times \mathbb{R}^n, \quad (5.9)$$

with the constant only depending on L and $\|g\|_\infty$ (and independent of $\varepsilon > 0$). Note that by definition (5.7) one has

$$\bar{w}(s, x) \geq 0, \quad \forall (s, x) \in [0, T] \times \mathbb{R}^n.$$

Therefore (5.4) follows. \square

Lemma 5.3. *Let \mathcal{A}_ε be given by (5.6) and let w be a bounded solution of the following:*

$$\begin{cases} w_s + \mathcal{A}_\varepsilon w + \lambda_0 w \geq -h_0, & (s, x) \in [0, T) \times \mathbb{R}^n, \\ w|_{s=T} \leq g_0, \end{cases} \quad (5.10)$$

for some constants $h_0, g_0 \geq 0$ and $\lambda_0 \in \mathbb{R}$, where the bound of w might depend on $\varepsilon > 0$, in general. Then, for any $\lambda > \lambda_0 \vee 0$,

$$w(s, x) \leq e^{\lambda T} \left[g_0 \vee \frac{h_0}{\lambda - \lambda_0} \right], \quad \forall (s, x) \in [0, T] \times \mathbb{R}^n. \quad (5.11)$$

Proof. Fix any $\lambda > \lambda_0 \vee 0$ and denote the right-hand side of (5.11) by C . Define $v(s, x) = e^{\lambda s} w(s, x) - C$, for $(s, x) \in [0, T] \times \mathbb{R}^n$. Then a direct computation shows that

$$v_s + \mathcal{A}_\varepsilon v - (\lambda - \lambda_0)v \geq -e^{\lambda s} L + (\lambda - \lambda_0)C \geq 0,$$

and $v(T, x) \leq e^{\lambda T} g_0 - C \leq 0$. Hence, by maximum principle (see [13], for example), we have $v(s, x) \leq 0$, for all $(s, x) \in [0, T] \times \mathbb{R}^n$. Thus (5.11) follows. \square

Proof of Theorem 5.1. We define (note (4.10)) $w^{\delta, \varepsilon}(s, x) \triangleq \tilde{V}^{\delta, \varepsilon}(s, x, \theta^\varepsilon(s, x)) \geq 0$, $\forall (s, x) \in [0, T] \times \mathbb{R}^n$. Then we obtain (using (4.11), (4.25) and (5.3))

$$\begin{aligned} 0 &= \tilde{V}_s^{\delta, \varepsilon} + \varepsilon \Delta \tilde{V}^{\delta, \varepsilon} + \frac{1}{2} \text{tr}[\sigma \sigma^T \tilde{V}_{xx}^{\delta, \varepsilon}] + \langle b, \tilde{V}_x^{\delta, \varepsilon} \rangle + \langle h, \tilde{V}_y^{\delta, \varepsilon} \rangle \\ &\quad + \frac{1}{2} \inf_{|z| \leq 1/\delta} \text{tr}[(\tilde{V}_{xy}^{\delta, \varepsilon})^T \sigma z^T + \tilde{V}_{xy}^{\delta, \varepsilon} z \sigma^T + \tilde{V}_{yy}^{\delta, \varepsilon} z z^T] \\ &= \{w_s^{\delta, \varepsilon} + \varepsilon \Delta w^{\delta, \varepsilon} + \frac{1}{2} \text{tr}[\sigma \sigma^T w_{xx}^{\delta, \varepsilon}] + \langle b, w_x^{\delta, \varepsilon} \rangle\} + \varepsilon \Delta_y \tilde{V}^{\delta, \varepsilon} \\ &\quad - \langle \tilde{V}_y^{\delta, \varepsilon}, \theta_s^\varepsilon + \varepsilon \Delta \theta^\varepsilon + \frac{1}{2} \text{tr}[\sigma \sigma^T \theta_{xx}^\varepsilon] + \theta_x^\varepsilon b - h \rangle \\ &\quad + \frac{1}{2} \inf_{|z| \leq 1/\delta} \text{tr}[2(z - \theta_x^\varepsilon \sigma) \sigma^T \tilde{V}_{xy}^{\delta, \varepsilon} + (z z^T - \theta_x^\varepsilon \sigma \sigma^T (\theta_x^\varepsilon)^T) \tilde{V}_{yy}^{\delta, \varepsilon}] \\ &\leq \{w_s^{\delta, \varepsilon} + \varepsilon \Delta w^{\delta, \varepsilon} + \frac{1}{2} \text{tr}[\sigma \sigma^T w_{xx}^{\delta, \varepsilon}] + \langle b, w_x^{\delta, \varepsilon} \rangle\} + \varepsilon C. \end{aligned} \quad (5.12)$$

The above is true for all $\varepsilon, \delta > 0$ such that $|\theta_x^\varepsilon(s, x)\sigma(s, x, \theta^\varepsilon(s, x))| \leq 1/\delta$, which is always possible for any fixed ε , and $\delta > 0$ sufficiently small. Then we obtain

$$\begin{cases} w_s^{\delta, \varepsilon} + \mathcal{A}_\varepsilon w^{\delta, \varepsilon} \geq -\varepsilon C, & \forall (s, x) \in [0, T] \times \mathbb{R}^n, \\ w^{\delta, \varepsilon}|_{s=T} = 0. \end{cases}$$

On the other hand, by (5.2) and (H1), we see that, corresponding to the control $Z_\delta(\cdot) = 0 \in \tilde{\mathcal{Z}}_\delta[s, T]$, we have (by Gronwall's inequality) $|Y(T)| \leq C(1 + |y|)$, almost surely. Thus, by the boundness of g , we obtain (using Lemma 5.3)

$$\begin{aligned} 0 \leq w^{\delta, \varepsilon}(s, x) &\equiv \tilde{V}^{\delta, \varepsilon}(s, x, \theta^\varepsilon(s, x)) \leq \tilde{J}^{\delta, \varepsilon}(s, x, \theta^\varepsilon(s, x); 0) \\ &\leq C(1 + |\theta^\varepsilon(s, x)|) \leq C. \end{aligned}$$

Next, by Lemma 5.3 (with $\lambda_0 = g_0 = 0$, $\lambda = 1$ and $h_0 = \varepsilon C$), we must have $w^{\delta, \varepsilon}(s, x) \leq \varepsilon C e^T$, $\forall (s, x) \in [0, T] \times \mathbb{R}^n$. Thus, we obtain the following conclusion: There exists a constant $C_0 > 0$, such that for any $\varepsilon > 0$, one can find a $\delta = \delta(\varepsilon)$ with the property that

$$0 \leq \tilde{V}^{\delta, \varepsilon}(s, x, \theta^\varepsilon(s, x)) \leq \varepsilon C_0, \quad \forall \delta \leq \delta(\varepsilon). \quad (5.13)$$

Then, by (4.12), (4.23) (with $\delta = 0$) and (5.13), we obtain

$$\begin{aligned} 0 \leq V(0, x, \theta^\varepsilon(0, x)) &\leq |\tilde{V}^{0,0}(0, x, \theta^\varepsilon(0, x)) - \tilde{V}^{\delta, \varepsilon}(0, x, \theta^\varepsilon(0, x))| + \varepsilon C_0 \\ &\leq \eta_0(\varepsilon + \delta, |x| + |\theta^\varepsilon(0, x)|) + \varepsilon C_0. \end{aligned}$$

Now, we let $\delta \rightarrow 0$ and then $\varepsilon \rightarrow 0$ to get the right-hand side of the above going to 0. This can be achieved due to (5.4). Finally, since $\theta^\varepsilon(s, x)$ is bounded, we can find a convergent subsequence. Thus, we obtain that $V(0, x, y) = 0$, for some $y \in \mathbb{R}^m$. This implies (3.4). \square

6. Construction of Approximate Adapted Solutions

In this section we turn our attention to the general scheme of constructing the approximate adapted solutions. We note that in Corollary 4.5 such a scheme was actually outlined, provided that one is able to start from the right initial position $(x, y) \in \mathcal{N}(\bar{V})$ (or, equivalently, $V(0, x, y) = 0$). A scheme based on such an assumption, however, is not practical, because we usually do not have a way to access the value function \bar{V} directly, because of the possible degeneracy of the forward diffusion coefficient σ and the non-compactness of the admissible control set $\mathcal{Z}[0, T]$; which in turn makes starting from the nodal set practically difficult. In a special case, as presented in Section 5, we were able to approximate the nodal set $\mathcal{N}(\bar{V})$ without using the information of \bar{V} , and showed that the correct initial position could be chosen as $y^\varepsilon = \theta^\varepsilon(0, x)$, where θ^ε is the classical solution of (5.3). However, unfortunately the applicability of such a method still seems to be very limited, because it involves some other subtleties such as, among others, the estimate (5.4).

In what follows we propose a scheme for general approximately solvable FBSDEs, which will overcome the difficulty mentioned above. The main idea is to try to start from

some initial state that is “close” to the nodal set $\mathcal{N}(\bar{V})$ in a certain sense. We note that the unique strong solution, $\tilde{V}^{\delta, \varepsilon}$, to the HJB equation (4.11) is the value function of a *regularized* control problem with the state equation being (4.8) which is non-degenerate and with compact control set, thus many standard methods can be applied to study its analytical and numerical properties, on which our scheme will rely.

For notational convenience, in this section we assume that all the processes involved are one-dimensional (i.e., $n = m = d = 1$). However, one should be able to extend the scheme to general higher-dimensional cases without substantial difficulties. Furthermore, throughout this section we assume that

(H4) Let $g \in C^2$. There exists a constant $L > 0$, such that for all $(t, x, y, z) \in [0, T] \times \mathbb{R}^3$, it holds that

$$\begin{cases} |b(t, x, y, z)| + |\sigma(t, x, y, z)| + |h(t, x, y, z)| \leq L(1 + |x|); \\ |g'(x)| + |g''(x)| \leq L. \end{cases} \quad (6.1)$$

We first give a lemma that will be useful in our discussion.

Lemma 6.1. *Let (H1) and (H4) hold. Then there exists a constant $C > 0$, depending only on L and T , such that for all $\delta, \varepsilon \geq 0$, and $(s, x, y) \in [0, T] \times \mathbb{R}^2$, it holds that*

$$\tilde{V}^{\delta, \varepsilon}(s, x, y) \geq f(x, y) - C(1 + |x|^2), \quad (6.2)$$

where $f(x, y)$ is defined by (2.4).

Proof. First, it is not hard to check that the function f is twice continuously differentiable, such that for all $(x, y) \in \mathbb{R}^2$, the following hold:

$$\begin{cases} |f_x(x, y)| \leq |g'(x)|, & |f_y(x, y)| \leq 1, \\ f_{xx}(x, y) = \frac{(g(x) - y)g''(x)}{[1 + (y - g(x))^2]^{1/2}} + \frac{g'(x)^2}{[1 + (y - g(x))^2]^{3/2}}, \\ f_{yy}(x, y) = \frac{1}{[1 + (y - g(x))^2]^{3/2}} > 0, & f_{xy}(x, y) = -g'(x)f_{yy}(x, y). \end{cases} \quad (6.3)$$

Now for any $\delta, \varepsilon \geq 0$, $(s, x, y) \in [0, T] \times \mathbb{R}^2$ and $Z \in \tilde{\mathcal{Z}}_\delta[s, T]$, let (X, Y) be the corresponding solution to the controlled system (4.8). Applying Itô's formula we have

$$\begin{aligned} \tilde{J}^{\delta, \varepsilon}(s, x, y; Z) &= Ef(X(T), Y(T)) \\ &= f(x, y) + E \int_s^T \Pi(t, X(t), Y(t), Z(t)) dt, \end{aligned} \quad (6.4)$$

where

$$\begin{aligned} \Pi(t, x, y, z) &= f_x(x, y)b(t, x, y, z) + f_y(x, y)h(t, x, y, z) \\ &\quad + \frac{1}{2}[f_{xx}(x, y)\sigma^2(t, x, y, z) + 2f_{xy}(x, y)\sigma(t, x, y, z)z + f_{yy}(x, y)z^2] \\ &\geq f_x(x, y)b(t, x, y, z) + f_y(x, y)h(t, x, y, z) \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{2} \left[f_{xx} - \frac{f_{xy}^2}{f_{yy}} \right] (x, y) \sigma^2(t, x, y, z) \\
& \geq -C(1 + |x|^2),
\end{aligned} \tag{6.5}$$

where $C > 0$ depends only on the constant L in (H4), thanks to the estimates in (6.3). Note that (H4) also implies, by a standard argument using Gronwall's inequality, that $E|X(t)|^2 \leq C(1 + |x|^2)$, $\forall t \in [0, T]$, uniformly in $Z(\cdot) \in \tilde{\mathcal{Z}}_\delta[s, T]$, $\delta \geq 0$. Thus we derive from (6.4) and (6.5) that

$$\begin{aligned}
\tilde{V}^{\delta, \varepsilon}(s, x, y) &= \inf_{Z \in \tilde{\mathcal{Z}}_\delta[s, T]} \tilde{J}^{\delta, \varepsilon}(s, x, y; Z) \\
&= f(x, y) + \inf_{Z \in \tilde{\mathcal{Z}}_\delta[s, T]} E \int_s^T \Pi(t, X(t), Y(t), Z(t)) dt \\
&\geq f(x, y) - C(1 + |x|^2),
\end{aligned}$$

proving the lemma. \square

Next, for any $x \in \mathbb{R}$ and $r > 0$, we define

$$Q_x(r) \triangleq \{y \in \mathbb{R}: f(x, y) \leq r + C(1 + |x|^2)\},$$

where $C > 0$ is the constant in (6.2). Since $\lim_{|y| \rightarrow \infty} f(x, y) = +\infty$, $Q_x(r)$ is a compact set for any $x \in \mathbb{R}$ and $r > 0$. Moreover, Lemma 6.1 shows that, for all $\delta, \varepsilon \geq 0$, one has

$$\{y \in \mathbb{R}: \tilde{V}^{\delta, \varepsilon}(0, x, y) \leq r\} \subseteq Q_x(r). \tag{6.6}$$

From now on we set $r = 1$. Recall that by Proposition 4.2 and Theorem 4.3, for any $\rho > 0$, and fixed $x \in \mathbb{R}$, we can first choose $\delta, \varepsilon > 0$ depending only on x and $Q_x(1)$, so that

$$0 \leq \tilde{V}^{\delta, \varepsilon}(0, x, y) < V(0, x, y) + \rho, \quad \text{for all } y \in Q_x(1). \tag{6.7}$$

Now suppose that the FBSDE (1.1) is approximately solvable, we have from Proposition 3.3 that $\inf_{y \in \mathbb{R}} V(0, x, y) = 0$ (note that (H4) implies (H2)). By (6.6), we have

$$0 = \inf_{y \in \mathbb{R}} V(0, x, y) = \min_{y \in Q_x(1)} V(0, x, y).$$

Thus, by (6.7), we conclude the following:

Lemma 6.2. *Assume (H1) and (H4), and assume that the FBSDE (1.1) is approximately solvable. Then for any $\rho > 0$, there exist $\delta, \varepsilon > 0$, depending only on ρ, x and $Q_x(1)$, such that*

$$0 \leq \inf_{y \in \mathbb{R}} \tilde{V}^{\delta, \varepsilon}(0, x, y) = \min_{y \in Q_x(1)} \tilde{V}^{\delta, \varepsilon}(0, x, y) < \rho.$$

Our scheme for finding the approximate adapted solution of (1.1) starting from $X(0) = x$ can now be described as follows: for any integer k , we want to find $\{y^{(k)}\} \subset Q_x(1)$ and $\{Z^{(k)}\} \subset \mathcal{Z}[0, T]$ such that

$$Ef(X^{(k)}(T), Y^{(k)}(T)) \leq \frac{C_x}{k}. \tag{6.8}$$

Here and below $C_x > 0$ denotes the generic constant depending only on L, T and x . To be more precise, we propose the following steps for each fixed k :

Step 1. Choose $0 < \delta < 1/k$ and $0 < \varepsilon < \delta^4$, such that

$$\inf_{y \in \mathbb{R}} \tilde{V}^{\delta, \varepsilon}(0, x, y) = \min_{y \in Q_x(1)} \tilde{V}^{\delta, \varepsilon}(0, x, y) < \frac{1}{k}.$$

Step 2. For the given δ and ε , choose $y^{(k)} \in Q_x(1)$ such that

$$\tilde{V}^{\delta, \varepsilon}(0, x, y^{(k)}) < \min_{y \in Q_x(1)} \tilde{V}^{\delta, \varepsilon}(0, x, y) + \frac{1}{k}.$$

Step 3. For the given δ, ε and $y^{(k)}$, find $Z^{(k)} \in \mathcal{Z}_\delta[0, T]$, such that

$$J(0, x, y^{(k)}; Z^{(k)}) = Ef(X^{(k)}(T), Y^{(k)}(T)) \leq \tilde{V}^{\delta, \varepsilon}(0, x, y^{(k)}) + \frac{C_x}{k},$$

where $(X^{(k)}, Y^{(k)})$ is the solution to (2.1) with $Y^{(k)}(0) = y^{(k)}$ and $Z = Z^{(k)}$; and C_x is a constant depending only on L, T and x .

It is obvious that a combination of the above three steps will serve our purpose (6.8). We would like to remark here that in the whole procedure we do not use the exact knowledge about the nodal set $\mathcal{N}(\tilde{V})$, nor do we have to solve any degenerate parabolic PDEs, which are the two most formidable parts in this problem. Now that the Step 1 is a consequence of Lemma 6.2 and Step 2 is a standard (non-linear) minimizing problem, we only briefly discuss Step 3. Note that $\tilde{V}^{\delta, \varepsilon}$ is the value function of a regularized control problem, and by standard methods of constructing ε -optimal strategies using the information of value functions (see, e.g., Chapter 5 of [12]), we can find a Markov type control $\hat{Z}^{(k)}(t) = \alpha^{(k)}(t, \hat{X}^{(k)}(t), \hat{Y}^{(k)}(t))$, where $\alpha^{(k)}$ is some smooth function satisfying $\sup_{t, x, y} |\alpha^{(k)}(t, x, y)| \leq 1/\delta$ and $(\hat{X}^{(k)}, \hat{Y}^{(k)})$ is the corresponding solution of (4.8) with $\hat{Y}^{(k)}(0) = y^{(k)}$, so that

$$\tilde{J}^{\delta, \varepsilon}(0, x, y^{(k)}; \hat{Z}^{(k)}) < \tilde{V}^{\delta, \varepsilon}(0, x, y^{(k)}) + \frac{1}{k}. \quad (6.9)$$

The last technical point is that (6.9) is only true if we use the state equation (4.8), which is different from (2.1), the original control problem that leads to the approximate solution that we need. However, if we denote $(X^{(k)}, Y^{(k)})$ to be the solutions to (2.1) with $Y^{(k)}(0) = y^{(k)}$ and the feedback control $Z^{(k)}(t) = \alpha^{(k)}(X^{(k)}(t), Y^{(k)}(t))$, then a simple calculation shows that

$$\begin{aligned} 0 &\leq J(0, x, y^{(k)}; Z^{(k)}) = Ef(X^{(k)}(T), Y^{(k)}(T)) \\ &< Ef(\hat{X}^{(k)}(T), \hat{Y}^{(k)}(T)) + C_\alpha \sqrt{2\varepsilon} < \tilde{V}^{\delta, \varepsilon}(0, x, y^{(k)}) + \frac{1}{k} + C_\alpha \sqrt{2\varepsilon}, \end{aligned} \quad (6.10)$$

thanks to (6.9), where C_α is some constant depending only on L, T and the Lipschitz constant of $\alpha^{(k)}$. However, on the other hand, in light of Lemma 5.1 of [12], the Lipschitz constant of $\alpha^{(k)}$ can be shown to depend only on the bounds of the coefficients of the system (2.1) (i.e., b, h, σ and $\hat{\sigma}(z) \equiv z$) and their derivatives. Therefore using assumptions (H1) and (H4), and noting that $\sup_t |Z^{(k)}(t)| \leq \sup |\alpha^{(k)}| \leq 1/\delta$, we see

that, for fixed δ , C_α is no more than $C(1 + |x| + 1/\delta)$ where C is some constant depending only on L . Consequently, noting the requirement we posed on ε and δ in Step 1, we have

$$C_\alpha \sqrt{2\varepsilon} < C \left(1 + |x| + \frac{1}{\delta} \right) \sqrt{2\delta^4} \leq 2\sqrt{2}C(1 + |x|)\delta \leq \frac{C_x - 1}{k}, \quad (6.11)$$

where $C_x \triangleq C(1 + |x|)2\sqrt{2} + 1$. Finally, we note that the process $Z^{(k)}(\cdot)$ obtained above is $\{\mathcal{F}_t\}$ -adapted and hence is in $\mathcal{Z}_\delta[0, T]$ (instead of $\tilde{\mathcal{Z}}_\delta[0, T]$). This, together with (6.10)–(6.11), fulfills Step 3.

References

1. Antonelli, F. (1993) Backward–forward stochastic differential equations, *Ann. Appl. Probab.* 3:777–793.
2. Bensoussan, A. (1983) Stochastic maximum principle for distributed parameter system, *J. Franklin Inst.*, 315:387–406.
3. Bismut, J. M. (1973) *Théorie Probabiliste du Contrôle des Diffusions*, Mem. Amer. Math. Soc. 176, AMS, Providence, Rhode Island.
4. Buckdahn, R., and Hu, Y. (1998) Hedging contingent claims for a large investor in an incomplete market, *Adv. in Appl. Probab.*, 30:239–255.
5. Crandall, M. G., Ishii, H., and Lions, P. L. (1992) User’s guide to viscosity solutions of second order partial differential equations, *Bull. Amer. Math. Soc. (NS)*, 27:1–67.
6. Cvitanic, J., and Ma, J. (1996) Hedging options for a large investor and forward–backward SDEs, *Ann. Appl. Probab.*, 6:370–398.
7. Duffie, D., and Epstein, L. (1992) Stochastic differential utility, *Econometrica*, 60:353–394.
8. Duffie, D., Ma, J., and Yong, J. (1995) Black’s consol rate conjecture, *Ann. Appl. Probab.*, 5:356–382.
9. El Karoui, N., Peng, S., and Quenez, M. C. (1997) Backward stochastic differential equations in finance, *Math. Finance*, 7:1–71.
10. Fleming, W. H., and Soner, H. M. (1992) *Controlled Markov Processes and Viscosity Solutions*, Springer-Verlag, New York.
11. Hu, Y., and Peng, S. (1995) Solution of forward–backward stochastic differential equations, *Probab. Theory Rel. Fields*, 103:273–283.
12. Krylov, N. (1980) *Controlled Diffusion Processes*, Springer-Verlag, New York.
13. Ladyzhenskaya, O. A., and Ural’tseva, N. N. (1968) *Linear and Quasilinear Elliptic Equations*, Academic Press, New York.
14. Ma, J., Protter, P., and Yong, J. (1994) Solving forward–backward stochastic differential equations explicitly—a four step scheme, *Probab. Theory Rel. Fields*, 98:339–359.
15. Ma, J., and Yong, J. (1995) Solvability of forward–backward SDEs and the nodal set of HJB equations, *Chinese Ann. Math. Ser. B*, 16:279–298.
16. Pardoux, E., and Peng, S. (1990) Adapted solution of a backward stochastic differential equation, *Systems Control Lett.* 14:55–61.
17. Pardoux, E., and Tang, S. (1999) The study of the forward–backward stochastic differential equation and its application in quasilinear PDEs, *Probab. Theory Rel. Fields*, 114:123–150.
18. Peng, S., and Wu, Z. (1999) Fully coupled forward–backward stochastic differential equations and applications to optimal control, *SIAM J. Control Optim.*, 37:825–843.
19. Protter, P. (1990) *Stochastic Integration and Differential Equations, A New Approach*, Springer-Verlag, Berlin.
20. Yong, J. (1997) Finding adapted solutions of forward–backward stochastic differential equations—method of continuation, *Probab. Theory Rel. Fields*, 107:537–572.
21. Yong, J. (1999) Linear forward–backward stochastic differential equations, *Appl. Math. Optim.*, 39:93–119.

This paper was completed in 1998. Since then there have been many further developments in the study of FBSDEs. The following references and the extensive literature cited therein might be useful to interested readers.

Ma, J., and Yong, J. (1999) *Forward–Backward Stochastic Differential Equations and Their Applications*, Lecture Notes in Mathematics, 1702, Springer-Verlag, Berlin.

Yong, J. (to appear) *Forward–backward stochastic differential equation—a useful tool for mathematical finance and other related fields*, *Survey Indust. Math.*

Accepted 17 April 2001. Online publication 14 August 2001.