We attempt to present a new numerical approach to solve nonlinear backward stochastic differential equations. First, we present some definitions and theorems to obtain the condition, from which we can approximate the nonlinear term of the backward stochastic differential equation (BSDE) and we get a continuous piecewise linear BSDE corresponding to the original BSDE. We use the relationship between backward stochastic differential equations and stochastic controls by interpreting BSDEs as some stochastic optimal control problems to solve the approximated BSDE and we prove that the approximated solution converges to the exact solution of the original nonlinear BSDE.

1. Introduction

The backward stochastic differential equation (BSDE) theory and its applications have been a focus of interest in recent years. Consider the nonlinear BSDE

\[ dx(t) = f(x(t), z(t), t) \, dt + z(t) \, dW(t), \]

\[ x(T) = \xi, \quad (1.1) \]

where \( \xi \) is a random variable that will become certain only at the terminal time \( T \). This type of equation, at least in nonlinear case, was first introduced by Pardoux and Peng [14], who proved the existence and uniqueness of a solution under suitable assumptions on \( f \) and \( \xi \). Their aim was to give a probabilistic interpretation of a solution to a second-order quasilinear partial differential equation. The BSDE theory has found wide applications in partial differential equations, stochastic controls, and especially, mathematical finance.

For example, the Black-Scholes formula for option pricing can be recovered via a system of forward-backward stochastic differential equations (FBSDEs). In this case, the random terminal condition is related to the price of the underlying stock at a given terminal date. Unlike a (forward) SDE, the solution of a BSDE is a pair of adapted processes \((x(\cdot), z(\cdot))\). The additional term \( z(\cdot) \) may be interpreted as a risk-adjustment factor and is required for the equation to have adapted solutions. This restriction of solutions to
the class of adapted processes is necessary if the insights gained from the study of BSDEs are to be useful in applications. The adapted processes depend on the past and present information but do not rely on the future knowledge. This is normal virtually in all applications; for example, the replicating portfolio for a contingent claim may depend at any particular time on the past and present stock price but not, quite normally, on the future stock prices [15].

The still-standing problem in the theory of BSDEs is to find an implementable numerical method. Many efforts have been made in this direction as well. For example, in the Markovian case, Douglas et al. [7] established a numerical method for a class of forward-backward SDEs—a more general version of the BSDEs (1.1)—based on a four-step scheme developed by Ma et al. [11]. Chevance, in his Ph.D. thesis [6], proposed a numerical method for BSDEs by using binomial approach to approximate the process \( x \). In the non-Markovian case, where the terminal value \( \xi \) is allowed to depend on the history of a forward diffusion \( x \), Bally [3] presented a discretization scheme and obtained its rate of convergence. Recently, Briand et al. [5] and Ma et al. [10] also proposed some numerical methods for BSDEs with path-dependent terminal values.

Zhang and Zheng [17] suggested another method using PDE approach, in which high regularity of coefficients are required. Finally, Zhang, in his Ph.D. thesis [16], presented a numerical method for a class of BSDEs, whose terminal value \( \xi \) takes the form \( \phi(x) \), where \( x \) is a diffusion process, and \( \phi(\cdot) \) is a so-called \( L^\infty \)-Lipschitz functional.

On the other hand, Kohlmann and Zhou [9] explored the relationship between linear BSDEs and stochastic controls from a different angle by interpreting BSDEs as a stochastic linear-quadratic (LQ) control problem which can be solved explicitly. As mentioned in [9], “To be precise, note that in (1.1) the terminal value is specified while the initial value is left open. Nevertheless, if the equation has a solution, then the initial value cannot be chosen arbitrarily; rather it is uniquely determined by the solution and is hence part of the solution. Therefore, solving (1.1) amounts to the following statement: start with a proper initial condition and choose an appropriate diffusion term to hit the given value \( \xi \) at the terminal. Then it will be very natural to modify the above statement and consider the following stochastic optional control problem. For the same dynamics of (1.1), starting with a given initial state \( x_0 \), choose a control \( z(\cdot) \) so that the terminal state \( x(T) \) stays as close to the given terminal value \( \xi \) as possible. Note that since now the initial value \( x_0 \) is given a priori, one in general cannot expect that \( x(T) \) will hit \( \xi \) exactly by choosing certain \( z(\cdot) \). Hence, it is reasonable to require that the difference between the two be minimized. Here, the “difference” may be measured by, say, the second moment of the algebraic difference between the two random variables. More interestingly, if we regard the initial state also as a decision variable, then the optimal state-control pair of the problem \((x(\cdot), z(\cdot))\) is exactly the solution of the original BSDE!”

In this paper, we will try to develop this method for nonlinear BSDEs in such a way that the nonlinear coefficient of BSDE (1.1) can be approximated by a piecewise linear function and therefore we obtain a piecewise linear BSDE associated with (1.1). Then, we will prove that the solution of the obtained piecewise linear BSDE converges to the original solution of the nonlinear BSDE (1.1). Afterwards, we will find the optimal control problem associated with the obtained piecewise linear BSDE.
2. Problem formulation and preliminaries

2.1. Notations. Throughout this paper $(\Omega, F, P, \{F_t\}_{t \geq 0})$ is a fixed filtered complete probability space on which a standard $F_t$-adapted $m$-dimensional Brownian motion $W(t) = (w^1(t), \ldots, w^m(t))^\prime$ with $W(0) = 0$, is defined. It is assumed that $F_t = \sigma\{W(t) : s \leq t\}$. Let $X$ be a given Hilbert space. The set of $X$-valued continuous functions is denoted by $C([0, T]; X)$. Suppose $\eta : \Omega \rightarrow \mathbb{R}^n$ is an $F_t$-random variable, we write $\eta \in L^2_F(\Omega; \mathbb{R}^n)$, if $\eta$ is square integrable (i.e., $E|\eta|^2 < \infty$). Now, consider the case when $f : [0, T] \times \Omega \rightarrow \mathbb{R}^n$ is an $\{F_t\}_{t \geq 0}$ adapted process. If $f(\cdot)$ is square integrable (i.e., $E \int_0^T |f(t)|^2 dt < \infty$), we will write $f(\cdot) \in L^2_F(0, T; \mathbb{R}^n)$. These definitions are general in the obvious way for the case when $f(\cdot)$ is $\mathbb{R}^{n \times m}$-valued.

Also, we make the following additional notations:

(i) $P^\prime$ denotes the transpose of any vector or matrix $P$;
(ii) $\langle a, b \rangle$ denotes the inner product of two vectors $a$ and $b$ (i.e., $\langle a, b \rangle = a \cdot b'$);
(iii) $S^n$ denotes the space of all $n \times n$ symmetric matrices;
(iv) $S^n_+$ denotes the subspace of all positive definite matrices of $S^n$;
(v) $H^{2,k}$ denotes the set of $P$-measurable processes $V = \{V_t\}_{t \leq 1}$ with values in $\mathbb{R}^k$ such that $E \int_0^1 |V_s|^2 ds < \infty$;
(vi) $S_+^{2,k}$ denotes the set of continuous $P$-measurable processes $V = \{V_t\}_{t \leq 1}$ with values in $\mathbb{R}^k$ such that $E(\sup_{t \leq 1} |V_t|^2) < \infty$.

2.2. Existence and uniqueness of solution. Consider the following nonlinear BSDE:

\[ dx(t) = f(x(t), z(t), t) dt - z(t) dW(t), \]
\[ x(T) = \xi, \quad (2.1) \]

where $x(t) \in \mathbb{R}^n$, $z(t) \in \mathbb{R}^{n \times m}$, $f$ is a mapping from $\mathbb{R}^n \times \mathbb{R}^{n \times m} \times [0, T]$ to $\mathbb{R}^n$, and $\xi \in L^2_F(\Omega; \mathbb{R}^n)$.

Note that $f$ can be random, here, we assume that $f$ is a deterministic function.

Using the definition of the stochastic differential, (2.1) is equivalent to the following stochastic integral equation:

\[ x(t) = \xi - \int_t^T f(x(s), z(s), s) dt + \int_t^T z(s) dW(s). \quad (2.2) \]

We now give a precise definition of a solution to the BSDE.

Definition 2.1. A pair of stochastic processes

\[ \{x(t), z(t)\}_{0 \leq t \leq T} \in L^2_F([0, T]; \mathbb{R}^n) \times L^2_F([0, T]; \mathbb{R}^{n \times m}) \quad (2.3) \]

is called an adapted solution of (2.1) if it has the following properties:

(i) $f(x(\cdot), z(\cdot), \cdot) \in L^2_F([0, T]; \mathbb{R}^n)$,
(ii) equation (2.2) holds for every $t \in [0, T]$ with probability 1.
A solution \( \{x(t), z(t)\} \) is said to be unique if for any other solution \( \{\bar{x}(t), \bar{z}(t)\} \), we have

\[
P\{x(t) = \bar{x}(t) \forall 0 \leq t \leq T\} = 1, \quad E \int_0^T |z(s) - \bar{z}(s)|^2 ds = 0. \tag{2.4}
\]

Pardoux and Peng proved the following existence and uniqueness theorem [14].

**Theorem 2.2.** Assume that

\[
f(0,0, \cdot) \in L^2_{\mathbb{P}}([0,T];\mathbb{R}^n). \tag{2.5}
\]

Also, assume that there exists a positive constant \( K > 0 \) such that

\[
|f(x,z,t) - f(\bar{x},\bar{z},t)|^2 \leq K(|x - \bar{x}|^2 + |z - \bar{z}|^2) \text{ a.s.} \tag{2.6}
\]

for all \( x, \bar{x} \in \mathbb{R}^n, z, \bar{z} \in \mathbb{R}^{n \times m}, \text{ and } t \in [0,T], \) then there exists a unique solution \( \{x(t), z(t)\} \) to (2.1) in \( L^2_{\mathbb{P}}([0,T];\mathbb{R}^n) \times L^2_{\mathbb{P}}([0,T];\mathbb{R}^{n \times m}). \)

For the proof of Theorem 2.2, see [13, Chapter 7] or [14].

The uniform Lipschitz condition is too strong to be applied. In [8], Hamadène proved the existence and uniqueness of solution to (2.1) in weaker conditions. To explore the theorem, which was proved by Hamadène, we need to present some assumptions.

**Assumption 2.3.** The process \( \{f(0,0,t)\}_{t \leq 1} \) belongs to \( H^{2,n} \) and, for any \( (x,z) \in \mathbb{R}^n \times \mathbb{R}^{n \times m}, \) \( \{f(x,z,t)\}_{t \leq 1} \) is \( P \)-measurable.

Let \( \Phi: \mathbb{R}^+ \to \mathbb{R}^+ \) be a continuous function of at most linear growth such that, \( \Phi(0) = 0 \) and \( \Phi(y) > 0 \) for all \( y > 0 \). We would like \( \Phi \) to satisfy the following assumption.

**Assumption 2.4.** \( u^\gamma(0) \to 0 \) as \( \gamma \to 0 \), where \( u^\gamma \) is the unique solution on \([0,1]\) of the deterministic backward differential equation (DBE)

\[
u^\gamma(t) = \gamma + \int \Phi(n \cdot u^\gamma(t)) dt \quad t \leq 1, \gamma > 0. \tag{2.7}
\]

This assumption implies, besides Assumption 2.3, the mapping \( (x,z,t) \to f(x,z,t) \) satisfies the following.

**Assumption 2.5.** (i) \( f \) is uniformly continuous in \( x \) uniformly with respect to \( (z,t) \), that is, there exists a continuous nondecreasing function \( \Phi \) from \( \mathbb{R}^+ \) into itself with at most linear growth and satisfying \( \Phi(0) = 0 \) and \( \Phi(y) > 0 \) for all \( y > 0 \) such that

\[
|f(x,z,t) - f(x',z',t)| \leq \Phi(|x - x'|) \quad \forall t,x,x',z \text{ a.s.,} \tag{2.8}
\]

moreover, \( \Phi \) satisfies Assumption 2.4.

(ii) \( f \) is uniformly continuous in \( z \), that is, there exists a continuous function \( \psi \) from \( \mathbb{R}^+ \) into itself with at most linear growth and satisfying \( \psi(0) = 0 \), such that

\[
|f(x,z,t) - f(x,z',t)| \leq \psi(|z - z'|) \quad \forall t,x,z,z' \text{ a.s.} \tag{2.9}
\]
Assumption 2.7. (i) For all such that \( t \in (0, 0, T) \) is continuous and nondecreasing, grows at most linearly and satisfies Theorem 2.10.

\[ x(t) = \xi + \int_t^1 f(x(s), z(s), s) ds - \int_t^1 z(s) dW(s) \quad \forall t \leq 1. \]  

(2.10)

For the proof of Theorem 2.6, see [8, proof of Theorem 3.1].

Assumption 2.7. (i) For all \( t, x, x', z, |f(x, z, t) - f(x', z, t)| \leq \Phi(|x - x'|) \), where \( \Phi \) is continuous and nondecreasing, grows at most linearly and satisfies \( \Phi(0) = 0, \Phi(y) > 0 \), for all \( y > 0 \), and \( \int_0^y [\Phi(y)] dy = \infty \).

(ii) The function \( z \to f(x, z, t) \) is Lipschitz uniformly with respect to \((x, t)\). In addition, for any \( i = 1, \ldots, n \), \( f_i(x, z, t) \) the \( i \)th component of \( f \), depends only on the \( i \)th row of the matrix \( z \).

Now, we have the following result.

Theorem 2.8. Under Assumptions 2.3 and 2.7, the solution \((x, z)\) of the BSDE (2.10) associated with \((f, \xi)\) is unique.

For the proof of Theorem 2.8, see [8, proof of Theorem 4.1].

The following theorem, which was proved by Mao in [12], is the especial case of Theorem 2.8.

Theorem 2.9. Assume that the following conditions are fulfilled:

(i) \( f(x(\cdot), z(\cdot), \cdot) \in L^2(|0, T|; \mathbb{R}^n) \),

(ii) for all \( x, \overline{x} \in \mathbb{R}^n, z, \overline{z} \in \mathbb{R}^{n \times m}, \text{ and } t \in [0, T] \),

\[ |f(x, z, t) - f(\overline{x}, \overline{z}, t)|^2 \leq \mathcal{K}(|x - \overline{x}|^2) + K(|z - \overline{z}|^2) \text{ a.s.,} \]  

(2.11)

where \( K \) is a positive constant and \( \mathcal{K}(\cdot) \) is a concave increasing function from \([0, \infty) \) to \([0, \infty)\) such that \( \mathcal{K}(0) = 0, \mathcal{K}(u) > 0 \) for \( u > 0 \) and

\[ \int_0^\infty \frac{du}{\mathcal{K}(u)} = \infty. \]  

(2.12)

Then there exists a unique solution \((x(t), z(t))\) to (2.1) in \( L^2([0, T]; \mathbb{R}^n) \times L^2([0, T]; \mathbb{R}^{n \times m}) \).

For the proof of Theorem 2.9, see [12] or [13, Chapter 7].

2.3. Regularities. In this subsection, we present two theorems, which guarantee the \( p \)th moment \((p \geq 2)\) of the solution to (2.1) to be finite.

Theorem 2.10. Let \( p \geq 2 \) and \( \xi \in L^p(F_0; \mathbb{R}^n) \). Let the following conditions hold:

(i) \( f(x(\cdot), z(\cdot), \cdot) \in L^p([0, T]; \mathbb{R}^n) \),

(ii) there exists a positive constant \( K > 0 \) such that

\[ |f(x, z, t) - f(0, 0, t)|^2 \leq K(1 + |x|^2 + |z|^2) \text{ a.s.} \]  

(2.13)
for all $x \in \mathbb{R}^n$, $z \in \mathbb{R}^{n \times m}$, and $t \in [0, T]$. Then the solution of (2.1) has the properties that
\[
E |x(t)|^p \leq (E|\xi|^p + C_1) \exp(2pT(4K + 1)) \quad \forall 0 \leq t \leq T,
\]
\[
E \int_0^T |x(t)|^{p-2} \left| z(t) \right|^2 dt \leq \frac{4}{p} (E|\xi|^p + C_1) \left[ 1 + \exp(2pT(4K + 1)) \right],
\]
where
\[
C_1 = \frac{2}{p} \int_0^T \left[ \frac{p}{8K} \left( |f(0,0,s)|^2 + K \right) + 4p(1+K) \right] ds < \infty.
\]
Note that (2.11) implies (2.13).

For the proof of Theorem 2.10, see [13, Chapter 7, proof of Theorem 5.1, page 259].

**Theorem 2.11.** Let $p \geq 2$ and $\xi \in L_{p(T)}^p(\Omega; \mathbb{R}^n)$. Let (2.11) and (2.12) hold. Then the solution of (2.1) has the properties that
\[
E \left( \sup_{0 \leq t \leq T} |x(t)|^p \right) < \infty, \quad E \left( \int_0^T |z(t)|^2 dt \right)^{p/2} < \infty.
\]

For the proof of Theorem 2.11, see [13, Chapter 7, proof of Theorem 5.2, page 261].

### 3. Continuous piecewise linear approximation

In this section, we try to find a way to be able to approximate the function $f$ in (2.1) by a continuous piecewise linear function.

By using Chebyshev’s inequality [2, Chapter 2, page 84] and Theorems 2.10 and 2.11, we can derive that, there exist two positive constants $M_1$ and $M_2$ such that
\[
\forall t \in [0, T], \omega \in \Omega \quad |x(t)|^2 \leq M_1 \text{ a.s.,}
\]
\[
\forall t \in [0, T], \omega \in \Omega \quad |z(t)|^2 \leq M_2 \text{ a.s.}
\]
Therefore, there exist the closed ball $B_1$ in $\mathbb{R}^n$ and the closed ball $B_2$ in $\mathbb{R}^{n \times m}$, respectively, for $x(\cdot)$ and $z(\cdot)$ such that
\[
\forall t \in [0, T], \omega \in \Omega, \quad x(t) \in B_1, \quad z(t) \in B_2 \text{ a.s.}
\]
Now, we define the closed set $B$ in $\mathbb{R}^n \times \mathbb{R}^{n \times m} \times [0, T]$ as follows:
\[
B = B_1 \times B_2 \times [0, T].
\]

The following lemma plays an important role in this section.

**Lemma 3.1.** Let $f : N \to M$ be a function from the metric space $N$ to another metric space $M$. Let also, $f$ be uniformly continuous on $A$ in $D_f$ (domain of $f$), and $M$ be a complete space. Then, there exists a uniformly continuous function $\overline{f}$ on $\overline{A}$ (closure of $A$), such that the restriction of $\overline{f}$ on $A$ is equal to the restriction of $f$ on $A$ (i.e., $\overline{f}|_A \equiv f|_A$).
For the proof of **Lemma 3.1**, see [1, Chapter 4].

Now, we can start the approximation process. First, we define the closed set $B_c \in \mathbb{R}^n \times \mathbb{R}^{n \times m} \times [0, T]$ as follows:

$$B_c = B \cap D_f,$$

(3.4)

where $D_f$ is the domain of $f$ in (2.1). Conditions (2.6) and (2.10) imply that $f$ in (2.1) is a uniformly continuous function on $B_c$ and so, **Lemma 3.1** holds for $f$. In other words, if $\bar{f}$ is the unique extension of $f$ on $B_c$, then $\bar{f}$ is uniformly continuous on $B_c$.

The next theorem implies that if we consider function $f$ as a deterministic function, then for a given $\varepsilon > 0$, we can find a uniform approximation of the function by the linear functions as $a_1 k x(t) + \sum_{j=1}^{m} a_{2k}^j z_j(t) + a_{3k}$, where $a_{1k}, a_{2k}^j, a_{3k} \in \mathbb{R}^n$, and $z_j(\cdot)$ is the $j$th column of $z(\cdot)$. Alternatively, for every $\varepsilon > 0$ given, we can divide the set $B_1$ to the finite disjoint cells $I_1, I_2, \ldots, I_q$, $q = q(\varepsilon)$, such that if the function $h_{q(\varepsilon)}$ is defined as follows:

$$\forall (x(t), z(t), t) \in I_k, \ h_{q(\varepsilon)}(x(t), z(t), t) = a_{1k} x(t) + \sum_{j=1}^{m} a_{2k}^j z_j(t) + a_{3k}, \ k = 1, 2, \ldots, q,$$

(3.5)

then, we have

$$| h_{q(\varepsilon)} - \bar{f} |_{B_c} < \varepsilon,$$

(3.6)

where

$$| h_{q(\varepsilon)} - \bar{f} |_{B_c} = \text{ess sup} \{ | h_{q(\varepsilon)}(x(t), z(t), t) - \bar{f}(x(t), z(t), t) | : (x(t), z(t), t) \in B_c \}.$$  

(3.7)

**Theorem 3.2.** Let $f$ be a continuous function, whose domain $D_f$ is a compact cell in $\mathbb{R}^p$ and whose values belong to $\mathbb{R}^q$. Then $f$ can be uniformly approximated on $D_f$ by continuous piecewise linear functions.

For the proof of **Theorem 3.2**, see [4, proof of Theorem 24.5, page 169].

Based on what we said, $\bar{f}(\cdot)$ satisfies the conditions of **Theorem 3.2**, therefore, since $f(\cdot)$ is the restriction of $\bar{f}(\cdot)$ on $B \cap D_f$, we can conclude that

$$| h_{q(\varepsilon)} - f |_{B \cap D_f} < \varepsilon,$$

(3.8)

and it means that, we can approximate function $f(\cdot)$ in (2.1) by a continuous piecewise linear function such $h_{q(\varepsilon)}(\cdot)$ on $B \cap D_f$, almost surely.

Hence, if we replace $h_{q(\varepsilon)}$ with coefficient $f$ in (2.1), we have

$$dx(t) \equiv h_{q(\varepsilon)}(x(t), z(t), t) dt - z(t) dW(t),$$

$$x(T) = \xi,$$

(3.9)
and by definition \{x_{q(\varepsilon)}(t), z_{q(\varepsilon)}(t)\} as the approximated solution to (2.1), we obtain a new BSDE as follows:

\[
dx_{q(\varepsilon)}(t) = h_{q(\varepsilon)}(x_{q(\varepsilon)}(t), z_{q(\varepsilon)}(t), t) dt - z_{q(\varepsilon)}(t) dW(t),
\]
\[
x_{q(\varepsilon)}(T) = \xi.
\]

(3.10)

In the next section, we will show that \{x_{q(\varepsilon)}(t), z_{q(\varepsilon)}(t)\} converges to \{x(t), z(t)\} in three different cases.

4. Convergence

In this section, we prove the convergence of the approximated solution \{x_{q(\varepsilon)}(t), z_{q(\varepsilon)}(t)\} to the exact solution to (2.1).

**Theorem 4.1.** Assume that the conditions of Theorems 2.6 and 2.8 are fulfilled. Also, assume that \{x_{q(\varepsilon)}(t), z_{q(\varepsilon)}(t)\} to be the solution of (3.10), when \(T = 1\). Then the approximated solution \{x_{q(\varepsilon)}(t), z_{q(\varepsilon)}(t)\} converges to \(x(t), z(t)\) in the sense that

\[
\lim_{\varepsilon \to 0} |x(t) - x_{q(\varepsilon)}(t)| = 0 \quad \forall t \in [0, 1].
\]

4.1.

**Proof.** Let \(\varepsilon > 0\) be given. First, we prove that the process \(x - x_{q}\) is uniformly bounded, that is, there exists a constant \(\tilde{C}\) such that \(|x(t) - x_{q}(t)| \leq \tilde{C}\), for all \(t \leq 1\). Indeed, using Ito’s formula we arrive, for all \(t \leq 1\),

\[
|x - x_{q}|^2 + \int_{t}^{1} |z - z_{q}|^2 ds = 2 \int_{t}^{1} \langle x - x_{q}, f(x, z, s) - h_{q}(x_{q}, z_{q}, s) \rangle ds \\
+ 2 \int_{t}^{1} \langle x - x_{q}, z - z_{q} \rangle dW(s).
\]

(4.2)

On the other hand, let \(f(x_{q}, z_{q}, s) - h_{q}(x_{q}, z_{q}, s) \equiv \varepsilon(x_{q}, z_{q}, s)\), again for the simplicity, we set \(\varepsilon(s) \equiv \varepsilon(x_{q}, z_{q}, s)\), therefore

\[
|x - x_{q}|^2 + \int_{t}^{1} |z - z_{q}|^2 ds \\
= 2 \int_{t}^{1} \langle x - x_{q}, f(x, z, s) - f(x_{q}, z_{q}, s) + \varepsilon(s) \rangle ds + 2 \int_{t}^{1} \langle x - x_{q}, z - z_{q} \rangle dW(s) \\
= 2 \int_{t}^{1} \langle x - x_{q}, f(x, z, s) - f(x_{q}, z_{q}, s) \rangle ds + 2 \int_{t}^{1} \langle x - x_{q}, \varepsilon(s) \rangle ds \\
+ 2 \int_{t}^{1} \langle x - x_{q}, z - z_{q} \rangle dW(s) \\
\leq 2 \int_{t}^{1} |x - x_{q}| \Phi(|x - x_{q}|) ds + 2k \int_{t}^{1} |x - x_{q}| |z - z_{q}| ds \\
+ 2 \int_{t}^{1} \langle x - x_{q}, \varepsilon(s) \rangle ds + 2 \int_{t}^{1} \langle x - x_{q}, z - z_{q} \rangle dW(s),
\]

where \(k\) is the Lipschitz constant. Since, for all \(a, b \in \mathbb{R}\) and \(\delta > 0\), \(|a \cdot b| \leq \delta a^2 + \delta^{-1} b^2\) and the growth of \(\Phi\) is at most linear, then \(\Phi(|y|)^2 \leq C_2 (1 + |y|^2)\) for all \(y \in \mathbb{R}\), we can
derive that
\[
|x - x_q|^2 + \int_t^1 |z - z_q|^2 ds \leq C_2 \int_t^1 |x - x_q|^2 ds + 2 \int_t^1 (1 + |x - x_q|^2) ds \\
+ 8k \int_t^1 |x - x_q|^2 ds + \frac{1}{2} \int_t^1 |z - z_q|^2 ds \\
+ \frac{2}{C_2} \int_t^1 |\epsilon(s)|^2 ds + 2 \int_t^1 \langle x - x_q, z - z_q \rangle dW(s). 
\]

(4.4)

On the other hand, since \((x, z)\) and \((x_q, z_q)\) belong to \(S^{2, \infty} \times H^{2, n \times m}\), then using the Burkholder-Davis-Gundy inequality, we deduce that \(\{f_t'(x - x_q, z - z_q) dW(s)\}_{t \leq 1}\) is an \(F_t\)-martingale. We thus have
\[
|x - x_q|^2 \leq C_3 \left\{ 1 + \epsilon^2 + \int_t^1 |x - x_q|^2 ds \right\} + \int_t^1 \langle x - x_q, z - z_q \rangle dW(s) \quad t \leq 1, 
\]

(4.5)

where \(\epsilon \geq \sup_{0 \leq t \leq T} |\epsilon(t)|\).

Then, for any \(r \geq t \geq 0\), we have
\[
E \left( \left| x(r) - x_q(r) \right|^2 \mid F_t \right) \leq C_3 \left\{ 1 + \epsilon^2 + \int_r^1 E \left( \left| x(s) - x_q(s) \right|^2 \mid F_t \right) ds \right\}. 
\]

(4.6)

Now by the Gronwall inequality we obtain \(E(|x(r) - x_q(r)|^2 \mid F_t) \leq C\), which yields the desired result after taking \(r = t\).

Now, we show the convergence of the approximated solution to the exact solution.

For any \(i = 1, \ldots, n\) and \(t \leq 1\), we have
\[
x^i(t) - x^i_q(t) = \int_t^1 \left( f^i(x, z^i, s) - h^i_q(x_q, z_q^i, s) \right) ds - \int_t^1 (z^i - z^i_q) dW(s), 
\]

(4.7)

where \(x^i, x^i_q, f^i, h^i_q, z^i,\) and \(z^i_q\) are the \(i\)th components and rows of, respectively, \(x, x_q, f, h_q, z\) and \(z_q\). Then, using Tanaka’s formula, we obtain
\[
|x^i - x^i_q| + 2(A^i_q(0) - A^i_q(0)) = \int_t^1 \text{sgn} \left( x^i - x^i_q \right) \left( f^i(x, z^i, s) - h^i_q(x_q, z^i_q, s) \right) ds \\
- \int_t^1 \text{sgn} \left( x^i - x^i_q \right) \langle z^i - z^i_q \rangle dW(s) \quad t \leq 1, 
\]

(4.8)

where \(\{A^i_q(0)\}_{t \leq 1}\) is the local time of \(x^i - x^i_q\) at 0.

Again, let \(f(x_q, z_q, s) - h_q(x_q, z_q, s) \equiv \epsilon(x_q, z_q, s)\), and we set \(\epsilon(s) \equiv \epsilon(x_q, z_q, s)\), therefore
\[
|x^i - x^i_q| + 2(A^i_q(0) - A^i_q(0)) = \int_t^1 \text{sgn} \left( x^i - x^i_q \right) \left( f^i(x, z^i, s) - f^i(x_q, z^i_q, s) \right) ds \\
+ \int_t^1 \text{sgn} \left( x^i - x^i_q \right) \epsilon(s) ds \\
- \int_t^1 \text{sgn} \left( x^i - x^i_q \right) \langle z^i - z^i_q \rangle dW(s) \quad t \leq 1. 
\]

(4.9)
Since the proof of Theorem 3.1.

Then, for all $\varepsilon$ where $\varepsilon > 0$ and $\delta > 0$, we have

$$
\int_{t}^{1} \Phi(x^i - x^i)'(f^i(x, z^i, s) - f^i(x_q, z^i_q, s)) ds + \varepsilon^2
$$

Thus, for any $\Phi$ such that $\delta > 0$, let $W$ be a Lipschitz function from $\mathbb{R}^d$ into itself such that, for all $y \in \mathbb{R}^d$, $\Phi_d(y) \rightarrow \Phi(y)$ as $d \rightarrow \infty$ (see [8, proof of Proposition 2.2] for the existence of $\Phi_d$).

For $d \geq 0$ and $\delta > 0$, let $v^d_{\delta}$ be the function such that

$$
v^d_{\delta}(t) = \delta + \int_{t}^{1} \Phi_d(n \cdot v^d_{\delta}(s)) ds \quad t \leq 1.
$$

Since $\{\Phi_d\}_{d \geq 0}$ is a nonincreasing sequence, $v^{d+1}_{\delta} \leq v^d_{\delta}$ for any $d \geq 0$. This implies that the sequence $\{v^d_{\delta}\}_{d \geq 0}$ converges pointwise to a function $v^\delta : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, which satisfies

$$
v^\delta(t) = \delta + \int_{t}^{1} \Phi(n \cdot v^\delta(s)) ds \quad t \leq 1.
$$

Now if $\delta \leq \lambda$ then $v^d_{\delta} \leq v^d_{\lambda}$ for any $d \geq 0$, and $v^\delta \leq v_\lambda$. It follows that $v^\delta \rightarrow \overline{v}$ as $\delta \rightarrow 0$, where for any $t \leq 1$, $\overline{v}_t = \int_{t}^{1} \Phi(n \cdot \overline{v}(s)) ds$, so that $\overline{v} \equiv 0$ (according to [8, Proposition 3.2]).

Now for $\delta$, $d$, and $k_1 \geq 0$, let $v^{d,k_1}_{\delta}$ be the function defined recursively as follows:

$$
v^{d,0}_{\delta} = \tilde{C},
$$

$$
v^{d,k_1}_{\delta}(t) = \delta + \int_{t}^{1} \Phi_d(n \cdot v^{d,k_1}_{\delta}(s)) ds \quad k_1 \geq 1, \ t \leq 1.
$$
Since $\Phi_d$ is Lipschitz, $v^{d,k_1}_\delta \to v^d_\delta$ as $k_1 \to +\infty$. On the other hand, it is easily seen by induction that for all $k_1 \geq 0$, $|x^i(t) - x^i_q(t)| \leq v^{d,k_1}_\delta(t)$, $t \leq 1$, $i = 1, \ldots, n$. Indeed, for $k_1 = 0$, the formula holds. Suppose it also holds for some $k_1 - 1$, then

$$\Phi(|x-x^i_q|) \leq \Phi(n \cdot v^{d,k_1-1}_\delta(t)) \leq \Phi_d(n \cdot v^{d,k_1-1}_\delta(t)) \quad \forall t \leq 1.$$  \hspace{1cm} (4.17)

Now, using (4.13) and (4.16), we have

$$|x^i - x^i_q| \leq v^{d,k_1}_\delta(t) \quad \forall t \leq 1, \quad i = 1, \ldots, n.$$  \hspace{1cm} (4.18)

Taking the limit as first $k_1 \to \infty$, then $d \to \infty$, $\delta \to 0$, and finally $\varepsilon \to 0$ (as $q \to \infty$), we obtain

$$|x^i - x^i_q| = 0 \quad \forall t \leq 1.$$  \hspace{1cm} (4.19)

\[\square\]

5. Metamorphosis

In the previous section, we showed that the solution (3.10) converges to exact solution to (2.1); therefore, we can use the problem (3.10) instead of the problem (2.1).

By using the definition of $h_q(x(t), z(t), t)$, which is a continuous piecewise linear function, we can substitute $h_q(x(t), z(t), t)$ by $A_q x(t) + \sum_{j=1}^m B_{qj} z_j(t) + C_q$ for all $t \in [0, T]$, where $A_q, B_{qj} \in \mathbb{R}^{n \times n}$, and $C_q \in \mathbb{R}^n$.

Hence, we have

$$dx(t) = \left[ A_q x(t) + \sum_{j=1}^m B_{qj} z_j(t) + C_q \right] dt - \sum_{j=1}^m z_j(t)dw^j(t) \quad t \in [0, T],$$  \hspace{1cm} (5.1)

$$x(T) = \xi.$$  

On the other hand, it should be noted that (5.1) is the especial case of the following linear BSDE:

$$dx(t) = \left[ A(t) x(t) + \sum_{j=1}^m B_j(t) z_j(t) + f(t) \right] dt - \sum_{j=1}^m z_j(t)dw^j(t) \quad t \in [0, T],$$  \hspace{1cm} (5.2)

$$x(T) = \xi,$$

where $x(t), \xi, z_j(t), f(t) \in \mathbb{R}^n$, and $A(t), B_j(t) \in \mathbb{R}^{n \times n}$. Also $A(t), B_j(t)$ are bounded deterministic functions and $f \in L^2_F(0, T; \mathbb{R}^n)$.

Kohlmann and Zhou in [9] explored the relationship between linear BSDEs and stochastic control by interpreting linear BSDEs as some optimal problems.

Now, we consider the controlled system

$$dp(t) = \left[ A_q p(t) + \sum_{j=1}^m B_{qj} u_j(t) + C_q \right] dt + \sum_{j=1}^m u_j(t)dw^j(t) \quad t \in [0, T],$$  \hspace{1cm} (5.3)

$$p(0) = p_0.$$
where \( p(t), p_0, u_j(t) \in \mathbb{R}^n \), and \( A_q, B_q \) are two bounded deterministic continuous piece-wise linear functions. For a given \( F_T \)-measurable square integrable random variable \( \xi \) (i.e., \( E|\xi|^2 < \infty \)), the problem is to select an \( F_T \)-adapted control process \( u(\cdot) \equiv (u_1(\cdot), \ldots, u_m(\cdot)) \in L^2_F(0, T; \mathbb{R}^{n \times m}) \) so as to minimize the cost functional

\[
J_q(p_0, u(\cdot)) = E \left( \frac{1}{2} | p(T) - \xi |^2 \right). \tag{5.4}
\]

To simplify the cost functional, it is normal to define

\[
y(t) = p(t) - E(\xi | F_t). \tag{5.5}
\]

Since \( E(\xi | F_t) \) is an \( F_T \)-martingale and \( F_t \) is generated by the standard Brownian motion \( W(t) \), by the martingale representation theorem [13], there is \( v(\cdot) \equiv (v_1(\cdot), \ldots, v_m(\cdot)) \in L^2_F(0, T; \mathbb{R}^{n \times m}) \) so that

\[
E(\xi | F_t) = E(\xi) + \sum_{j=1}^{m} \int_{0}^{T} v_j(s) dw_j(t). \tag{5.6}
\]

By (5.3), (5.5), and (5.6), with the new state variable \( y(\cdot) \) the controlled system becomes

\[
dy(t) = \left[ A_q y(t) + \sum_{j=1}^{m} B_q u_j(t) + g_q \right] dt + \sum_{j=1}^{m} [u_j(t) - v_j(t)] dw_j(t) \quad t \in [0, T],
\]

\[
y(0) = x_0 - E\xi \equiv y, \tag{5.7}
\]

where \( g_q = C_q + A_q E(\xi | F_t) \), and the cost functional reduces to \( J_q(y, u(\cdot)) = E(1/2) | y(T) |^2 \). Notice that the above problem is a stochastic LQ control problem.

Now, by considering the defined notations, we derive that

\[
\text{Min} J_q(y, u(\cdot)) = E \left( \frac{1}{2} | y(T) |^2 \right)
\]

\[
s.t. \quad dy(t) = \left[ A_q y(t) + \sum_{j=1}^{m} B_q u_j(t) + g_q \right] dt + \sum_{j=1}^{m} [u_j(t) - v_j(t)] dw_j(t), \tag{5.8}
\]

\[
y(0) = y.
\]

Here, we introduce the following stochastic Riccati equation (SRE):

\[
\begin{align*}
\dot{P}_q + P_q A_q + A_q' P_q - \sum_{j=1}^{m} P_q B_{q_j} P_q^{-1} B_{q_j} P_q & = 0, \\
P_q(T) & = I, \\
P_q(t) > 0 & \quad \forall t \in [0, T],
\end{align*} \tag{5.9}
\]
along with a BSDE

\[
d\varphi(t) = -\left[ \left( A'_q - \sum_{j=1}^{m} P_q B_{qj} P_q^{-1} B'_{qj} \right) \varphi - \sum_{j=1}^{m} P_q B_{qj} P_q^{-1} \beta_j \\
+ P_q \left( g_q + \sum_{j=1}^{m} B_{qj} v_j \right) \right] dt + \sum_{j=1}^{m} \beta_j(t) dw_j(t),
\]

\( \varphi(T) = 0. \) (5.10)

Note that the SRE (5.9) is different from the traditional Riccati equation in that (5.9) involves the inverse of the unknown. Moreover, the third constraint of (5.9) must also be satisfied by any solution. In general, an equation like (5.9) does not automatically admit a solution. (The solvability of this equation is considered by Kohlmann and Zhou in [9, Section 4].) Besides, if (5.9) has a solution \( P_q(\cdot) \in C([0, T]; \mathbb{R}^{n \times m}) \), then (5.10) must admit an \( F_t \)-adapted solution \((\varphi(\cdot), \beta_j(\cdot), j = 1, \ldots, m)\) as (5.10) is a piecewise linear BSDE with bounded coefficients and square integrable nonhomogeneous terms. (For more details, see [15].)

The two following theorems are the especial cases of [9, Theorems 3.1 and 3.2], which were proved by Kohlmann and Zhou.

**Theorem 5.1.** If (5.9) and (5.10) admit solutions \( P_q \in C([0, T]; \mathbb{R}^{n \times n}) \) and \((\varphi(\cdot), \beta_j(\cdot), j = 1, \ldots, m)\) as (5.10) is a piecewise linear BSDE with bounded coefficients and square integrable nonhomogeneous terms, then the problem (5.8) has an optimal feedback control 

\[ u^*_j(t) = -P_q^{-1}(t) B'_{qj} \left[ P_q(t) y^*(t) + \varphi(t) \right] - P_q^{-1}(t) \beta_j(t) + v'_j(t), \quad j = 1, \ldots, m. \] (5.11)

Moreover, the optimal cost value under the above control is

\[
J^*_q(y) = \frac{1}{2} y' P(0) y + y' \varphi(0) + \frac{1}{2} E \int_0^T \left[ 2 \varphi' g_q - 2 \sum_{j=1}^{m} \beta_j v_j + \sum_{j=1}^{m} v'_j P v_j \\
- \sum_{j=1}^{m} \left( P_q^{-1} B'_{qj} \varphi + P_q^{-1} \beta_j - v_j \right)' \left( P_q^{-1} B'_{qj} \varphi + P_q^{-1} \beta_j - v_j \right) \right](t) dt.
\] (5.12)

For the proof of Theorem 5.1, see [9, proof of Theorem 3.1].

**Theorem 5.2.** Under the same assumptions of Theorem 5.1, the control problem (5.3)-(5.4) has an optimal feedback control 

\[ u^*_j(t) = -P_q^{-1}(t) B'_{qj} P_q(t) [p^*(t) - x_q(t)] - z_{qj}(t), \quad j = 1, \ldots, m, \] (5.13)
where \((x_q(\cdot), z_{qj}(\cdot), j = 1, \ldots, m) \in L^2_F(0, T; \mathbb{R}^n) \times L^2_F(0, T; \mathbb{R}^{n \times m})\) is the unique \(F_t\)-adapted solution of the following BSDE:

\[
 dx(t) = \left[ A_q x(t) + \sum_{j=1}^m B_{qj} z_j(t) + C_q \right] dt - \sum_{j=1}^m z_j(t) dw^j(t) \quad t \in [0, T],
\]

\[ x(T) = \xi. \tag{5.14} \]

Notice that (5.14) yields

\[
 x_q(t) = E(\xi | F_t) - P_q(t)^{-1} \varphi(t),
\]

\[
 z_{qj}(t) = v_j(t) - P_q(t)^{-1} \beta_j(t) \quad j = 1, \ldots, m. \tag{5.15} \]

For the proof of Theorem 5.2, see [9, proof of Theorem 3.2].

Remark 5.3. Equation (5.12) also gives the optimal cost functional value as a function of the initial value \(y \equiv p_0 - E\xi\), which turns out to be quadratic. If the controller has the choice of selecting the initial value \(y\) so as to minimize \(J^*_q(y)\), then the “best” initial value would be obtained by setting \((d/dy)J^*_q(y)|_{y=y^*} = 0\). This yields \(y^* = -P_q(0)^{-1} \varphi(0)\). Returning to the original variable \(p_0\), we get that the best initial value for \(p(\cdot)\) will be

\[
 p^*_0 = y^* + E\xi = -P_q(0)^{-1} \varphi(0) + E\xi = x(0), \tag{5.16} \]

where the last equality is due to (5.15). This certainly makes perfect sense, as it implies that one should choose the initial value \(x_q(0)\) so as to minimize the difference between the terminal state value and the given value \(\xi\). (Of course, in this case the minimum difference is zero since starting with \(x_q(0)\) one can hit \(\xi\) exactly at the end, by the BSDE theory.) Therefore, the solution pair \((x_q(\cdot), z_{q}(\cdot))\) of the BSDE (5.14) may be regarded as the optimal state-control pair of minimizing \(J_q(p_0, u(\cdot))\) (as given by (5.4)) over \((p_0, u(\cdot))\) subject to the dynamics of (5.3). This gives an interpretation of \((x_q(\cdot), z_{q}(\cdot))\) via a stochastic control problem. In this perspective, if a BSDE does not have an adapted solution (e.g., when the underlying filtration is not generated by the Brownian motion involvement), we may still define a “pseudosolution” via the corresponding stochastic control problem.

Remark 5.4. Under the optimal feedback (5.13), the optimal trajectory \(p^*(\cdot)\) evolves as

\[
 dp^*(t) = \left\{ A_q - \sum_{j=1}^m B_{qj} P^{-1}_q B'_{qj} P_q \right\} p^*(t) dt + \sum_{j=1}^m \left[ -P^{-1}_q B'_{qj} P_{q} p^*_q + P^{-1}_q B'_{qj} P_q x_q + z_q \right](t) dw^j(t), \tag{5.17} \]

\[ p^*(0) = p_0. \]
Moreover, the difference $\Delta(t) = p^*(t) - x_q(t)$ satisfies

$$d\Delta(t) = \left[A_q - \sum_{j=1}^{m} B_{qj} P_q^{-1} B_{qj}' P_q \right](t) dt - \sum_{j=1}^{m} \left[-P_{qj}^{-1} B_{qj}' P_q \right](t) \Delta(t) dw^j(t),$$

$$\Delta(0) = p_0 - x_q(0).$$

(5.18)

Notice that $\Delta(\cdot)$ satisfies a homogeneous linear SDE, and hence must be identically zero if the initial value is zero, namely, if $p_0 = x_q(0)$. In this case, by (5.13), the optimal control is $u_j^*(t) = z_{qj}(t)$. This is exactly in line with the observation in Remark 5.3. (See [9].)

### 6. Numerical results

In this section, we use the results of previous sections to solve a nonlinear BSDE as (1.1) in two different cases.

**Example 6.1.** In this example, we set $f(x,z,t) = \sqrt{x(t)} + z(t)$, $\xi = 3$ and $T = 1$, then we have the following nonlinear BSDE:

$$dx(t) = \left[\sqrt{x(t)} + z(t)\right] dt + z(t) dW(t),$$

$$x(1) = 3.$$  

(6.1)

First, we approximate the nonlinear term (i.e., $f(x,z,t) = \sqrt{x(t)} + z(t)$) by a continuous piecewise linear function as $A_q x(t) + z(t)$. We define a stochastic optimal control problem.
corresponding with the original problem as below:

$$J_q(p_0, u(\cdot)) = E \frac{1}{2} | p(1) - 3 |^2$$

s.t. \[ dp(t) = [A_q(t) + u(t)] dt + u(t) dW \quad t \in [0, 1], \]

\[ p(0) = p_0. \] (6.2)

In this example, we set \( p_0 = 0 \). By solving the stochastic optimal control problem, we obtain a sample path with \( x_q^+ (1) = 2.922 \). In Figure 6.1, we plot the obtained result.

**Example 6.2.** In this example, again we set \( f(x, z, t) = \sqrt{x(t)} + z(t) \), \( \xi = 3 \) and \( T = 1 \), then we have the same nonlinear BSDE in Example 6.1. But in this example we consider \( p_0 \) as a decision variable. By solving the stochastic optimal control problem, we obtain a sample path with \( x_q^+ (0) = -0.058 \) and \( x_q^+ (1) = 3.060 \). Numerical result is in Figure 6.2.

**Acknowledgments**

The authors wish to thank S. Hamadène, N. Ghal-eh and N. Tavallaee for inspiring discussions. Also, the authors would like to thank the referee for his/her careful reading of the manuscript and his/her constructive comments that led to an improved version of the paper.

**References**


Omid. S. Fard: Department of Mathematics, Damghan University of Basic Sciences, Damghan, Iran

E-mail address: osfard@dubs.ac.ir

Ali. V. Kamyad: Department of Mathematics, Ferdowsi University of Mashhad, Mashhad, Iran

E-mail address: kamyad@math.um.ac.ir