Available online at www.sciencedirect.com
ScienceDirect
Nonlinear Analysis 67 (2007) 1260–1274
Nonlinear Analysis
www.elsevier.com/locate/na

# On backward stochastic evolution equations in Hilbert spaces and optimal control

N.I. Mahmudov[a,*], M.A. McKibben[b]

[a] Department of Mathematics, Eastern Mediterranean University, Gazimagusa, TRNC, Mersin 10, Turkey
[b] Goucher College, Mathematics and Computer Science Department Baltimore, MD 21204, USA

Received 23 June 2005; accepted 25 July 2006

**Abstract**

In this paper a new result on the existence and uniqueness of the adapted solution to a backward stochastic evolution equation in Hilbert spaces under a non-Lipschitz condition is established. The applicability of this result is then illustrated in a discussion of a concrete backward stochastic partial differential equation. Furthermore, a stochastic maximum principle for optimal control problems of stochastic systems governed by backward stochastic evolution equations in Hilbert spaces is obtained.


*Keywords:* Stochastic evolution systems; Semilinear systems; Maximum principle

## 1. Introduction

Backward stochastic differential equations (BSDEs for short) have important applications in stochastic control and financial markets. Since the publication of the work of Pardoux and Peng [13], many papers have been dedicated to the study of backward stochastic differential equations. Several of these papers (see [6–14,16–18]) have been devoted to the case of BSDE in infinite dimensional spaces. Hu and Peng [7,8] have considered two cases of semilinear backward stochastic evolution equations (BSEEs): in the first case the existence of a so-called "mild solution" was established, and in the second case semilinear backward stochastic partial differential equations (BSPDEs) were considered. Such equations appear, for example, in the theory of optimal control and controllability for stochastic partial differential equations (see [2]). Maximum principles for stochastic control systems in infinite dimensional spaces were studied by Bensoussan [1,2], Mahmudov [4], Hu and Peng [8]. Maximum principles for backward stochastic equations in finite dimensional space were studied by Dokuchaev and Zhou [5]. In the present paper, we first establish a result concerning the existence and uniqueness of a mild solution for a class of BSEEs with non-Lipschitzian coefficients in Hilbert space that generalizes some of the results in [7,18,8]. Secondly, we formulate a stochastic maximum principle for optimal control problems of stochastic systems governed by BSEEs in Hilbert spaces and solve a backward linear quadratic stochastic control.

* Corresponding author. Tel.: +90 392 630 1227; fax: +90 392 365 1604.
*E-mail addresses:* nazim.mahmudov@emu.edu.tr (N.I. Mahmudov), mmckibben@goucher.edu (M.A. McKibben).


doi:10.1016/j.na.2006.07.013

## 2. Preliminaries

In this section we introduce the notation needed to establish our results.

$(\Omega, \mathfrak{F}_T, \mathbf{P})$ is a probability space together with a normal filtration $\{\mathfrak{F}_t, 0 \le t \le T\}$, $X, U$ and $E$ are three separable Hilbert spaces. $W$ is a $Q$-Wiener process on $(\Omega, \mathfrak{F}_T, \mathbf{P})$ with a linear bounded covariance operator such that $tr\, Q < \infty$. We assume that there exists a complete orthonormal system $\{e_k\}$ in $E$, a bounded sequence of nonnegative real numbers $\lambda_k$ such that $Qe_k = \lambda_k e_k$, $k = 1, 2, \ldots$, and a sequence $\{\beta_k\}$ of independent Brownian motions such that

$$\langle w(t), e\rangle = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\, \langle e_k, e\rangle\, \beta_k(t), \quad e \in E,\ t \in [0, T],$$

where $\langle\cdot,\cdot\rangle$ is the inner product in $E$. Moreover we assume that $\mathfrak{F}_t$ is generated by $w(t)$. Let $L_2^0 = L_2\left(Q^{1/2}E, X\right)$ be the space of all Hilbert–Schmidt operators from $Q^{1/2}E$ to $X$ with the inner product $\langle \Psi, \Phi\rangle_{L_2^0} = tr\left[\Psi Q \Phi^*\right]$. $L^2(\Omega, \mathfrak{F}_T, X)$ is the Hilbert space of all $\mathfrak{F}_T$-measurable square integrable variables with values in a Hilbert space $X$. $L^2_{\mathfrak{F}}([0, T], X)$ is the Hilbert space of all square integrable and $\mathfrak{F}_t$-adapted processes with values in $X$. We recall that $f$ is said to be $\mathfrak{F}_t$-adapted if $f(t,\cdot) : \Omega \to X$ is $\mathfrak{F}_t$-measurable, a.e. $t \in [0, T]$.

For any $\beta \in \mathbb{R}$, define $M_\beta[t, T]$ to be the Banach space

$$M_\beta[t, T] = L^2_{\mathfrak{F}}(\Omega, C([t, T], X)) \times L^2_{\mathfrak{F}}\left([t, T], L_2^0\right)$$

equipped with the norm

$$\|(y, z)\|^2_{\beta,t} = \mathbf{E} \sup_{t\le s\le T} e^{2\beta s}\|y(s)\|^2 + \mathbf{E}\int_t^T e^{2\beta s}\|z(s)\|^2 ds.$$

Since $0 < T < \infty$, all the norms $\|\cdot\|_{\beta,t}$ with different $\beta \in \mathbb{R}$ are equivalent. $M[0, T] = M_0[0, T]$ is the Banach space endowed with the norm

$$\|(y, z)\|^2 = \mathbf{E}\sup_{0\le s\le T}\|y(s)\|^2 + \mathbf{E}\int_0^T \|z(s)\|^2 ds.$$

## 3. Backward stochastic evolution equations

In this section we study the existence and uniqueness of solutions to the following class of backward stochastic evolution equations in a Hilbert space $X$

$$\begin{cases} dy(t) = -\left[Ay(t) + F(t, y(t), z(t))\right]dt - \left[G(t, y(t)) + z(t)\right]dw(t), \\ y(T) = \xi, \end{cases} \tag{1}$$

where $A : D(A) \subset X \to X$ is a linear operator which generates a $C_0$-semigroup $\{S(t), 0 \le t \le T\}$ on $X$, $F : [0, T] \times X \times L_2^0 \to X$ and $G : [0, T] \times X \to L_2^0$ are given measurable mappings, see [3], and $\xi \in L^2(\Omega, \mathfrak{F}_T, X)$.

**Definition 1.** A pair of adapted processes $(y, z) \in L^2_{\mathfrak{F}}(\Omega, C([0, T], X)) \times L^2_{\mathfrak{F}}\left(\Omega \times [0, T], L_2^0\right)$ is a mild solution of (1) if for all $t \in [0, T]$ they satisfy

$$\begin{aligned} y(t) = {} & S(T-t)\xi + \int_t^T S(s-t)F(s, y(s), z(s))\, ds \\ & + \int_t^T S(s-t)\left[G(s, y(s)) + z(s)\right]dw(s), \quad P\text{-a.s.} \end{aligned}$$

*3.1. Lipschitz case*

In this subsection we study the existence and uniqueness of mild solutions to the equation

$$y(t) = S(T-t)\,\xi + \int_t^T S(s-t)\,F(s, y(s), z(s))\,\mathrm{d}s + \int_t^T S(s-t)\,[g(s) + z(s)]\,\mathrm{d}w(s), \quad P\text{-a.s.} \tag{2}$$

We make the following assumptions for the function $F : [0,T] \times X \times L_2^0 \to X$.

(L1) There exists an $L > 0$ such that

$$\|F(t,y,z) - F(t,\overline{y},\overline{z})\| \le L\left(\|y - \overline{y}\| + \|z - \overline{z}\|\right),$$

for all $t \in [0,T]$, $y, \overline{y} \in X$, $z, \overline{z} \in L_2^0$.

(L2) $F(\cdot, 0, 0) \in L^2([0,T], X)$.

**Lemma 2.** *For any $(f,g) \in L^2_{\mathfrak{F}}([0,T], X) \times L^2_{\mathfrak{F}}\left([0,T], L_2^0\right)$ the equation*

$$y(t) = S(T-t)\,\xi + \int_t^T S(s-t)\,f(s)\mathrm{d}s + \int_t^T S(s-t)\,[g(s) + z(s)]\,\mathrm{d}w(s), \quad P\text{-}a.s. \tag{3}$$

*has a unique solution in $M_\beta[0,T]$, and moreover*

$$\begin{aligned}
&\mathbf{E} \sup_{0\le s\le t} \mathrm{e}^{2\beta s}\|y(s)\|^2 + \mathbf{E}\int_t^T \mathrm{e}^{2\beta s}\|z(s)\|^2\mathrm{d}s \\
&\quad \le 24M_S^2\left(\mathrm{e}^{2\beta T}\mathbf{E}\|\xi\|^2 + \frac{1}{2\beta}\int_t^T \mathrm{e}^{2\beta r}\mathbf{E}\|f(r)\|^2\mathrm{d}r\right) + 2\mathbf{E}\int_t^T \mathrm{e}^{2\beta r}\|g(r)\|^2\mathrm{d}r,
\end{aligned} \tag{4}$$

*where $M_S = \sup\left\{\|S(t)\|_{\mathfrak{B}(X)}, 0 \le t \le T\right\}$ and $\mathfrak{B}(X)$ is the space of bounded, linear operators on $X$.*

**Proof.** Eq. (3) is a linear BSEE. As such, by Lemma 2.1 [7], it admits a unique mild solution $(y,z) \in M_\beta[0,T]$ given by

$$y(t) = S(T-t)\,\mathbf{E}\{\xi \mid F_t\} + \int_t^T S(s-t)\,\mathbf{E}\{f(s) \mid \mathfrak{F}_t\}\,\mathrm{d}s, \tag{5}$$

$$\widetilde{z}(t) = S(T-t)\,L(t) - \int_t^T S(s-t)\,K(t,s)\,\mathrm{d}s, \qquad z(t) = \widetilde{z}(t) - g(t), \tag{6}$$

where, by the martingale representation theorem (see [7], [4]) the processes $L \in L^2_{\mathfrak{F}}\left([0,T], L_2^0\right)$ and $K \in L^2_{\mathfrak{F}}\left([0,T]\times[0,T], L_2^0\right)$ satisfy the following relations:

$$\mathbf{E}\{\xi \mid \mathfrak{F}_t\} = \mathbf{E}\xi + \int_0^t L(\theta)\mathrm{d}w(\theta),$$

$$\mathbf{E}\{f(s) \mid \mathfrak{F}_t\} = \mathbf{E}f(s) + \int_0^t K(s,\theta)\,\mathrm{d}w(\theta).$$

Now, we estimate the solution $(y,z)$ given by (5) and (6) in $M_\beta[t,T]$ for $\beta > 0$. From (5) it follows that

$$\begin{aligned}
\mathbf{E}\sup_{t\le s\le T} \mathrm{e}^{2\beta s}\|y(s)\|^2 &\le 2M_S^2\mathbf{E}\sup_{t\le s\le T}\mathrm{e}^{2\beta s}\|\mathbf{E}\{\xi \mid \mathfrak{F}_s\}\|^2 \\
&\quad + 2M_S^2\mathbf{E}\sup_{t\le s\le T}\mathrm{e}^{2\beta s}\left(\int_s^T \mathbf{E}\{\|f(r)\| \mid \mathfrak{F}_s\}\,\mathrm{d}r\right)^2 = I_1 + I_2.
\end{aligned} \tag{7}$$

Standard computations yield

$$
\begin{aligned}
I_2 &\le 2M_S^2 \mathbf{E} \sup_{t\le s\le T} \left( \mathbf{E}\left\{ \mathrm{e}^{\beta s} \int_s^T \|f(r)\|\mathrm{d}r \mid \mathfrak{F}_s \right\}\right)^2 \\
&\le 2M_S^2 \mathbf{E} \sup_{t\le s\le T} \left( \mathbf{E}\left\{ \sup_{t\le\tau\le T} \mathrm{e}^{\beta \tau} \int_\tau^T \|f(r)\|\mathrm{d}r \mid \mathfrak{F}_s \right\}\right)^2 \\
&\le 8M_S^2 \mathbf{E} \left( \sup_{t\le\tau\le T} \mathrm{e}^{\beta \tau} \int_\tau^T \|f(r)\|\mathrm{d}r \right)^2 \\
&\le 8M_S^2 \mathbf{E} \sup_{t\le\tau\le T} \mathrm{e}^{2\beta \tau} \int_\tau^T \mathrm{e}^{-2\beta r}\mathrm{d}r \int_\tau^T \mathrm{e}^{2\beta r}\|f(r)\|^2\mathrm{d}r \\
&\le 8M_S^2 \mathbf{E} \sup_{t\le\tau\le T} \mathrm{e}^{2\beta \tau} \frac{1}{2\beta}\left[\mathrm{e}^{-2\beta\tau} - \mathrm{e}^{-2\beta T}\right] \int_\tau^T \mathrm{e}^{2\beta r}\|f(r)\|^2\mathrm{d}r \\
&\le 8M_S^2 \frac{1}{2\beta} \int_t^T \mathrm{e}^{2\beta r}\mathbf{E}\|f(r)\|^2\mathrm{d}r.
\end{aligned}
$$

Consequently, by (7)

$$
\mathbf{E} \sup_{t\le s\le T} \mathrm{e}^{2\beta s}\|y(s)\|^2 \le 8M_S^2 \mathrm{e}^{2\beta T}\mathbf{E}\|\xi\|^2 + 8M_S^2 \frac{1}{2\beta}\int_t^T \mathrm{e}^{2\beta r}\mathbf{E}\|f(r)\|^2\mathrm{d}r. \tag{8}
$$

Next, we estimate $z$. We have

$$
\|\widetilde{z}(s)\|^2 \le 2M_S^2\|L(s)\|^2 + 2M_S^2 \frac{\mathrm{e}^{-2\beta s}}{2\beta}\int_s^T \mathrm{e}^{2\beta\theta}\|K(\theta,s)\|^2\mathrm{d}\theta.
$$

From here it follows that

$$
\begin{aligned}
\mathbf{E}\int_t^T \mathrm{e}^{2\beta s}\|\widetilde{z}(s)\|^2\mathrm{d}s &\le 2M_S^2\mathbf{E}\int_t^T \mathrm{e}^{2\beta s}\|L(s)\|^2\mathrm{d}s + 2M_S^2\frac{1}{2\beta}\mathbf{E}\int_t^T\int_s^T \mathrm{e}^{2\beta\theta}\|K(\theta,s)\|^2\mathrm{d}\theta\mathrm{d}s \\
&\le 8M_S^2\mathrm{e}^{2\beta T}\mathbf{E}\|\xi\|^2 + 2M_S^2\frac{1}{2\beta}\mathbf{E}\int_t^T\int_0^\theta \mathrm{e}^{2\beta\theta}\|K(\theta,s)\|^2\mathrm{d}s\mathrm{d}\theta \\
&\le 8M_S^2\mathrm{e}^{2\beta T}\mathbf{E}\|\xi\|^2 + 4M_S^2\frac{1}{\beta}\int_t^T \mathrm{e}^{2\beta\theta}\mathbf{E}\|f(\theta)\|^2\mathrm{d}\theta.
\end{aligned} \tag{9}
$$

The inequalities (8) and (9), together, imply that

$$
\begin{aligned}
&\mathbf{E}\sup_{t\le s\le T}\mathrm{e}^{2\beta s}\|y(s)\|^2 + 2\mathbf{E}\int_t^T \mathrm{e}^{2\beta s}\|\widetilde{z}(s)\|^2\mathrm{d}s \\
&\quad\le 24M_S^2\left(\mathrm{e}^{2\beta T}\mathbf{E}\|\xi\|^2 + \frac{1}{2\beta}\int_t^T \mathrm{e}^{2\beta r}\mathbf{E}\|f(r)\|^2\mathrm{d}r\right),
\end{aligned}
$$

which in turn implies (4). ■

**Lemma 3.** *For any $(f,g) \in L^2_{\mathfrak{F}}([0,T],X) \times L^2_{\mathfrak{F}}\big([0,T],L_2^0\big)$, the associated mild solution of (3) satisfies the following estimate:*

$$
\begin{aligned}
&\mathbf{E}\sup_{0\le s\le t}\|y(s)\|^2 + \mathbf{E}\int_t^T \|z(s)\|^2\mathrm{d}s \\
&\quad\le 24M_S^2\left(\mathbf{E}\|\xi\|^2 + (T-t)\int_t^T \mathbf{E}\|f(r)\|^2\mathrm{d}r\right) + 2\mathbf{E}\int_t^T \|g(r)\|^2\mathrm{d}r.
\end{aligned} \tag{10}
$$

**Proof.** See [7]. ■

**Theorem 4.** *BSEE (2) admits a unique solution* $(y, z) \in M_\beta [0, T]$.

**Proof.** For any fixed $(\overline{y}, \overline{z}) \in M_\beta [0, T]$, it follows from (L2) that

$$f(\cdot) = F(\cdot, \overline{y}(\cdot), \overline{z}(\cdot)) \in L^2_{\mathfrak{F}}([0, T], X).$$

By Lemma 2, the equation

$$y(t) = S(T-t)\xi + \int_t^T S(s-t) F(s, \overline{y}(s), \overline{z}(s)) \, ds + \int_t^T S(s-t) [g(s) + z(s)] \, dw(s), \quad P\text{-a.s.} \tag{11}$$

has a unique solution in $M_\beta [0, T]$.

Thus, the operator $\Phi : M_\beta [0, T] \to M_\beta [0, T]$ defined by

$$\Phi(\overline{y}, \overline{z}) = (y, z),$$

where $(y, z)$ is the solution of (11), is well defined. Moreover, the inequality (4) implies that

$$\begin{aligned}
\| \Phi(\overline{y}, \overline{z}) - \Phi(\widetilde{y}, \widetilde{z}) \|_0^2 &\le 12 M_S^2 \frac{1}{\beta} \int_0^T e^{2\beta s} \mathbf{E} \| F(s, \overline{y}(s), \overline{z}(s)) - F(s, \widetilde{y}(s), \widetilde{z}(s)) \|^2 ds \\
&\le 12 M_S^2 L \frac{1}{\beta} \int_0^T e^{2\beta s} \mathbf{E} \left( \| \overline{y}(s) - \widetilde{y}(s) \|^2 + \| \overline{z}(s) - \widetilde{z}(s) \|^2 \right) ds \\
&= 12 M_S^2 L \frac{1}{\beta} T \| (\overline{y}, \overline{z}) - (\widetilde{y}, \widetilde{z}) \|_0^2.
\end{aligned}$$

We can choose $\beta > 0$ large enough to get the contractivity of the operator $\Phi$ on $M_\beta [0, T]$, which in turn implies the existence and uniqueness of the solution to (2). ■

### 3.2. Approximation

We now construct an approximate sequence using a Picard-type iteration. Let $y_0(t) = 0$, and let $\{y_n, z_n\}$ be a sequence in $L^2_{\mathfrak{F}}([0, T], X) \times L^2_{\mathfrak{F}}\left([0, T], L_2^0\right)$ defined recursively by

$$\begin{aligned}
y_n(t) = {} & S(T-t)\xi + \int_t^T S(s-t) F(s, y_{n-1}(s), z_n(s)) \, ds \\
& + \int_t^T S(s-t) \left[ G(s, y_{n-1}(s)) + z_n(s) \right] dw(s),
\end{aligned} \tag{12}$$

on $0 \le t \le T$. We remark that by Theorem 4, Eq. (12) has a unique solution $(y_n, z_n)$.

To state our main result, we impose the following assumptions on the functions $F$ and $G$.

(N1) $F(\cdot, 0, 0) \in L^2([0, T], X)$, $G(\cdot, 0) \in L^2\left([0, T], L_2^0\right)$.

(N2) There exists an $l > 0$ such that

$$\| F(t, y, z) - F(t, \overline{y}, \overline{z}) \|^2 \le \rho\left( \| y - \overline{y} \|^2 \right) + l \| z - \overline{z} \|^2,$$

$$\| G(t, y) - G(t, \overline{y}) \|^2 \le \rho\left( \| y - \overline{y} \|^2 \right)$$

for all $t \in [0, T]$, $y, \overline{y} \in X$, $z, \overline{z} \in L_2^0$. Here $\rho$ is a concave increasing function from $[0, \infty)$ to $[0, \infty)$ such that $\rho(0) = 0$, $\rho(u) > 0$ for $u > 0$ and

$$\int_{0^+} \frac{du}{\rho(u)} = \infty.$$

Since $\rho$ is concave and $\rho(0) = 0$, there is a pair of positive numbers $a$ and $b$ such that

$$\rho(u) \le a + bu \tag{13}$$

for all $u \ge 0$. Therefore, under assumptions (N1) and (N2), $F(\cdot, y(\cdot), z(\cdot)) \in L^2_{\mathfrak{F}}([0,T], X)$ and $G(\cdot, y(\cdot)) \in L^2_{\mathfrak{F}}\big([0,T], L^0_2\big)$ whenever $y(\cdot) \in L^2_{\mathfrak{F}}([0,T], X)$ and $z(\cdot) \in L^2_{\mathfrak{F}}\big([0,T], L^0_2\big)$.

Now we introduce some important constants used throughout the paper.

$$\begin{aligned} C_1 \;&=\; 24M_S^2\left(e^{2\beta T}\mathbf{E}\|\xi\|^2 + \frac{1}{2\beta}\int_0^T e^{2\beta s}\left(2\|F(s,0,0)\|^2 + 2a\right)\mathrm{d}s\right) \\ &\quad + 2\int_0^T e^{2\beta s}\left(2\|G(s,0)\|^2 + 2a\right)\mathrm{d}s, \quad C_1 \ge 4aT, \\ C_2 &= 24M_S^2\frac{1}{\beta}b + 4b, \\ C_3 &= \left(12M_S^2\frac{1}{\beta} + 2\right)e^{2\beta T}, \qquad C_4 = C_3\rho\left(4C_1\exp\left(C_2T\right)\right). \end{aligned} \tag{14}$$

**Lemma 5.** *Under hypotheses* (N1) *and* (N2)*, for all* $0 \le t \le T$ *and* $n \ge 1$*,*

$$\mathbf{E}\left(\sup_{t\le s\le T} e^{2\beta s}\|y_n(s)\|^2\right) \le C_1\exp\left(C_2(T-t)\right), \tag{15}$$

$$\mathbf{E}\int_t^T e^{2\beta s}\|z_n(s)\|^2\mathrm{d}s \le 2C_1\left(1 + C_2(T-t)\exp\left(C_2(T-t)\right)\right), \tag{16}$$

*where $C_1$ and $C_2$ are both positive constants defined in (14).*

**Proof.** It follows from Lemma 2 that

$$\begin{aligned} &\mathbf{E}\sup_{t\le s\le T} e^{2\beta s}\|y_n(s)\|^2 + \mathbf{E}\int_t^T e^{2\beta s}\|z_n(s)\|^2\mathrm{d}s \\ &\quad \le 24M_S^2\left(e^{2\beta T}\mathbf{E}\|\xi\|^2 + \frac{1}{2\beta}\int_t^T e^{2\beta r}\mathbf{E}\|F(r, y_{n-1}(r), z_n(r))\|^2\mathrm{d}r\right) \\ &\qquad + 2\mathbf{E}\int_t^T e^{2\beta r}\|G(r, y_{n-1}(r))\|^2\mathrm{d}r. \end{aligned} \tag{17}$$

Using hypotheses (N1) and (N2) with (13) yields

$$\begin{aligned} &\|F(s, y_{n-1}(s), z_n(s))\|^2 \le 2\|F(s,0,0)\|^2 + 2a + 2b\|y_{n-1}(s)\|^2 + 2l\|z_n(s)\|^2, \\ &\|G(s, y_{n-1}(s))\|^2 \le 2\|G(s,0)\|^2 + 2a + 2b\|y_{n-1}(s)\|^2. \end{aligned}$$

Substituting these into (17) gives

$$\begin{aligned} &\mathbf{E}\sup_{t\le s\le T} e^{2\beta s}\|y_n(s)\|^2 + \mathbf{E}\int_t^T e^{2\beta s}\|z_n(s)\|^2\mathrm{d}s \\ &\quad \le 24M_S^2 e^{2\beta T}\mathbf{E}\|\xi\|^2 + 24M_S^2\frac{1}{2\beta}\mathbf{E}\int_t^T e^{2\beta s}\left(2\|F(s,0,0)\|^2 + 2a + 2b\|y_{n-1}(s)\|^2 + 2l\|z_n(s)\|^2\right)\mathrm{d}s \\ &\qquad + 2\mathbf{E}\int_t^T e^{2\beta s}\left(2\|G(s,0)\|^2 + 2a + 2b\|y_{n-1}(s)\|^2\right)\mathrm{d}s. \end{aligned}$$

Thus

$$\begin{aligned} &\mathbf{E}\sup_{t\le s\le T} e^{2\beta s}\|y_n(s)\|^2 + \left(1 - 24M_S^2\frac{1}{\beta}l\right)\mathbf{E}\int_t^T e^{2\beta s}\|z_n(s)\|^2\mathrm{d}s \\ &\quad \le C_1 + C_2\mathbf{E}\int_t^T \sup_{s\le r\le T}\left(e^{2\beta r}\|y_{n-1}(r)\|^2\right)\mathrm{d}s, \end{aligned} \tag{18}$$

where $C_1$ and $C_2$ are defined in (14). Choosing $\beta > 0$ such that $1 - 24M_S^2 \frac{1}{\beta} l = \frac{1}{2}$, we obtain

$$\sup_{1\le n\le m} \mathbf{E}\left(\sup_{t\le s\le T} \mathrm{e}^{2\beta s}\|y_n(s)\|^2\right) \le C_1 + C_2 \int_t^T \sup_{1\le n\le m} \mathbf{E} \sup_{s\le r\le T}\left(\mathrm{e}^{2\beta r}\|y_{n-1}(r)\|^2\right)\mathrm{d}s.$$

An application of the Gronwall inequality now implies

$$\sup_{1\le n\le m} \mathbf{E}\left(\sup_{t\le s\le T} \mathrm{e}^{2\beta s}\|y_n(s)\|^2\right) \le C_1 \exp\left(C_2\left(T-t\right)\right).$$

Since $m$ was arbitrary, the inequality (15) follows. Finally it follows from (18) that for $\beta = 48M_S^2 l$ we have

$$\begin{aligned}\mathbf{E}\int_t^T \mathrm{e}^{2\beta s}\|z_n(s)\|^2\mathrm{d}s &\le 2C_1 + 2C_2\int_t^T C_1\exp\left(C_2\left(T-s\right)\right)\mathrm{d}s \\ &\le 2C_1\left(1 + C_2\left(T-t\right)\exp\left(C_2\left(T-t\right)\right)\right). \quad \blacksquare\end{aligned}$$

**Lemma 6.** *Under hypotheses* (N1) *and* (N2)*, there exists a constant $C_3 > 0$ defined in* (14) *such that*

$$\mathbf{E}\sup_{t\le s\le T} \mathrm{e}^{2\beta s}\|y_{n+m}(s) - y_n(s)\|^2 \le C_3\int_t^T \rho\left(\mathbf{E}\sup_{s\le r\le T}\mathrm{e}^{2\beta r}\|y_{n+m-1}(r) - y_{n-1}(r)\|^2\right)\mathrm{d}s$$

*for all $0 \le t \le T$ and $n, m \ge 1$.*

**Proof.** Applying Lemma 2 we have

$$\begin{aligned}&\mathbf{E}\sup_{t\le s\le T}\mathrm{e}^{2\beta s}\|y_{n+m}(s) - y_n(s)\|^2 + \mathbf{E}\int_t^T \mathrm{e}^{2\beta s}\|z_{n+m}(s) - z_n(s)\|^2\mathrm{d}s \\ &\le 12M_S^2\frac{1}{\beta}\mathbf{E}\int_t^T \mathrm{e}^{2\beta s}\|F(s, y_{n+m-1}(s), z_{n+m}(s)) - F(s, y_{n-1}(s), z_n(s))\|^2\mathrm{d}s \\ &\quad + 2\mathbf{E}\int_t^T \mathrm{e}^{2\beta s}\|G(s, y_{n+m-1}(s)) - G(s, y_{n-1}(s))\|^2\mathrm{d}s \\ &\le \left(12M_S^2\frac{1}{\beta} + 2\right)\mathrm{e}^{2\beta T}\mathbf{E}\int_t^T \rho\left(\mathrm{e}^{2\beta s}\|y_{n+m-1}(s) - y_{n-1}(s)\|^2\right) \\ &\quad + 12M_S^2\frac{1}{\beta}\mathbf{E}\int_t^T \mathrm{e}^{2\beta s}\|z_{n+m}(s) - z_n(s)\|^2\mathrm{d}s.\end{aligned}$$

For sufficiently large $\beta > 0$ we have

$$\begin{aligned}&\mathbf{E}\sup_{t\le s\le T}\mathrm{e}^{2\beta s}\|y_{n+m}(s) - y_n(s)\|^2 + \left(1 - 12M_S^2\frac{1}{\beta}\right)\mathbf{E}\int_t^T \mathrm{e}^{2\beta s}\|z_{n+m}(s) - z_n(s)\|^2\mathrm{d}s \\ &\le C_3\int_t^T \rho\left(\mathbf{E}\sup_{s\le r\le T}\mathrm{e}^{2\beta r}\|y_{n+m-1}(r) - y_{n-1}(r)\|^2\right)\mathrm{d}s. \quad \blacksquare\end{aligned} \tag{19}$$

**Lemma 7.** *Under hypotheses* (N1) *and* (N2)*, there exists a constant $C_4 > 0$ defined in* (14) *such that*

$$\mathbf{E}\sup_{t\le s\le T}\mathrm{e}^{2\beta s}\|y_{n+m}(s) - y_n(s)\|^2 \le C_4\left(T-t\right)$$

*for all $0 \le t \le T$ and for all $n, m \ge 1$.*

**Proof.** By Lemmas 5 and 6 we have

$$\begin{aligned}\mathbf{E}\sup_{t\le s\le T}\mathrm{e}^{2\beta s}\|y_{n+m}(s)-y_n(s)\|^2 &\le C_3\int_t^T\rho\left(\mathbf{E}\sup_{s\le r\le T}\mathrm{e}^{2\beta r}\|y_{n+m-1}(r)-y_{n-1}(r)\|^2\right)\mathrm{d}s\\ &\le C_3\int_t^T\rho\left(4C_1\exp\left(C_2\left(T-s\right)\right)\right)\mathrm{d}s\\ &\le C_3\rho\left(4C_1\exp\left(C_2T\right)\right)\left(T-t\right)=C_4\left(T-t\right).\end{aligned}$$

The proof is complete. ■

Define

$$\begin{aligned}&\varphi_1(t)=C_4\left(T-t\right),\\ &\varphi_{n+1}(t)=C_3\int_t^T\rho\left(\varphi_n(s)\right)\mathrm{d}s,\quad n\ge 1,\\ &\widetilde{\varphi}_{n,m}(t)=\mathbf{E}\sup_{t\le s\le T}\mathrm{e}^{2\beta s}\|y_{n+m}(s)-y_n(s)\|^2,\quad n\ge 1,\ m\ge 1.\end{aligned}$$

**Lemma 8.** *There exists $0\le T_1<T$ such that for all $n,m\ge 1$*

$$0\le\widetilde{\varphi}_{n,m}(t)\le\varphi_n\left(t\right)\le\varphi_{n-1}(t)\le\cdots\le\varphi_1\left(t\right)\quad \textit{for all } t\in\left[T_1,T\right]. \tag{20}$$

**Proof.** We prove this lemma by induction in $n$.

By Lemma 7, we have

$$\widetilde{\varphi}_{1,m}(t)=\mathbf{E}\sup_{t\le s\le T}\mathrm{e}^{2\beta s}\|y_{1+m}(s)-y_1(s)\|^2\le C_4\left(T-t\right)=\varphi_1(t).$$

By Lemma 6

$$\begin{aligned}\widetilde{\varphi}_{2,m}(t)&=\mathbf{E}\sup_{t\le s\le T}\mathrm{e}^{2\beta s}\|y_{2+m}(s)-y_2(s)\|^2\\ &\le C_3\int_t^T\rho\left(\mathbf{E}\sup_{s\le r\le T}\mathrm{e}^{2\beta r}\|y_{1+m}(r)-y_1(r)\|^2\right)\mathrm{d}s\\ &=C_3\int_t^T\rho\left(\widetilde{\varphi}_{1,m}(s)\right)\mathrm{d}s\le C_3\int_t^T\rho\left(\varphi_1(s)\right)\mathrm{d}s=\varphi_2(t).\end{aligned}$$

We must show that there exists $T_1>0$ such that for all $t\in[T_1,T]$ the inequality

$$\varphi_2(t)=C_3\int_t^T\rho\left(C_4\left(T-s\right)\right)\mathrm{d}s\le C_4\left(T-t\right)=\varphi_1\left(t\right) \tag{21}$$

holds.

To this end, note that this inequality holds provided that

$$C_3\rho\left(C_4\left(T-t\right)\right)\le C_4=C_3\rho\left(4C_1\exp\left(C_2T\right)\right)$$

or

$$C_3\rho\left(4C_1\exp\left(C_2T\right)\right)=C_4\left(T-t\right)\le 4C_1\exp\left(C_2T\right)=4A.$$

On the other hand, this holds if

$$C_3\left\{A+4bA\right\}\left(T-t\right)\le 4A.$$

Since $A=C_1\exp\left(C_2T\right)\ge C_1\ge 4aT$ the above inequality holds if

$$T-t\le\frac{4}{C_3\left[\frac{1}{4T}+4b\right]}\le\frac{4}{C_3\left[\frac{a}{A}+4b\right]}.$$

Thus, (21) holds for any $t$ satisfying

$$T - t \leq \frac{4}{C_3\left[\frac{1}{4T} + 4b\right]}.$$

Clearly, such a $t$ does not depend on the final value $\xi$. Thus, there exists $T_1 > 0$ such that

$$\varphi_2(t) \leq \varphi_1(t)$$

for all $t \in [T_1, T]$. Now, assume that (20) holds for some $n \geq 2$. Then, using the same inequalities as above yields

$$\begin{aligned}\widetilde{\varphi}_{n+1,m}(t) &\leq C_3 \int_t^T \rho\left(\mathbf{E}\sup_{s\leq r\leq T} e^{2\beta r}\|y_{n+m}(r) - y_n(r)\|^2\right) ds \\ &= C_3\int_t^T \rho\left(\widetilde{\varphi}_{n,m}(s)\right) ds \leq C_3 \int_t^T \rho\left(\varphi_n(s)\right) ds = \varphi_{n+1}(t)\end{aligned}$$

for all $t \in [T_1, T]$. On the other hand, we have

$$\varphi_{n+1}(t) = C_3 \int_t^T \rho\left(\varphi_n(s)\right) ds \leq C_3 \int_t^T \rho\left(\varphi_{n-1}(s)\right) ds = \varphi_n(t)$$

for all $t \in [T_1, T]$. This completes the proof. ■

**Theorem 9.** *Assume that* (N1) *and* (N2) *hold. Then, there exists a unique mild solution* $(y, z)$ *to Eq.* (1).

**Proof.** *Uniqueness:* To show the uniqueness, let both $(y, z)$ and $(\widetilde{y}, \widetilde{z})$ be solutions of the Eq. (1). Then, Lemma 2 implies

$$\begin{aligned}&\mathbf{E}\sup_{t\leq s\leq T} e^{2\beta s}\|y(s) - \widetilde{y}(s)\|^2 + \mathbf{E}\int_t^T e^{2\beta s}\|z(s) - \widetilde{z}(s)\|^2 ds \\ &\leq 12M_S^2\frac{1}{\beta}\int_t^T e^{2\beta s}\|F(s, y(s), z(s)) - F(s, \widetilde{y}(s), \widetilde{z}(s))\|^2 ds + 2\mathbf{E}\int_t^T e^{2\beta s}\|G(s, y(s)) - G(s, \widetilde{y}(s))\|^2 ds \\ &\leq C\int_t^T \rho\left(\mathbf{E}\sup_{s\leq r\leq T} e^{2\beta r}\|y(r) - \widetilde{y}(r)\|^2\right) ds + 12M_S^2\frac{1}{\beta}\mathbf{E}\int_t^T e^{2\beta s}\|z(s) - \widetilde{z}(s)\|^2 ds.\end{aligned}$$

Therefore, one can apply the Bihari inequality to (4) to obtain

$$\mathbf{E}\sup_{t\leq s\leq T} e^{2\beta s}\|y(s) - \widetilde{y}(s)\|^2 = 0.$$

So, $y(t) = \widetilde{y}(t)$ for all $0 \leq t \leq T$ almost surely. It then follows from (4) that $z(t) = \widetilde{z}(t)$ for all $0 \leq t \leq T$ almost surely as well. This establishes the uniqueness.

*Existence:* We claim that

$$\mathbf{E}\sup_{t\leq s\leq T} e^{2\beta s}\|y_{n+m}(s) - y_n(s)\|^2 \to 0, \quad \text{for all } T_1 \leq t \leq T, \tag{22}$$

as $n, m \to \infty$. Note that, by definition, $\varphi_n$ is continuous on $[T_1, T]$. Note also that for each $n \geq 1$, $\varphi_n(\cdot)$ is decreasing on $[T_1, T]$, and for each $t$, $\varphi_n(t)$ is a nonincreasing sequence. Therefore, we can define the function $\varphi(t)$ by $\varphi_n(t) \downarrow \varphi(t)$. It is easy to verify that $\varphi(t)$ is continuous and nonincreasing on $[T_1, T]$. By the definitions of $\varphi_n(t)$ and $\varphi(t)$ we get

$$\varphi(t) = \lim_{n\to\infty} C_3 \int_t^T \rho\left(\varphi_n(s)\right) ds = C_3 \int_t^T \rho\left(\varphi(s)\right) ds$$

for each $t \in [T_1, T]$. Since

$$\int_{0^+} \frac{du}{\rho(u)} = \infty$$

the Bihari inequality implies $\varphi(t) = 0$ for all $t \in [T_1, T]$. Consequently, $\lim_{n\to\infty} \varphi_n(T_1) = 0$. By Lemma 8

$$\mathbf{E} \sup_{t\le s\le T} e^{2\beta s} \|y_{n+m}(s) - y_n(s)\|^2 \le \sup_{T_1\le t\le T} \widetilde{\varphi}_{n,m}(t) \le \sup_{T_1\le t\le T} \varphi_n(t) = \varphi_n(T_1) \overset{n\to\infty}{\to} 0.$$

So, (22) must hold. Applying (22) to (19) we see that $\{y_n, z_n\}$ is a Cauchy (hence convergent) sequence in $M_\beta[T_1, T]$; denote the limit by $(y, z)$. Now letting $n \to \infty$ in (12) we obtain

$$y(t) = S(T-t)\xi + \int_t^T S(s-t) F(s, y(s), z(s))\,ds + \int_t^T S(s-t)[G(s, y(s)) + z(s)]\,dw(s),$$

on $[T_1, T]$. Since the value of $T_1$ depends only on the function $\rho$, one can deduce by iteration the existence on $[T - k(T - T_1), T]$ for each $k$, and therefore the existence on the entire interval $[0, T]$.

The theorem has been proved. ■

As an illustration of the applicability of this general existence and uniqueness result, we consider examples of concrete backward stochastic partial differential equations.

**Example A.** Let $\mathcal{D}$ be a bounded domain in $\mathbb{R}^N$ with smooth boundary $\partial\mathcal{D}$. Consider the following initial boundary value problem:

$$\begin{aligned} \partial y(t,z) &= (\Delta_z z(t,x) + F(t,x,y(t,x),z(t,x)))\,\partial t \\ &\quad + [G(t,x,y(t,x)) + z(t,x)]\,d\beta(t), \quad \text{a.e. on } (0,T)\times\mathcal{D} \\ y(t,x) &= 0, \quad \text{a.e. on } (0,T)\times\partial\mathcal{D}, \\ y(T,x) &= \xi(T,x), \quad \text{a.e. on } \mathcal{D}, \end{aligned} \tag{23}$$

where $y : [0,T]\times\mathcal{D} \to \mathbb{R}$, $z : [0,T]\times\mathcal{D} \to L_2^0(\mathbb{R}^N; L^2(\mathcal{D}))$ $F : [0,T]\times\mathcal{D}\times\mathbb{R}\times L_2^0(\mathbb{R}^N; L^2(\mathcal{D})) \to \mathbb{R}$, $G : [0,T]\times\mathcal{D}\times\mathbb{R} \to L_2^0(\mathbb{R}^N, L^2(\mathcal{D}))$, $\beta$ is a standard $N$-dimensional Brownian motion (equipped with a normal filtration $\{\mathfrak{F}_t\}$, and $\xi : [0,T]\times\mathcal{D} \to \mathbb{R}$ is an $\mathfrak{F}_T$-measurable random variable.

We impose the following conditions:

(E1) $F$ satisfies the Caratheodory conditions (i.e., measurable in $(t, x, y)$ and continuous in the fourth variable) and there exists $M_F > 0$ such that

$$|F(t,x,w_1,y_1) - F_1(t,x,w_2,y_2)| \le M_F[|w_1 - w_2| + \|y_1 - y_2\|_{L_2^0(\mathbb{R}^N, L^2(\mathcal{D}))}],$$

for all $0 \le t \le T$, $x \in \mathcal{D}$, $w_1, w_2 \in \mathbb{R}$, $y_1, y_2 \in L_2^0(\mathbb{R}^N, L^2(\mathcal{D}))$.

(E2) $G$ satisfies the Caratheodory conditions and there exists $M_G > 0$ such that

$$\|G(t,x,w_1) - G(t,x,w_2)\|_{L_2^0(\mathbb{R}^N, L^2(\mathcal{D}))} \le M_G |w_1 - w_2|,$$

for all $0 \le t \le T$, $z \in \mathcal{D}$, $w_1, w_2 \in \mathbb{R}$.

We have the following theorem:

**Theorem 10.** *If* (E1) *and* (E2) *are satisfied, then* (23) *has a unique mild solution* $(y, z) \in L^2\left(0, T; L^2(\Omega, L^2(\mathcal{D}))\right) \times L_F^2\left(0, T; L^2(\mathbb{R}^N, L^2(\Omega, L^2(\mathcal{D})))\right)$.

**Proof.** Let $X = L^2(\mathcal{D})$ and $K = \mathbb{R}^N$. Also, denote $\frac{\partial y}{\partial t}$ by $y'(t)$, and define the operator $A$ by

$$Ay(t,\cdot) = \Delta_x y(t,\cdot), \quad y \in H^2(\mathcal{D}) \cap H_0^1(\mathcal{D}). \tag{24}$$

It is known that $A$ generates a strongly continuous semigroup $\{S(t)\}$ on $L^2(\mathcal{D})$ (see [15]). Define the maps $f : [0,T]\times X\times L_2^0(K; X) \to X$ and $g : [0,T]\times X \to L_2^0(K, X)$ by

$$f(t, y(t), z(t))(x) = F(t, x, y(t,x), z(t,x)) \tag{25}$$

$$g(t, y(t))(x) = G(t, x, y(t,x)), \tag{26}$$

for all $0 \leq t \leq T, x \in \mathcal{D}$. With these identifications, we observe that (23) can be written in the abstract form (1). Clearly, $f$ and $g$ as defined in (25) and (26) satisfy (N1) and (N2), respectively. Hence, we can invoke Theorem 9 to conclude that (23) has a unique mild solution $(y, z) \in L^2\left(0, T; L^2(\Omega, L^2(\mathcal{D}))\right) \times L^2_F\left(0, T; L^2(\mathbb{R}^N, L^2(\Omega, L^2(\mathcal{D})))\right)$. ■

## 4. Stochastic maximum principle

In this section we consider the following stochastic controlled system:

$$y(t) + \int_t^T S(s-t) f(s, y(s), z(s), u(s))\,\mathrm{d}s + \int_t^T S(s-t) z(s)\mathrm{d}w(s) = S(T-t)\xi \tag{27}$$

with the cost functional

$$J(u) = \mathbf{E}h(y(0)) + \mathbf{E}\int_0^T l(t, y(t), z(t), u(t))\,\mathrm{d}t. \tag{28}$$

Here, $f : [0,T] \times X \times L_2^0 \times U \to X, h : X \to X, l : [0,T] \times X \times L_2^0 \times U \to \mathbb{R}$ are measurable functions, $\xi \in L^2(\Omega, \mathfrak{F}_T, X)$ and $u : [0,T] \times \Omega \to U$.

We impose the following assumptions.

(A1) $f, l, h$ are continuously differentiable with respect to $(y, z)$.

(A2) The derivatives of $f$ with respect to $y, z$ are uniformly bounded:

$$\|f_y\| + \|f_z\| \leq C$$

and

$$\|h_y\| + \|l_y\| \leq C(1 + \|y\|), \qquad \|l_z\| \leq C(1 + \|z\|),$$

where $C$ is a positive constant.

Now we define

$$U_{ad} = \{u \in L^2_{\mathfrak{F}}([0,T], U) : u(t, \omega) \in U\}.$$

It is clear that under (A1)–(A2) for any $u \in U_{ad}$ the state Eq. (27) admits a unique solution $(y, z) = (y(\cdot, u(\cdot)), z(\cdot, u(\cdot)))$ and the cost functional (28) is well defined. We call $(y, z, u)$ an admissible triple, and $(y, z)$ an admissible state process. An optimal control problem can be stated as follows.

**Problem A.** Find a control $u^0(\cdot) \in U_{ad}$ such that

$$J\left(u^0\right) = \inf_{u \in U_{ad}} J(u). \tag{29}$$

Any control $u^0$ satisfying the equality (29) is called an optimal control. The corresponding $\left(y^0, z^0\right) = \left(y\left(\cdot, u^0(\cdot)\right), z\left(\cdot, u^0(\cdot)\right)\right)$ and $\left(y^0, z^0, u^0\right)$ are called an optimal state process and optimal triple, respectively.

Assume that

$$\left(y^0(\cdot), z^0(\cdot), u^0(\cdot)\right)$$

is an optimal solution of the control problem (28) and (27). Consider the following forward stochastic equation:

$$\psi(t) = S(t)h_y\left(y^0(0)\right) + \int_0^t S(s-t)\{f_y^*[s]\,\psi(s) + l_y[s]\}\mathrm{d}s + \int_0^t S(s-t)\left\{f_z^*[s]\,\psi(s) + l_z[s]\right\}\mathrm{d}w(s). \tag{30}$$

In what follows we shall use the following notation.

$$F[t] = F\left(t, y^0(t), z^0(t), u^0(t)\right),$$
$$\Delta_u F(t) = F\left(t, y^0(t), z^0(t), u(t)\right) - F[t],$$
$$\Delta_y F(t) = F\left(t, y(t), z^0(t), u^0(t)\right) - F[t],$$
$$\Delta_z F(t) = F\left(t, y^0(t), z(t), u^0(t)\right) - F[t].$$

Let $H$ be a Hamiltonian function given by

$$H(t, v) = \left\langle f\left(t, y^0(t), z^0(t), v\right), \psi(t)\right\rangle - l\left(t, y^0(t), z^0(t), v\right).$$

For any given

$$v \in U_{ad}, t_0 \in [0, T), 0 < \varepsilon \le T - t_0,$$

define a spike variational control by

$$u^\varepsilon(t) = \begin{cases} v, & t \in [t_0, t_0 + \varepsilon], \\ u^0(t), & \text{otherwise.} \end{cases}$$

Let $(y^\varepsilon(\cdot), z^\varepsilon(\cdot))$ be the solution of (27) corresponding to the admissible control $u^\varepsilon(\cdot)$, and let $(p^\varepsilon(\cdot), q^\varepsilon(\cdot))$ be the solution of the following linear BSDE:

$$p^\varepsilon(t) + \int_t^T S(s-t) f_y[s] p^\varepsilon(s) \mathrm{d}s + \int_t^T S(s-t) f_z[s] q^\varepsilon(s) \mathrm{d}s + \int_t^T S(s-t) q^\varepsilon(s) \mathrm{d}w(s) + \int_t^T S(s-t) \Delta_{u^\varepsilon} f[s] \mathrm{d}s = 0. \tag{31}$$

We have the following theorem.

**Theorem 11.** *Let* (A1) *and* (A2) *hold. Then*

$$\sup_{0\le t\le T} \mathbf{E}\|p^\varepsilon(t)\|^2 + \mathbf{E}\int_0^T \|q^\varepsilon(t)\|^2 \mathrm{d}t = O\left(\varepsilon^2\right), \tag{32}$$

$$\sup_{0\le t\le T} \mathbf{E}\|p^\varepsilon(t)\|^4 + \mathbf{E}\int_0^T \|q^\varepsilon(t)\|^4 \mathrm{d}t = O\left(\varepsilon^4\right), \tag{33}$$

$$\sup_{0\le t\le T} \mathbf{E}\|y^\varepsilon(t) - y^0(t) - p^\varepsilon(t)\|^2 + \mathbf{E}\int_0^T \|z^\varepsilon(t) - z^0(t) - q^\varepsilon(t)\|^2 \mathrm{d}t = o\left(\varepsilon^2\right). \tag{34}$$

*Moreover, the following formula holds:*

$$J(u^\varepsilon) - J\left(u^0\right) = \mathbf{E}\left\langle h_y\left(y^0(0)\right), p^\varepsilon(0)\right\rangle + \mathbf{E}\int_0^T \left\langle l_y[s], p^\varepsilon(s)\right\rangle \mathrm{d}s + \mathbf{E}\int_0^T \left\langle l_z[s], q^\varepsilon(s)\right\rangle \mathrm{d}s + \mathbf{E}\int_0^T \Delta l_u[s] \mathrm{d}s + o(\varepsilon). \tag{35}$$

**Proof.** By the Taylor formula we have

$$\begin{aligned}
&J\left(u^{\varepsilon}\right)-J\left(u^{0}\right) \\
&\quad=\mathbf{E}\left[h\left(y^{\varepsilon}(0)\right)-h\left(y^{0}(0)\right)\right] \\
&\qquad+\mathbf{E}\int_{0}^{T}\left[l\left(s, y^{\varepsilon}(s), z^{\varepsilon}(s), u^{\varepsilon}(s)\right)-l\left(s, y^{0}(s), z^{0}(s), u^{0}(s)\right)\right]\mathrm{d}s \\
&\quad=\mathbf{E}\int_{0}^{1}\left\langle h_{y}\left(y^{0}(0)+\theta\left(y^{\varepsilon}(0)-y^{0}(0)\right)\right), \Delta y^{0}(0)\right\rangle \mathrm{d}\theta \\
&\qquad+\mathbf{E}\int_{0}^{T}\left\langle l_{y}\left(s, y^{0}(s)+\theta\left(y^{\varepsilon}(s)-y^{0}(s)\right), z^{\varepsilon}(s), u^{\varepsilon}(s)\right), \Delta y^{0}(s)\right\rangle \mathrm{d}s \\
&\qquad+\mathbf{E}\int_{0}^{T}\left\langle l_{z}\left(s, y^{0}(s), z^{0}(s)+\theta\left(z^{\varepsilon}(s)-z^{0}(s)\right), u^{\varepsilon}(s)\right), \Delta z^{0}(s)\right\rangle \mathrm{d}s,
\end{aligned}$$

where $\Delta y^{0}(s)=y^{\varepsilon}(s)-y^{0}(s)$ and $\Delta z^{0}(s)=z^{\varepsilon}(s)-z^{0}(s)$. Let $Y^{\varepsilon}(t)=y^{\varepsilon}(t)-y^{0}(t)-p^{\varepsilon}(t)$ and $Z^{\varepsilon}(t)=z^{\varepsilon}(t)-z^{0}(t)-q^{\varepsilon}(t)$. Then

$$\begin{aligned}
&J\left(u^{\varepsilon}\right)-J\left(u^{0}\right) \\
&\quad=\mathbf{E}\left\langle h_{y}\left(y^{0}(0)\right), p^{\varepsilon}(0)\right\rangle+\mathbf{E}\left\langle h_{y}\left(y^{0}(0)\right), Y^{\varepsilon}(0)\right\rangle \\
&\qquad+\mathbf{E}\int_{0}^{1}\left\langle h_{y}\left(y^{0}(0)+\theta\left(y^{\varepsilon}(0)-y^{0}(0)\right)\right)-h_{y}\left(y^{0}(0)\right), p^{\varepsilon}(0)+Y^{\varepsilon}(0)\right\rangle \mathrm{d}\theta \\
&\qquad+\mathbf{E}\int_{0}^{T}\left\langle l_{y}[s], p^{\varepsilon}(s)+Y^{\varepsilon}(s)\right\rangle \mathrm{d}s \\
&\qquad+\mathbf{E}\int_{0}^{T}\left\langle l_{z}[s], q^{\varepsilon}(s)+Z^{\varepsilon}(s)\right\rangle \mathrm{d}s+\mathbf{E}\int_{0}^{T}\Delta l_{u}[s]\,\mathrm{d}s \\
&\qquad+\mathbf{E}\int_{0}^{T}\left\langle l_{y}\left(s, y^{0}(s)+\theta\left(y^{\varepsilon}(s)-y^{0}(s)\right), z^{\varepsilon}(s), u^{\varepsilon}(s)\right)-l_{y}[s], p^{\varepsilon}(s)+Y^{\varepsilon}(s)\right\rangle \mathrm{d}s \\
&\qquad+\mathbf{E}\int_{0}^{T}\left\langle l_{z}\left(s, y^{0}(s), z^{0}(s)+\theta\left(z^{\varepsilon}(s)-z^{0}(s)\right), u^{\varepsilon}(s)\right)-l_{z}[s], q^{\varepsilon}(s)+Z^{\varepsilon}(s)\right\rangle \mathrm{d}s.
\end{aligned}$$

Then by Theorem 11, we can obtain (35). ■

**Theorem 12** (*Stochastic Maximum Principle*). *Assume that* (A1) *and* (A2) *hold, and let* $\left(y^{0}, z^{0}, u^{0}\right)$ *be an optimal triple of Problem* A. *Then, there is a process* $\psi$ *satisfying* (30) *such that*

$$H(t, v) \leq H\left(t, u^{0}(t)\right), \tag{36}$$

*for all* $v \in U$, *a.e.* $t \in[0, T]$, $P$*-a.s.*

**Proof.** By the formula (35) we have

$$\begin{aligned}
J\left(u^{\varepsilon}\right)-J\left(u^{0}\right)=&\;\mathbf{E}\left\langle h_{y}\left(y^{0}(0)\right), p^{\varepsilon}(0)\right\rangle+\mathbf{E}\int_{0}^{T}\left\langle l_{y}[s], p^{\varepsilon}(s)\right\rangle \mathrm{d}s \\
&+\mathbf{E}\int_{0}^{T}\left\langle l_{z}[s], q^{\varepsilon}(s)\right\rangle \mathrm{d}s+\mathbf{E}\int_{0}^{T}\Delta l_{u}[s]\,\mathrm{d}s+o(\varepsilon).
\end{aligned}$$

On the other hand

$$\begin{aligned}
&\mathbf{E}\left\langle h_{y}\left(y^{0}(0)\right), p^{\varepsilon}(0)\right\rangle+\mathbf{E}\int_{0}^{T}\left\langle l_{y}[s], p^{\varepsilon}(s)\right\rangle \mathrm{d}s+\mathbf{E}\int_{0}^{T}\left\langle l_{z}[s], q^{\varepsilon}(s)\right\rangle \mathrm{d}s \\
&\quad=\mathbf{E}\int_{0}^{T}\left\langle\Delta_{u^{\varepsilon}} f[s], p^{\varepsilon}(s)\right\rangle \mathrm{d}s.
\end{aligned}$$

Thus

$$0 \le J\left(u^{\varepsilon}\right) - J\left(u^{0}\right) = \mathbf{E}\int_0^T \left\langle \Delta_{u^{\varepsilon}} f\,[s]\,,\, p^{\varepsilon}(s)\right\rangle \mathrm{d}s + \mathbf{E}\int_0^T \Delta l_u\,[s]\,\mathrm{d}s + o\,(\varepsilon)$$

and from here we can easily obtain the variational inequality (36). ■

## 5. A backward linear quadratic problem

In this section we apply Theorem 12 to a linear quadratic problem as a particular case of Problem A.

Consider the following problem:

$$J\,(u) = \mathbf{E}\,\langle Gy\,(0)\,,\, y\,(0)\rangle + \mathbf{E}\int_0^T \langle \Gamma\,(t)\,u(t), u(t)\rangle\,\mathrm{d}t \to \min \tag{37}$$

subject to

$$\begin{cases} \mathrm{d}y(t) = [Ay(t) + Bu(t) + Cz(t)]\,\mathrm{d}t + z(t)\mathrm{d}w(t), \\ y\,(T) = \xi. \end{cases} \tag{38}$$

Here $B : U \to X$, $C : L_2^0 \to X$, $\Gamma : [0, T] \to L\,(U)$. We assume that $G = G^*$, $\Gamma(t) = \Gamma^*(t) \ge \gamma I$.

Let $u^0$ be an optimal control, and $\left(y^0, z^0\right)$ be the corresponding state process. The adjoint process $\psi$ is the solution of

$$\begin{aligned} &-\mathrm{d}\psi(t) = A^*\psi(t)\mathrm{d}t + C^*\psi(t)\mathrm{d}w(t), \\ &\psi\,(0) = Gz^0\,(0)\,. \end{aligned} \tag{39}$$

**Theorem 13.** *There exists a unique optimal control $u^0$ for the problem (37) and (38) in the class $U_{ad}$. Moreover, $u^0$ has the following representation:*

$$u^0(t) = \Gamma^{-1}(t)B^*\psi\,(t)\,. \tag{40}$$

**Proof.** It is clear that (37) is a positive quadratic functional of control because of the assumptions on $G$ and $\Gamma(t)$. Hence an optimal control exists. Furthermore, the BSDE has the form (39). By Theorem 12

$$\left\langle Bu^0(t), \psi(t)\right\rangle - \left\langle \Gamma(t)u^0(t), u^0\,(t)\right\rangle \ge \langle Bv, \psi(t)\rangle - \langle \Gamma(t)v, v\rangle$$

for all $v \in U$. This in turn implies (40). Hence the control (40) is the only control which satisfies the stochastic maximum principle. It then must be the optimal control. This completes the proof. ■

**Theorem 14.** *The optimal control $u^0$ for the problem (37) and (38) can be represented as*

$$u^0(t) = \Gamma^{-1}(t)B^*P\,(t)\,y(t), \tag{41}$$

*where $y(t)$ is the solution of*

$$\begin{aligned} &y'(t) = \left[A + B\Gamma^{-1}(t)B^*P(t)\right]y(t), \\ &y\,(0) = y^0\,(0)\,, \end{aligned} \tag{42}$$

*and $P(t)$ is the mild solution of*

$$\mathrm{d}P(t) = -\left(P(t)A + A^*P\,(t) + P(t)B\Gamma^{-1}B^*P(t)\right)\mathrm{d}t + C^*P(t)\mathrm{d}w(t).$$

**Proof.** Let $\widetilde{\psi}(t) = P(t)y(t)$. We have

$$\begin{aligned}
\mathrm{d}\widetilde{\psi}(t) &= \mathrm{d}P(t)y(t) + P(t)\mathrm{d}y(t) \\
&= -\left(P(t)A + A^*P(t) + P\,(t)\,B\Gamma^{-1}(t)B^*P(t)\right)y(t)\mathrm{d}t \\
&\quad - C^*P(t)\mathrm{d}w(t)y(t) + P(t)\left(A + B\Gamma^{-1}(t)B^*\right)y(t)\mathrm{d}t \\
&= -\left(A^*P(t) - C^*P(t)\mathrm{d}w\,(t)\right)y(t) \\
&= -A^*\widetilde{\psi}(t)\mathrm{d}t - C^*\widetilde{\psi}(t)\mathrm{d}w(t)
\end{aligned}$$

So $\widetilde{\psi}(t)$ satisfies the same equation as $\psi(t)$. Hence $\widetilde{\psi}(t) = \psi(t)$ by uniqueness. The theorem is proved. ■

## References

[1] A. Bensoussan, Stochastic maximum principle for distributed parameter systems, J. Franklin Inst. 315 (1983) 387–406.
[2] A. Bensoussan, Lectures on stochastic control, in: S.K. Mitter, A. Moro (Eds.), Nonlinear Liltering and Stochastic Control, in: Lecture Notes in Mathematics, vol. 972, Springer, 1982.
[3] G. DaPrato, J. Zabczyk, Stochastic Equations in Infinite Dimensions, Cambridge University Press, Cambridge, 1992.
[4] J.P. Dauer, N.I. Mahmudov, Controllability of stochastic semilinear functional differential equations in Hilbert spaces, J. Math. Anal. Appl. 290 (2) (2004) 373–394.
[5] N. Dokuchaev, X.Y. Zhou, Stochastic controls with terminal contingent conditions, J. Math. Anal. Appl. 238 (1) (1999) 143–165.
[6] M. Hassani, Y. Ouknine, Infinite dimensional BSDE with jumps, Stochastic Anal. Appl. 20 (3) (2002) 519–565.
[7] Y. Hu, S. Peng, Adapted solutions of a backward semilinear stochastic evolution equation, Stochastic Anal. Appl. 9 (4) (1991) 445–459.
[8] Y. Hu, S. Peng, Maximum principle for semilinear stochastic evolution control systems, Stoch. Stoch. Rep. 33 (1990) 159–180.
[9] Y. Hu, J. Ma, J. Yong, On semi-linear degenerate backward stochastic partial differential equations, Probab. Theory Related Fields 123 (3) (2002) 381–411.
[10] J. Ma, J. Yong, On linear degenerate backward stochastic partial differential equations, Probab. Theory Related Fields 113 (2) (1999) 135–170.
[11] N.I. Mahmudov, The maximum principle for stochastic evolution systems in Hilbert spaces, Int. J. Pure Appl. Math. 2 (3) (2002) 287–298.
[12] X. Mao, Adapted solutions of backward stochastic differential equations with non-Lipschitz coefficients, Stochastic Process Appl. 58 (1995) 281–292.
[13] E. Pardoux, S. Peng, Adapted solution of a backward stochastic differential equation, Systems Control Lett. 14 (1990) 55–61.
[14] E. Pardoux, A. Rascanu, Backward stochastic variational inequalities, Stoch. Stoch. Rep. 67 (3–4) (1999) 159–167.
[15] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Springer-Verlag, New York, NY, 1983.
[16] S. Peng, Backward stochastic differential equations and applications to optimal control, Appl. Math. Optim. 27 (1993) 125–144.
[17] S. Rong, On solutions of backward stochastic differential equations with jumps and with non-Lipschitzian coefficients in Hilbert spaces and stochastic control, Statist. Probab. Lett. 60 (3) (2002) 279–288.
[18] G. Tessitore, Existence, uniqueness, and space regularity of the adapted solutions of a backward SPDE, Stochastic Anal. Appl. 14 (4) (1996) 461–486.