Fully Coupled Forward-Backward Stochastic Functional Differential Equations and Applications to Quadratic Optimal Control*

XU Xiaoming

DOI: 10.1007/s11424-020-9027-x
Received: 24 January 2019 / Revised: 26 May 2019

Abstract This paper considers the fully coupled forward-backward stochastic functional differential equations (FBSFDEs) with stochastic functional differential equations as the forward equations and the generalized anticipated backward stochastic differential equations as the backward equations. The authors will prove the existence and uniqueness theorem for FBSFDEs. As an application, we deal with a quadratic optimal control problem for functional stochastic systems, and get the explicit form of the optimal control by virtue of FBSFDEs.

Keywords Forward-backward stochastic functional differential equation, functional stochastic system, generalized anticipated backward stochastic differential equation, quadratic optimal control, stochastic functional differential equation.

1 Introduction

The general form of backward stochastic differential equation (BSDE) was considered the first time by Pardoux and Peng[1] in 1990. From then on, the theory of BSDEs has been studied with great interest (see, e.g., [2–4]). One hot topic is the forward-backward stochastic differential equation (FBSDE) (see, e.g., [5–7]), due to its wide applications in the pricing/hedging problem, in the stochastic control and game theory (see, e.g., [3, 8–11]).

Especially the fully coupled FBSDE of the form

$$\begin{cases} dX_t = b(t, X_t, Y_t, Z_t)dt + \sigma(t, X_t, Y_t, Z_t)dB_t; \\ -dY_t = f(t, X_t, Y_t, Z_t)dt - Z_t dB_t; \\ X_0 = a, \quad Y_T = \Phi(X_T), \end{cases}$$

XU Xiaoming
Institute of Finance and Statistics, School of Mathematical Sciences, Nanjing Normal University, Nanjing 210023, China. Email: xmxu@njnu.edu.cn.
*This research was supported by the Program of Natural Science Research of Jiangsu Higher Education Institutions of China under Grant No. 17KJB110009.
◇ *This paper was recommended for publication by Editor LIU Yungang.*

with the forward SDE as the state equation and the BSDE as the dual equation, plays an important role in the linear-quadratic (LQ) optimal control problem (see e.g. [3, 10, 12]), which is one of the most important class of optimal control and game problems. Next some efforts has been made to generalize this model. For example, motivated by the work of Mohammed[13] and the work of Peng and Yang[4] together, Chen and Wu[8] studied the following general FBSDE:

$$\begin{cases} dX_t = b(t, X_t, Y_t, Z_t, X_{t-\delta})dt + \sigma(t, X_t, Y_t, Z_t, X_{t-\delta})dB_t, & t \in [0, T]; \\ -dY_t = f(t, X_t, Y_t, Z_t, Y_{t+\delta}, Z_{t+\delta})dt - Z_t dB_t, & t \in [0, T]; \\ X_t = \rho_t, & t \in [-\delta, 0]; \\ Y_T = \Phi(X_T), \quad Y_t = \xi_t & t \in (T, T+\delta]; \\ Z_t = \eta_t, & t \in [T, T+\delta], \end{cases}$$

where $\delta \geq 0$. Easily we can find that here the value of the state (or the dual process) at time t depends not only on its value at t but also on the value at $t - \delta$ (or $t + \delta$), that is to say, the influence brought by the other time intervals is ignored.

However, many natural and social phenomena shows that the state process at time t depends not only on its present state but also its whole history of the past. Similarly, for the dual process, its value at time t depends not only on its present value but also its full information of the future. Thus in this paper we will study the following general case:

$$\begin{cases} dX_t = b(t, \{X_r\}_{r\in[-M,t]}, Y_t, Z_t)dt + \sigma(t, \{X_r\}_{r\in[-M,t]}, Y_t, Z_t)dB_t, & t \in [0, T]; \\ -dY_t = f(t, X_t, \{Y_r\}_{r\in[t,T+K]}, \{Z_r\}_{r\in[t,T+K]})dt - Z_t dB_t, & t \in [0, T]; \\ X_t = \rho_t, & t \in [-M, 0]; \\ Y_t = \xi_t, \quad Z_t = \eta_t, & t \in [T, T+K], \end{cases}$$

where $M \geq 0$ and $K \geq 0$. It is clearly that here the state process is given in the form of stochastic functional differential equations (see, e.g., Mohammed[13]), and the dual process in the form of generalized anticipated BSDEs, which is just the new type of BSDEs studied by Yang[14] (see also Yang and Elliott[15]).

We prove that under proper assumptions, the solution of the above equation exists uniquely (see Section 3). Then in Section 4, as an application, we deal with an optimal control problem for the following functional stochastic system:

$$\begin{cases} dX_t = \left(A_t \int_{-M}^{t} X_s ds + C_t v_t\right)dt + \left(D_t \int_{-M}^{t} X_s ds + F_t v_t\right)dB_t, & t \in [0, T]; \\ X_t = \rho_t, & t \in [-M, 0], \end{cases}$$

where $v.$ is a control process. Our aim is to minimize the classical quadratic optimal control cost function. For this problem, we can get the explicit unique optimal control by virtue of the results obtained in the previous section. Finally conclusions and future works are given in Section 5.

Next we first make some preliminaries.

2 Preliminaries

Let $\{B_t; t \geq 0\}$ be a d-dimensional standard Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ and $\{\mathcal{F}_t; t \geq 0\}$ be its natural filtration. Denote by $|\cdot|$ the norm in $\mathbb{R}^n$, and $\langle \cdot, \cdot \rangle$ denotes the inner product. Given $T > 0$, we will use the following notations:

- $C(-M, 0; \mathbb{R}^n) := \{\varphi. : [-M, 0] \to \mathbb{R}^n \mid \varphi. \text{ satisfies } \sup_{-M \leq t \leq 0} |\varphi_t| < +\infty\}$;
- $L^2(\mathcal{F}_T; \mathbb{R}^n) := \{\xi \in \mathbb{R}^n \mid \xi \text{ is an } \mathcal{F}_T\text{-measurable random variable such that } E|\xi|^2 < +\infty\}$;
- $L^2_{\mathcal{F}}(0, T; \mathbb{R}^n) := \{\varphi. : \Omega \times [0, T] \to \mathbb{R}^n \mid \varphi. \text{ is an } \mathcal{F}_t\text{-progressively measurable process such that } E\int_0^T |\varphi_t|^2 dt < +\infty\}$.

2.1 Generalized Anticipated Backward Stochastic Differential Equations

Consider the following generalized anticipated backward stochastic differential equation (GABSDE):

$$\begin{cases} -dY_t = f(t, \{Y_r\}_{r\in[t,T+K]}, \{Z_r\}_{r\in[t,T+K]})dt - Z_t dB_t, & t \in [0, T]; \\ Y_t = \xi_t, & t \in [T, T+K]; \\ Z_t = \eta_t, & t \in [T, T+K]. \end{cases} \tag{1}$$

For the generator $f(\omega, t, \{y_r\}_{r\in[t,T+K]}, \{z_r\}_{r\in[t,T+K]}) : \Omega \times [0, T] \times L^2_{\mathcal{F}}(t, T+K; \mathbb{R}^m) \times L^2_{\mathcal{F}}(t, T+K; \mathbb{R}^{m\times d}) \to L^2(\mathcal{F}_t; \mathbb{R}^m)$, we use several hypotheses (see Yang[14]):

(A1) There exists a constant $L > 0$ such that for each $t \in [0, T]$, $y., y'. \in L^2_{\mathcal{F}}(0, T+K; \mathbb{R}^m)$, $z., z'. \in L^2_{\mathcal{F}}(0, T+K; \mathbb{R}^{m\times d})$, the following holds:

$$\begin{aligned} &E\bigg[\int_t^T |f(s, \{y_r\}_{r\in[s,T+K]}, \{z_r\}_{r\in[s,T+K]}) - f(s, \{y'_r\}_{r\in[s,T+K]}, \{z'_r\}_{r\in[s,T+K]})|^2 ds\bigg] \\ \leq &LE\bigg[\int_t^{T+K} (|y_s - y'_s|^2 + |z_s - z'_s|^2) ds\bigg]; \end{aligned}$$

(A1′) There exists a constant $L' > 0$ such that for each $t \in [0, T]$, $y., y'. \in L^2_{\mathcal{F}}(0, T+K; \mathbb{R}^m)$, $z., z'. \in L^2_{\mathcal{F}}(0, T+K; \mathbb{R}^{m\times d})$, the following holds:

$$\begin{aligned} &E\bigg[\int_t^T \mathrm{e}^{\theta s}|f(s, \{y_r\}_{r\in[s,T+K]}, \{z_r\}_{r\in[s,T+K]}) - f(s, \{y'_r\}_{r\in[s,T+K]}, \{z'_r\}_{r\in[s,T+K]})|^2 ds\bigg] \\ \leq &L'E\bigg[\int_t^{T+K} \mathrm{e}^{\theta s}(|y_s - y'_s|^2 + |z_s - z'_s|^2) ds\bigg], \end{aligned}$$

where $\theta \geq 0$ is an arbitrary constant;

(A2) $E[\int_0^T |f(s, 0, 0)|^2 ds] < +\infty$.

Remark 2.1 In fact, (A1) $\Leftrightarrow$ (A1′), see Remark 2.2.1 of Yang[14].

By using the fixed point theorem, Yang[14] (see also Yang and Elliott[15]) proved the following existence and uniqueness theorem for GABSDEs:

Theorem 2.2 *Assume that f satisfies* (A1) *and* (A2)*, then for arbitrary given terminal conditions* $(\xi_\cdot, \eta_\cdot) \in L^2_{\mathcal{F}}(T, T+K; \mathbb{R}^m) \times L^2_{\mathcal{F}}(T, T+K; \mathbb{R}^{m\times d})$*, the GABSDE* (1) *has a unique solution, i.e., there exists a unique pair of $\mathcal{F}_t$-adapted processes* $(Y_\cdot, Z_\cdot) \in L^2_{\mathcal{F}}(0, T+K; \mathbb{R}^m) \times L^2_{\mathcal{F}}(0, T+K; \mathbb{R}^{m\times d})$ *satisfying* (1).

Remark 2.3 It should be mentioned here that, in fact condition (A1) can be weaken to (A1″), which says

(*A*1″) There exists a constant $L'' > 0$ such that for each $y_\cdot, y'_\cdot \in L^2_{\mathcal{F}}(0, T+K; \mathbb{R}^m)$, $z_\cdot, z'_\cdot \in L^2_{\mathcal{F}}(0, T+K; \mathbb{R}^{m\times d})$, the following holds:

$$\begin{aligned} &E\bigg[\int_0^T |f(s, \{y_r\}_{r\in[s,T+K]}, \{z_r\}_{r\in[s,T+K]}) - f(s, \{y'_r\}_{r\in[s,T+K]}, \{z'_r\}_{r\in[s,T+K]})|^2 ds\bigg] \\ \le &L'' E\bigg[\int_0^{T+K} (|y_s - y'_s|^2 + |z_s - z'_s|^2) ds\bigg]. \end{aligned}$$

This can be easily checked from the detailed proofs of the theorem.

Remark 2.4 Let us give some examples of generator functions satisfying (A1). Assume that $h(\omega, t, y, z) : \Omega \times [0, T] \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^m$ is $\mathcal{F}_t$-adapted and Lipschitz in (y, z), i.e., there exists a constant $L_h > 0$ such that $|h(t, y, z) - h(t, y', z')| \le L_h(|y - y'| + |z - z'|)$ for any $(y, z), (y', z') \in \mathbb{R}^m \times \mathbb{R}^{m\times d}$. Then we can easily check that f_i $(i = 1, 2, 3, 4)$ defined below will satisfy (A1):

$$\begin{aligned} f_1(t, \{y_r\}_{r\in[t,T+K]}, \{z_r\}_{r\in[t,T+K]}) &:= h\bigg(t, E^{\mathcal{F}_t}\bigg[\int_t^{T+K} y_r dr\bigg], E^{\mathcal{F}_t}\bigg[\int_t^{T+K} z_r dr\bigg]\bigg), \\ f_2(t, \{y_r\}_{r\in[t,T+K]}, \{z_r\}_{r\in[t,T+K]}) &:= E^{\mathcal{F}_t}\bigg[h\bigg(t, \int_t^{T+K} y_r dr, \int_t^{T+K} z_r dr\bigg)\bigg]; \\ f_3(t, \{y_r\}_{r\in[t,T+K]}, \{z_r\}_{r\in[t,T+K]}) &:= h\bigg(t, \mathcal{E}_g^{\mathcal{F}_t}\bigg[\int_t^{T+K} y_r dr\bigg], \mathcal{E}_g^{\mathcal{F}_t}\bigg[\int_t^{T+K} z_r dr\bigg]\bigg), \\ f_4(t, \{y_r\}_{r\in[t,T+K]}, \{z_r\}_{r\in[t,T+K]}) &:= \mathcal{E}_g^{\mathcal{F}_t}\bigg[h\bigg(t, \int_t^{T+K} y_r dr, \int_t^{T+K} z_r dr\bigg)\bigg], \end{aligned}$$

where $\mathcal{E}_g^{\mathcal{F}_t}$ is the conditional g-expection, a nonlinear expectation, introduced by a BSDE (see Peng[16] for details).

2.2 Stochastic Functional Differential Equations

For each $t \in [0, T]$, let

$$\begin{aligned} &b(t, \{x_r\}_{r\in[-M,t]}) : \Omega \times [0, T] \times L^2_{\mathcal{F}}(-M, t; \mathbb{R}^n) \to L^2(\mathcal{F}_t; \mathbb{R}^n), \\ &\sigma(t, \{x_r\}_{r\in[-M,t]}) : \Omega \times [0, T] \times L^2_{\mathcal{F}}(-M, t; \mathbb{R}^n) \to L^2(\mathcal{F}_t; \mathbb{R}^{n\times d}). \end{aligned}$$

Consider the following stochastic functional differential equation (SFDE):

$$\begin{cases} dX_t = b(t, \{X_r\}_{r\in[-M,t]})dt + \sigma(t, \{X_r\}_{r\in[-M,t]})dB_t, & t \in [0, T]; \\ X_t = \rho_t, & t \in [-M, 0], \end{cases} \tag{2}$$

where $\rho_\cdot \in C(-M, 0; \mathbb{R}^n)$.

Definition 2.5 A process $X_\cdot : \Omega \times [-M,T] \to \mathbb{R}^n$ is called an adapted solution of SFDE (2) if $X_\cdot \in L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n)$ and it satisfies (2).

It should be mentioned that, Mohammed[13] has considered several types of SFDEs, and got the existence and uniqueness result by using Picard iterations. Here in order to make the paper self-contained, we will provide a proof by applying the fixed point theorem rather than Picard iterations.

We impose the following assumption:

(A3) There exists a constant $L > 0$ such that for each $x_\cdot, x'_\cdot \in L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n)$, the following hold:

$$E\bigg[\int_0^T \mathrm{e}^{-\theta s}|b(s,\{x_r\}_{r\in[-M,s]}) - b(s,\{x'_r\}_{r\in[-M,s]})|^2 ds\bigg] \le LE\int_{-M}^T \mathrm{e}^{-\theta s}|x_s - x'_s|^2 ds,$$
$$E\bigg[\int_0^T \mathrm{e}^{-\theta s}|\sigma(s,\{x_r\}_{r\in[-M,s]}) - \sigma(s,\{x'_r\}_{r\in[-M,s]})|^2 ds\bigg] \le LE\int_{-M}^T \mathrm{e}^{-\theta s}|x_s - x'_s|^2 ds,$$

where $\theta \ge 0$ is an arbitrary constant;

(A4) $b(t,0) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ and $\sigma(t,0) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$.

Remark 2.6 Let us give some examples of coefficients satisfying (A3). Assume that $p(\omega,t,x) : \Omega \times [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ and $q(\omega,t,x) : \Omega \times [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times d}$ are $\mathcal{F}_t$-adapted and Lipschitz w.r.t. x, i.e., there exist constants $L_p > 0$, $L_q > 0$ such that $|p(t,x) - p(t,x')| \le L_p|x - x'|$, $|q(t,x) - q(t,x')| \le L_q|x - x'|$ for any $x, x' \in \mathbb{R}^n$. Then we can easily check that b_1, b_2, σ_1 and σ_2 defined below will satisfy (A3):

$$b_1(t,\{x_r\}_{r\in[-M,t]}) := p\bigg(t, \int_{-M}^t x_r dr\bigg), \quad b_2(t,\{x_r\}_{r\in[-M,t]}) := \int_{-M}^t p(r,x_r)dr,$$
$$\sigma_1(t,\{x_r\}_{r\in[-M,t]}) := q\bigg(t, \int_{-M}^t x_r dr\bigg), \quad \sigma_2(t,\{x_r\}_{r\in[-M,t]}) := \int_{-M}^t q(r,x_r)dr.$$

We now give the existence and uniqueness result for SFDE (2).

Theorem 2.7 *Assume that* (A3) *and* (A4) *hold. Then SFDE* (2) *has a unique adapted solution.*

Proof Let θ be a nonnegative constant. Now we use the following norm in $L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n)$:

$$\|v(\cdot)\|_{-\theta} := \bigg(E\int_{-M}^T \mathrm{e}^{-\theta s}|v(s)|^2 ds\bigg)^{\frac{1}{2}},$$

which is equivalent to the original norm of $L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n)$. Henceforth we will find that this new norm is more convenient for us to construct a contraction mapping.

Let $X_\cdot$ be the unique solution of

$$\begin{cases} dX_t = b(t,\{x_r\}_{r\in[-M,t]})dt + \sigma(t,\{x_r\}_{r\in[-M,t]})dB_t, & t \in [0,T]; \\ X_t = \rho_t, & t \in [-M,0], \end{cases}$$

where $x_\cdot \in L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n)$. Now introduce a mapping I from $L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n)$ into itself by $X_\cdot = I(x_\cdot)$.

Let $x'_\cdot$ be another element of $L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n)$, and define $X'_\cdot = I(x'_\cdot)$. We make the following notations:

$$\begin{aligned}&\widehat{x}_\cdot = x_\cdot - x'_\cdot, \quad \widehat{X}_\cdot = X_\cdot - X'_\cdot,\\&\widehat{b}_t = b(t, \{x_r\}_{r\in[-M,t]}) - b(t, \{x'_r\}_{r\in[-M,t]}),\\&\widehat{\sigma}_t = \sigma(t, \{x_r\}_{r\in[-M,t]}) - \sigma(t, \{x'_r\}_{r\in[-M,t]}).\end{aligned}$$

Then for any $\theta \geq 0$, applying Itô's formula to $\mathrm{e}^{-\theta t}|\widehat{X}_t|^2$, and taking expectation, we have

$$E\mathrm{e}^{-\theta t}|\widehat{X}_t|^2 = E\int_0^t (-\theta)\mathrm{e}^{-\theta s}|\widehat{X}_s|^2 ds + E\int_0^t \mathrm{e}^{-\theta s}|\widehat{\sigma}_s|^2 ds + 2E\int_0^t \mathrm{e}^{-\theta s}\widehat{X}_s\widehat{b}_s ds.$$

This, together with $(A3)$, yields

$$E\int_0^T \theta\mathrm{e}^{-\theta s}|\widehat{X}_s|^2 ds \leq E\int_{-M}^T \mathrm{e}^{-\theta s}\Big(L^2 + \frac{2L^2}{\theta}\Big)|\widehat{x}_s|^2 ds + E\int_0^T \mathrm{e}^{-\theta s}\frac{\theta}{2}|\widehat{X}_s|^2 ds.$$

Thus if we choose $\theta = 2L^2 + 2L\sqrt{L^2+2}$, and note that $\widehat{X}_s \equiv 0$ for $s \in [-M,0]$, then we deduce

$$E\int_{-M}^T \mathrm{e}^{-\theta s}|\widehat{X}_s|^2 ds \leq \frac{1}{2}E\int_{-M}^T \mathrm{e}^{-\theta s}|\widehat{x}_s|^2 ds,$$

so that I is a strict contraction on $L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n)$. It follows by the fixed point theorem that SFDE (2) has a unique solution $X_\cdot \in L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n)$. ∎

At the end of this part, for the following SFDE, with the same form as in Chapter II of [13]:

$$\begin{cases} dX'_t = b'(t, \{X'_r\}_{r\in[t-M,t]})dt + \sigma'(t, \{X'_r\}_{r\in[t-M,t]})dB_t, & t \in [0,T];\\ X'_t = \rho_t, & t \in [-M,0],\end{cases} \tag{3}$$

we also give an existence and uniqueness theorem. Since the method to prove it is similar to Theorem 2.7, we omit here. For

$$\begin{aligned}&b'(t, \{x_r\}_{r\in[t-M,t]}) : \Omega \times [0,T] \times L^2_{\mathcal{F}}(t-M,t;\mathbb{R}^n) \to L^2(\mathcal{F}_t;\mathbb{R}^n),\\&\sigma'(t, \{x_r\}_{r\in[t-M,t]}) : \Omega \times [0,T] \times L^2_{\mathcal{F}}(t-M,t;\mathbb{R}^n) \to L^2(\mathcal{F}_t;\mathbb{R}^{n\times d}),\end{aligned}$$

we assume that

(A3′) There exists a constant $L' > 0$ such that for each $x_\cdot, x'_\cdot \in L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n)$, the following hold:

$$E\bigg[\int_0^T \mathrm{e}^{-\theta s}|b'(s, \{x_r\}_{r\in[s-M,s]}) - b'(s, \{x'_r\}_{r\in[s-M,s]})|^2 ds\bigg] \leq L'E\int_{-M}^T \mathrm{e}^{-\theta s}|x_s - x'_s|^2 ds,$$

$$E\bigg[\int_0^T \mathrm{e}^{-\theta s}|\sigma'(s, \{x_r\}_{r\in[s-M,s]}) - \sigma'(s, \{x'_r\}_{r\in[s-M,s]})|^2 ds\bigg] \leq L'E\int_{-M}^T \mathrm{e}^{-\theta s}|x_s - x'_s|^2 ds,$$

where $\theta \geq 0$ is an arbitrary constant;

$(A4')$ $b'(t,0) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ and $\sigma'(t,0) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$.

Theorem 2.8 *Assume that* (A3′) *and* (A4′) *hold. Then SFDE* (3) *has a unique adapted solution.*

3 Fully Coupled Forward-Backward Stochastic Functional Differential Equations

In this section, we consider the following fully coupled forward-backward stochastic functional differential equation (FBSFDE):

$$\begin{cases} dX_t = b(t, \{X_r\}_{r\in[-M,t]}, Y_t, Z_t)dt + \sigma(t, \{X_r\}_{r\in[-M,t]}, Y_t, Z_t)dB_t, & t \in [0,T]; \\ -dY_t = f(t, X_t, \{Y_r\}_{r\in[t,T+K]}, \{Z_r\}_{r\in[t,T+K]})dt - Z_t dB_t, & t \in [0,T]; \\ X_t = \rho_t, & t \in [-M, 0]; \\ Y_T = \Phi(X_T), \quad Y_t = \xi_t, & t \in (T, T+K]; \\ Z_t = \eta_t, & t \in [T, T+K], \end{cases} \tag{4}$$

where

$$b(t,\cdot,\cdot,\cdot) : \Omega \times [0,T] \times L^2_{\mathcal{F}}(-M,t;\mathbb{R}^n) \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to L^2(\mathcal{F}_t;\mathbb{R}^n),$$
$$\sigma(t,\cdot,\cdot,\cdot) : \Omega \times [0,T] \times L^2_{\mathcal{F}}(-M,t;\mathbb{R}^n) \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to L^2(\mathcal{F}_t;\mathbb{R}^{n\times d}),$$
$$f(t,\cdot,\cdot,\cdot) : \Omega \times [0,T] \times \mathbb{R}^n \times L^2_{\mathcal{F}}(t,T+K;\mathbb{R}^m) \times L^2_{\mathcal{F}}(t,T+K;\mathbb{R}^{m\times d}) \to L^2(\mathcal{F}_t,\mathbb{R}^m),$$
$$\Phi : \Omega \times \mathbb{R}^n \to \mathbb{R}^m, \ \rho_\cdot \in C(-M,0;\mathbb{R}^n), \ \xi_\cdot \in L^2_{\mathcal{F}}(T,T+K;\mathbb{R}^m), \ \eta_\cdot \in L^2_{\mathcal{F}}(T,T+K;\mathbb{R}^{m\times d}).$$

Definition 3.1 A triple of processes $(X_\cdot, Y_\cdot, Z_\cdot) : \Omega \times [-M,T] \times [0,T+K] \times [0,T+K] \to \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$ is called an adapted solution of FBSFDE (4) if $(X_\cdot, Y_\cdot, Z_\cdot) \in L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n) \times L^2_{\mathcal{F}}(0,T+K;\mathbb{R}^m) \times L^2_{\mathcal{F}}(0,T+K;\mathbb{R}^{m\times d})$ and it satisfies FBSFDE (4).

Given an $m \times n$ full-rank matrix G, we use the following notations:

$$u = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} \{x_r\}_{r\in[-M,\ \cdot]} \\ \{y_r\}_{r\in[\cdot,\ T+K]} \\ \{z_r\}_{r\in[\cdot,\ T+K]} \end{pmatrix}, \quad A(t,u,\alpha,\beta,\gamma) = \begin{pmatrix} -G^{\mathrm{T}} f(t,x,\beta,\gamma) \\ Gb(t,\alpha,y,z) \\ G\sigma(t,\alpha,y,z) \end{pmatrix},$$

where G^{T} denotes the transpose of G and $G\sigma = (G\sigma_1, G\sigma_2, \cdots, G\sigma_d)$.

Now we introduce the following assumptions:

(H1) $E\int_0^T |A(s,u,\alpha,\beta,\gamma)|^2 ds < +\infty$ for each $(u,\ \alpha,\ \beta,\ \gamma)$;

(H2) There exists a constant $L > 0$ such that for each $x_\cdot, x'_\cdot \in L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n)$, $y_\cdot, y'_\cdot \in L^2_{\mathcal{F}}(0,T+K;\mathbb{R}^m)$, $z_\cdot, z'_\cdot \in L^2_{\mathcal{F}}(0,T+K;\mathbb{R}^{m\times d})$, the following hold:

$$\begin{aligned} &E\int_0^T \mathrm{e}^{-\theta s}|b(s,\{x_r\}_{r\in[-M,s]},y_s,z_s) - b(s,\{x'_r\}_{r\in[-M,s]},y'_s,z'_s)|^2 ds \\ &+ E\int_0^T \mathrm{e}^{-\theta s}|\sigma(s,\{x_r\}_{r\in[-M,s]},y_s,z_s) - \sigma(s,\{x'_r\}_{r\in[-M,s]},y'_s,z'_s)|^2 ds \\ \le &LE\int_{-M}^T \mathrm{e}^{-\theta s}|x_s - x'_s|^2 ds + LE\int_0^T \mathrm{e}^{-\theta s}(|y_s - y'_s|^2 + |z_s - z'_s|^2) ds, \\ &E\int_0^T \mathrm{e}^{\theta s}|f(s,x_s,\{y_r\}_{r\in[s,T+K]},\{z_r\}_{r\in[s,T+K]}) - f(s,x'_s,\{y'_r\}_{r\in[s,T+K]},\{z'_r\}_{r\in[s,T+K]})|^2 ds \\ \le &LE\int_0^T \mathrm{e}^{\theta s}|x_s - x'_s|^2 ds + LE\int_0^{T+K} \mathrm{e}^{\theta s}(|y_s - y'_s|^2 + |z_s - z'_s|^2) ds, \end{aligned}$$

where $\theta \geq 0$ is an arbitrary constant;

(H3) $\Phi(x) \in L^2(\mathcal{F}_T; \mathbb{R}^m)$ and it is uniformly Lipschitz w.r.t. $x \in \mathbb{R}^n$;

(H4) $A(\cdot,\cdot,\cdot,\cdot,\cdot)$ and $\Phi(\cdot)$ satisfy

$$\begin{aligned}&E\int_0^T \langle A(s,u_s,\alpha_s,\beta_s,\gamma_s) - A(s,u'_s,\alpha'_s,\beta'_s,\gamma'_s), u_s - u'_s\rangle ds\\ \leq& -\lambda_1 E\int_{-M}^{\mathrm{T}} |G\widehat{x}_s|^2 ds - \lambda_2 E\int_0^{T+K} (|G^{\mathrm{T}}\widehat{y}_s|^2 + |G^{\mathrm{T}}\widehat{z}_s|^2) ds,\\ &\langle \Phi(x) - \Phi(x'), G(x-x')\rangle \geq \mu|G\widehat{x}|^2,\end{aligned}$$

for all $(u,\ \alpha,\ \beta,\ \gamma)$, $(u',\ \alpha',\ \beta',\ \gamma')$, x and x', $\widehat{x} = x - x'$, $\widehat{y} = y - y'$, $\widehat{z} = z - z'$, where λ_1, λ_2 and μ are given nonnegative constants with $\lambda_1 + \lambda_2 > 0$, $\lambda_2 + \mu > 0$. Moreover, we have $\lambda_1 > 0$, $\mu > 0$ (resp. $\lambda_2 > 0$) when $m > n$ (resp. $n > m$).

We first give the uniqueness theorem.

Theorem 3.2 *Assume that* (H1)–(H4) *hold. Then FBSFDE* (4) *has at most one adapted solution.*

Proof Suppose that $U_\cdot = (X_\cdot, Y_\cdot, Z_\cdot)$ and $U'_\cdot = (X'_\cdot, Y'_\cdot, Z'_\cdot)$ are two solutions of FBSFDE (4). We denote $\widehat{U}_\cdot = (\widehat{X}_\cdot, \widehat{Y}_\cdot, \widehat{Z})_\cdot = (X_\cdot - X'_\cdot, Y_\cdot - Y'_\cdot, Z_\cdot - Z'_\cdot)$. Applying Itô's formula to $\langle G\widehat{X}_t, \widehat{Y}_t\rangle$ and noting $(H4)$, we have

$$\begin{aligned}&E\langle \Phi(X_T) - \Phi(X'_T), G\widehat{X}_T\rangle\\ =&E\int_0^T \langle A(s,U_s,\alpha_s,\beta_s,\gamma_s) - A(s,U'_s,\alpha'_s,\beta'_s,\gamma'_s), \widehat{U}_s\rangle ds\\ \leq& -\lambda_1 E\int_{-M}^{\mathrm{T}} |G\widehat{X}_s|^2 ds - \lambda_2 E\int_0^{T+K} (|G^{\mathrm{T}}\widehat{Y}_s|^2 + |G^{\mathrm{T}}\widehat{Z}_s|^2) ds\\ =& -\lambda_1 E\int_0^T |G\widehat{X}_s|^2 ds - \lambda_2 E\int_0^T (|G^{\mathrm{T}}\widehat{Y}_s|^2 + |G^{\mathrm{T}}\widehat{Z}_s|^2) ds,\end{aligned}$$

where the last equality is due to $X_s = X'_s = \rho_s$ for $s \in [-M, 0]$ and $(Y_s, Z_s) = (Y'_s, Z'_s) = (\xi_s, \eta_s)$ for $s \in (T, T+K]$.

Together with (H4) again, we obtain

$$\lambda_1 E\int_0^T |G\widehat{X}_s|^2 ds + \lambda_2 E\int_0^T (|G^{\mathrm{T}}\widehat{Y}_s|^2 + |G^{\mathrm{T}}\widehat{Z}_s|^2) ds + \mu|G\widehat{X}_T|^2 \leq 0.$$

For the case when $m > n$, we note that $\lambda_1 > 0$ and $\mu > 0$. Then it is easy to get that for $s \in [0,T]$, $|G\widehat{X}_s|^2 \equiv 0$, which implies $\widehat{X}_s \equiv 0$. Thus $X_s \equiv X'_s$, for $s \in [0,T]$. Then according to Theorem 2.2 together with Remark 2.3, we know $(Y_s, Z_s) = (Y'_s, Z'_s)$ for $s \in [0,T]$.

For the case when $n > m$, we note that $\lambda_2 > 0$. Then for $s \in [0,T]$, $|G^{\mathrm{T}}\widehat{Y}_s|^2 \equiv 0$ and $|G^{\mathrm{T}}\widehat{Z}_s|^2 \equiv 0$, which implies $(Y_s, Z_s) \equiv (Y'_s, Z'_s)$. Finally, from the uniqueness of SFDEs (see Theorem 2.7), it follows that $X_s \equiv X'_s$ for $s \in [0,T]$.

Similar to the above, for the case when $m = n$, the result can be easily obtained. ∎

From now on, we will mainly study the existence of the solution to FBSFDE (4). For this, we first consider the following family of FBSFDEs parameterized by $\varepsilon \in [0,1]$:

$$\begin{cases} dX_t^\varepsilon = [(1-\varepsilon)\lambda_2(-G^{\mathrm{T}}Y_t^\varepsilon) + \varepsilon b(t,\{X_r^\varepsilon\}_{r\in[-M,t]}, Y_t^\varepsilon, Z_t^\varepsilon) + \varphi_t]dt \\ \qquad +[(1-\varepsilon)\lambda_2(-G^{\mathrm{T}}Z_t^\varepsilon) + \varepsilon\sigma(t,\{X_r^\varepsilon\}_{r\in[-M,t]}, Y_t^\varepsilon, Z_t^\varepsilon) + \phi_t]dB_t, \quad t\in[0,T]; \\ -dY_t^\varepsilon = [(1-\varepsilon)\lambda_1 GX_t^\varepsilon + \varepsilon f(t, X_t^\varepsilon, \{Y_r^\varepsilon\}_{r\in[t,T+K]}, \{Z_r^\varepsilon\}_{r\in[t,T+K]}) + \psi_t]dt - Z_t^\varepsilon dB_t, \\ X_t^\varepsilon = \rho_t, \qquad t\in[-M,0]; \\ Y_T^\varepsilon = \varepsilon\Phi(X_T^\varepsilon) + (1-\varepsilon)GX_T^\varepsilon + \zeta, \quad Y_t^\varepsilon = \xi_t, \qquad t\in(T,T+K]; \\ Z_t^\varepsilon = \eta_t, \qquad t\in[T,T+K], \end{cases} \tag{5}$$

where $\zeta \in L^2(\mathcal{F}_T;\mathbb{R}^m)$, $\varphi_\cdot \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$, $\phi_\cdot \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^{n\times d})$ and $\psi_\cdot \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^m)$. It is obvious that the existence of (4) just follows from that of (5) when $\varepsilon = 1$.

Lemma 3.3 *Assume that* (H1)–(H4) *hold. If for an* $\varepsilon_0 \in [0,1)$, *there exists a solution* $(X_\cdot^{\varepsilon_0}, Y_\cdot^{\varepsilon_0}, Z_\cdot^{\varepsilon_0})$ *of FBSFDE* (5), *then there exists a positive constant* δ_0, *such that for each* $\delta\in[0,\delta_0]$ *there exists a solution* $(X_\cdot^{\varepsilon_0+\delta}, Y_\cdot^{\varepsilon_0+\delta}, Z_\cdot^{\varepsilon_0+\delta})$ *of FBSFDE* (5) *for* $\varepsilon = \varepsilon_0+\delta$.

Proof Let $u_\cdot = (x_\cdot, y_\cdot, z_\cdot) \in L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n) \times L^2_{\mathcal{F}}(0,T+K;\mathbb{R}^m) \times L^2_{\mathcal{F}}(0,T+K;\mathbb{R}^{m\times d})$. Then it follows that there exists a unique triple $U_\cdot = (X_\cdot, Y_\cdot, Z_\cdot) \in L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n) \times L^2_{\mathcal{F}}(0,T+K;\mathbb{R}^m) \times L^2_{\mathcal{F}}(0,T+K;\mathbb{R}^{m\times d})$ satisfying the following FBSFDE:

$$\begin{cases} dX_t = [(1-\varepsilon_0)\lambda_2(-G^{\mathrm{T}}Y_t) + \varepsilon_0 b(t,\{X_r\}_{r\in[-M,t]}, Y_t, Z_t) + \varphi_t]dt \\ \qquad +\delta(\lambda_2 G^{\mathrm{T}}y_t + b(t,\{x_r\}_{r\in[-M,t]}, y_t, x_t))dt \\ \qquad +[(1-\varepsilon_0)\lambda_2(-G^{\mathrm{T}}Z_t) + \varepsilon_0\sigma(t,\{X_r\}_{r\in[-M,t]}, Y_t, Z_t) + \phi_t]dB_t \\ \qquad +\delta\lambda_2(G^{\mathrm{T}}z_t + \sigma(t,\{x_r\}_{r\in[-M,t]}, y_t, z_t))dB_t, \quad t\in[0,T]; \\ -dY_t = [(1-\varepsilon_0)\lambda_1 GX_t + \varepsilon_0 f(t, X_t, \{Y_r\}_{r\in[t,T+K]}, \{Z_r\}_{r\in[t,T+K]}) + \psi_t]dt \\ \qquad +\delta(-\lambda_1 Gx_t + f(t, x_t, \{y_r\}_{r\in[t,T+K]}, \{z_r\}_{r\in[t,T+K]}))dt - Z_t dB_t, \quad t\in[0,T]; \\ X_t = \rho_t, \qquad t\in[-M,0]; \\ Y_T = \varepsilon_0\Phi(X_T) + (1-\varepsilon_0)GX_T + \delta(\Phi(x_T) - Gx_T) + \zeta, \ \ Y_t = \xi_t, \qquad t\in(T,T+K]; \\ Z_t = \eta_t, \qquad t\in[T,T+K]. \end{cases}$$

Our objective is to prove that for sufficiently small δ, the mapping $I_{\varepsilon_0+\delta}$, defined by $U_\cdot = I_{\varepsilon_0+\delta}(u_\cdot)$ from $L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n) \times L^2_{\mathcal{F}}(0,T+K;\mathbb{R}^m) \times L^2_{\mathcal{F}}(0,T+K;\mathbb{R}^{m\times d})$ into itself, is a contraction mapping.

Let $u'_\cdot = (x'_\cdot, y'_\cdot, z'_\cdot)$ be another element of $L^2_{\mathcal{F}}(-M,T;\mathbb{R}^n) \times L^2_{\mathcal{F}}(0,T+K;\mathbb{R}^m) \times L^2_{\mathcal{F}}(0,T+$

$K;\mathbb{R}^{m\times d})$ and define $U'_\cdot = I_{\varepsilon_0+\delta}(u'_\cdot)$. We make the following notations:

$$\begin{aligned}
&\widehat{u}_\cdot = (\widehat{x}_\cdot, \widehat{y}_\cdot, \widehat{z}_\cdot) = (x_\cdot - x'_\cdot, y_\cdot - y'_\cdot, z_\cdot - z'_\cdot),\\
&\widehat{U}_\cdot = (\widehat{X}_\cdot, \widehat{Y}_\cdot, \widehat{Z}_\cdot) = (X_\cdot - X'_\cdot, Y_\cdot - Y'_\cdot, Z_\cdot - Z'_\cdot),\\
&\widehat{b}_t = b(t, \{x_r\}_{r\in[-M,t]}, y_t, z_t) - b(t, \{x'_r\}_{r\in[-M,t]}, y_t, z_t),\\
&\widehat{\sigma}_t = \sigma(t, \{x_r\}_{r\in[-M,t]}, y_t, z_t) - \sigma(t, \{x'_r\}_{r\in[-M,t]}, y_t, z_t),\\
&\widehat{f}_t = f(t, x_t, \{y_r\}_{r\in[t,T+K]}, \{z_r\}_{r\in[t,T+K]}) - f(t, x'_t, \{y'_r\}_{r\in[t,T+K]}, \{z'_r\}_{r\in[t,T+K]}).
\end{aligned}$$

Apply Itô's formula to $\langle G\widehat{X}_t, \widehat{Y}_t\rangle$, and take expectation,

$$\begin{aligned}
&\varepsilon_0 E\langle \Phi(X_T) - \Phi(X'_T), G\widehat{X}_T\rangle + (1-\varepsilon_0)E|G\widehat{X}_T|^2 + \delta E\langle \Phi(x_T) - \Phi(x'_T) - G\widehat{x}_T, G\widehat{X}_T\rangle ds\\
=&E\int_0^T \varepsilon_0\langle A(s, U_s, \alpha_s, \beta_s, \gamma_s) - A(s, U'_s, \alpha'_s, \beta'_s, \gamma'_s), \widehat{U}_s\rangle ds\\
&-(1-\varepsilon_0)E\int_0^T (\lambda_1\langle G\widehat{X}_s, G\widehat{X}_s\rangle + \lambda_2\langle G^{\mathrm{T}}\widehat{Y}_s, G^{\mathrm{T}}\widehat{Y}_s\rangle + \lambda_2\langle G^{\mathrm{T}}\widehat{Z}_s, G^{\mathrm{T}}\widehat{Z}_s\rangle)ds\\
&+\delta E\int_0^T (\lambda_1\langle G\widehat{X}_s, G\widehat{x}_s\rangle + \lambda_2\langle G^{\mathrm{T}}\widehat{Y}_s, G^{\mathrm{T}}\widehat{y}_s\rangle + \lambda_2\langle G^{\mathrm{T}}\widehat{Z}_s, G^{\mathrm{T}}\widehat{z}_s\rangle\\
&+\langle \widehat{X}_s, -G^{\mathrm{T}}\widehat{f}_s\rangle + \langle G^{\mathrm{T}}\widehat{Y}_s, \widehat{b}_s\rangle + \langle \widehat{Z}_s, G\widehat{\sigma}_s\rangle)ds.
\end{aligned}$$

From (H1)–(H4), we have

$$\begin{aligned}
&(\varepsilon_0\mu + (1-\varepsilon_0))E|G\widehat{X}_T|^2 + \lambda_1 E\int_{-M}^{\mathrm{T}} |G\widehat{X}_s|^2 ds + \lambda_2 E\int_0^{T+K} (|G^{\mathrm{T}}\widehat{Y}_s|^2 + |G^{\mathrm{T}}\widehat{Z}_s|^2)ds\\
\le& C_1\delta E\int_{-M}^{\mathrm{T}} (|\widehat{X}_s|^2 + |\widehat{x}_s|^2)ds + C_1\delta E\int_0^{T+K} (|\widehat{Y}_s|^2 + |\widehat{y}_s|^2 + |\widehat{Z}_s|^2 + |\widehat{z}_s|^2)ds\\
&+C_1\delta E|\widehat{X}_T|^2 + C_1\delta E|\widehat{x}_T|^2. \qquad (6)
\end{aligned}$$

Here the constant C_1 depends on G, L, λ_1, λ_2.

Next we will give two other estimates. On the one hand, similarly to the proof of Theorem 2.7, for any $\theta \ge 0$, by applying Itô's formula to $\mathrm{e}^{-\theta t}|\widehat{X}_t|^2$, we have

$$\begin{aligned}
E\mathrm{e}^{-\theta T}|\widehat{X}_T|^2 =&E\int_0^T (-\theta)\mathrm{e}^{-\theta s}|\widehat{X}_s|^2 ds + E\int_0^T \mathrm{e}^{-\theta s}|\widehat{\sigma}_s|^2 ds + 2E\int_0^T \mathrm{e}^{-\theta s}\widehat{X}_s\widehat{b}_s ds\\
\le&E\int_0^T \Big(-\frac{\theta}{2}\Big)\mathrm{e}^{-\theta s}|\widehat{X}_s|^2 ds + E\int_0^T \mathrm{e}^{-\theta s}|\widehat{\sigma}_s|^2 ds + \frac{2}{\theta}E\int_0^T \mathrm{e}^{-\theta s}|\widehat{b}_s|^2 ds,
\end{aligned}$$

where

$$\begin{aligned}
\widehat{b}_s =&(1-\varepsilon_0)\lambda_2(-G^{\mathrm{T}}\widehat{Y}_s) + \varepsilon_0(b(s, \{X_r\}_{r\in[-M,s]}, Y_s, Z_s) - b(s, \{X'_r\}_{r\in[-M,s]}, Y'_s, Z'_s))\\
&+\delta(\lambda_2 G^{\mathrm{T}}\widehat{y}_s + b(s, \{x_r\}_{r\in[-M,s]}, y_s, z_s) - b(s, \{x'_r\}_{r\in[-M,s]}, y'_s, z'_s)),\\
\widehat{\sigma}_s =&(1-\varepsilon_0)\lambda_2(-G^{\mathrm{T}}\widehat{Z}_s) + \varepsilon_0(\sigma(s, \{X_r\}_{r\in[-M,s]}, Y_s, Z_s) - \sigma(s, \{X'_r\}_{r\in[-M,s]}, Y'_s, Z'_s))\\
&+\delta(\lambda_2 G^{\mathrm{T}}\widehat{z}_s + \sigma(s, \{x_r\}_{r\in[-M,s]}, y_s, z_s) - \sigma(s, \{x'_r\}_{r\in[-M,s]}, y'_s, z'_s)).
\end{aligned}$$

According to (H2),

$$
\begin{aligned}
E\int_0^T \mathrm{e}^{-\theta s}|\widehat{b}_s|^2 ds \leq & 4(1-\varepsilon_0)^2\lambda_2^2 E\int_0^T \mathrm{e}^{-\theta s}|G^{\mathrm{T}}\widehat{Y}_s|^2 ds \\
& + 4\varepsilon_0^2 L\Big(E\int_{-M}^{\mathrm{T}} \mathrm{e}^{-\theta s}|\widehat{X}_s|^2 ds + E\int_0^T \mathrm{e}^{-\theta s}|\widehat{Y}_s|^2 ds + E\int_0^T \mathrm{e}^{-\theta s}|\widehat{Z}_s|^2 ds\Big) \\
& + 4\delta^2\lambda_2^2 E\int_0^T \mathrm{e}^{-\theta s}|G^{\mathrm{T}}\widehat{y}_s|^2 ds \\
\leq & 4\lambda_2^2 E\int_0^T \mathrm{e}^{-\theta s}|G^{\mathrm{T}}\widehat{Y}_s|^2 ds \\
& + 4L\Big(E\int_{-M}^{\mathrm{T}} \mathrm{e}^{-\theta s}|\widehat{X}_s|^2 ds + E\int_0^T \mathrm{e}^{-\theta s}|\widehat{Y}_s|^2 ds + E\int_0^T \mathrm{e}^{-\theta s}|\widehat{Z}_s|^2 ds\Big) \\
& + 4\delta^2\lambda_2^2 E\int_0^T \mathrm{e}^{-\theta s}|G^{\mathrm{T}}\widehat{y}_s|^2 ds \\
& + 4\delta^2 L\Big(E\int_{-M}^{\mathrm{T}} \mathrm{e}^{-\theta s}|\widehat{x}_s|^2 ds + E\int_0^T \mathrm{e}^{-\theta s}|\widehat{y}_s|^2 ds + E\int_0^T \mathrm{e}^{-\theta s}|\widehat{z}_s|^2 ds\Big),
\end{aligned}
$$

and similarly,

$$
\begin{aligned}
E\int_0^T \mathrm{e}^{-\theta s}|\widehat{\sigma}_s|^2 ds \leq & 4\lambda_2^2 E\int_0^T \mathrm{e}^{-\theta s}|G^{\mathrm{T}}\widehat{Z}_s|^2 ds \\
& + 4L\Big(E\int_{-M}^{\mathrm{T}} \mathrm{e}^{-\theta s}|\widehat{X}_s|^2 ds + E\int_0^T \mathrm{e}^{-\theta s}|\widehat{Y}_s|^2 ds + E\int_0^T \mathrm{e}^{-\theta s}|\widehat{Z}_s|^2 ds\Big) \\
& + 4\delta^2\lambda_2^2 E\int_0^T \mathrm{e}^{-\theta s}|G^{\mathrm{T}}\widehat{z}_s|^2 ds \\
& + 4\delta^2 L\Big(E\int_{-M}^{\mathrm{T}} \mathrm{e}^{-\theta s}|\widehat{x}_s|^2 ds + E\int_0^T \mathrm{e}^{-\theta s}|\widehat{y}_s|^2 ds + E\int_0^T \mathrm{e}^{-\theta s}|\widehat{z}_s|^2 ds\Big).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\frac{\theta}{2}E\int_0^T \mathrm{e}^{-\theta s}|\widehat{X}_s|^2 ds \leq & E\int_0^T \mathrm{e}^{-\theta s}|\widehat{\sigma}_s|^2 ds + \frac{2}{\theta}E\int_0^T \mathrm{e}^{-\theta s}|\widehat{b}_s|^2 ds \\
\leq & 4L\Big(1+\frac{2}{\theta}\Big)E\int_{-M}^{\mathrm{T}} \mathrm{e}^{-\theta s}|\widehat{X}_s|^2 ds \\
& + C_2 E\int_0^{T+K} (\mathrm{e}^{-\theta s}|\widehat{Y}_s|^2 + \mathrm{e}^{-\theta s}|\widehat{Z}_s|^2) ds \\
& + C_2\delta^2 E\int_{-M}^{\mathrm{T}} \mathrm{e}^{-\theta s}|\widehat{x}_s|^2 ds + C_2\delta^2 E\int_0^{T+K} \mathrm{e}^{-\theta s}(|\widehat{y}_s|^2 + |\widehat{z}_s|^2) ds.
\end{aligned}
$$

Choosing θ sufficiently large, we can easily get the following estimate:

$$
\begin{aligned}
& E\int_{-M}^{\mathrm{T}} |\widehat{X}_s|^2 ds \\
\leq & C_2\delta^2 E\int_{-M}^{\mathrm{T}} |\widehat{x}_s|^2 ds + C_2\delta^2 E\int_0^{T+K} (|\widehat{y}_s|^2 + |\widehat{z}_s|^2) ds + C_2 E\int_0^{T+K} (|\widehat{Y}_s|^2 + |\widehat{Z}_s|^2) ds. \quad (7)
\end{aligned}
$$

Here the constant C_2 depends on G, L, λ_2.

On the other hand, for $(\widehat{Y}., \widehat{Z}.)$, thanks to the estimate of BSDEs, together with $(H2)$, we can easily derive

$$\begin{aligned}&E\int_0^{T+K}(|\widehat{Y}_s|^2+|\widehat{Z}_s|^2)ds\\ \le\ & C_3\delta^2 E\int_{-M}^{\mathrm{T}}|\widehat{x}_s|^2ds + C_3\delta^2 E\int_0^{T+K}(|\widehat{y}_s|^2+|\widehat{z}_s|^2)ds + C_3\delta^2 E|\widehat{x}_T|^2\\ &+C_3E\int_{-M}^{\mathrm{T}}|\widehat{X}_s|^2ds + C_3E|\widehat{X}_T|^2. \qquad (8)\end{aligned}$$

Here the constant C_3 depends on G, L, λ_1.

Now combining the above three estimates (6)–(8), and noting the fact that $\mu > 0$ implies $\varepsilon_0\mu + (1-\varepsilon_0) > 0$, we can easily check that, whenever $\lambda_1 > 0$, $\mu > 0$, $\lambda_2 \ge 0$ or $\lambda_1 \ge 0$, $\mu \ge 0$, $\lambda_2 > 0$, the following always holds:

$$\begin{aligned}&E\int_{-M}^{\mathrm{T}}|\widehat{X}_s|^2ds + E\int_0^{T+K}(|\widehat{Y}_s|^2+|\widehat{Z}_s|^2)ds + E|\widehat{X}_T|^2\\ \le\ & C(\delta+\delta^2)\left(E\int_{-M}^{\mathrm{T}}|\widehat{x}_s|^2ds + E\int_0^{T+K}(|\widehat{y}_s|^2+|\widehat{z}_s|^2)ds + E|\widehat{x}_T|^2\right),\end{aligned}$$

where the constant C depends on C_1, C_2, C_3, λ_1, λ_2, μ.

Thus if we choose $\delta_0 = \min\{1, \frac{1}{4C}\}$, we can clearly see that, for each $\delta \in [0, \delta_0]$, the mapping $I_{\varepsilon_0+\delta}$ is a strict contraction on $L^2_{\mathcal{F}}(-M, T; \mathbb{R}^n) \times L^2_{\mathcal{F}}(0, T+K; \mathbb{R}^m) \times L^2_{\mathcal{F}}(0, T+K; \mathbb{R}^{m\times d})$ in the sense that

$$\begin{aligned}&E\int_{-M}^{\mathrm{T}}|\widehat{X}_s|^2ds + E\int_0^{T+K}(|\widehat{Y}|^2+|\widehat{Z}_s|^2)ds + E|\widehat{X}_T|^2\\ \le\ & \frac{1}{2}\left(E\int_{-M}^{\mathrm{T}}|\widehat{x}_s|^2ds + E\int_0^{T+K}(|\widehat{y}|^2+|\widehat{z}_s|^2)ds + E|\widehat{x}_T|^2\right).\end{aligned}$$

Then it follows by the fixed point theorem that the mapping $I_{\varepsilon_0+\delta}$ has a unique fixed point $U^{\varepsilon_0+\delta}_{\cdot} = (X^{\varepsilon_0+\delta}_{\cdot}, Y^{\varepsilon_0+\delta}_{\cdot}, Z^{\varepsilon_0+\delta}_{\cdot})$, which is the unique solution of (5) for $\varepsilon = \varepsilon_0 + \delta$. ∎

Now we give the main result of this part.

Theorem 3.4 *Assume that* (H1)–(H4) *hold. Then there exists a unique adapted solution* $(X., Y., Z.)$ *of FBSFDE* (4).

Proof The uniqueness is an immediate result from Theorem 3.2. Next we prove the existence.

Note that FBSFDE (5) for $\varepsilon = 0$ admits a unique solution (see Theorem 2.6 in [3]). Thus from Lemma 3.3, there exists a positive constant δ_0 such that, for each $\delta \in [0, \delta_0]$, (5) for $\varepsilon = \varepsilon_0 + \delta$ admits a unique solution. Repeat this process for N times with $1 \le N\delta_0 < 1 + \delta_0$, then we can obtain that particularly for $\varepsilon = 1$ with $\phi. = 0$, $\varphi. = 0$, $\psi_0 = 0$ and $\zeta = 0$, (5) has a unique solution, i.e., FBSFDE (4) has a unique solution. ∎

4 Quadratic Optimal Control Problem for Functional Stochastic Systems

Let $\rho_\cdot \in C(-M,0;\mathbb{R}^n)$, and let $v_\cdot$ be an admissible control process, i.e., an $\mathcal{F}_t$-adapted square integrable process taking values in a given subset of $\mathbb{R}^k$. Then we consider the following control system:

$$\begin{cases} dX_t = \left(A_t \int_{-M}^{t} X_s ds + C_t v_t\right)dt + \left(D_t \int_{-M}^{t} X_s ds + F_t v_t\right)dB_t, & t \in [0,T]; \\ X_t = \rho_t, & t \in [-M,0], \end{cases} \tag{9}$$

where $A_\cdot$, $C_\cdot$, $D_\cdot$ and $F_\cdot$ are bounded progressively measurable matrix-valued processes with appropriate dimensions. Then according to Theorem 2.7 and Remark 2.6, SFDE (9) admits a unique solution.

The classical quadratic optimal control problem is to minimize the cost function

$$J(v_\cdot) = \frac{1}{2}E\left[\int_0^T (\langle R_t X_t, X_t\rangle + \langle N_t v_t, v_t\rangle)dt + \langle Q X_T, X_T\rangle\right],$$

where Q is an $\mathcal{F}_T$-measurable nonnegative symmetric bounded matrix, $R_\cdot$ is an $n\times n$ nonnegative symmetric bounded progressively measurable matrix-valued process, $N_\cdot$ is an $k\times k$ positive symmetric bounded progressively measurable matrix-valued process and its inverse $N_\cdot^{-1}$ is also bounded.

The following theorem tells us that, for the above optimal control problem, we can find the explicit form of the optimal control $u_\cdot$ satisfying

$$J(u_\cdot) = \inf_{v_\cdot} J(v_\cdot),$$

by means of the fully coupled forward-backward stochastic functional differential equations.

Theorem 4.1 *The process*

$$u_t = -N_t^{-1}(C_t^{\mathrm{T}} Y_t + F_t^{\mathrm{T}} Z_t), \quad t \in [0,T]$$

is the unique optimal control which satisfies

$$J(u_\cdot) = \inf_{v_\cdot} J(v_\cdot),$$

where $(X_\cdot, Y_\cdot, Z_\cdot)$ *is the unique solution of the following FBSFDE:*

$$\begin{cases} dX_t = \Big[A_t \int_{-M}^{t} X_s ds - C_t N_t^{-1}(C_t^{\mathrm{T}} Y_t + F_t^{\mathrm{T}} Z_t)\Big] dt \\ \qquad + \Big[D_t \int_{-M}^{t} X_s ds - F_t N_t^{-1}(C_t^{\mathrm{T}} Y_t + F_t^{\mathrm{T}} Z_t)\Big] dB_t, & t \in [0,T]; \\ -dY_t = \Big[E^{\mathcal{F}_t}\Big(\int_t^{T+K} A_s^{\mathrm{T}} Y_s ds\Big) + E^{\mathcal{F}_t}\Big(\int_t^{T+K} C_s^{\mathrm{T}} Z_s ds\Big) + R_t X_t\Big] dt \\ -Z_t dB_t, & t \in [0,T]; \\ X_t = \rho_t, & t \in [-M,0]; \\ Y_T = QX_T, \quad Y_t = 0, & t \in (T, T+K]; \\ Z_t = 0, & t \in [T, T+K]. \end{cases} \tag{10}$$

Proof From Theorem 3.4, we know that FBSFDE (10) admits a unique solution $(X_\cdot, Y_\cdot, Z_\cdot)$. Denote the unique solution of SFDE (9) by $X_\cdot^v$ for the control $v_\cdot$.

Applying Itô's formula to $\langle X_t^v - X_t, Y_t\rangle$, and taking expectation, we have

$$\begin{aligned} & E\langle X_T^v - X_T, Y_T\rangle \\ =& -E\int_0^T \Big[\Big\langle E^{\mathcal{F}_t}\Big(\int_t^{T+K} A_s^{\mathrm{T}} Y_s ds\Big) + E^{\mathcal{F}_t}\Big(\int_t^{T+K} D_s^{\mathrm{T}} Z_s ds\Big) + R_t X_t, X_t^v - X_t\Big\rangle\Big] dt \\ & + E\int_0^T \Big(\Big\langle A_t \int_{-M}^t (X_s^v - X_s) ds, Y_t\Big\rangle + \langle C_t(v_t - u_t), Y_t\rangle\Big) dt \\ & + E\int_0^T \Big(\Big\langle D_t \int_{-M}^t (X_s^v - X_s) ds, Z_t\rangle + \langle F_t(v_t - u_t, Z_t)\Big\rangle\Big) dt. \end{aligned}$$

Note that

$$\begin{aligned} & E\int_0^T \Big(\Big\langle A_t \int_{-M}^t (X_s^v - X_s) ds, Y_t\Big\rangle - \Big\langle E^{\mathcal{F}_t}\Big(\int_t^{T+K} A_s^{\mathrm{T}} Y_s ds\Big), X_t^v - X_t\Big\rangle\Big) dt \\ =& E\int_0^T \Big\langle A_t \int_{-M}^t (X_s^v - X_s) ds, Y_t\Big\rangle dt - E\int_0^T \Big\langle \int_t^{T+K} A_s^{\mathrm{T}} Y_s ds, X_t^v - X_t\Big\rangle dt \\ =& E\int_0^T \Big\langle A_t \int_{-M}^t (X_s^v - X_s) ds, Y_t\Big\rangle dt \\ & - E\int_0^T \Big\langle A_t \int_0^t (X_s^v - X_s) ds, Y_t\Big\rangle dt - E\int_T^{T+K} \Big\langle A_t \int_0^T (X_s^v - X_s) ds, Y_t\Big\rangle dt \\ =& E\int_0^T \Big\langle A_t \int_{-M}^0 (X_s^v - X_s) ds, Y_t\Big\rangle dt - E\int_T^{T+K} \Big\langle A_t \int_0^T (X_s^v - X_s) ds, Y_t\Big\rangle dt \\ =& 0, \end{aligned}$$

and similarly,

$$E\int_0^T \Big(\Big\langle D_t \int_{-M}^t (X_s^v - X_s) ds, Z_t\Big\rangle - \Big\langle E^{\mathcal{F}_t}\Big(\int_t^{T+K} D_s^{\mathrm{T}} Z_s ds\Big), X_t^v - X_t\Big\rangle\Big) dt = 0.$$

Combining the above three equalities, we have

$$E\langle X_T^v - X_T, Y_T\rangle = E\int_0^T (\langle -R_tX_t, X_t^v - X_t\rangle + \langle C_t(v_t-u_t), Y_t\rangle + \langle F_t(v_t-u_t), Z_t\rangle)dt,$$

which implies

$$E\langle X_T^v - X_T, Y_T\rangle + E\int_0^T \langle R_tX_t, X_t^v - X_t\rangle = E\int_0^T (\langle C_t(v_t-u_t), Y_t\rangle + \langle F_t(v_t-u_t), Z_t\rangle)dt. \quad (11)$$

On the other hand,

$$\begin{aligned}
&J(v.) - J(u.)\\
=&\frac{1}{2}E\Big[\int_0^T (\langle R_tX_t^v, X_t^v\rangle - \langle R_tX_t, X_t\rangle + \langle N_tv_t, v_t\rangle - \langle N_tu_t, u_t\rangle)dt\\
&\quad + \langle QX_T^v, X_T^v\rangle - \langle QX_T, X_T\rangle\Big]\\
=&\frac{1}{2}E\Big[\int_0^T (\langle R_t(X_t^v - X_t), X_t^v - X_t\rangle + 2\langle R_tX_t, X_t^v - X_t\rangle\\
&\quad + \langle N_t(v_t-u_t, v_t-u_t\rangle + 2\langle N_tu_t, v_t-u_t\rangle)dt\\
&\quad + \langle Q(X_T^v - X_T), X_T^v - X_T\rangle + 2\langle QX_T, X_T^v - X_T\rangle\Big]\\
\geq& E\int_0^T (\langle R_tX_t, X_t^v - X_t\rangle + \langle N_tu_t, v_t-u_t\rangle)dt + \langle QX_T, X_T^v - X_T\rangle,
\end{aligned}$$

where the last inequality is due to the positivity of $N.$, and the nonnegativity of $R.$ and Q. Then together with (11), and noting that $Y_T = QX_T$, we obtain

$$\begin{aligned}
&J(v.) - J(u.)\\
\geq& E\int_0^T (\langle R_tX_t, X_t^v - X_t\rangle + \langle N_tu_t, v_t-u_t\rangle)dt + \langle QX_T, X_T^v - X_T\rangle]\\
=&E\int_0^T (\langle C_t(v_t-u_t), Y_t\rangle + \langle F_t(v_t-u_t), Z_t\rangle + \langle N_tu_t, v_t-u_t\rangle)dt\\
=&0.
\end{aligned}$$

Hence, $u_t = -N_t^{-1}(C_t^{\mathrm{T}}Y_t + F_t^{\mathrm{T}}Z_t)$ is an optimal control.

Moreover, the optimal control is unique. In fact, assume that $u.$ and $u'_.$ are both optimal controls, and denote $J(u.) = J(u'_.) \triangleq J \geq 0$. The corresponding trajectories are $X.$ and $X'_.$. It is easy to check that for the control $\frac{u_\cdot+u'_\cdot}{2}$, the trajectory is $\frac{X_\cdot+X'_\cdot}{2}$. Then,

$$\begin{aligned}
2J &= J(u_\cdot) + J(u'_\cdot)\\
&= \frac{1}{2}E\bigg[\int_0^T (\langle R_tX_t, X_t\rangle + \langle N_tu_t, u_t\rangle)dt + \langle QX_T, X_T\rangle\bigg]\\
&\quad + \frac{1}{2}E\bigg[\int_0^T (\langle R_tX'_t, X'_t\rangle + \langle N_tu'_t, u'_t\rangle)dt + \langle QX'_T, X'_T\rangle\bigg]\\
&= 2J\Big(\frac{u_\cdot+u'_\cdot}{2}\Big) + E\bigg[\int_0^T \Big(\Big\langle R_t\frac{X_t-X'_t}{2}, \frac{X_t-X'_t}{2}\Big\rangle + \Big\langle N_t\frac{u_t-u'_t}{2}, \frac{u_t-u'_t}{2}\Big\rangle\Big)dt\bigg]\\
&\quad + E\Big\langle Q\frac{X_T-X'_T}{2}, \frac{X_T-X'_T}{2}\Big\rangle\\
&\geq 2J + E\int_0^T \Big\langle N_t\frac{u_t-u'_t}{2}, \frac{u_t-u'_t}{2}\Big\rangle dt,
\end{aligned}$$

where the last inequality is due to the nonnegativity of $R_\cdot$ and Q.

Therefore $u_\cdot = u'_\cdot$, thanks to the positivity of $N_\cdot$. ∎

Remark 4.2 It should be mentioned here that, the method we applied to prove the uniqueness above is in fact a classical method, readers are referred to [8] or [10].

5 Conclusions and Future Works

In this paper, we proved the existence and uniqueness theorem for fully coupled forward-backward stochastic functional differential equations (FBSFDEs). Moreover, as an application, we studied a quadratic optimal control problem for functional stochastic systems, and got the optimal control by virtue of FBSFDEs. Motivated by the work of [10], [12], [17], [18], etc., in the future we could further study FBSFDEs with Lévy noise, or study the other applications of these equations (such as the application to optimal control problems, the stochastic differential games, and so on). Besides, it is known that observer design, as well as controller design, is the core problem of control theory. It has been widely used in many aspects. Compared with classical linear systems, the design of observers for the general systems is more complex and difficult. Observer design for the system in this current work will also be discussed in our future research.

References

[1] Pardoux E and Peng S G, Adapted solution of a backward stochastic differential equation, *Systems Control Lett.*, 1990, **14**: 55–61.

[2] El Karoui N, Peng S G, and Quenez M C, Backward stochastic differential equations in finance, *Math. Finance*, 1997, **7**: 1–71.

[3] Peng S G and Wu Z, Fully coupled forward-backward stochastic differential equations and applications to optimal control, *SIAM J. Control Optim.*, 1999, **37**(3): 825–843.

[4] Peng S G and Yang Z, Anticipated backward stochastic differential equations, *Ann. Probab.*, 2009, **37**: 877–902.

[5] Hu Y and Peng S G, Solution of forward-backward stochastic differential equations, *Prob. Theory Rel. Fields*, 1995, **103**(2): 273–283.

[6] Ma J, Protter P, and Yong J M, Solving forward-backward stochastic differential equations explicitly-a four step scheme, *Prob. Theory Rel. Fields*, 1994, **98**: 339–359.

[7] Yong J M, Finding adapted solution of forward backward stochastic differential equations-method of contituation, *Prob. Theory Rel. Fields*, 1997, **107**(4): 537–572.

[8] Chen L and Wu Z, A type of general forward-backward stochastic differential equations and applications, *Chin. Ann. Math.*, 2011, **32**B(2): 279–292.

[9] Cvitanic J and Ma J, Hedging options for a large investor and forward-backward SDE's, *Ann. Appl. Probab.*, 1996, **6**(2): 370–398.

[10] Wu Z, Forward-backward stochastic differential equations, linear quadratic stochastic optimal control and nonzero sum differential games, *Journal of Systems Science & Complexity*, 2005, **18**(2): 179–192.

[11] Yu Z Y and Ji S L, Linear-quadratic nonzero-sum differential game of backward stochastic differential equations, *Proceedings of the* 27*th Chinese Control Conference*, Kunming, Yunnan, 2008.

[12] Wu Z, Maximum principle for optimal control problem of fully coupled forward-backward stochastic systems, *Systems Sci. Mathe. Sci.*, 1998, **11**(3): 249–259.

[13] Mohammed S E A, *Stochastic Functional Differential Equations*, Longman, New York, 1986.

[14] Yang Z, Anticipated BSDEs and related results in SDEs, Doctoral Dissertation, Shandong University, Jinan, 2007.

[15] Yang Z and Elliott R J, Some properties of generalized anticipated backward stochastic differential equations, *Electron. Commun. Probab.*, 2013, **18**(63): 1–10.

[16] Peng S G, Backward SDE and related g-expectation, *Backward Stochastic Dierential Equations* (ed. by El Karoui N and Mazliak L), Pitman Res. Notes Math. Ser., Longman, Harlow, 1997, **364**: 141–159.

[17] Zhu Q X, Razumikhin-type theorem for stochastic functional differential equations with Lévy noise and Markov switching, *International Journal of Control*, 2017, **90**(8): 1703–1712.

[18] Zhu Q X, Stability analysis of stochastic delay differential equations with Lévy noise, *Systems and Control Lett.*, 2018, **118**: 62–68.