Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2014, Article ID 361259, 12 pages
http://dx.doi.org/10.1155/2014/361259

Research Article

Stochastic Maximum Principle of Near-Optimal Control of Fully Coupled Forward-Backward Stochastic Differential Equation

Maoning Tang

Department of Mathematical Sciences, Huzhou University, Huzhou, Zhejiang 313000, China

Correspondence should be addressed to Maoning Tang; tangmaoning@hutc.zj.cn

Received 30 December 2013; Revised 29 January 2014; Accepted 6 February 2014; Published 16 March 2014

Academic Editor: Shen Yin

This paper first makes an attempt to investigate the near-optimal control of systems governed by fully nonlinear coupled forward-backward stochastic differential equations (FBSDEs) under the assumption of a convex control domain. By Ekeland's variational principle and some basic estimates for state processes and adjoint processes, we establish the necessary conditions for any ε-near optimal control in a local form with an error order of exact $\varepsilon^{1/2}$. Moreover, under additional convexity conditions on Hamiltonian function, we prove that an ε-maximum condition in terms of the Hamiltonian in the integral form is sufficient for near-optimality of order $\varepsilon^{1/2}$.

1. Introduction

Bismut [1] first investigated linear backward stochastic differential equations (BSDEs in short) as the adjoint equation of the forward stochastic system. The existence and uniqueness of BSDEs with nonlinear generators under Lipschitz condition were first proved by Pardoux and Peng [2] in 1990. Since then, the theory of BSDEs has extensive applications in both mathematical finance and stochastic control. Forward-backward stochastic differential equations (FBSDEs in short) consist of forward stochastic differential equations (SDEs in short) of Itô type and BSDEs of Pardoux-Peng. Forward-backward stochastic equations (FBSDEs) not only are widely used in stochastic control and differential games but also have profound applications in mathematical economics and mathematical finance. Therefore, it is natural to investigate control problems for systems governed by this kind of stochastic equations. In mathematical finance, FBSDEs can be formulated as the price equations of financial assets under model uncertainty. In the stochastic optimal control problem, FBSDEs arise as the Hamilton system which is composed of the optimality conditions, the adjoint equation, and the state equation and which completely characterizes the optimal control.

A classical approach for optimal control problems is to derive necessary conditions satisfied by an optimum, such as Pontryagin's maximum principle. Now the maximum principles for optimal controls of FBSDEs have rich literatures which can be referred to [3–12] and references therein.

The references stated in the above are all concerned with (exact) optimal control. But, in fact, the (exact) optimal control may not exist in many situations. So it becomes very important to study near-optimal controls which are more available and much easier to be obtained than optimal ones, both analytically and numerically. The near-optimal deterministic control problems have been investigated in [13–15]. Near-optimal control problems for SDEs with controlled diffusion coefficients were first investigated in 1998 by Zhou [16], where necessary and sufficient conditions are established by introducing second adjoint equation, for all near-optimal controls. Inspired by Zhou [16], we refer to [16–20] on the near-optimal control of other forward stochastic systems.

For forward-backward stochastic systems, Huang et al. [21] in 2010 and Bahlali et al. [22] in 2009, respectively, established the corresponding stochastic maximum principle for the near-optimal control of linear systems and nonlinear systems, where diffusion coefficients and control variables are each independently based on Ekeland's principle and spike

variation. In 2011, Hui et al. [23] studied the near-optimal control of nonlinear FBSDEs, where diffusion coefficients can be dependent on the control variable, with the assumption that the control domain is convex. In 2012, for linear FBSDEs, Zhang et al. [24] extended the results of [21–23] to the general case of control domains based on the Ekeland's principle, spike variation, reduction technique developed recently by Yong [25], and the methodology recently introduced by Wu [26].

The control systems of FBSDEs studied in references [21–24] are nonfully coupled which are only coupled in BSDE and not in SDE. For the control systems of fully coupled FBSDEs, the existing literatures mainly focused on exact optimal control problems and few on near-optimal control problems. The purpose of the present paper is to make the first attempt to discuss the near-optimal control for fully coupled FBSDEs. Its main contribution is the developments of maximum principle and verification theorem of the near-optimal control in a uniform manner by Ekeland's variational principle. Compared with references [21–24], this paper mainly has three advantages as follows. Firstly, our systems studied are fully coupled FBSDEs, which are coupled not only in BSDEs but also in SDEs. Secondly, we get necessary optimality conditions for near-optimal control with an error order of exact $\varepsilon^{1/2}$, which is better than all in the existing literature on the cases of FBSDEs, where the error orders are almost $\varepsilon^{1/3}$. In fact, by Ekeland's variational principle, we know that the error order of exact $\varepsilon^{1/2}$ for the near-optimal control is the best error order. Thirdly, different from [21–24], by continuous dependence theorem of FBSDEs (see Lemma 4), we obtain directly the basic estimates for state processes and adjoint processes (see Lemmas 10, 11, 12, and 14) which play a very important role in proving our main results. Therefore, our approach is simpler and more quickly.

The paper is organized as follows. In Section 2, we present the notations and give main theory on FBSDEs. In Section 3, the problem studied is formulated and basic assumptions are given. In Section 4, we prove some prior estimates for state trajectories and adjoint equation. In Section 5, we obtain a variational formula for the performance functional. Sections 6 and 7 are devoted to deriving verification theorem and stochastic maximum principle by Ekeland's variational principle. In Section 8, we conclude our paper.

2. Preliminary Notations and Basic Theory for FBSDEs

Now we first introduce some preliminary notations which will be used throughout this paper. Let $(\Omega, \mathscr{F}, P)$ be a probability space. Let $\{W(t), 0 \le t \le T\}$ be a d-dimensional Brownian motion. Let $\{\mathscr{F}_t\}_{0\le t\le T}$ be P-completed natural filtration generated by $\{W(t), 0 \le t \le T\}$. Let E be a Euclidean space, where the inner product and norm are denoted by $(\cdot,\cdot)$ and $|\cdot|$, respectively. For a given function, $\phi : \mathbb{R}^n \to \mathbb{R}$, we denote its gradient and Hessian by ϕ_x and ϕ_{xx}, respectively. If $\phi : \mathbb{R}^n \to \mathbb{R}^k$ (with $k \ge 2$), then $\phi_x = (\partial\phi_i/\partial x_j)$ is the corresponding $(k \times n)$ Jacobian matrix. By $\mathscr{P}$ we denote the predictable σ field on $\Omega \times [0, T]$ and by $\mathscr{B}(\Lambda)$ the Borel σ-algebra of any topological space Λ. Denote by $M^2_{\mathscr{F}}(0, T; E)$ the space of all $\mathscr{P}$-measurable E-valued stochastic processes $f = \{f(t,\omega), (t,\omega) \in [0,T] \times \Omega\}$ satisfying $\|f\|_{M^2_{\mathscr{F}}(0,T;E)} \triangleq \sqrt{E\int_0^T |f(t)|^2 dt} < \infty$, by $S^2_{\mathscr{F}}(0, T; E)$, the space of all $\mathscr{F}_t$-adapted E-valued stochastic càdlàg processes $f = \{f(t,\omega), (t,\omega) \in [0,T] \times \Omega\}$ such that $\|f\|_{S^2_{\mathscr{F}}(0,T;E)} \triangleq \sqrt{E\sup_{0\le t\le T}|f(t)|^2 dt} < +\infty$, by $L^2(\Omega, \mathscr{F}, P; E)$, and the set of all E-valued random variables ξ on $(\Omega, \mathscr{F}, P)$ such that $\|\xi\|_{L^2(\Omega,\mathscr{F},P;E)} \triangleq \sqrt{E|\xi|^2} < \infty$. Finally, we define the space

$$\mathbb{M}^2[0,T] := S^2_{\mathscr{F}}(0,T;\mathbb{R}^n) \times S^2_{\mathscr{F}}(0,T;\mathbb{R}^m) \times M^2_{\mathscr{F}}\left(0,T;\mathbb{R}^{m\times d}\right). \tag{1}$$

Then $\mathbb{M}^2[0,T]$ is a Banach space with respect to the norm $\|\cdot\|_{\mathbb{M}^2}$ given by

$$\|\Theta(\cdot)\|^2_{\mathbb{M}^2} = E\sup_{0\le t\le T}|x(t)|^2 + E\sup_{0\le t\le T}|y(t)|^2 + E\int_0^T |z(t)|^2 dt, \tag{2}$$

for $\Theta(\cdot) = (x(\cdot), y(\cdot), z(\cdot)) \in \mathbb{M}^2[0,T]$.

Now we are in position to present the preliminary results of fully coupled FBSDEs. Consider a general FBSDE as follows:

$$\begin{aligned} dx(t) &= b(t, x(t), y(t), z(t))\,dt \\ &\quad + \sigma(t, x(t), y(t), z(t))\,dB(t), \\ y(t) &= -f(t, x(t), y(t), z(t))\,dt + z(t)\,dB(t), \\ x(0) &= a, \\ y(T) &= h(x(T)). \end{aligned} \tag{3}$$

Here $f : [0,T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^m$, $b : [0,T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^n$, $h : \Omega \times \mathbb{R}^n \to \mathbb{R}^m$, and $\sigma : [0,T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^{n\times d}$ are given mappings and $a \in \mathbb{R}^n$. For a given full-rank $m \times n$ matrix G, we use the notations $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ and $A(t,u) = \begin{pmatrix} -G^* f \\ Gb \\ G\sigma \end{pmatrix}$, where G^* is the transpose matrix of G.

Definition 1. A stochastic process $(x(\cdot), y(\cdot), z(\cdot)) \in \mathbb{M}^2[0,T]$ for the coefficients (a, b, σ, f, h) is said to be an adapted solution of (3) if, for any $t \in [0,T]$, it follows almost surely

$$\begin{aligned} x(t) &= a + \int_0^t b(r, x(r), y(r), z(r))\,dr \\ &\quad + \int_0^t \sigma(r, x(r), y(r), z(r))\,dB(r), \\ y(t) &= h(x(T)) + \int_t^T f(r, x(r), y(r), z(r))\,dr \\ &\quad - \int_t^T z(r)\,dB(r). \end{aligned} \tag{4}$$

Furthermore, FBSDE (3) is said to be solvable if it has an adapted solution. An FBSDE is said to be nonsolvable if it is not solvable.

In order to get the solvability of FBSDE (3), we make the basic assumptions as follows.

Assumption 2. (i) The random mappings b, σ, and f are $\mathscr{P}\otimes\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(\mathbb{R}^m)\times\mathscr{B}(\mathbb{R}^{m\times d})$ measurable with $b(\cdot,0,0,0)\in M^2_{\mathscr{F}}(0,T;R^n)$, $\sigma(\cdot,0,0,0)\in M^2_{\mathscr{F}}(0,T;\mathbb{R}^{n\times d})$, and $f(\cdot,0,0,0)\in M^2_{\mathscr{F}}(0,T;\mathbb{R}^m)$. And h is $\mathscr{F}_T\times\mathscr{B}(\mathbb{R}^m)$ measurable with $h(0)\in L^2(\Omega,\mathscr{F}_T,P;\mathbb{R}^m)$. Moreover, b, σ, and f are uniformly Lipschitz continuous in (x,y,z) and h is uniformly Lipschitz continuous in x.

(ii) Monotonicity conditions

$$\begin{aligned}&\langle A(t,u)-A(t,\overline{u}),u-\overline{u}\rangle\\&\quad\le-\gamma_1|G\widehat{x}|^2-\gamma_2\left(\left|G^*\widehat{y}\right|^2+\left|G^*\widehat{z}\right|^2\right),\\&\langle x-\overline{x},h(x)-h(\overline{x})\rangle\ge\theta_1|G\widehat{x}|^2,\end{aligned}\tag{5}$$

or

$$\begin{aligned}&\langle A(t,u)-A(t,\overline{u}),u-\overline{u}\rangle\\&\quad\ge\gamma_1|G\widehat{x}|^2+\gamma_2\left(\left|G^*\widehat{y}\right|^2+\left|G^*\widehat{z}\right|^2\right),\\&\langle h(x)-h(\overline{x}),x-\overline{x}\rangle\le-\theta_1|G\widehat{x}|^2,\end{aligned}\tag{6}$$

for all $u=(x,y,z)$ and $\overline{u}=(\overline{x},\overline{y},\overline{z})$, $\widehat{x}=x-\overline{x}$, $\widehat{y}=y-\overline{y}$, $\widehat{z}=z-\overline{z}$, where γ_1, γ_2 and θ_1 are nonnegative constants with $\gamma_1+\gamma_2\ge0$, $\gamma_2+\theta_1\ge0$. Moreover, we have $\gamma_1>0$, $\theta_1>0$ (resp., $\gamma_2>0$), if $m>n$ (resp., $m<n$).

The following two lemmas present the solvability results and continuous dependence theorem of FBSDE (3), respectively, which will be used to demonstrate the basic estimates for the state equation and adjoint equation (see Lemmas 10, 11, 12, and 14).

Lemma 3. *Let Assumption 2 be satisfied. Then* (3) *admits a unique solution* $(x(\cdot),y(\cdot),z(\cdot))\in\mathbb{M}^2[0,T]$.

The proof can be found in Peng and Wu [27].

Lemma 4. *Let* $(x(\cdot),y(\cdot),z(\cdot))$ *and* $(\overline{x}(\cdot),\overline{y}(\cdot),\overline{z}(\cdot))$ *be the solutions of the FBSDE* (3) *corresponding to two given coefficients* (b,σ,f,h,a) *and* $(\overline{b},\overline{\sigma},\overline{f},\overline{h},\overline{a})$ *which both satisfy Assumption 2, respectively. Then there exists a constant such that*

$$\begin{aligned}&\left\|\left(\widehat{x}(\cdot),\widehat{y}(\cdot),\widehat{z}(\cdot)\right)\right\|_{\mathbb{M}^2[0,T]}\\&\le K\Bigg[|a-\overline{a}|^2\\&\quad+E\int_0^T\left|b\left(r,\overline{x}(r),\overline{y}(r),\overline{z}(r)\right)-\overline{b}\left(r,\overline{x}(r),\overline{y}(r),\overline{z}(r)\right)\right|^2dr\\&\quad+E\int_0^T\left|\sigma\left(r,\overline{x}(r),\overline{y}(r),\overline{z}(r)\right)-\overline{\sigma}\left(r,\overline{x}(r),\overline{y}(r),\overline{z}(r)\right)\right|^2dr\\&\quad+E\int_0^T\left|f\left(r,\overline{x}(r),\overline{y}(r),\overline{z}(r)\right)-\overline{f}\left(r,\overline{x}(r),\overline{y}(r),\overline{z}(r)\right)\right|^2dr\\&\quad+E\left|h(\overline{x}(T))-\overline{h}(\overline{x}(T))\right|^2\Bigg].\end{aligned}\tag{7}$$

Particularly, if $(\overline{b},\overline{\sigma},\overline{f},\overline{h},\overline{a})=(0,0,0,0)$, *we have*

$$\begin{aligned}&\left\|\left(x(\cdot),y(\cdot),z(\cdot)\right)\right\|_{\mathbb{M}^2[0,T]}\\&\le K\Bigg[|a|^2+E\int_0^T|b(t,0,0,0)|^2dt+E\int_0^T|\sigma(t,0,0,0)|^2dt\\&\quad+E\int_0^T\left|f(t,0,0,0)\right|^2dt+E|h(0)|^2\Bigg].\end{aligned}\tag{8}$$

The proof can be found in Lin [28].

3. Statement of the Problem and Basic Assumptions

Suppose that U is a given compact convex subset of $\mathbb{R}^k$. The stochastic process $u(\cdot):[0,T]\times\Omega\to\mathbb{R}^k$ is said to be admissible, if it is an $\mathscr{F}_t$-adopted process taking values in U. We denote all admissible controls by the set $\mathscr{A}$.

For any admissible control $u(\cdot)\in\mathscr{A}$, we consider the following controlled FBSDE:

$$\begin{aligned}dx(s)&=b\left(r,x(r),y(r),z(r),u(r)\right)dr\\&\quad+\sigma\left(r,x(r),y(r),z(r),u(r)\right)dB(r),\\dy(r)&=-f\left(r,x(r),y(r),z(r),u(r)\right)dr+z(r)\,dB(r),\\x(0)&=a\in\mathbb{R}^n,\\y(T)&=h(x(T)),\end{aligned}\tag{9}$$

with the performance functional

$$\begin{aligned}J(u(\cdot))=E\Bigg[&\int_0^T l\left(t,x(t),y(t),z(t),u(t)\right)dt\\&+\phi(x(T))+\gamma\left(y(0)\right)\Bigg].\end{aligned}\tag{10}$$

In the above, b,σ,f,h,l,ϕ, and γ are given random mappings. $b:[0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times U\to\mathbb{R}^n$, $\sigma:[0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times U\to\mathbb{R}^{n\times d}$, $f:[0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times U\to\mathbb{R}^m$, $l:[0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times U\to\mathbb{R}^1$, $h:\Omega\times\mathbb{R}^n\times\mathbb{R}^n$, $\gamma:\Omega\times\mathbb{R}^m\to\mathbb{R}^1$, and $\phi:\Omega\times\mathbb{R}^n\to\mathbb{R}^1$ are given measurable mappings.

The basic assumptions on coefficients $(b, \sigma, f, h, l, \phi, \gamma)$ are given as follows.

Assumption 5. (i) For any $u \in U$, (a, b, σ, f, h) satisfy Assumption 2. Moreover, b, f, and σ are differentiable in (x, y, z, u), h is differentiable in x, and the corresponding derivatives are continuous and uniformly bounded for all $(t, \omega) \in [0, T] \times \Omega$.

(ii) $l : [0, T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \times U \to \mathbb{R}^1$ is continuous differentiable in (x, y, z, u), $\phi : \Omega \times \mathbb{R}^n \to \mathbb{R}^1$ is continuous differentiable in x, and $\gamma : \Omega \times \mathbb{R}^m \to \mathbb{R}^1$ is continuous differentiable in y. For all $(t, \omega) \in [0, T] \times \Omega$, there is a constant C such that, for all $(x, y, z, u) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \times U$,

$$
\begin{aligned}
|l| &\le C\left(1 + |y|^2 + |x|^2 + |z|^2 + |u|^2\right), \\
|\gamma| &\le C\left(|y|^2 + 1\right), \qquad |\phi| \le C\left(|x|^2 + 1\right), \\
|l_x| + |l_y| + |l_z| + |l_u| &\le C\left(1 + |z| + |x| + |y| + |u|\right), \\
|\phi_x| &\le C\left(|x| + 1\right), \\
|\gamma_y| &\le C\left(|y| + 1\right).
\end{aligned} \tag{11}
$$

Under Assumption 5, from Lemma 3, we know that, for every $u(\cdot) \in \mathscr{A}$, (9) has a unique solution. The corresponding strong solution is denoted by $(x^u(\cdot), y^u(\cdot), z^u(\cdot))$ or $(x(\cdot), y(\cdot), z(\cdot))$. Then $(x(\cdot), y(\cdot), z(\cdot))$ is said to be the state processes associated with the admissible control $u(\cdot)$ and $(u(\cdot); x(\cdot), y(\cdot), z(\cdot))$ is called the admissible control pair. Moreover, under Assumption 5, using a priori estimates (8), we can deduce the fact that

$$
|J(u(\cdot))| < \infty. \tag{12}
$$

The so-called stochastic optimal control problem is to minimize the cost function $J(u(\cdot))$, over all $u(\cdot) \in \mathscr{A}$. The corresponding value function is defined as

$$
V(a) = \inf_{u(\cdot) \in \mathscr{A}} J(u(\cdot)). \tag{13}
$$

We denote the above problem (9)–(13) by *PRO*. Any $\overline{u}(\cdot) \in \mathscr{A}$ is said to be an optimal control of Problem *PRO*, if $\overline{u}(\cdot)$ achieves the infimum of $J(u(\cdot))$ over $\mathscr{A}$. The state process $(\overline{x}(\cdot), \overline{y}(\cdot), \overline{z}(\cdot))$ is said to be the optimal state. And $(\overline{u}(\cdot); \overline{x}(\cdot), \overline{y}(\cdot), \overline{z}(\cdot))$ is called an optimal pair of Problem *PRO*.

Since this paper is devoted to discussing the near-optimal problem of FBSDEs, we recall the definition of the near-optimal control, following [16].

Definition 6. An admissible control pair $(u^\varepsilon(\cdot); x^\varepsilon(\cdot), y^\varepsilon(\cdot), z^\varepsilon(\cdot))$ is said to be an ε-optimal control for some $\varepsilon \ge 0$, if

$$
\left|J\left(u^\varepsilon(\cdot)\right) - V(a)\right| \le \varepsilon. \tag{14}
$$

Definition 7. The set of parameterized admissible control pairs $\{(u^\varepsilon(\cdot); x^\varepsilon(\cdot), y^\varepsilon(\cdot), z^\varepsilon(\cdot))\}$ is said to be near-optimal for sufficient small ε, if

$$
\left|J\left(u^\varepsilon(\cdot)\right) - V(a)\right| \le r(\varepsilon). \tag{15}
$$

Here r is a function with respect to ε satisfying $r(\varepsilon) \to 0$ as $\varepsilon \to 0$. We call the estimate $r(\varepsilon)$ an error bound. If $r(\varepsilon) = c\varepsilon^\delta$ for some $\delta > 0$ independent of the constant c, then we call $u^\varepsilon(\cdot)$ the near-optimal control with order ε^δ.

Before we conclude this section, let us recall the definition of the Clarke generalized gradient as well as Ekeland's variational principle which will be used to prove our main results.

Definition 8 (see Zhou [16]). Let X be a convex set in $\mathbb{R}^d$ and let $\eta(\cdot) : \to R$ be a locally Lipschitz function. At any given $x \in X$, we define the generalized gradient of η as a set given by

$$
\partial_x \eta = \left\{ \xi : \langle \xi, \beta \rangle \le \lim_{y \to x, y \in X, h \downarrow 0} \sup \frac{\eta(y + h\beta) - \eta(x)}{h}, \right.
$$
$$
\left. \text{for any } \beta \in \mathbb{R}^d \right\}. \tag{16}
$$

Lemma 9 (Ekeland's variational principle [29]). *Suppose that (S, d) is a complete metric space and $\rho(\cdot) : S \to \mathbb{R}$ is bounded from below and lower-semi-continuous. For $\varepsilon > 0$, let $u^\varepsilon \in S$ satisfy the following inequality:*

$$
\rho\left(u^\varepsilon\right) \le \inf_{u \in S} \rho(u) + \varepsilon. \tag{17}
$$

Then, for any $\lambda > 0$, there exists u^λ such that

(1) $\rho(u^\lambda) \le \rho(u^\varepsilon)$,

(2) $d(u^\lambda, u^\varepsilon) \le \lambda$,

(3) $\rho(u^\lambda) \le \rho(u) + (\varepsilon/\lambda) d(u^\lambda, u)$, *for all $u \in S$.*

4. Some Prior Estimates for State Trajectories and Adjoint Equations

In order to apply Ekeland's variational principle to obtain our main result, we must define a distance d on the space of admissible controls such that $(\mathscr{A}, d)$ is a complete metric space. For any given $v(\cdot), u(\cdot) \in \mathscr{A}$, we define

$$
d(v(\cdot), u(\cdot)) = \left[E \int_0^T |v(r) - u(r)|^2 dr \right]^{1/2}. \tag{18}
$$

To simplify our notation, for any admissible control pair $(u(\cdot); x^u(\cdot), y^u(\cdot), z^u(\cdot))$, we set

$$
\Theta^u(t) := \left(x^u(t), y^u(t), z^u(t)\right). \tag{19}
$$

The following is devoted to proving the boundedness and continuity of the state and adjoint processes with the control processes under the metric (18). Note that, in the following, C is a generic constants, which may change from line to line.

Lemma 10. *Let Assumption 5 be satisfied. Then there exists a constant C s.t. that, for every admissible pair $(u(\cdot); \Theta^u(\cdot)) = (u(\cdot); x^u(\cdot), y^u(\cdot), z^u(\cdot))$,*

$$
\left\|\Theta^u(\cdot)\right\|_{\mathbb{M}^2} \le C. \tag{20}
$$

Proof. Under Assumption 5, by the estimate (8), we have

$$\begin{aligned}\left\|\Theta^u(\cdot)\right\|_{\mathbb{M}^2} &\le C\Bigg[E\int_0^T |b(t,0,0,0,u(t))|^2dt \\ &\quad + E\int_0^T |\sigma(t,0,0,0,u(t))|^2dt \\ &\quad + E\int_0^T \left|f(t,0,0,0,u(t))\right|^2dt \\ &\quad + E|h(0)|^2 + |a|^2\Bigg] \\ &\le C\Bigg[E\int_0^T |b(t,0,0,0,0)|^2dt \\ &\quad + E\int_0^T |\sigma(t,0,0,0,0)|^2dt \\ &\quad + E\int_0^T \left|f(t,0,0,0,0)\right|^2dt \\ &\quad + E|h(0)|^2 + 1 + |a|^2\Bigg],\end{aligned} \tag{21}$$

where the last inequality is obtained by the boundedness of the control domain U. The proof is complete. □

Lemma 11. *Let Assumption 5 be satisfied. Then there is a positive constant C s.t. for any given two admissible pairs $(u(\cdot);\Theta^u(\cdot)) = (u(\cdot);x^u(\cdot),y^u(\cdot),z^u(\cdot))$ and $(v(\cdot);\Theta^v(\cdot)) = (v(\cdot);x^v(\cdot),y^v(\cdot),z^v(\cdot))$,*

$$\left\|\Theta^u(\cdot)-\Theta^v(\cdot)\right\|^2_{\mathbb{M}^2} \le Cd(u(\cdot),v(\cdot))^2. \tag{22}$$

Proof. Under Assumption 5, from the estimate (7), we have

$$\begin{aligned}&\left\|\Theta^u(\cdot)-\Theta^v(\cdot)\right\|^2_{\mathbb{M}^2} \\ &\le C\Bigg[E\int_0^T \left|b(r,\Theta^u(r),u(r))-b(r,\Theta^u(r),v(r))\right|^2dr \\ &\quad + E\int_0^T \left|\sigma(r,\Theta^u(r),u(r))-\sigma(r,\Theta^u(r),v(r))\right|^2dr \\ &\quad + E\int_0^T \left|f(r,\Theta^u(r),u(r))-f(r,\Theta^u(r),v(r))\right|^2dr\Bigg] \\ &\le CE\int_0^T |u(r)-v(r)|^2dr \\ &= Cd(u(\cdot),v(\cdot))^2,\end{aligned} \tag{23}$$

where the second inequality is obtained by the mean value theorem and the boundedness of b_u, σ_u, and f_u. The proof is complete. □

We know that the adjoint process plays a key role in establishing stochastic maximum principle. In the following, we will study certain boundedness and continuity of adjoint processes with the control variable under the metric d.

For a given admissible pair $(u(\cdot);\Theta^u(\cdot))$, corresponding adjoint process $\Lambda^u(\cdot) = (k^u(\cdot),p^u(\cdot),q^u(\cdot))$ is defined as the solution to the following FBSDE:

$$\begin{aligned}dk_t &= -\Big[b_y^*(t,\Theta^u(t),u(t))\,p_t + \sigma_y^*(t,\Theta^u(t),u(t))\,q_t \\ &\quad - f_y^*(t,\Theta^u(t),u(t))\,k_t \\ &\quad + l_y(t,\Theta^u(t),u(t))\Big]dt \\ &\quad - \left[b_z^*(t,\Theta^u(t),u(t))\,p_t\right. \\ &\quad + \sigma_z^*(t,\Theta^u(t),u(t))\,q_t \\ &\quad \left. - f_z^*(t,\Theta^u(t),u(t))\,k_t + l_z(t,\Theta^u(t),u(t))\right]dB_t, \\ dp_t &= -\left[b_x^*(t,\Theta^u(t),u(t))\,p_t\right. \\ &\quad + \sigma_x^*(t,\Theta^u(t),u(t))\,q_t - f_x^*(t,\Theta^u(t),u(t))\,k_t \\ &\quad \left. + l_x(t,\Theta^u(t),u(t))\right]dt + q_t dB_t, \\ k_0 &= -\gamma_y(y^u(0)), \\ p_T &= -h_x^*(x^u(T))\,k(T) + \phi_x(x^u(T)), \\ &\qquad 0 \le t \le T.\end{aligned} \tag{24}$$

The adjoint equation (24) is a linear FBSDE whose solution consists of $(p^u(\cdot),q^u(\cdot),k^u(\cdot))$. Under Assumption 5, by Lemma 3, the adjoint equation has a unique solution $\Lambda^u(\cdot) = (k^u(\cdot),p^u(\cdot),q^u(\cdot)) \in \mathbb{M}^2$.

Next, the Hamiltonian $H : [0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times U\times\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}^{n\times d} \to \mathbb{R}$ is defined as follows:

$$\begin{aligned}&H(t,x,y,z,u,k,p,q) \\ &= (k,-f(t,x,y,z,u)) + (p,b(t,x,y,z,u)) \\ &\quad + (q,\sigma(t,x,y,z,u)) + l(t,x,y,z,u).\end{aligned} \tag{25}$$

Then (24) can be rewritten in Hamiltonian system as follows:

$$\begin{aligned}dk(t) &= -H_y(t,\Theta^u(t),u(t),\Lambda^u(t))\,dt \\ &\quad - H_z(t,\Theta^u(t),u(t),\Lambda^u(t))\,dB(t), \\ dp(t) &= -H_x(t,\Theta^u(t),u(t)\,\Lambda^u(t))\,dt + q(t)\,dB(t), \\ k_0 &= -\gamma_y(y^u(0)), \\ p(t) &= -h_x^*(x^u(T))\,k(T) + \phi_x(x(T)).\end{aligned} \tag{26}$$

Lemma 12. *Let Assumption 5 be satisfied. Then there is a constant C s.t. for all control pairs $(u(\cdot);\Theta(\cdot)) = (u(\cdot);x^u(\cdot),y^u(\cdot),z^u(\cdot))$; it holds*

$$\left\|\Lambda^u(\cdot)\right\|_{\mathbb{M}^2} \le C. \tag{27}$$

Proof. Under Assumption 5, by the estimate (8), we have

$$\begin{aligned}\left\|\Lambda^u(\cdot)\right\|_{\mathbb{M}^2}^2 \le C\Bigg\{&E\int_0^T \left|l_x\left(r,\Theta^u(r),u(r)\right)\right|^2 dr\\&+E\int_0^T \left|l_y\left(r,\Theta^u(r),u(r)\right)\right|^2 dr\\&+E\int_0^T \left|l_z\left(r,\Theta^u(r),u(r)\right)\right|^2 dr\\&+E\left|\phi_x\left(x^u(T)\right)\right|^2+E\left|\gamma_y\left(y^u(0)\right)\right|^2\Bigg\}\\ \le C\left\{\left\|\Theta^u(\cdot)\right\|_{\mathbb{M}^2}^2+1\right\}\le C,\end{aligned}\tag{28}$$

where the last inequality is obtained by Lemma 10. The proof is complete. □

Assumption 13. There is a constant $C > 0$ s.t. for every $(t,\theta,u)=(t,x,y,z,u)$, $(t,\overline{\theta},\overline{u})=(t,\overline{x},\overline{y},\overline{z},\overline{u})\in[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times U$ and a.s. $\omega\in\Omega$,

$$\begin{aligned}\left|l_\alpha(t,\theta,u)-l_\alpha\left(t,\overline{\theta},\overline{u}\right)\right|&\le C\left(|\widehat{x}|+|\widehat{z}|+\left|\widehat{y}\right|+|\widehat{u}|\right),\\ \left|\phi_x(x)-\phi_x(\overline{x})\right|&\le C|\widehat{x}|,\\ \left|\gamma_y(y)-\gamma_y(\overline{y})\right|&\le C\left|\widehat{y}\right|,\end{aligned}\tag{29}$$

where $\alpha = x, y, z, u$.

Lemma 14. *Let Assumptions 5 and 13 be satisfied. Let $\Lambda^u(\cdot)=(k^u(\cdot),p^u(\cdot),q^u(\cdot))$ and $\Lambda^v(\cdot)=(k^v(\cdot),p^v(\cdot),q^v(\cdot))$ be adjoint processes corresponding to two admissible pairs $(u(\cdot);\Theta^u(\cdot))=(u(\cdot);x^u(\cdot),y^u(\cdot),z^u(\cdot))$ and $(v(\cdot);\Theta^v(\cdot))=(v(\cdot);x^v(\cdot),y^v(\cdot),z^v(\cdot))$, respectively. Then we have*

$$\left\|\Lambda^u(\cdot)-\Lambda^v(\cdot)\right\|_{\mathbb{M}^2}^2\le Cd(u(\cdot),v(\cdot))^2.\tag{30}$$

Proof. Under Assumptions 5 and 13, from the estimate (7), we have

$$\begin{aligned}&\left\|\Lambda^u(\cdot)-\Lambda^v(\cdot)\right\|_{\mathbb{M}^2}^2\\&\le C\Bigg\{E\int_0^T\Big|l_x\left(r,\Theta^u(r),u(r)\right)\\&\qquad+l_y\left(r,\Theta^u(r),u(r)\right)+l_z\left(r,\Theta^u(r),u(r)\right)\\&\qquad-l_x\left(r,\Theta^v(r),v(r)\right)-l_y\left(r,\Theta^v(r),v(r)\right)\\&\qquad-l_z\left(r,\Theta^u(r),v(r)\right)\Big|^2dr\\&\quad+E\left|\phi_x\left(x^u(T)\right)-\phi_x\left(x^v(T)\right)\right|^2\\&\quad+E\left|\gamma_x\left(y^u(0)\right)-\gamma_x\left(y^v(0)\right)\right|^2\Bigg\}\\&\le C\left\{\left\|\Theta(\cdot)-\Theta(\cdot)\right\|_{\mathbb{M}_2}^2+E\int_0^T|u(r)-v(r)|^2dr\right\}\\&\le Cd(u(\cdot),v(\cdot))^2,\end{aligned}\tag{31}$$

where the last inequality is obtained by Lemma 11 directly. □

5. A Variational Formula

The purpose of this section is to obtain a variational formula for the cost functional (10). For any two given control pairs $(\overline{u}(\cdot);\Theta^{\overline{u}}(\cdot))=(\overline{u}(\cdot);x^{\overline{u}}(\cdot),y^{\overline{u}}(\cdot),z^{\overline{u}}(\cdot))$ and $(u(\cdot);\Theta^u(\cdot))=(u(\cdot);x^u(\cdot),y^u(\cdot),z^u(\cdot))$, from the convex property of the control domain U, we can define an admissible control process as follows:

$$u^\delta(\cdot)=\overline{u}(\cdot)+\delta\left(u(\cdot)-\overline{u}(\cdot)\right),\quad 0\le\delta\le 1.\tag{32}$$

We denote the corresponding state process by $\Theta^{u^\delta}(\cdot)=(x^{u^\delta}(\cdot),y^{u^\delta}(\cdot),z^{u^\delta}(\cdot))$.

In the following, using the Hamiltonian H (see (25)) and adjoint process $\Lambda^{\overline{u}}(\cdot)=(k^{\overline{u}}(\cdot),p^{\overline{u}}(\cdot),q^{\overline{u}}(\cdot))$ associated with the admissible control pair $(\overline{u}(\cdot);\Theta^{\overline{u}}(\cdot))$, we will state and prove a presentation for the difference $J(u^\delta(\cdot))-J(\overline{u}(\cdot))$.

Lemma 15. *Let Assumption 5 be satisfied. Then we get*

$$\begin{aligned}&J\left(u^\delta(\cdot)\right)-J\left(\overline{u}(\cdot)\right)\\&=E\int_0^T\Big[H\left(r,\Theta^{u^\delta}(r),u^\delta(r),\Lambda^{\overline{u}}\right)\\&\qquad-H\left(r,\Theta^{\overline{u}}(r),\overline{u}(r),\Lambda^{\overline{u}}(r)\right)\\&\qquad-H_x\left(r,\Theta^{\overline{u}}(r),\overline{u}(r),\Lambda^{\overline{u}}(r)\right)\\&\qquad\times\left(x^{u^\delta}(r)-x^{\overline{u}}(r)\right)\\&\qquad-H_y\left(r,\Theta^{\overline{u}}(r),\overline{u}(r),\Lambda^{\overline{u}}(r)\right)\\&\qquad\times\left(y^{u^\delta}(r)-y^{\overline{u}}(r)\right)\\&\qquad-H_z\left(r,\Theta^{\overline{u}}(r),\overline{u}(r),\Lambda^{\overline{u}}(r)\right)\\&\qquad\times\left(z^{u^\delta}(r)-z^{\overline{u}}(r)\right)\Big]dr\\&\quad+E\Big[\phi\left(x^{u^\delta}(T)\right)-\phi\left(x^{\overline{u}}(T)\right)\\&\qquad-\phi_x\left(x^{\overline{u}}(T)\right)\cdot\left(x^{u^\delta}(T)-x^{\overline{u}}(T)\right)\Big]\\&\quad+E\Big[\gamma\left(y^{u^\delta}(0)\right)-\gamma\left(y^{\overline{u}}(0)\right)\\&\qquad-\gamma_y\left(y^{\overline{u}}(0)\right)\cdot\left(y^{u^\delta}(0)-y^{\overline{u}}(0)\right)\Big]\\&\quad-E\Big[\left(h\left(x^{u^\delta}(T)\right)-h\left(x^{\overline{u}}(T)\right),k^{\overline{u}}(T)\right)\\&\qquad-\left(h_x\left(x^{\overline{u}}(T)\right)\cdot\left(x^{u^\delta}(T)-x^{\overline{u}}(T)\right),k^{\overline{u}}(T)\right)\Big]\\&\triangleq\beta^{(u^\delta,\overline{u})}.\end{aligned}\tag{33}$$

Proof. Applying the definitions of $J(u(\cdot))$ and Hamilton H, we obtain

$$\begin{aligned}
&J\left(u^{\delta}(\cdot)\right)-J\left(\overline{u}(\cdot)\right)\\
&\quad=E\int_0^T\Big[H\left(r,\Theta^{u^{\delta}}(r),u^{\delta}(r),\Lambda^{\overline{u}}(r)\right)\\
&\qquad-H\left(r,\Theta^{\overline{u}}(r),\overline{u}(r),\Lambda^{\overline{u}}(r)\right)\\
&\qquad-\Big(b\left(r,\Theta^{u^{\delta}}(r),u^{\delta}(r)\right)\\
&\qquad\quad-b\left(r,\Theta^{\overline{u}}(r),\overline{u}(r)\right),p^{\overline{u}}(r)\Big)\\
&\qquad-\Big(\sigma\left(r,\Theta^{u^{\delta}}(r),u^{\delta}(r)\right)\\
&\qquad\quad-\sigma\left(r,\Theta^{\overline{u}}(r),\overline{u}(r)\right),q^{\overline{u}}(r)\Big)\\
&\qquad+\Big(f\left(r,\Theta^{u^{\delta}}(r),u^{\delta}(r)\right)\\
&\qquad\quad-f\left(r,\Theta^{\overline{u}}(r),\overline{u}(r)\right),k^{\overline{u}}(r)\Big)\Big]dr\\
&\quad+E\left[\phi\left(x^{u^{\delta}}(T)\right)-\phi\left(x^{\overline{u}}(T)\right)\right]\\
&\quad+E\left[\gamma\left(y^{u^{\delta}}(0)\right)-\gamma\left(y^{\overline{u}}(0)\right)\right].
\end{aligned}\tag{34}$$

Applying Itô formula to $(p^{\overline{u}}(r),x^{u^{\delta}}(r)-x^{\overline{u}}(r))+(k^{\overline{u}}(r),y^{u^{\delta}}(r)-y^{\overline{u}}(r))$, we have

$$\begin{aligned}
&E\int_0^T\Big[\Big(b\left(r,\Theta^{u^{\delta}}(r),u^{\delta}(r)\right)-b\left(r,\Theta^{\overline{u}}(r),\overline{u}(r)\right),p^{\overline{u}}(r)\Big)\\
&\qquad+\Big(\sigma\left(r,\Theta^{u^{\delta}}(r),u^{\delta}(r)\right)\\
&\qquad\quad-\sigma\left(r,\Theta^{\overline{u}}(r),\overline{u}(r)\right),q^{\overline{u}}(r)\Big)\\
&\qquad-\Big(f\left(r,\Theta^{u^{\delta}}(r),u^{\delta}(r)\right)\\
&\qquad\quad-f\left(r,\Theta^{\overline{u}}(r),\overline{u}(r)\right),k^{\overline{u}}(r)\Big)\Big]dr\\
&\quad=E\int_0^T\Big[H_x\left(r,\Theta^{\overline{u}}(r),\overline{u}(r),\Lambda^{\overline{u}}(r)\right)\left(x^{u^{\delta}}(r)-x^{\overline{u}}(r)\right)\\
&\qquad+H_y\left(r,\Theta^{\overline{u}}(r),\overline{u}(r),\Lambda^{\overline{u}}(r)\right)\\
&\qquad\times\left(y^{u^{\delta}}(r)-y^{\overline{u}}(r)\right)\\
&\qquad+H_z\left(r,\Theta^{\overline{u}}(r),\overline{u}(r),\Lambda^{\overline{u}}(r)\right)\\
&\qquad\times\left(z^{u^{\delta}}(r)-z^{\overline{u}}(r)\right)\Big]dr\\
&\quad+E\left[\phi_x\left(x^{\overline{u}}(T)\right)\cdot\left(x^{u^{\delta}}(T)-x^{\overline{u}}(T)\right)\right]\\
&\quad+E\left[\gamma_y\left(y^{\overline{u}}(0)\right)\cdot\left(y^{u^{\delta}}(0)-y^{\overline{u}}(0)\right)\right]\\
&\quad+E\Big[\Big(h\left(x^{u^{\delta}}(T)\right)-h\left(x^{\overline{u}}(T)\right),k^{\overline{u}}(T)\Big)\\
&\qquad-\Big(h_x\left(x^{\overline{u}}(T)\right)\cdot\left(x^{u^{\delta}}(T)-x^{\overline{u}}(T)\right),k^{\overline{u}}(T)\Big)\Big].
\end{aligned}\tag{35}$$

Now putting (35) into (34), we deduce the fact that (33) holds. The proof is complete. □

Remark 16. According to the above proof, it is easy to check that $u^{\delta}(\cdot)$ can be changed as any admissible control and need not have the form of the convex variation $u^{\delta}(\cdot)=\overline{u}(\cdot)+\delta(u(\cdot)-\overline{u}(\cdot))$.

Now we state and prove the variational formula for the cost functional (10) as follows.

Theorem 17. *Suppose that Assumption 5 holds. Let $\overline{u}(\cdot)$ be any given admissible control. Then we have*

$$\begin{aligned}
&\frac{d}{d\delta}J\left(\overline{u}(\cdot)+\delta\left(u(\cdot)-\overline{u}(\cdot)\right)\right)\Big|_{\delta=0}\\
&\quad:=\lim_{\delta\to0}\frac{J\left(\overline{u}(\cdot)+\delta\left(v(\cdot)-\overline{u}(\cdot)\right)\right)-J\left(\overline{u}(\cdot)\right)}{\delta}\\
&\quad=E\int_0^T\left(H_u\left(r,\Theta^{\overline{u}}(r),\overline{u}(r),\Lambda^{\overline{u}}(r)\right),u(r)-\overline{u}(r)\right)dr,
\end{aligned}\tag{36}$$

where $u(\cdot)$ is any given admissible control and $\delta>0$.

Proof. Define $u^{\delta}(\cdot)=\overline{u}(\cdot)+\delta(v(\cdot)-\overline{u}(\cdot))$; by Lemma 15, we have

$$\begin{aligned}
&J\left(u^{\delta}(\cdot)\right)-J\left(\overline{u}(\cdot)\right)\\
&\quad=J\left(\overline{u}(\cdot)+\delta\left(u(\cdot)-\overline{u}(\cdot)\right)\right)-J\left(\overline{u}(\cdot)\right)\\
&\quad=E\int_0^T H_u\left(r,\Theta^{\overline{u}}(r),\overline{u}(r),\Lambda^{\overline{u}}(r)\right)\left(u^{\delta}(r)-\overline{u}(r)\right)dr\\
&\qquad+\beta^{(u^{\delta},\overline{u})}-E\int_0^T H_u\left(r,\Theta^{\overline{u}}(r),\overline{u}(r),\Lambda^{\overline{u}}(r)\right)\\
&\qquad\qquad\times\left(u^{\delta}(r)-\overline{u}(r)\right)dr\\
&\quad=\delta E\int_0^T H_u\left(r,\Theta^{\overline{u}}(r),\overline{u}(r),\Lambda^{\overline{u}}(r)\right)\left(v(r)-\overline{u}(r)\right)dr\\
&\qquad+\beta^{(u^{\delta},\overline{u})}-E\int_0^T H_u\left(r,\Theta^{\overline{u}}(r),\overline{u}(r),\Lambda^{\overline{u}}(r)\right)\\
&\qquad\qquad\times\left(u^{\delta}(r)-\overline{u}(r)\right)dr.
\end{aligned}\tag{37}$$

Applying Lemma 11 and Assumption 5, we get

$$\begin{aligned}
&\beta^{(u^{\delta},\overline{u})}-E\int_0^T H_u\left(r,\Theta^{\overline{u}}(r),\overline{u}(r),\Lambda^{\overline{u}}(r)\right)\\
&\qquad\times\left(u^{\delta}(r)-\overline{u}(r)\right)dr=o(\delta).
\end{aligned}\tag{38}$$

Hence, by (38) and (37), we get

$$\lim_{\delta\to 0}\frac{J\left(u^{\delta}(\cdot)\right)-J\left(\overline{u}(\cdot)\right)}{\delta} = E\int_0^T \left(H_u\left(r,\Theta^{\overline{u}}(r),\overline{u}(r),\Lambda^{\overline{u}}(r)\right),v(r)-\overline{u}(r)\right)dr. \tag{39}$$

The proof is complete. □

6. Necessary Conditions for Near-Optimality

In this section, we will state and prove our main results, that is, the stochastic maximum principle of the near-optimal control of Problem *PRO*. Moreover, we give the additional assumption as follows.

Assumption 18. There is a constant $C > 0$ s.t. for all $(t,x,y,z,\overline{x},\overline{y},\overline{z},u,\overline{u})$ and a.s. $\omega\in\Omega$,

$$\left|l(t,x,z,y,u)-l(t,\overline{x},\overline{z},\overline{y},u)\right| \le C\left(|\widehat{x}|+|\widehat{z}|+\left|\widehat{y}\right|+|\widehat{u}|\right). \tag{40}$$

Theorem 19. *Suppose that Assumptions 5 and 13 hold. Let $(u^{\varepsilon}(\cdot);\Theta^{\varepsilon}(\cdot)) = (u^{\varepsilon}(\cdot);x^{\varepsilon}(\cdot),y^{\varepsilon}(\cdot),z^{\varepsilon}(\cdot))$ be ε-optimal pair of problem PRO. Then, for any given $\varepsilon > 0$, there is a positive constant C s.t.:*

$$H_u\left(r,\Theta^{\varepsilon}(r),u^{\varepsilon}(r),\Lambda^{\epsilon}(r)\right)\cdot\left(u-u^{\varepsilon}(r)\right)\ge -C\varepsilon^{1/2},\quad \forall u\in U,\ a.e.\ (r,\omega)\in[0,T]\times\Omega, \tag{41}$$

where $\Lambda^{\epsilon}(\cdot) = (k^{\epsilon}(\cdot),p^{\epsilon}(\cdot),q^{\epsilon}(\cdot))$ is the adjoint process corresponding to $(u^{\varepsilon}(\cdot);\Theta^{\varepsilon}(\cdot))$.

Proof. By Lemma 11 and Assumption 13, we can deduce the fact that $J(u(\cdot))$ is continuous on $\mathscr{A}$ with respect to the metric (18). Using Ekeland's variational principle (see [16]) with $\delta=\varepsilon^{1/2}$, there exists an admissible pair $(\overline{u}^{\varepsilon}(\cdot);\overline{\Theta}^{\varepsilon}(\cdot)) = (\overline{u}^{\varepsilon}(\cdot);\overline{x}^{\varepsilon}(\cdot),\overline{y}^{\varepsilon}(\cdot),\overline{z}^{\varepsilon}(\cdot))$ such that

$$d\left(u^{\varepsilon}(\cdot),\overline{u}^{\varepsilon}(\cdot)\right)\le\varepsilon^{1/2}, \tag{42}$$

$$J(u(\cdot))-J\left(\overline{u}^{\varepsilon}(\cdot)\right)\ge-\varepsilon^{1/2}d\left(u(\cdot),\overline{u}^{\varepsilon}(\cdot)\right),\quad\forall u(\cdot)\in\mathscr{A}. \tag{43}$$

Now we define a convex perturbed control $u^{\varepsilon,h}(\cdot)$ of $\overline{u}^{\varepsilon}(\cdot)$ as

$$u^{\varepsilon,h}(\cdot)=\overline{u}^{\varepsilon}(\cdot)+h\left(\overline{u}^{\varepsilon}(\cdot)-u(\cdot)\right), \tag{44}$$

where $u(\cdot)\in\mathscr{A}$ is an arbitrary given admissible control and $0\le h\le 1$.

Then by the variational formula (36), (43), and the fact that

$$d\left(u^{\varepsilon,h}(\cdot),\overline{u}^{\varepsilon}(\cdot)\right)\le Ch, \tag{45}$$

we have

$$\begin{aligned}
&E\int_0^T H_u\left(t,\overline{\Theta}^{\varepsilon}(r),\overline{u}^{\varepsilon}(r),\overline{\Lambda}^{\epsilon}(r)\right)\cdot\left(u(r)-\overline{u}^{\varepsilon}(t)\right)dt\\
&=\lim_{\varepsilon\to0^+}\frac{J\left(u^{\varepsilon,h}(\cdot)\right)-J\left(\overline{u}^{\varepsilon}(\cdot)\right)}{h}\\
&\ge\lim_{\varepsilon\to0^+}\frac{-\varepsilon^{1/2}d\left(u^{\varepsilon,h}(\cdot),\overline{u}^{\varepsilon}(\cdot)\right)}{h}\\
&\ge-C\varepsilon^{1/2},
\end{aligned} \tag{46}$$

where $\overline{\Lambda}^{\epsilon} = (\overline{p}^{\varepsilon}(\cdot),\overline{q}^{\varepsilon}(\cdot),\overline{k}^{\epsilon}(\cdot))$ is the adjoint process corresponding to $(\overline{u}^{\varepsilon}(\cdot);\overline{\Theta}^{\varepsilon}(\cdot))$.

Now in order to obtain the optimal condition (41), we now have to estimate the following formula:

$$\begin{aligned}
I^{\varepsilon}:=&E\int_0^T H_u\left(t,\Theta^{\varepsilon}(r),u^{\epsilon}(r),\Lambda^{\epsilon}(r)\right)\cdot\left(u(r)-u^{\varepsilon}(r)\right)dr\\
&-E\int_0^T H_u\left(t,\overline{\Theta}^{\epsilon}(r),\overline{u}^{\epsilon}(r),\overline{\Lambda}^{\epsilon}(r)\right)\cdot\left(u(r)-\overline{u}^{\varepsilon}(r)\right)dr.
\end{aligned} \tag{47}$$

First, by adding and subtracting $E\int_0^T H_u(t,\Theta^{\epsilon}(r),u^{\epsilon}(r),\Lambda^{\epsilon}(r))\cdot(u(r)-\overline{u}^{\varepsilon}(r))dr$, we have

$$\begin{aligned}
I^{\varepsilon}=&E\int_0^T H_u\left(t,\Theta^{\varepsilon}(r),u^{\epsilon}(r),\Lambda^{\epsilon}(r)\right)\cdot\left(\overline{u}^{\varepsilon}(r)-u^{\varepsilon}(r)\right)dr\\
&+E\int_0^T\Big(H_u\left(r,\Theta^{\varepsilon}(r),u^{\epsilon}(r),\Lambda^{\epsilon}(r)\right)\\
&\qquad-H_u\left(t,\overline{\Theta}^{\varepsilon}(r),\overline{u}^{\epsilon}(r),\overline{\Lambda}^{\epsilon}(r)\right)\Big)\\
&\qquad\cdot\left(u(r)-\overline{u}^{\varepsilon}(r)\right)dr\\
=&I_1^{\varepsilon}+I_2^{\varepsilon}.
\end{aligned} \tag{48}$$

Next, using Lemmas 11 and 14 and (42), we have

$$\begin{aligned}
\left|I_2^{\varepsilon}\right|\le&CE\int_0^T\Big|H_u\left(t,\Theta^{\varepsilon}(r),u^{\epsilon}(r),\Lambda^{\epsilon}(r)\right)\\
&\qquad-H_u\left(t,\overline{\Theta}^{\varepsilon}(r),\overline{u}^{\epsilon}(r),\overline{\Lambda}^{\epsilon}(r)\right)\Big|dr\\
\le&CE\int_0^T\Big|H_u\left(t,\Theta^{\varepsilon}(r),u^{\epsilon}(r),\Lambda^{\epsilon}(r)\right)\\
&\qquad-H_u\left(t,\overline{\Theta}^{\varepsilon}(r),\overline{u}^{\epsilon}(r),\Lambda^{\epsilon}(r)\right)\Big|dr\\
&+CE\int_0^T\Big|H_u\left(t,\overline{\Theta}^{\varepsilon}(r),\overline{u}^{\epsilon}(r),\Lambda^{\epsilon}(r)\right)\\
&\qquad-H_u\left(t,\overline{\Theta}^{\varepsilon}(r),\overline{u}^{\epsilon}(r),\overline{\Lambda}^{\epsilon}(r)\right)\Big|dr
\end{aligned}$$

$$\begin{aligned}
&\le C\left(E\sup_{0\le r\le T}|\overline{x}^{\varepsilon}(r)-x^{\varepsilon}(r)|^{2}\right)^{1/2}\\
&\quad+C\left(E\sup_{0\le r\le T}|\overline{p}^{\varepsilon}(r)-p^{\varepsilon}(r)|^{2}\right)^{1/2}\\
&\quad+C\left(E\int_{0}^{T}|\overline{q}^{\varepsilon}(r)-q^{\varepsilon}(r)|^{2}dr\right)^{1/2}\\
&\quad+C\left(E\int_{0}^{T}\left|\overline{u}^{\varepsilon}(r)-u^{\varepsilon}(r)\right|^{2}dr\right)^{1/2}\\
&\quad+\left(E\sup_{0\le r\le T}|\overline{k}^{\varepsilon}(r)-k^{\varepsilon}(r)|^{2}\right)^{1/2}\\
&\le C\left(E\int_{0}^{T}\left|\overline{u}^{\varepsilon}(r)-u^{\varepsilon}(r)\right|^{2}dr\right)^{1/2}\\
&\le C\varepsilon^{1/2}.
\end{aligned}\tag{49}$$

Then, combining Schwarz's inequality and Lemmas 10 and 12 and (42), we have

$$\begin{aligned}
\left|I_{1}^{\varepsilon}\right|&\le E\left(\int_{0}^{T}\left|H_{u}\left(t,\Theta^{\varepsilon}(r),u^{\varepsilon}(r),\Lambda^{\epsilon}(r)\right)\right|^{2}dr\right)^{1/2}\\
&\quad\times\left(E\int_{0}^{T}\left|\overline{u}^{\varepsilon}(r)-u^{\varepsilon}(r)\right|^{2}dr\right)^{1/2}\\
&\le C\left(E\int_{0}^{T}\left|\overline{u}^{\varepsilon}(t)-u^{\varepsilon}(r)\right|^{2}dr\right)^{1/2}\\
&\le C\varepsilon^{1/2}.
\end{aligned}\tag{50}$$

Therefore, combining (46), (47), (49), and (50), we have

$$\begin{aligned}
&E\int_{0}^{T}H_{u}\left(r,\Theta^{\epsilon}(r),u^{\varepsilon}(t),\Lambda^{\epsilon}(r)\right)\cdot\left(u(r)-u^{\varepsilon}(r)\right)dr\\
&\quad=E\int_{0}^{T}H_{u}\left(t,\overline{\Theta}^{\epsilon}(r),\overline{u}^{\varepsilon}(r),\overline{\Lambda}^{\epsilon}(r)\right)\\
&\qquad\cdot\left(u(r)-\overline{u}^{\varepsilon}(r)\right)dr+I^{\varepsilon}\\
&\quad=E\int_{0}^{T}H_{u}\left(r,\overline{\Theta}^{\epsilon}(r),\overline{u}^{\varepsilon}(r),\overline{\Lambda}^{\epsilon}(r)\right)\\
&\qquad\cdot\left(u(r)-\overline{u}^{\varepsilon}(r)\right)dr+I_{1}^{\varepsilon}+I_{2}^{\varepsilon}\\
&\quad\ge-C\varepsilon^{1/2}
\end{aligned}\tag{51}$$

which implies that (41) holds. The proof is complete. □

7. Sufficient Optimality Conditions

In this section, we will show that, under certain convex conditions, the near-maximum condition of the Hamiltonian function in the integral form is sufficient for near-optimality.

Theorem 20. *Under Assumption 5, let* $(u^{\varepsilon}(\cdot);\Theta^{\epsilon}(\cdot))=(u^{\varepsilon}(\cdot);x^{\epsilon}(\cdot),y^{\epsilon}(\cdot),z^{\epsilon}(\cdot))$ *be an admissible pair with* $\overline{y}(T)=Mx^{\varepsilon}(T)$, $M\in L^{2}(\Omega,\mathscr{F}_{T},P;\mathbb{R}^{m\times n})$. *Let* $\Lambda^{\epsilon}(\cdot)=(p^{\epsilon}(\cdot),q^{\epsilon}(\cdot),k^{\varepsilon}(\cdot))$ *be the adjoint process associated with* $(u^{\varepsilon}(\cdot);\Theta^{\epsilon}(\cdot))$. *Assume that for almost all* $(t,\omega)\in[0,T]\times\Omega$, $H(t,x,y,z,u,\Lambda^{\epsilon}(t))$ *is convex in* (x,y,z,u), $\gamma(y)$ *is convex in* y, *and* $\phi(x)$ *is convex in* x, *respectively, and for some* ε, *the optimality conditions,*

$$\begin{aligned}
&E\int_{0}^{T}H\left(r,\Theta^{\epsilon}(r),u^{\varepsilon}(r),\Lambda^{\varepsilon}(r)\right)dr\\
&\quad\le E\inf_{u(\cdot)\in\mathscr{A}}\int_{0}^{T}H\left(r,\Theta^{\epsilon}(r),u(r),\Lambda^{\varepsilon}(r)\right)dr+\varepsilon,
\end{aligned}\tag{52}$$

hold. Then

$$J\left(u^{\varepsilon}(\cdot)\right)\le\inf_{u(\cdot)\in\mathscr{A}}J(u(\cdot))+C_{1}\varepsilon^{1/2},\tag{53}$$

where C_1 *is a constant independent of* ε.

Proof. In the following, C_1 is a constant which may change from line to line and is independent of ε.

According to Lemma 15, we deduce the fact that

$$\begin{aligned}
&J(u(\cdot))-J\left(u^{\varepsilon}(\cdot)\right)\\
&\quad=E\int_{0}^{T}\left[H\left(t,\Theta^{u}(r),u(r),\Lambda^{\varepsilon}(r)\right)\right.\\
&\qquad-H\left(r,\Theta^{\epsilon}(r),u^{\varepsilon}(r),\Lambda^{\varepsilon}(r)\right)\\
&\qquad-H_{x}\left(r,\Theta^{\epsilon}(r),u^{\varepsilon}(r),\Lambda^{\varepsilon}(r)\right)\\
&\qquad\times\left(x^{u}(r)-x^{\epsilon}(r)\right)\\
&\qquad-H_{y}\left(r,\Theta^{\epsilon}(r),u^{\varepsilon}(r),\Lambda^{\varepsilon}(r)\right)\\
&\qquad\times\left(y^{u}(r)-y^{\epsilon}(r)\right)\\
&\qquad-H_{z}\left(r,\Theta^{\epsilon}(r),u^{\varepsilon}(r),\Lambda^{\varepsilon}(r)\right)\\
&\qquad\left.\times\left(z^{u}(r)-z^{\epsilon}(r)\right)\right]dr\\
&\quad+E\left[\phi\left(x^{u}(T)\right)-\phi\left(x^{\epsilon}(T)\right)\right.\\
&\qquad\left.-\Phi_{x}\left(x^{\epsilon}(T)\right)\cdot\left(x^{u}(R)-x^{\epsilon}(T)\right)\right]\\
&\quad+E\left[\gamma\left(y^{u}(0)\right)-\gamma\left(y^{\epsilon}(0)\right)\right.\\
&\qquad\left.-\gamma_{y}\left(y^{\epsilon}(0)\right)\cdot\left(y^{u}(0)-y^{\epsilon}(0)\right)\right],
\end{aligned}\tag{54}$$

where $(u(\cdot);x^{u}(\cdot),y^{u}(\cdot),z^{u}(\cdot))$ are any given admissible control pairs. By the convexity of H, ϕ, and γ, we have

$$\begin{aligned}
&H\left(t,\Theta^{u}(r),u(r),\Lambda^{\varepsilon}(r)\right)\\
&\qquad-H\left(r,\Theta^{\epsilon}(r),u^{\varepsilon}(r),\Lambda^{\varepsilon}(r)\right)\\
&\quad\ge H_{x}\left(r,\Theta^{\epsilon}(r),u^{\varepsilon}(r),\Lambda^{\varepsilon}(r)\right)\left(x^{u}(r)-x^{\epsilon}(r)\right)
\end{aligned}$$

$$
\begin{aligned}
&+ H_y\left(r, \Theta^\varepsilon(r), u^\varepsilon(r), \Lambda^\varepsilon(r)\right)\left(y^u(r) - y^\epsilon(r)\right)\\
&+ H_z\left(r, \Theta^\varepsilon(r), u^\varepsilon(r), \Lambda^\varepsilon(r)\right)\left(z^u(r) - z^\epsilon(r)\right)\\
&+ H_u\left(r, \Theta^\varepsilon(r), u^\varepsilon(r), \Lambda^\varepsilon(r)\right)\left(u(r) - u^\varepsilon(r)\right),\\
E\big[&\phi\left(x^u(T)\right) - \phi\left(x^\epsilon(T)\right)\\
&-\phi_x\left(x^\epsilon(T)\right)\cdot\left(x^u(T) - x^\epsilon(T)\right)\big]\\
&+ E\big[\gamma\left(y^u(0)\right) - \gamma\left(y^\epsilon(0)\right)\\
&\quad -\gamma_y\left(y^u(0)\right)\cdot\left(y^u(0) - y^\epsilon(0)\right)\big] \geq 0.
\end{aligned}
\tag{55}
$$

Putting (55) into (54), we have

$$
\begin{aligned}
&J\left(u^\varepsilon(\cdot)\right) - J\left(u(\cdot)\right)\\
&\leq -E\int_0^T H_u\left(r, \Theta^\varepsilon(r), u^\varepsilon(r), \Lambda^\varepsilon(r)\right)\left(u(r) - u^\varepsilon(r)\right)dr.
\end{aligned}
\tag{56}
$$

Therefore, the rest of the proof is only to estimate the term $H_u(r, \Theta^\varepsilon(r), u^\varepsilon(r), \Lambda^\varepsilon(r))(u(r)-u^\varepsilon(r))$. To this end, for a given $\varepsilon > 0$, let us introduce a new metric $\widetilde{d}$ on $\mathscr{A}$ as follows:

$$
\widetilde{d}\left(u(\cdot), u'(\cdot)\right) = E\int_0^T \nu^\varepsilon(r)\left|u(r) - u'(r)\right|dr, \tag{57}
$$

where

$$
\nu^\varepsilon(r) = 1 + \left|p^\varepsilon(r)\right| + \left|q^\varepsilon(r)\right| + \left|k^\varepsilon(r)\right| \geq 1. \tag{58}
$$

Now on $\mathscr{A}$ we define a new functional F by

$$
F\left(u(\cdot)\right) = E\int_0^T H\left(r, \Theta^\varepsilon(r), u(r), \Lambda^\varepsilon(r)\right)dr. \tag{59}
$$

It is easy to check that

$$
\left|F\left(u(\cdot)\right) - F\left(u'(\cdot)\right)\right| \leq CE\int_0^T \nu^\varepsilon(r)\left|u(r) - u'(r)\right|dr. \tag{60}
$$

Therefore F is continuous on $\mathscr{A}$ with respect to metric $\widetilde{d}$. Using (52) and Ekeland's variational principle, we can find an admissible control $\overline{u}^\varepsilon(\cdot) \in \mathscr{A}$ such that

$$
\widetilde{d}\left(u^\varepsilon(\cdot), \overline{u}^\varepsilon(\cdot)\right) \leq \varepsilon^{1/2}, \tag{61}
$$

$$
\begin{aligned}
&E\int_0^T \widetilde{H}\left(r, \Theta^\varepsilon(r), \overline{u}^\epsilon(r), \Lambda^\varepsilon(r)\right)\\
&\quad = \max_{u(\cdot)\in\mathscr{A}} E\int_0^T \widetilde{H}\left(t, \Theta^\varepsilon(r), u(r), \Lambda^\varepsilon(r)\right),
\end{aligned}
\tag{62}
$$

where

$$
\begin{aligned}
&\widetilde{H}\left(t, \Theta^\varepsilon(r), u(r), \Lambda^\varepsilon(r)\right)\\
&\quad = H\left(t, \Theta^\varepsilon(r), u(r), \Lambda^\varepsilon(r)\right) - \varepsilon^{1/2}\nu^\varepsilon(r)\left|u(r) - \overline{u}^\varepsilon(r)\right|.
\end{aligned}
\tag{63}
$$

By standard methods, the maximum condition (52) implies that

$$
\begin{aligned}
&\widetilde{H}\left(t, \Theta^\varepsilon(r), \overline{u}^\epsilon(r), \Lambda^\varepsilon(r)\right)\\
&\quad = \max_{u(\cdot)\in\mathscr{A}} \widetilde{H}\left(t, \Theta^\varepsilon(r), u(r), \Lambda^\varepsilon(r)\right),\\
&\qquad \text{a.e. } (r, \omega) \in [0, T]\times\Omega.
\end{aligned}
\tag{64}
$$

Applying Proposition 2.3.2 in [30], we have

$$
\begin{aligned}
0 &\in \partial_u H\left(t, \Theta^\varepsilon(r), \overline{u}^\epsilon(r), \Lambda^\varepsilon(r)\right)\\
&\subset \partial_u H\left(t, \Theta^\varepsilon(r), \overline{u}^\epsilon(r), \Lambda^\varepsilon(r)\right)\\
&\quad + \left[-\varepsilon^{1/2}\nu^\varepsilon(r), \varepsilon^{1/2}\nu^\varepsilon(r)\right],
\end{aligned}
\tag{65}
$$

which implies that there exists $\beta^\varepsilon(r) \in [-\varepsilon^{1/2}\nu^\varepsilon(r), \varepsilon^{1/2}\nu^\varepsilon(r)]$ such that

$$
H_u\left(r, \Theta^\varepsilon(r), \overline{u}^\epsilon(r), \Lambda^\varepsilon(r)\right) + \beta^\varepsilon(r) = 0. \tag{66}
$$

Therefore, under Assumptions 13 and 5,

$$
\begin{aligned}
&\left|H_u\left(r, \Theta^\varepsilon(r), u^\epsilon(r), \Lambda^\varepsilon(r)\right)\right.\\
&\quad \left. - H_u\left(r, \Theta^\varepsilon(r), \overline{u}^\epsilon(r), \Lambda^\varepsilon(r)\right)\right|\\
&\quad + \left|H_u\left(r, \Theta^\varepsilon(r), \overline{u}^\epsilon(r), \Lambda^\varepsilon(r)\right)\right|\\
&\leq C\left(1 + \left|p^\varepsilon(r)\right| + \left|q^\varepsilon(r)\right| + \left|k^\varepsilon(r)\right|\right)\\
&\quad \times \left|u^\varepsilon(t) - \overline{u}^\varepsilon(t)\right| + \varepsilon^{1/2}\nu^\varepsilon\\
&\leq C\left(1 + \left|p^\varepsilon(r)\right| + \left|q^\varepsilon(r)\right| + \left|k^\varepsilon(r)\right|\right)\\
&\quad \times \left(\left|u^\varepsilon(r) - \overline{u}^\varepsilon(r)\right| + \varepsilon^{1/2}\right).
\end{aligned}
\tag{67}
$$

Then, applying Holder's inequality and Lemma 12 and (61), we deduce

$$
\begin{aligned}
&E\int_0^T \left|H_u\left(t, \Theta^\varepsilon(r), u^\epsilon(r), \Lambda^\varepsilon(r)\right)\right|dr\\
&\quad \leq d\left(u^\varepsilon(\cdot), \overline{u}^\varepsilon(\cdot)\right) + C\varepsilon^{1/2}\\
&\quad \leq C\varepsilon^{1/2}.
\end{aligned}
\tag{68}
$$

By (56) and (68), we get

$$
J\left(u^\varepsilon(\cdot)\right) \leq J\left(u(\cdot)\right) + C\varepsilon^{1/2}. \tag{69}
$$

Since $u(\cdot)$ is arbitrary, $u(\cdot)$ is a near-optimal control with order $\varepsilon^{1/2}$. □

8. Conclusion

This paper is the near-optimal control problem for a stochastic system driven by fully coupled FBSDEs. Stochastic maximum principle and verification theory of the near-optimal control are obtained. The control variable appears in both drift and diffusion coefficients of the FBSDEs. The control domain is assumed to be convex. The reviewers suggest that the data-driven control has extensive applications in industry and finance (see, e.g., [31–33] and the references therein) and the model discussed in this present paper may has the potential to achieve more practical oriented results under data-driven framework. Some investigations on this topic will be studied and carried out in our future publications.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

Financial supports by the National Natural Science Foundation of China (Grant nos. 11101140 and 11301177), China Postdoctoral Science Foundation (2011M500721 and 2012T50391), and the Natural Science Foundation of Zhejiang (Grant nos. Y6110775 and Y6110789) are gratefully acknowledged.

References

[1] J.-M. Bismut, "Conjugate convex functions in optimal stochastic control," *Journal of Mathematical Analysis and Applications*, vol. 44, no. 3, pp. 384–404, 1973.

[2] É. Pardoux and S. G. Peng, "Adapted solution of a backward stochastic differential equation," *Systems & Control Letters*, vol. 14, no. 1, pp. 55–61, 1990.

[3] S. Peng, "Backward stochastic differential equations and applications to optimal control," *Applied Mathematics and Optimization*, vol. 27, no. 2, pp. 125–144, 1993.

[4] W. S. Xu, "Stochastic maximum principle for optimal control problem of forward and backward system," *Journal of the Australian Mathematical Society B*, vol. 37, no. 2, pp. 172–185, 1995.

[5] Z. Wu, "Maximum principle for optimal control problem of fully coupled forward-backward stochastic systems," *Systems Science and Mathematical Sciences*, vol. 11, no. 3, pp. 249–259, 1998.

[6] J. T. Shi and Z. Wu, "Maximum principle for partially-observed optimal control of fully-coupled forward-backward stochastic systems," *Journal of Optimization Theory and Applications*, vol. 145, no. 3, pp. 543–578, 2010.

[7] J.-T. Shi and Z. Wu, "The maximum principle for fully coupled forward-backward stochastic control system," *Acta Automatica Sinica*, vol. 32, no. 2, pp. 161–169, 2006.

[8] G. Wang and Z. Wu, "The maximum principles for stochastic recursive optimal control problems under partial information," *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1230–1242, 2009.

[9] G. Wang, Z. Wu, and J. Xiong, "Maximum principles for forward-backward stochastic control systems with correlated state and observation noises," *SIAM Journal on Control and Optimization*, vol. 51, no. 1, pp. 491–524, 2013.

[10] S. Ji and Q. Wei, "A maximum principle for fully coupled forward-backward stochastic control systems with terminal state constraints," *Journal of Mathematical Analysis and Applications*, vol. 407, no. 2, pp. 200–210, 2013.

[11] Q. Meng, "A maximum principle for optimal control problem of fully coupled forward-backward stochastic systems with partial information," *Science in China A*, vol. 52, no. 7, pp. 1579–1588, 2009.

[12] M. Tang and Q. Meng, "Stochastic differential games of fully coupled forward-backward stochastic systems under partial information," in *Proceedings of the 29th Chinese Control Conference (CCC '10)*, pp. 1150–1155, IEEE, July 2010.

[13] X. Y. Zhou, "Deterministic near-optimal control. I. Necessary and sufficient conditions for near-optimality," *Journal of Optimization Theory and Applications*, vol. 85, no. 2, pp. 473–488, 1995.

[14] X. Y. Zhou, "Deterministic near-optimal controls. II. Dynamic programming and viscosity solution approach," *Mathematics of Operations Research*, vol. 21, no. 3, pp. 655–674, 1996.

[15] I. Ekeland, "On the variational principle," *Journal of Mathematical Analysis and Applications*, vol. 47, pp. 324–353, 1974.

[16] X. Y. Zhou, "Stochastic near-optimal controls: necessary and sufficient conditions for near-optimality," *SIAM Journal on Control and Optimization*, vol. 36, no. 3, pp. 929–947, 1998.

[17] F. Chighoub and B. Mezerdi, "Near optimality conditions in stochastic control of jump diffusion processes," *Systems & Control Letters*, vol. 60, no. 11, pp. 907–916, 2011.

[18] M. Hafayed, P. Veverka, and S. Abbas, "On maximum principle of near-optimality for diffusions with jumps, with application to consumption-investment problem," *Differential Equations and Dynamical Systems*, vol. 20, no. 2, pp. 111–125, 2012.

[19] M. Hafayed and S. Abbas, "On near-optimal mean-field stochastic singular controls: necessary and sufficient conditions for near-optimality," *Journal of Optimization Theory and Applications*, 2013.

[20] M. Hafayed, S. Abbas, and P. Veverka, "On necessary and sufficient conditions for near-optimal singular stochastic controls," *Optimization Letters*, vol. 7, no. 5, pp. 949–966, 2013.

[21] J. Huang, X. Li, and G. Wang, "Near-optimal control problems for linear forward-backward stochastic systems," *Automatica*, vol. 46, no. 2, pp. 397–404, 2010.

[22] K. Bahlali, N. Khelfallah, and B. Mezerdi, "Necessary and sufficient conditions for near-optimality in stochastic control of FBSDEs," *Systems & Control Letters*, vol. 58, no. 12, pp. 857–864, 2009.

[23] E. Hui, J. Huang, X. Li, and G. Wang, "Near-optimal control for stochastic recursive problems," *Systems & Control Letters*, vol. 60, no. 3, pp. 161–168, 2011.

[24] L. Zhang, J. Huang, and X. Li, "Stochastic maximum principle for nearoptimal control of fbsdes," 2012, http://arxiv.org/pdf/1203.1774.pdf.

[25] J. Yong, "Optimality variational principle for controlled forward-backward stochastic differential equations with mixed initial-terminal conditions," *SIAM Journal on Control and Optimization*, vol. 48, no. 6, pp. 4119–4156, 2010.

[26] Z. Wu, "A general maximum principle for optimal control of forward-backward stochastic systems," *Automatica*, vol. 49, no. 5, pp. 1473–1480, 2013.

[27] S. Peng and Z. Wu, "Fully coupled forward-backward stochastic differential equations and applications to optimal control," *SIAM Journal on Control and Optimization*, vol. 37, no. 3, pp. 825–843, 1999.

[28] Q. Lin, "Optimal control of coupled forward-backward stochastic system with jumps and related hamilton-jacobi-bellman equations," 2011, http://arxiv.org/pdf/1111.4642.pdf.

[29] J. Yong and X. Y. Zhou, *Stochastic Controls*, vol. 43 of *Applications of Mathematics*, Springer, New York, NY, USA, 1999.

[30] F. H. Clarke, *Optimization and Nonsmooth Analysis*, vol. 5 of *Classics in Applied Mathematics*, SIAM, Philadelphia, Pa, USA, Second edition, 1990.

[31] S. Yin, S. X. Ding, A. H. A. Sari, and H. Hao, "Data-driven monitoring for stochastic systems and its application on batch process," *International Journal of Systems Science*, vol. 44, no. 7, pp. 1366–1376, 2013.

[32] S. Yin, S. X. Ding, A. Haghani, H. Hao, and P. Zhang, "A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark tennessee eastman process," *Journal of Process Control*, vol. 22, no. 9, pp. 1567–1581, 2012.

[33] S. Yin, H. Luo, and S. Ding, "Real-time implementation of fault-tolerant control systems with performance optimization," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 5, pp. 2402–2411, 2013.