Hindawi Publishing Corporation
International Journal of Stochastic Analysis
Volume 2012, Article ID 258674, 50 pages
doi:10.1155/2012/258674

Research Article

Necessary Conditions for Optimal Control of Forward-Backward Stochastic Systems with Random Jumps

Jingtao Shi

School of Mathematics, Shandong University, Jinan 250100, China

Correspondence should be addressed to Jingtao Shi, shijingtao@sdu.edu.cn

Received 28 September 2011; Revised 28 December 2011; Accepted 3 January 2012

Academic Editor: Jiongmin Yong

This paper deals with the general optimal control problem for fully coupled forward-backward stochastic differential equations with random jumps (FBSDEJs). The control domain is not assumed to be convex, and the control variable appears in both diffusion and jump coefficients of the forward equation. Necessary conditions of Pontraygin's type for the optimal controls are derived by means of spike variation technique and Ekeland variational principle. A linear quadratic stochastic optimal control problem is discussed as an illustrating example.

1. Introduction

1.1. Basic Notations

Throughout this paper, we denote by $\mathbf{R}^n$ the space of n-dimensional Euclidean space, by $\mathbf{R}^{n\times d}$ the space of $n \times d$ matrices, and by $\mathcal{S}^n$ the space of $n \times n$ symmetric matrices. $\langle \cdot, \cdot \rangle$ and $|\cdot|$ denote the scalar product and norm in the Euclidean space, respectively. $\top$ appearing in the superscripts denotes the transpose of a matrix.

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$ be a complete filtered probability space satisfying the usual conditions, where the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is generated by the following two mutually independent processes:

(i) a one-dimensional standard Brownian motion $\{W(t)\}_{t\geq 0}$;

(ii) a Poisson random measure N on $\mathbf{E} \times \mathbf{R}_+$, where $\mathbf{E} \subset \mathbf{R}$ is a nonempty open set equipped with its Borel field $\mathcal{B}(\mathbf{E})$, with compensator $\widehat{N}(dedt) = \pi(de)dt$, such that $\widetilde{N}(A \times [0,t]) = (N - \widehat{N})(A \times [0,t])_{t\geq 0}$ is a martingale for all $A \in \mathcal{B}(\mathbf{E})$ satisfying

$\pi(A) < \infty$. π is assumed to be a σ-finite measure on $(\mathbf{E}, \mathcal{B}(\mathbf{E}))$ and is called the characteristic measure.

Let $T > 0$ be fixed and $\mathbf{U}$ be a nonempty subset of $\mathbf{R}^k$. We denote $\mathbb{M} = \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^m \times \mathbf{R}^m$. Any generic point in $\mathbb{M}$ is denoted by $\theta \equiv (x, y, z, c)$. Let $\mathcal{U}[0,T]$ be the set of all $\mathcal{F}_t$-predictable processes $u : [0,T] \times \Omega \to \mathbf{U}$ such that $\sup_{0\leq t\leq T} \mathbb{E}|u(t)|^i < +\infty$, for all $i = 1, 2, \ldots$. Any $u(\cdot) \in \mathcal{U}[0,T]$ is called an admissible control process. We denote by $\mathcal{L}^2(\mathbf{E}, \mathcal{B}(\mathbf{E}), \pi; \mathbf{R}^m)$ or $\mathcal{L}^2$ the set of integrable functions $c : \mathbf{E} \to \mathbf{R}^m$ with norm $\|c(e)\|^2_{\mathcal{L}^2} := \int_{\mathbf{E}} |c(e)|^2 \pi(de) < \infty$. We define

$$
\begin{aligned}
L^2_{\mathcal{F}_T}(\Omega; \mathbf{R}^m) &:= \left\{ \xi : \Omega \longrightarrow \mathbf{R}^m \mid \xi \text{ is } \mathcal{F}_T\text{-measurable}, \mathbb{E}|\xi|^2 < \infty \right\}, \\
L^2_{\mathcal{F}}([0,T]; \mathbf{R}^m) &:= \left\{ \varphi : [0,T] \times \Omega \longrightarrow \mathbf{R}^m \mid \varphi \text{ is } \mathcal{F}_t\text{-adapted}, \mathbb{E}\int_0^T \left|\varphi(t)\right|^2 dt < \infty \right\}, \\
L^2_{\mathcal{F},p}([0,T]; \mathbf{R}^m) &:= \left\{ \widehat{\varphi} : [0,T] \times \Omega \longrightarrow \mathbf{R}^m \mid \widehat{\varphi} \text{ is } \mathcal{F}_t\text{-predictable}, \mathbb{E}\int_0^T \left|\widehat{\varphi}(t)\right|^2 dt < \infty \right\}, \\
F^2_p([0,T]; \mathbf{R}^m) &:= \Bigg\{ \widetilde{\varphi} : [0,T] \times \Omega \times \mathbf{E} \longrightarrow \mathbf{R}^m \mid \\
&\qquad \widetilde{\varphi} \text{ is } \mathcal{F}_t\text{-predictable, such that, } \mathbb{E}\int_0^T \int_{\mathbf{E}} \left|\widetilde{\varphi}(t,e)\right|^2 \pi(de) dt < \infty \Bigg\}.
\end{aligned}
\tag{1.1}
$$

We denote

$$
\mathcal{M}^2[0,T] := L^2_{\mathcal{F}}([0,T]; \mathbf{R}^n) \times L^2_{\mathcal{F}}([0,T]; \mathbf{R}^m) \times L^2_{\mathcal{F},p}([0,T]; \mathbf{R}^m) \times F^2_p([0,T]; \mathbf{R}^m). \tag{1.2}
$$

Clearly, $\mathcal{M}^2[0,T]$ is a Banach space. Any process in $\mathcal{M}^2[0,T]$ is denoted by $\Theta(\cdot) \equiv (x(\cdot), y(\cdot), z(\cdot), c(\cdot,\cdot))$, whose norm is defined by

$$
\|\Theta(\cdot)\|^2_{\mathcal{M}^2[0,T]} := \mathbb{E}\left[\sup_{t\in[0,T]} |x(t)|^2 + \sup_{t\in[0,T]} \left|y(t)\right|^2 + \int_0^T |z(t)|^2 dt + \mathbb{E}\int_0^T \int_{\mathbf{E}} |c(t,e)|^2 \pi(de) dt \right]. \tag{1.3}
$$

1.2. Formulation of the Optimal Control Problem and Basic Assumptions

For any $u(\cdot) \in \mathcal{U}[0,T]$ and $a \in \mathbf{R}^n$, we consider the following fully coupled forward-backward stochastic control system:

$$
\begin{aligned}
dx(t) &= b\big(t, x(t), y(t), z(t), c(t,\cdot), u(t)\big) dt + \sigma\big(t, x(t), y(t), z(t), c(t,\cdot), u(t)\big) dW(t) \\
&\quad + \int_{\mathbf{E}} g\big(t, x(t-), y(t-), z(t), c(t,\cdot), u(t), e\big) \widetilde{N}(dedt), \\
-dy(t) &= f\big(t, x(t), y(t), z(t), c(t,\cdot), u(t)\big) dt - z(t) dW(t) - \int_{\mathbf{E}} c(t,e) \widetilde{N}(dedt), \\
x(0) &= a, \qquad y(T) = h(x(T)),
\end{aligned}
\tag{1.4}
$$

with the cost functional given by

$$J(u(\cdot)) = \mathbb{E}\left[\int_0^T l(t, x(t), y(t), z(t), c(t,\cdot), u(t))dt + \phi(x(T)) + \gamma(y(0))\right]. \tag{1.5}$$

Here, $\Theta(\cdot) \equiv (x(\cdot), y(\cdot), z(\cdot), c(\cdot,\cdot))$ takes value in $\mathbb{M} = \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^m \times \mathbf{R}^m$ and

$$\begin{gathered} b, \sigma : [0,T] \times \mathbb{M} \times \mathbf{U} \longrightarrow \mathbf{R}^n, \qquad g : [0,T] \times \mathbb{M} \times \mathbf{U} \times \mathbf{E} \longrightarrow \mathbf{R}^n, \\ f : [0,T] \times \mathbb{M} \times \mathbf{U} \longrightarrow \mathbf{R}^m, \qquad h : \mathbf{R}^n \longrightarrow \mathbf{R}^m, \\ \phi : \mathbf{R}^n \longrightarrow \mathbf{R}, \qquad \gamma : \mathbf{R}^m \longrightarrow \mathbf{R}, \qquad l : [0,T] \times \mathbb{M} \times \mathbf{U} \longrightarrow \mathbf{R}. \end{gathered} \tag{1.6}$$

For any $u(\cdot) \in \mathcal{U}[0,T]$ and $a \in \mathbf{R}^n$, we refer to $\Theta(\cdot) \equiv (x(\cdot), y(\cdot), z(\cdot), c(\cdot,\cdot))$ as the state process corresponding to the admissible control $u(\cdot)$ if FBSDEJ (1.4) admits a unique adapted solution. For controlled FBSDEJ (1.4) with cost functional (1.5), we consider the following.

Problem C. Find $\overline{u}(\cdot) \in \mathcal{U}[0,T]$, such that

$$J(\overline{u}(\cdot)) = \inf_{u(\cdot)\in\mathcal{U}[0,T]} J(u(\cdot)). \tag{1.7}$$

Any $\overline{u}(\cdot) \in \mathcal{U}[0,T]$ satisfying (1.7) is called an optimal control process of Problem C, and the corresponding state process, denoted by $\overline{\Theta}(\cdot) \equiv (\overline{x}(\cdot), \overline{y}(\cdot), \overline{z}(\cdot), \overline{c}(\cdot,\cdot))$, is called optimal state process. We also refer to $(\overline{x}(\cdot), \overline{y}(\cdot), \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot))$ as an optimal 5-tuple of Problem C.

Our main goal in this paper is to derive some necessary conditions for the optimal control of Problem C, which is called the stochastic maximum principle of Pontraygin's type. For this target, we first introduce the following basic assumption throughout this paper.

(H0) *For any* $u(\cdot) \in \mathcal{U}[0,T]$ *and* $a \in \mathbf{R}^n$, FBSDEJ (1.4) *admits a unique adapted solution* $\Theta(\cdot) \equiv (x(\cdot), y(\cdot), z(\cdot), c(\cdot,\cdot)) \in \mathcal{M}^2[0,T]$. *Moreover, the following estimate holds:*

$$\begin{aligned} \|\Theta(\cdot)\|_{\mathcal{M}^2[0,T]} &\equiv \mathbb{E}\left[\sup_{t\in[0,T]} |x(t)|^2 + \sup_{t\in[0,T]} |y(t)|^2 + \int_0^T |z(t)|^2 dt + \mathbb{E}\int_0^T \int_{\mathbf{E}} |c(t,e)|^2 \pi(de) dt\right] \\ &\le C\mathbb{E}\Bigg[|a|^2 + |h(0)|^2 + \left(\int_0^T |b(t,0,u(t))| dt\right)^2 + \left(\int_0^T |f(t,0,u(t))| dt\right)^2 \\ &\quad + \int_0^T |\sigma(t,0,u(t))|^2 dt + \int_0^T \int_{\mathbf{E}} |g(t,0,u(t),e)|^2 \pi(de) dt\Bigg]. \end{aligned} \tag{1.8}$$

Further, if $\widetilde{\Theta}(\cdot) \equiv (\widetilde{x}(\cdot), \widetilde{y}(\cdot), \widetilde{z}(\cdot), \widetilde{c}(\cdot,\cdot)) \in \mathcal{M}^2[0,T]$ *is the unique adapted solution of* (1.4) *with* $h(\cdot)$ *and* $u(\cdot)$ *replaced by* $\widetilde{h} : \mathbf{R}^n \to \mathbf{R}^m$ *and* $\widetilde{u}(\cdot) \in \mathcal{U}[0,T]$, *respectively, then the following stability estimate holds:*

$$\begin{aligned}
\left\|\Theta(\cdot) - \widetilde{\Theta}(\cdot)\right\|_{\mathcal{M}^2[0,T]} \equiv{}& \mathbb{E}\Bigg[\sup_{t\in[0,T]} |x(t) - \widetilde{x}(t)|^2 + \sup_{t\in[0,T]} \left|y(t) - \widetilde{y}(t)\right|^2 \\
&+ \int_0^T |z(t) - \widetilde{z}(t)|^2 dt + \mathbb{E}\int_0^T \int_{\mathbf{E}} |c(t,e) - \widetilde{c}(t,e)|^2 \pi(de)dt\Bigg] \\
\le{}& C\mathbb{E}\Bigg[|h(x(T)) - h(\widetilde{x}(T))|^2 + \left(\int_0^T \left|b(t,\Theta(t),u(t)) - b\left(t,\widetilde{\Theta}(t),\widetilde{u}(t)\right)\right|dt\right)^2 \\
&+ \left(\int_0^T \left|f(t,\Theta(t),u(t)) - f\left(t,\widetilde{\Theta}(t),\widetilde{u}(t)\right)\right|dt\right)^2 \\
&+ \int_0^T \left|\sigma(t,\Theta(t),u(t)) - \sigma\left(t,\widetilde{\Theta}(t),\widetilde{u}(t)\right)\right|^2 dt \\
&+ \int_0^T \int_{\mathbf{E}} \left|g(t,\Theta(t),u(t),e) - g\left(t,\widetilde{\Theta}(t),\widetilde{u}(t),e\right)\right|^2 \pi(de)dt\Bigg].
\end{aligned} \tag{1.9}$$

By adopting the ideas from Wu [1], we know that under certain G-monotonicity conditions for the coefficients, the existence and uniqueness of FBSDEJ (1.4) guaranteed which leads to hypothesis (H0). Our main goal in this paper is to derive some necessary conditions for the optimal control of Problem C. Hence, we impose the well-posedness of the state equation (1.4) as an assumption to avoid some technicalities not closely to our main results.

1.3. Developments of Stochastic Optimal Control and Contributions of the Paper

It is well known that the optimal control problem is one of the central themes of modern control sciences. Necessary conditions for the optimal control of the (forward) continuous stochastic control system, that is, the so-called stochastic maximum principle of Pontraygin's type, were extensively studied since early 1960s. When Brownian motion is the unique noise source, Peng [2] (see also Yong and Zhou [3]) obtained the maximum principle for the general case, that is, the control variable entering the diffusion coefficient and control domain being not convex.

Forward-backward stochastic control systems where the controlled systems are described by *forward-backward stochastic differential equations* (FBSDEs) are widely used in mathematical economics and mathematical finance, which includes the usual forward SDEs as a special case. They are encountered in stochastic recursive utility optimization problems (see [4–8]) and principal-agent problems (see [9, 10]). Peng [11] first obtained the necessary conditions for optimal control for the partially coupling case when the control domain is convex. And then Xu [12] studied the nonconvex control domain case and obtained the corresponding necessary conditions. But he needs to assume that the diffusion coefficient in the forward control system does not contain the control variable. Ji and Zhou [8] applied the Ekeland variational principle to establish a maximum principle for a partially coupled forward-backward stochastic control system, while the forward state is constrained in a convex set at the terminal time. Wu [13] recently established a general maximum principle for optimal control problems derived by forward-backward stochastic systems, where control domains are nonconvex and forward diffusion coefficients explicitly depend on control variables. Moreover, some financial optimization problems for large investors (see [14–16]) and some asset pricing problems with forward-backward differential utility (see [7]) directly lead to fully coupled FBSDEs. Wu [17] first (see also Meng [18]) obtained the necessary conditions for optimal control of fully coupled forward-backward stochastic control systems when the control domain is convex. And then Shi and Wu [19] studied the nonconvex control domain case and obtained the corresponding necessary conditions under some G-monotonic assumptions. But they also (similar to Xu [12]) need to assume that the control variable does not appear in the diffusion coefficient of the forward equation. Very recently, Yong [20] completely solved the problem of finding necessary conditions for optimal control of fully coupled FBSDEs. He considered an optimal control problem for general coupled FBSDEs with mixed initial-terminal conditions and derived the necessary conditions for the optimal controls when the control domain is not assumed to be convex and the control variable appears in the diffusion coefficient of the forward equation.

However, recently more and more research attentions are drawn towards the optimal control problem for *discontinuous* stochastic systems or stochastic systems with random jumps. The reason is clear for its applicable aspect. For example, there is compelling evidence that the dynamics of prices of financial instruments exhibit jumps that cannot be adequately captured solely by diffusion processes (i.e., processes satisfying some Itô-type *stochastic differential equations* (SDE for short), see Merton [21] and Cont and Tankov [22]. Several empirical studies demonstrate the existence of jumps in stock markets, the foreign exchange market, and bond markets. Jumps constitute also a key feature in the description of credit risk sensitive instruments. Therefore, models that incorporate jumps have become increasingly popular in finance and several areas of science and engineering, this leads to paying more attention to *stochastic differential equations with jumps* (SDEJs for short). As consequences, optimal control problems involving systems of SDEJs are widely studied. Situ [23] first obtained the maximum principle for (forward) stochastic control system with jumps but the jump coefficient does not contain the control variable. Tang and Li [24] discussed a more general case, where the control is allowed into both diffusion and jump coefficients, and the control domain is not necessarily convex; also some general state constraints are imposed. Maximum principles for forward-backward stochastic control systems with random jumps are studied in Øksendal and Sulem [25], Shi and Wu [26], where the FBSDEJs are partially coupled and the control domains are convex. Recently, Shi [27] obtained the necessary condition of optimal control as well as a sufficient condition of optimality under the assumption that the diffusion and jump coefficients do not contain the control variable and the control domain need not be convex. Necessary conditions for fully coupled forward-backward stochastic control systems with random jumps were studied in Shi and Wu [28] (see also Meng and Sun [29]) where the control domains are convex.

In this paper, we consider the general optimal control problem for fully coupled FBSDEJ (1.4). Here, by the word "general" we mean the allowance of the control variable into both diffusion and jump coefficients of the forward equation and the control domain is not assumed to be convex. It is well worth mentioning that the idea of second-order spike variation technique, developed by Peng [2], plays an important role in deriving the necessary conditions of general stochastic optimal control of jump-diffusion process in Tang and Li [24].

Following the standard idea of deriving necessary conditions for optimal control processes, due to the fact the control domain $\mathbf{U}$ is not assumed to be convex, one needs to use spike variation for the control process and then to try obtaining a Taylor-type expansion for the state process $(x(\cdot), y(\cdot), z(\cdot), c(\cdot,\cdot))$ and the cost functional (1.5) with respect to the spike variation of the control process, followed by some suitable duality relations to get a maximum principle of Pontryagin-type. However, the derivation of Taylor expansion of the state process $(x(\cdot), y(\cdot), z(\cdot), c(\cdot,\cdot))$ with respect to the spike variation of the control process is technically difficult. The main reasons are that both $(z(\cdot), c(\cdot,\cdot))$ and $u(\cdot)$ appear in the diffusion and jump coefficients of the forward equation of (1.4) and that the regularity/integrability of the continuous martingale process $z(\cdot)$ and the discontinuous martingale process $c(\cdot,\cdot)$ (as part of the state process) seems to be not enough in the case when a second-order expansion is necessary. Note that in [25–29], due to the special structure of the problems, the second-order expansion is not necessary when one tries to derive the necessary conditions for optimal controls.

We overcome the above difficulty by the newly developed *reduction method* by Wu [13] and Yong [20] in the continuous case, independently. In fact, some ideas have been proposed early in Kohlmann and Zhou [30] and Ma and Yong [31]. Let us make it more precise. We first introduce the controlled initial value problem for a system of SDEJs, where $(x(\cdot), y(\cdot))$ is regarded as the state process and $(z(\cdot), c(\cdot,\cdot), u(\cdot))$ is regarded as the control process. Next, we regard the original terminal condition $y(T) = h(x(T))$ as terminal state constraint and then we translate Problem C into a high-dimensional reduced optimal control problem described by standard SDEJ with state constraint (see Problem $\widehat{\mathrm{C}}$ in Section 3). The advantage of this reduced optimal control problem is that one needs not much regularity/integrability of processes $(z(\cdot), c(\cdot,\cdot))$ since it is treated as a control process. We apply the Ekeland variational principle to deal with this high-dimensional reduced optimal control problem with state constraint. Finally, necessary conditions for the optimal control of Problem C are derived by the equivalence of Problems C and $\widehat{\mathrm{C}}$.

The rest of this paper is organized as follows. In Section 2, under some suitable assumptions, we give the main result of this paper, together with some discussions on special cases. In Section 3, after we make the reduction of our optimal control problems for FBSDEJs Problem C, a proof of our main result will be presented. Section 4 is devoted to a linear quadratic stochastic optimal control problem as an illustrating example. Finally, in Section 5 we give the concluding remarks and compare our theorem with some existing results.

2. The Main Result and Some Special Cases

In this section, under some suitable assumptions, we will state the necessary conditions for the optimal control of our Problem C. Also, some interesting special cases will be discussed.

Let us introduce the following further assumptions beyond (H0).

(H1) *For any* $(\theta, u) \in \mathbb{M} \times \mathbf{U}$, *maps* $t \mapsto (b(t,\theta,u), \sigma(t,\theta,u), g(t,\theta,u,e), f(t,\theta,u))$ *is* $\mathcal{F}_t$*-progressively measurable and there exists a constant* $L > 0$*such that*

$$
\begin{aligned}
&\left|b(t,x,y,z,c,u) - b(t,\tilde{x},\tilde{y},z,c,u)\right|^2 + \left|\sigma(t,x,y,z,c,u) - \sigma(t,\tilde{x},\tilde{y},z,c,u)\right|^2 \\
&\quad + \left\|g(t,x,y,z,c,u,e) - g(t,\tilde{x},\tilde{y},z,c,u,e)\right\|^2_{\mathcal{L}^2} + \left|f(t,x,y,z,c,u) - f(t,\tilde{x},\tilde{y},z,c,u)\right|^2 \\
&\leq L\left(|x-\tilde{x}|^2 + |y-\tilde{y}|^2\right), \quad \forall (t,z,c,u) \in [0,T] \times \mathbf{R}^m \times \mathbf{R}^n \times \mathbf{U},\ (x,y), (\tilde{x},\tilde{y}) \in \mathbf{R}^n \times \mathbf{R}^m,\ a.s.,
\end{aligned}
$$

$$\begin{aligned}&\left|b(t,x,y,z,c,u)\right|^2+\left|\sigma(t,x,y,z,c,u)\right|^2+\|g(t,x,y,z,c,u,e)\|_{\mathcal{L}^2}^2+\left|f(t,x,y,z,c,u)\right|^2\\&\quad\le L\left(1+|x|^2+|y|^2\right),\quad\forall(t,z,c,u)\in[0,T]\times\mathbf{R}^m\times\mathbf{R}^n\times\mathbf{U},\ (x,y)\in\mathbf{R}^n\times\mathbf{R}^m,\ a.s.\end{aligned}\tag{2.1}$$

(H2) *Maps $\theta\mapsto(b(t,\theta,u),\sigma(t,\theta,u),g(t,\theta,u,e),f(t,\theta,u))$ is twice continuously differential, with the (partial) derivatives up to the second order being uniformly bounded, Lipschitz continuous in $\theta\in\mathbb{M}$, and continuous in $u\in\mathbf{U}$.*

(H3) *The map h is twice differentiable with the derivatives up to the second order being uniformly bounded and uniformly Lipschitz continuous; the maps ϕ,γ are twice differentiable with the derivatives up to the second order being Lipschitz continuous; the map $\theta\mapsto l(t,\theta,u)$ is twice differentiable with the derivatives up to the second order being Lipschitz continuous in $\theta\in\mathbb{M}$, and continuous in $u\in\mathbf{U}$.*

Next, to simplify our presentation, we now introduce some abbreviation notations. First, we make the following convention: for any differentiable map $a\equiv(a^1,a^2,\ldots,a^l)^\top:\mathbf{R}^n\to\mathbf{R}^l$, let

$$a_x=\begin{pmatrix}a^1_{x_1}&a^1_{x_2}&\cdots&a^1_{x_n}\\a^2_{x_1}&a^2_{x_2}&\cdots&a^2_{x_n}\\\vdots&\vdots&\ddots&\vdots\\a^l_{x_1}&a^l_{x_2}&\cdots&a^l_{x_n}\end{pmatrix}:\mathbf{R}^n\longrightarrow\mathbf{R}^{l\times n}.\tag{2.2}$$

In particular, for $l=1$, $a_x=(a_{x_1},a_{x_2},\ldots,a_{x_n})\in\mathbf{R}^{1\times n}$ is an n-dimensional row vector. Also, for any twice differentiable function $a:\mathbf{R}^n\to\mathbf{R}$, the Hessian matrix is given by

$$a_{xx}\equiv\left(a_x^\top\right)_x=\begin{pmatrix}a_{x_1x_1}&a_{x_1x_2}&\cdots&a_{x_1x_n}\\a_{x_2x_1}&a_{x_2x_2}&\cdots&a_{x_2x_n}\\\vdots&\vdots&\ddots&\vdots\\a_{x_nx_1}&a_{x_nx_2}&\cdots&a_{x_nx_n}\end{pmatrix}:\mathbf{R}^n\longrightarrow\mathcal{S}^n,\tag{2.3}$$

hereafter, $\mathcal{S}^n$ stands for the set of all real $n\times n$ symmetric matrices. On the other hand, for a twice differentiable function $a:\mathbf{R}^n\times\mathbf{R}^m\to\mathbf{R}$ (denoting $(x,y)\in\mathbf{R}^n\times\mathbf{R}^m$), we denote

$$a_{xy}\equiv\left(a_x^\top\right)_y=\begin{pmatrix}a_{x_1y_1}&a_{x_1y_2}&\cdots&a_{x_1y_m}\\a_{x_2y_1}&a_{x_2y_2}&\cdots&a_{x_2y_m}\\\vdots&\vdots&\ddots&\vdots\\a_{x_ny_1}&a_{x_ny_2}&\cdots&a_{x_ny_m}\end{pmatrix}:\mathbf{R}^n\times\mathbf{R}^m\longrightarrow\mathbf{R}^{n\times m}.\tag{2.4}$$

Now, let $(\overline{x}(\cdot),\overline{y}(\cdot),\overline{z}(\cdot),\overline{c}(\cdot,\cdot),\overline{u}(\cdot))$ be an optimal 5-tuple of Problem C. For $a(\cdot) = b(\cdot),\sigma(\cdot),f(\cdot),l(\cdot)$, we denote

$$\begin{aligned}
\overline{a}_x(t,\cdot) &= a_x\big(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)\big),\\
\overline{a}_y(t,\cdot) &= a_y\big(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)\big),\\
\overline{a}_z(t,\cdot) &= a_z\big(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)\big),\\
\overline{a}_c(t,\cdot) &= a_c\big(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)\big),
\end{aligned} \tag{2.5}$$

and let

$$\begin{aligned}
\overline{g}_x(t,e) &= g_x\big(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,e),\overline{u}(t),e\big),\\
\overline{g}_y(t,e) &= g_y\big(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,e),\overline{u}(t),e\big),\\
\overline{g}_z(t,e) &= g_z\big(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,e),\overline{u}(t),e\big),\\
\overline{g}_c(t,e) &= g_c\big(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,e),\overline{u}(t),e\big).
\end{aligned} \tag{2.6}$$

We also introduce

$$\begin{aligned}
\overline{B}_X(t,\cdot) &= \begin{pmatrix} \overline{b}_x(t,\cdot) & \overline{b}_y(t,\cdot) \\ -\overline{f}_x(t,\cdot) & -\overline{f}_y(t,\cdot) \end{pmatrix},\\
\overline{\Sigma}_X(t,\cdot) &= \begin{pmatrix} \overline{\sigma}_x(t,\cdot) & \overline{\sigma}_y(t,\cdot) \\ 0 & 0 \end{pmatrix},\\
\overline{\Upsilon}_X(t,e) &= \begin{pmatrix} \overline{g}_x(t,e) & \overline{g}_y(t,e) \\ 0 & 0 \end{pmatrix}.
\end{aligned} \tag{2.7}$$

For $a(\cdot) = b^i(\cdot),\sigma^i(\cdot),f^j(\cdot),l(\cdot), 1 \le i \le n, 1 \le j \le m$, let

$$\begin{aligned}
\overline{a}_{xx}(t,\cdot) &= a_{xx}\big(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)\big),\\
\overline{a}_{yy}(t,\cdot) &= a_{yy}\big(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)\big),\\
\overline{a}_{xy}(t,\cdot) &= a_{xy}\big(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)\big),\\
\overline{a}_{yx}(t,\cdot) &= a_{yx}\big(\mathrm{t},\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)\big),
\end{aligned} \tag{2.8}$$

and let

$$\begin{aligned}
\overline{g}_{xx}(t,e) &= g_{xx}\big(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t),e\big),\\
\overline{g}_{yy}(t,e) &= g_{yy}\big(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t),e\big),\\
\overline{g}_{xy}(t,e) &= g_{xy}\big(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t),e\big),\\
\overline{g}_{yx}(t,e) &= g_{yx}\big(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t),e\big).
\end{aligned} \tag{2.9}$$

Our main result in this paper is the following theorem for Problem C.

Theorem 2.1. *Suppose that (H0)–(H3) hold, and let* $(\overline{x}(\cdot),\overline{y}(\cdot),\overline{z}(\cdot),\overline{c}(\cdot,\cdot),\overline{u}(\cdot))$ *be an optimal 5-tuple of Problem C. Then there exists a unique adapted solution* $(p(\cdot),q(\cdot),k(\cdot),r(\cdot,\cdot)) \in L^2_{\mathcal{F}}([0,T];\mathbf{R}^n)\times L^2_{\mathcal{F}}([0,T];\mathbf{R}^m)\times L^2_{\mathcal{F},p}([0,T];\mathbf{R}^n)\times F^2_p([0,T];\mathbf{R}^n)$ *to the following FBSDEJ:*

$$\begin{aligned}
-dp(t) &= \Big[\overline{b}_x(t,\cdot)^\top p(t) - \overline{f}_x(t,\cdot)^\top q(t) + \overline{\sigma}_x(t,\cdot)^\top k(t) + \int_{\mathbf{E}} \overline{g}_x(t,e)^\top r(t,e)\pi(de) - \overline{l}_x(t,\cdot)\Big]dt \\
&\quad - k(t)dW(t) - \int_{\mathbf{E}} r(t,e)\widetilde{N}(dedt), \\
dq(t) &= \Big[-\overline{b}_y(t,\cdot)^\top p(t) + \overline{f}_y(t,\cdot)^\top q(t) - \overline{\sigma}_y(t,\cdot)^\top k(t) - \int_{\mathbf{E}} \overline{g}_y(t,e)^\top r(t,e)\pi(de) + \overline{l}_y(t,\cdot)\Big]dt \\
&\quad + \Big[- \overline{b}_z(t,\cdot)^\top p(t) + \overline{f}_z(t,\cdot)^\top q(t) - \overline{\sigma}_z(t,\cdot)^\top k(t) \\
&\qquad - \int_{\mathbf{E}} \overline{g}_z(t,e)^\top r(t,e)\pi(de) + \overline{l}_z(t,\cdot)\Big]dW(t) \\
&\quad + \int_{\mathbf{E}} \Big[-\overline{b}_c(t,e)^\top p(t-) + \overline{f}_c(t,e)^\top q(t-) \\
&\qquad -\overline{\sigma}_c(t,e)^\top k(t) - \overline{g}_c(t,e)^\top r(t,e) + \overline{l}_c(t,e)\Big]\widetilde{N}(dedt), \\
p(T) &= -\phi_x(\overline{x}(T)) - h_x(\overline{x}(T))^\top q(T), \qquad q(0) = \mathbb{E}\gamma_y(\overline{y}_0).
\end{aligned} \tag{2.10}$$

Let $(P(\cdot),Q(\cdot),K(\cdot,\cdot)) \in L^2_{\mathcal{F}}([0,T];\mathcal{S}^{n+m})\times L^2_{\mathcal{F},p}([0,T];\mathcal{S}^{n+m})\times F^2_p([0,T];\mathcal{S}^{n+m})$ *be the unique adapted solution to the following matrix-valued BSDEJ:*

$$\begin{aligned}
-dP(t) &= \Big\{\overline{B}_X(t,\cdot)^\top P(t) + P(t)\overline{B}_X(t,\cdot) + \overline{\Sigma_X}(t,\cdot)^\top P(t)\overline{\Sigma_X}(t,\cdot) + \overline{\Sigma_X}(t,\cdot)^\top Q(t) + Q(t)\overline{\Sigma_X}(t,\cdot) \\
&\quad + \int_{\mathbf{E}} \Big[\overline{\Upsilon}_X(t,e)^\top P(t)\overline{\Upsilon}_X(t,e) + \overline{\Upsilon}_X(t,e)^\top K(t,e)\overline{\Upsilon}_X(t,e) \\
&\qquad + \overline{\Upsilon}_X(t,e)^\top K(t,e) + K(t,e)\overline{\Upsilon}_X(t,e)\Big]\pi(de) + \overline{H}_{XX}(t,\cdot)\Big\}dt \\
&\quad - Q(t)dW(t) - \int_{\mathbf{E}} K(t,e)\widetilde{N}(dedt), \\
P(T) &= \begin{pmatrix} -\phi_{xx}(\overline{x}(T)) - h_{xx}(\overline{x}(T))^\top q(T) & 0 \\ 0 & 0 \end{pmatrix},
\end{aligned} \tag{2.11}$$

where $\overline{H}_{XX}(t,\cdot) = \overline{H}_{XX}(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t),p(t),q(t),k(t),r(t,\cdot))$*, and the Hamiltonian function* $\overline{H} : [0,T]\times\mathbb{M}\times\mathbf{U}\times\mathbf{R}^n\times\mathbf{R}^m\times\mathbf{R}^n\times\mathbf{R}^n \to \mathbf{R}$ *is defined as follows:*

$$\begin{aligned}
&\overline{H}(t,x,y,z,c(\cdot),u,p,q,k,r(\cdot)) \\
&\quad := \langle p, b(t,x,y,z,c(\cdot),u)\rangle + \langle q, -f(t,x,y,z,c(\cdot),u)\rangle \\
&\qquad + \langle k, \sigma(t,x,y,z,c(\cdot),u)\rangle + \int_{\mathbf{E}} \langle r(e), g(t,x,y,z,c(e),u,e)\rangle \pi(de) - l(t,x,y,z,c(\cdot),u).
\end{aligned} \tag{2.12}$$

Then,

$$\begin{aligned}
&\overline{H}(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t),p(t),q(t),k(t),r(t,\cdot))\\
&\quad-\overline{H}(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),u,p(t),q(t),k(t),r(t,\cdot))\\
&\quad-\frac{1}{2}\left(\sigma(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t))-\sigma(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),u)\right)^{\top}P_1(t)\\
&\quad\times\left(\sigma(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t))-\sigma(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),u)\right)\\
&\quad-\frac{1}{2}\int_{\mathbf{E}}\left(g(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,e),\overline{u}(t),e)-g(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,g),u,e)\right)^{\top}\\
&\quad\times[P_1(t)+K_1(t,e)]\\
&\quad\times\left(g(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,e),\overline{u}(t),e)-g(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,e),u,e)\right)\pi(de)\geq 0,\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\forall u\in\mathbf{U},\ a.e.,\ a.s.,
\end{aligned}\tag{2.13}$$

with $P(\cdot)\equiv\begin{pmatrix}P_1(\cdot) & P_2(\cdot)\\ P_2(\cdot)^{\top} & P_3(\cdot)\end{pmatrix}$, *and* $K(\cdot,\cdot)\equiv\begin{pmatrix}K_1(\cdot,\cdot) & K_2(\cdot,\cdot)\\ K_2(\cdot,\cdot)^{\top} & K_3(\cdot,\cdot)\end{pmatrix}$.

Remark 2.2. In fact, the second-order adjoint equation (2.11) can be split into the following three BSDEJs if we add that $Q(\cdot)\equiv\begin{pmatrix}Q_1(\cdot) & Q_2(\cdot)\\ Q_2(\cdot)^{\top} & Q_3(\cdot)\end{pmatrix}$:

$$\begin{aligned}
-dP_1(t)=&\Big\{\overline{b}_x(t,\cdot)^{\top}P_1(t)+P_1(t)\overline{b}_x(t,\cdot)+\overline{\sigma}_x(t,\cdot)^{\top}P_1(t)\overline{\sigma}_x(t,\cdot)\\
&\quad+\overline{\sigma}_x(t,\cdot)^{\top}Q_1(t)+Q_1(t)\overline{\sigma}_x(t,\cdot)-\overline{f}_x(t,\cdot)^{\top}P_2(t)-P_2(t)\overline{f}_x(t,\cdot)\\
&\quad+\int_{\mathbf{E}}\Big[\overline{g}_x(t,e)^{\top}P_1(t)\overline{g}_x(t,e)+\overline{g}_x(t,e)^{\top}K_1(t,e)\overline{g}_x(t,e)\\
&\qquad+\overline{g}_x(t,e)^{\top}K_1(t,e)+K_1(t,e)\overline{g}_x(t,e)\Big]\pi(de)+\overline{H}_{xx}(t,\cdot)\Big\}dt\\
&-Q_1(t)dW(t)-\int_{\mathbf{E}}K_1(t,e)\widetilde{N}(dedt),\\
P_1(T)=&-\phi_{xx}(\overline{x}(T))-h_{xx}(\overline{x}(T))^{\top}q(T),
\end{aligned}\tag{2.14}$$

$$\begin{aligned}
-dP_2(t)=&\Big\{P_1(t)\overline{b}_y(t,\cdot)+\overline{b}_x(t,\cdot)P_2(t)+\overline{\sigma}_x(t,\cdot)^{\top}P_1(t)\overline{\sigma}_y(t,\cdot)^{\top}\\
&\quad+\overline{\sigma}_x(t,\cdot)Q_2(t)+Q_1(t)^{\top}\overline{\sigma}_y(t,\cdot)-\overline{f}_x(t,\cdot)P_3(t)-P_2(t)^{\top}\overline{f}_y(t,\cdot)\\
&\quad+\int_{\mathbf{E}}\Big[\overline{g}_x(t,e)^{\top}P_1(t)\overline{g}_y(t,e)^{\top}+\overline{g}_x(t,e)^{\top}K_1(t,e)\overline{g}_y(t,e)^{\top}\\
&\qquad+\overline{g}_x(t,e)K_2(t,e)+K_1(t,e)\overline{g}_y(t,e)^{\top}\Big]\pi(de)+\overline{H}_{xy}(t,\cdot)\Big\}dt\\
&-Q_2(t)dW(t)-\int_{\mathbf{E}}K_2(t,e)\widetilde{N}(dedt),\\
P_2(T)=&\,0,
\end{aligned}\tag{2.15}$$

$$
\begin{aligned}
-dP_3(t) = \Big\{ & \overline{b}_y(t,\cdot)^\top P_2(t) + P_2(t)\overline{b}_y(t,\cdot) + \overline{\sigma}_y(t,\cdot)P_1(t)\overline{\sigma}_y(t,\cdot)^\top \\
& + \overline{\sigma}_y(t,\cdot)^\top Q_2(t) + Q_2(t)\overline{\sigma}_y(t,\cdot) - \overline{f}_y(t,\cdot)^\top P_3(t) - P_3(t)\overline{f}_y(t,\cdot) \\
& + \int_{\mathbf{E}} \Big[\overline{g}_y(t,e)P_1(t)\overline{g}_y(t,e)^\top + \overline{g}_y(t,e)K_1(t,e)\overline{g}_y(t,e)^\top \\
& \qquad + \overline{g}_y(t,e)^\top K_2(t,e) + K_2(t,e)\overline{g}_y(t,e)\Big]\pi(de) + \overline{H}_{yy}(t,\cdot)\Big\} dt \\
& - Q_3(t)dW(t) - \int_{\mathbf{E}} K_3(t,e)\widetilde{N}(dedt), \\
P_3(T) = 0. &
\end{aligned}
\tag{2.16}
$$

Note that this kind of three second-order adjoint equations appears in Wu [13] whereas not in Yong [20].

Let us now look at some special cases. It can be seen that our theorem recovers the known results.

(1) *A Classical Stochastic Optimal Control Problem with Random Jumps*

Consider a controlled SDEJ:

$$
\begin{aligned}
dx(t) &= b(t,x(t),u(t))dt + \sigma(t,x(t),u(t))dW(t) + \int_{\mathbf{E}} g(t,x(t-),u(t),e)\widetilde{N}(dedt), \\
x(0) &= a,
\end{aligned}
\tag{2.17}
$$

with the cost functional

$$
J(u(\cdot)) = \mathbb{E}\left[\int_0^T l(t,x(t),u(t))dt + \phi(x(T))\right]. \tag{2.18}
$$

In this case,

$$
\begin{gathered}
b(t,x,y,z,c(\cdot),u) = b(t,x,u), \qquad \sigma(t,x,y,z,c(\cdot),u) = \sigma(t,x,u), \\
g(t,x,y,z,c(\cdot),u,\cdot) = g(t,x,u,\cdot), \qquad f(t,x,y,z,c(\cdot),u) = 0, \\
l(t,x,y,z,c(\cdot),u,\cdot) = l(t,x,u), \qquad \gamma(y) = 0.
\end{gathered}
\tag{2.19}
$$

Hence, we have, by some direct computation/observation,

$$
\begin{gathered}
y(\cdot) = z(\cdot) = c(\cdot,\cdot) = 0, \qquad q(\cdot) = 0, \\
P_2(\cdot) = Q_2(\cdot) = K_2(\cdot,\cdot) = 0, \qquad P_3(\cdot) = Q_3(\cdot) = K_3(\cdot,\cdot) = 0.
\end{gathered}
\tag{2.20}
$$

Consequently, (2.10) and (2.11) (or equivalently, (2.14)) become

$$\begin{aligned} -dp(t) &= \left[\bar{b}_x(t,\cdot)^\top p(t) + \bar{\sigma}_x(t,\cdot)^\top k(\mathrm{t}) + \int_{\mathbf{E}} \bar{g}_x(t,e)^\top r(t,e)\pi(de) - \bar{l}_x(t,\cdot)\right] dt \\ &\quad - k(t)dW(t) - \int_{\mathbf{E}} r(t,e)\widetilde{N}(dedt), \\ p(T) &= \phi_x(\bar{x}(T)), \end{aligned} \tag{2.21}$$

$$\begin{aligned} -dP_1(t) &= \Big\{ \bar{b}_x(t,\cdot)^\top P_1(t) + P_1(t)\bar{b}_x(t,\cdot) + \bar{\sigma}_x(t,\cdot)^\top P_1(t)\bar{\sigma}_x(t,\cdot) \\ &\quad + \bar{\sigma}_x(t,\cdot)^\top Q_1(t) + Q_1(t)\bar{\sigma}_x(t,\cdot) \\ &\quad + \int_{\mathbf{E}} \Big[\bar{g}_x(t,e)^\top P_1(t)\bar{g}_x(t,e) + \bar{g}_x(t,e)^\top K_1(t,e)\bar{g}_x(t,e) \\ &\qquad + \bar{g}_x(t,e)^\top K_1(t,e) + K_1(t,e)\bar{g}_x(t,e)\Big]\pi(de) + \overline{H}_{xx}(t,\cdot)\Big\} dt \\ &\quad - Q_1(t)dW(t) - \int_{\mathbf{E}} K_1(t,e)\widetilde{N}(dedt), \\ P_1(T) &= -\phi_{xx}(\bar{x}(T)), \end{aligned} \tag{2.22}$$

where the Hamiltonian function $\overline{H} : [0,T] \times \mathbf{R}^n \times \mathbf{U} \times \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^n \rightarrow \mathbf{R}$ is defined as

$$\begin{aligned} \overline{H}(t,x,u,p,k,r(\cdot)) &:= \langle p, b(t,x,u)\rangle + \langle k, \sigma(t,x,u)\rangle \\ &\quad + \int_{\mathbf{E}} \langle r(e), g(t,x,u,e)\rangle \pi(de) - l(t,x,u), \text{ a.s.} \end{aligned} \tag{2.23}$$

The maximum condition (2.13) is reduced to the following:

$$\begin{aligned} &\overline{H}(t,\bar{x}(t),\bar{u}(t),p(t),k(t),r(t,\cdot)) - \overline{H}(t,\bar{x}(t),u,p(t),k(t),r(t,\cdot)) \\ &\quad - \frac{1}{2}(\sigma(t,\bar{x}(t),\bar{u}(t)) - \sigma(t,\bar{x}(t),u))^\top P_1(t)(\sigma(t,\bar{x}(t),\bar{u}(t)) - \sigma(t,\bar{x}(t),u)) \\ &\quad - \frac{1}{2}\int_{\mathbf{E}} (g(t,\bar{x}(t),\bar{u}(t),e) - g(t,\bar{x}(t),u,e))^\top [P_1(t) + K_1(t,e)] \\ &\quad \times (g(t,\bar{x}(t),\bar{u}(t),e) - g(t,\bar{x}(t),u,e))\pi(de) \geq 0, \quad \forall u \in \mathbf{U}, \text{ a.e., a.s.} \end{aligned} \tag{2.24}$$

These are the necessary conditions for the stochastic optimal control problem with random jumps (see (2.37) of Tang and Li [24]). When the jump coefficient g is independent of $u(\cdot)$, our result is reduced to Situ [23]. Where there are no random jumps, our result recovers the well-known maximum principle for the classical stochastic optimal control problem of Peng [2] and Yong and Zhou [3]. And when $\mathbf{U}$ is convex, the classical result of Bensoussan [32] is recovered.

(2) *A Stochastic Optimal Control Problem for BSDEJs with Random Jumps*

Consider a controlled BSDEJ:

$$\begin{aligned} -dy(t) &= f(t, y(t), z(t), c(t,\cdot), u(t))dt - z(t)dW(t) - \int_{\mathbf{E}} c(t,e)\widetilde{N}(dedt), \\ y(T) &= h_T, \end{aligned} \tag{2.25}$$

with the cost functional

$$J(u(\cdot)) = \mathbb{E}\left[\int_0^T l(t, y(t), z(t), c(t,\cdot), u(t))dt + \gamma(y(0))\right]. \tag{2.26}$$

In this case,

$$\begin{gathered} b(t,x,y,z,c(\cdot),u) = \sigma(t,x,y,z,c(\cdot),u) = g(t,x,y,z,c(\cdot),u,\cdot) = 0, \\ f(t,x,y,z,c(\cdot),u) = f(t,y,z,c(\cdot),u), \quad l(t,x,y,z,c(\cdot),u) = g(t,y,z,c(\cdot),u,\cdot), \quad \phi(x) = 0. \end{gathered} \tag{2.27}$$

Hence, we have

$$\begin{gathered} x(\cdot) = 0, \qquad p(\cdot) = k(\cdot) = r(\cdot,\cdot) = 0, \\ P_1(\cdot) = Q_1(\cdot) = K_1(\cdot,\cdot) = 0, \qquad P_2(\cdot) = Q_2(\cdot) = K_2(\cdot,\cdot) = 0. \end{gathered} \tag{2.28}$$

Consequently, (2.10) and (2.11) (or equivalently, (2.16)) become

$$\begin{aligned} dq(t) &= \left[\overline{f}_y(t,\cdot)^\top q(t) + \bar{l}_y(t,\cdot)\right]dt - \left[\overline{f}_z(t,\cdot)^\top q(t) + \bar{l}_z(t,\cdot)\right]dW(t) \\ &\quad - \int_{\mathbf{E}}\left[\overline{f}_c(t,e)^\top q(t) + \bar{l}_c(t,e)\right]\widetilde{N}(dedt), \\ q(0) &= \mathbb{E}\gamma_y(\overline{y}_0); \\ -dP_3(t) &= \Big\{ -\overline{f}_y(t,\cdot)^\top P_3(t) - P_3(t)\overline{f}_y(t,\cdot) \\ &\quad + \int_{\mathbf{E}}\Big[\overline{g}_y(t,e)P_1(t)\overline{g}_y(t,e)^\top + \overline{g}_y(t,e)K_1(t,e)\overline{g}_y(t,e)^\top + \overline{g}_y(t,e)^\top K_2(t,e) \\ &\qquad + K_2(t,e)\overline{g}_y(t,e)\Big]\pi(de) - q(t)^\top\overline{f}_{yy}(t,\cdot) - \bar{l}_{yy}(t,\cdot)\Big\}dt - Q_3(t)dW(t) \\ &\quad - \int_{\mathbf{E}} K_3(t,e)\widetilde{N}(dedt), \\ P_3(T) &= 0. \end{aligned} \tag{2.29}$$

The maximum condition (2.13) is reduced to the following:

$$\begin{aligned} &\langle q(t), f(t,\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) - f(t,\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),u)\rangle \\ &\quad + l(t,\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) - l(t,\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),u) \le 0, \quad \forall u \in \mathbf{U}, \text{ a.e., a.s.} \end{aligned} \tag{2.30}$$

As we know, this new result has not been published elsewhere. When there are no random jumps, our result partially recovers that of Dokuchaev and Zhou [33].

(3) *A Stochastic Optimal Control Problem for FBSDEJs with Random Jumps*

Consider a controlled FBSDEJ:

$$\begin{aligned}
dx(t) &= b(t,x(t),u(t))dt + \sigma(t,x(t),u(t))dW(t) + \int_{\mathbf{E}} g(t,x(t-),u(t),e)\widetilde{N}(dedt),\\
-dy(t) &= f\left(t,x(t),y(t),z(t),c(t,\cdot),u(t)\right)dt - z(t)dW(t) - \int_{\mathbf{E}} c(t,e)\widetilde{N}(dedt),\\
x(0) &= a, \qquad y(T) = h(x(T)),
\end{aligned} \tag{2.31}$$

with the cost functional

$$J(u(\cdot)) = \mathbb{E}\left[\int_0^T l\left(t,x(t),y(t),z(t),c(t,\cdot),u(t)\right)dt + \phi(x(T)) + \gamma\left(y(0)\right)\right]. \tag{2.32}$$

In this case, we have

$$\begin{gathered}
b\left(t,x,y,z,c(\cdot),u\right) = b(t,x,u), \qquad \sigma\left(t,x,y,z,c(\cdot),u\right) = \sigma(t,x,u),\\
g\left(t,x,y,z,c(\cdot),u,\cdot\right) = g(t,x,u,\cdot).
\end{gathered} \tag{2.33}$$

Consequently, (2.10) becomes

$$\begin{aligned}
-dp(t) &= \left[\overline{b}_x(t,\cdot)^\top p(t) - \overline{f}_x(t,\cdot)^\top q(t) + \overline{\sigma}_x(t,\cdot)^\top k(t) + \int_{\mathbf{E}} \overline{g}_x(t,e)^\top r(t,e)\pi(de) - \overline{l}_x(t,\cdot)\right]dt\\
&\quad - k(t)dW(t) - \int_{\mathbf{E}} r(t,e)\widetilde{N}(dedt),\\
dq(t) &= \left[\overline{f}_y(t,\cdot)^\top q(t) + \overline{l}_y(t,\cdot)\right]dt - \left[\overline{f}_z(t,\cdot)^\top q(t) + \overline{l}_z(t,\cdot)\right]dW(t)\\
&\quad - \int_{\mathbf{E}} \left[\overline{f}_c(t,e)^\top q(t) + \overline{l}_c(t,e)\right]\widetilde{N}(dedt),\\
p(T) &= \phi_x(\overline{x}(T)) + h_x(\overline{x}(T))^\top q(T), \qquad q(0) = \mathbb{E}\gamma_y\left(\overline{y}_0\right),
\end{aligned} \tag{2.34}$$

and (2.14), (2.15), (2.16) become

$$\begin{aligned}
-dP_1(t) &= \Big\{\overline{b}_x(t,\cdot)^\top P_1(t) + P_1(t)\overline{b}_x(t,\cdot) + \overline{\sigma}_x(t,\cdot)^\top P_1(t)\overline{\sigma}_x(t,\cdot) \\
&\qquad + \overline{\sigma}_x(t,\cdot)^\top Q_1(t) + Q_1(t)\overline{\sigma}_x(t,\cdot) - \overline{f}_x(t,\cdot)^\top P_2(t) - P_2(t)\overline{f}_x(t,\cdot) \\
&\qquad + \int_{\mathbf{E}}\Big[\overline{g}_x(t,e)^\top P_1(t)\overline{g}_x(t,e) + \overline{g}_x(t,e)^\top K_1(t,e)\overline{g}_x(t,e) \\
&\qquad\qquad + \overline{g}_x(t,e)^\top K_1(t,e) + K_1(t,e)\overline{g}_x(t,e)\Big]\pi(de) + \overline{H}_{xx}(t,\cdot)\Big\}dt \\
&\qquad - Q_1(t)dW(t) - \int_{\mathbf{E}} K_1(t,e)\widetilde{N}(dedt), \\
P_1(T) &= -\phi_{xx}(\overline{x}(T)) - h_{xx}(\overline{x}(T))^\top q(T), \\
-dP_2(t) &= \Big\{\overline{b}_x(t,\cdot)P_2(t) + \overline{\sigma}_x(t,\cdot)Q_2(t) - \overline{f}_x(t,\cdot)P_3(t) - P_2(t)^\top\overline{f}_y(t,\cdot) \\
&\qquad + \int_{\mathbf{E}}\Big[\overline{g}_x(t,e)^\top P_1(t)\overline{g}_y(t,e)^\top + \overline{g}_x(t,e)^\top K_1(t,e)\overline{g}_y(t,e)^\top \\
&\qquad\qquad + \overline{g}_x(t,e)K_2(t,e) + K_1(t,e)\overline{g}_y(t,e)^\top\Big]\pi(de) + \overline{H}_{xy}(t,\cdot)\Big\}dt \\
&\qquad - Q_2(t)dW(t) - \int_{\mathbf{E}} K_2(t,e)\widetilde{N}(dedt), \\
P_2(T) &= 0, \\
-dP_3(t) &= \Big\{-\overline{f}_y(t,\cdot)^\top P_3(t) - P_3(t)\overline{f}_y(t,\cdot) - q(t)^\top\overline{f}_{yy}(t,\cdot) - \overline{l}_{yy}(t,\cdot)\Big\}dt \\
&\qquad - Q_3(t)dW(t) - \int_{\mathbf{E}} K_3(t,e)\widetilde{N}(dedt), \\
P_3(T) &= 0,
\end{aligned} \tag{2.35}$$

where the Hamiltonian function $\overline{H} : [0,T]\times\mathbb{M}\times\mathbf{U}\times\mathbf{R}^n\times\mathbf{R}^m\times\mathbf{R}^n\times\mathbf{R}^n \rightarrow \mathbf{R}$ is defined as follows:

$$\begin{aligned}
\overline{H}(t,x,y,z,c(\cdot),u,p,q,k,r(\cdot)) &:= \langle p, b(t,x,u)\rangle + \langle q, -f(t,x,y,z,c(\cdot),u)\rangle + \langle k, \sigma(t,x,u)\rangle \\
&\quad + \int_{\mathbf{E}}\langle r(e), g(t,x,u,e)\rangle\pi(de) - l(t,x,y,z,c(\cdot),u).
\end{aligned} \tag{2.36}$$

The maximum condition (2.13) remains the same. When U is convex, we essentially recover the results in Shi and Wu [26] and Øksendal and Sulem [25] (partial information case), and when σ, g are independent of $u(\cdot)$ and $\mathbf{U}$ is not assumed to be convex, our result is reduced to that of Shi [27]. Note that in both of these cases, our second adjoint equations are new.

When there are no random jumps, our result partially recovers that of Wu [13]. Note that in (3.27) of [13], some additional parameters have to be involved and to be determined. And when $\mathbf{U}$ is convex, our result is reduced to that of Peng [11] and Wang and Wu [34] (partial observation case). Result of Xu [12] is recovered when σ is independent of $u(\cdot)$ and $\mathbf{U}$ is not assumed to be convex.

(4) *A Stochastic Optimal Control Problem for Fully Coupled FBSDEJs with Random Jumps*

Consider the controlled fully coupled FBSDEJ (1.4) with the cost functional (1.5). When $\mathbf{U}$ is convex, we essentially recover the results in Shi and Wu [28] and Meng and Sun [29] (partial information case). When there are no random jumps, our result becomes a special case of Yong [20], because in [20] the author considered the mixed initial-terminal conditions, and some additional necessary conditions for the optimal control are derived. When σ is independent of $u(\cdot)$, our result recovers those of Shi and Wu [19, 35] (partial observation case). And when $\mathbf{U}$ is convex, our result is reduced to that of Wu [17].

3. Problem Reduction and the Proof of the Main Theorem

This section is devoted to the proof of our main theorem. The proof is lengthy and technical. Therefore, we divide it into several steps to make the idea clear.

Step 1 (problem reduction). Consider the following initial value problem for a control system of SDEJs:

$$\begin{aligned}
dx(t) &= b\big(t,x(t),y(t),z(t),c(t,\cdot),u(t)\big)dt + \sigma\big(t,x(t),y(t),z(t),c(t,\cdot),u(t)\big)dW(t)\\
&\qquad + \int_{\mathbf{E}} g\big(t,x(t-),y(t-),z(t),c(t,\cdot),u(t),e\big)\widetilde{N}(dedt),\\
dy(t) &= -f\big(t,x(t),y(t),z(t),c(t,\cdot),u(t)\big)dt + z(t)dW(t) + \int_{\mathbf{E}} c(t,e)\widetilde{N}(dedt),\\
x(0) &= a, \qquad y(0) = y_0,
\end{aligned} \tag{3.1}$$

where $(x(\cdot),y(\cdot))$ is regarded as the state process and $(z(\cdot),c(\cdot,\cdot),u(\cdot))$ as the control process.

It is standard that (see [36]), for any $(z(\cdot),c(\cdot,\cdot),u(\cdot)) \in L^2_{\mathcal{F},p}([0,T];\mathbf{R}^m) \times F^2_p([0,T];\mathbf{R}^m) \times \mathcal{U}[0,T]$ and $y_0 \in \mathbf{R}^m$, there exists a unique strong solution

$$\begin{aligned}
(x(\cdot),y(\cdot)) &\equiv \big(x(\cdot,y_0,z(\cdot),c(\cdot,\cdot),u(\cdot)),y(\cdot,y_0,z(\cdot),c(\cdot,\cdot),u(\cdot))\big) \in L^2_{\mathcal{F}}([0,T];\mathbf{R}^n)\\
&\qquad \times L^2_{\mathcal{F}}([0,T];\mathbf{R}^m)
\end{aligned} \tag{3.2}$$

to (3.1), depending on the 4-tuple $(y_0,z(\cdot),c(\cdot,\cdot),u(\cdot))$. Next, we regard the original terminal condition as the terminal state constraint:

$$y(T) = h(x(T)). \tag{3.3}$$

Let $\mathcal{A}$ be the set of all 4-tuples $(y_0,z(\cdot),c(\cdot,\cdot),u(\cdot)) \in \mathbf{R}^m \times L^2_{\mathcal{F},p}([0,T];\mathbf{R}^m) \times F^2_p([0,T];\mathbf{R}^m) \times \mathcal{U}[0,T]$ such that the unique corresponding state process $(x(\cdot),y(\cdot))$ satisfies the constraint (3.3). Note that hypothesis (H0) implies that, for any $u(\cdot) \in \mathcal{U}[0,T]$, there exists a unique 4-tuple $(y_0,z(\cdot),c(\cdot,\cdot)) \in \mathbf{R}^m \times L^2_{\mathcal{F},p}([0,T];\mathbf{R}^m) \times F^2_p([0,T];\mathbf{R}^m)$ such that state equation (3.1) admits a unique state process $(x(\cdot),y(\cdot)) \in L^2_{\mathcal{F}}([0,T];\mathbf{R}^n) \times L^2_{\mathcal{F}}([0,T];\mathbf{R}^m)$ satisfying the state constraint (3.3). Hence, (H0) and (H1) imply $\mathcal{A} \neq \emptyset$. We rewrite the cost functional (1.5) as follows:

$$J(y_0, z(\cdot), c(\cdot,\cdot), u(\cdot)) = \mathbb{E}\left[\int_0^T l(t, x(t), y(t), z(t), c(t,\cdot), u(t))dt + \phi(x(T)) + \gamma(y_0)\right]. \quad (3.4)$$

Now, we can pose the following.

Problem $\widehat{C}$. Find $(\overline{y}_0, \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot)) \in \mathcal{A}$, such that

$$J(\overline{y}_0, \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot)) = \inf_{(y_0, z(\cdot), c(\cdot,\cdot), u(\cdot)) \in \mathcal{A}} J(y_0, z(\cdot), c(\cdot,\cdot), u(\cdot)). \quad (3.5)$$

We, respectively, refer to $(\overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot))$ as an optimal control process, to $(\overline{x}(\cdot), \overline{y}(\cdot))$ as the corresponding optimal state process, and to $(\overline{y}_0, \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot))$ as an optimal 4-tuple of Problem $\widehat{C}$.

It is clear that Problems C and $\widehat{C}$ are equivalent. The advantage of Problem $\widehat{C}$ is that one does not need much regularity/integrability on $(z(\cdot), c(\cdot,\cdot))$ since it is treated as part of a control process; the disadvantage is that one has to treat terminal constraint (3.3).

Step 2 (*applying Ekeland's variational principle*). For convenience, let us cite Ekeland's variational principle, whose proof can be found in Ekeland [37] or Yong and Zhou [3].

Lemma 3.1. *Let $(V, d(\cdot,\cdot))$ be a complete metric space, and let $f(\cdot) : V \to \mathbf{R}$ be lower-semicontinuous, bounded below. If for all $\rho > 0$ there exists $u \in V$ satisfying*

$$f(u) \leq \inf_{v \in V} f(v) + \rho, \quad (3.6)$$

then there exists $u_\rho \in V$, satisfying the following:

(i) $f(u_\rho) \leq f(u)$,

(ii) $d(u, u_\rho) \leq \sqrt{\rho}$,

(iii) *for all* $v \in V$, $f(v) + \sqrt{\rho} d(v, u_\rho) \geq f(u_\rho)$.

Let $(\overline{y}_0, \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot))$ be an optimal 4-tuple of Problem $\widehat{C}$, with the corresponding optimal state process $(\overline{x}(\cdot), \overline{y}(\cdot))$. For any $\delta > 0$, we define

$$\begin{aligned} &J^\delta(y_0, z(\cdot), c(\cdot,\cdot), u(\cdot)) \\ &\quad \doteq \left\{ \left[J(y_0, z(\cdot), c(\cdot,\cdot), u(\cdot)) - J(\overline{y}_0, \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot)) + \delta\right]^2 + \mathbb{E}\left|y(T) - h(x(T))\right|^2 \right\}^{1/2} \\ &\qquad \forall (y_0, z(\cdot), c(\cdot,\cdot), u(\cdot)) \in \mathbf{R}^m \times L^2_{\mathcal{F},p}([0,T]; \mathbf{R}^m) \times F^2_p([0,T]; \mathbf{R}^m) \times \mathcal{U}[0,T]. \end{aligned} \quad (3.7)$$

If we define the Ekeland's distance on $\mathcal{U}[0,T]$ by

$$d_E(u(\cdot), \widetilde{u}(\cdot)) \doteq \int_\Omega |\{t \in [0,T],\ u(t) \neq \widetilde{u}(t)\}| d\mathbf{P}, \quad \forall u(\cdot), \widetilde{u}(\cdot) \in \mathcal{U}[0,T], \quad (3.8)$$

with $|A|$ being the Lebesgue measure of set $A \subseteq [0,T]$, then $L^2_{\mathcal{F},p}([0,T];\mathbf{R}^m) \times F^2_p([0,T];\mathbf{R}^m) \times \mathcal{U}[0,T]$ is a complete metric space under the following metric:

$$\begin{aligned}
&\left[\|z(\cdot)-\tilde{z}(\cdot)\|^2 + \|c(\cdot,\cdot)-\tilde{c}(\cdot,\cdot)\|^2_{\pi} + d_E(u(\cdot),\tilde{u}(\cdot))^2\right]^{1/2}, \\
&\quad \forall (z(\cdot),c(\cdot,\cdot),u(\cdot)),\ (\tilde{z}(\cdot),\tilde{c}(\cdot,\cdot),\tilde{u}(\cdot)) \in L^2_{\mathcal{F},p}([0,T];\mathbf{R}^m) \times F^2_p([0,T];\mathbf{R}^m) \times \mathcal{U}[0,T],
\end{aligned} \tag{3.9}$$

where $\|z(\cdot)\|^2 \doteq \mathbb{E}\int_0^T |z(t)|^2 dt$, $\|c(\cdot,\cdot)\|^2_{\pi} \doteq \mathbb{E}\int_0^T \int_{\mathbf{E}} |c(t,e)|^2 \pi(de)dt$, and $(y_0, z(\cdot), c(\cdot,\cdot), u(\cdot)) \mapsto J^{\delta}(y_0, z(\cdot), c(\cdot,\cdot), u(\cdot))$ is continuous. Also it is clear that

$$\begin{aligned}
&J^{\delta}(y_0, z(\cdot), c(\cdot,\cdot), u(\cdot)) \\
&\quad > 0, \quad \forall (y_0, z(\cdot), c(\cdot,\cdot), u(\cdot)) \in \mathbf{R}^m \times L^2_{\mathcal{F},p}([0,T];\mathbf{R}^m) \times F^2_p([0,T];\mathbf{R}^m) \times \mathcal{U}[0,T], \\
&J^{\delta}(\overline{y}_0, \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot)) \\
&\quad = \delta \le \inf_{(y_0,z(\cdot),c(\cdot,\cdot),u(\cdot)) \in \mathbf{R}^m \times L^2_{\mathcal{F},p}([0,T];\mathbf{R}^m) \times F^2_p([0,T];\mathbf{R}^m) \times \mathcal{U}[0,T]} J^{\delta}(y_0, z(\cdot), c(\cdot,\cdot), u(\cdot)) + \delta.
\end{aligned} \tag{3.10}$$

Hence, by Lemma 3.1, there exists a 4-tuple $(y_0^{\delta}, z^{\delta}(\cdot), c^{\delta}(\cdot,\cdot), u^{\delta}(\cdot)) \in \mathbf{R}^m \times L^2_{\mathcal{F},p}([0,T];\mathbf{R}^m) \times F^2_p([0,T];\mathbf{R}^m) \times \mathcal{U}[0,T]$ such that

$$\begin{aligned}
&\text{(i) } J^{\delta}\left(y_0^{\delta}, z^{\delta}(\cdot), c^{\delta}(\cdot,\cdot), u^{\delta}(\cdot)\right) \le J^{\delta}(\overline{y}_0, \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot)) = \delta, \\
&\text{(ii) } \left\|z^{\delta}(\cdot) - \overline{z}(\cdot)\right\|^2 + \left\|c^{\delta}(\cdot,\cdot) - \overline{c}(\cdot,\cdot)\right\|^2_{\pi} + d_E\left(u^{\delta}(\cdot), \overline{u}(\cdot)\right)^2 \le \delta, \\
&\text{(iii) } -\sqrt{\delta}\left[\left\|z^{\delta}(\cdot) - \overline{z}(\cdot)\right\|^2 + \left\|c^{\delta}(\cdot,\cdot) - \overline{c}(\cdot,\cdot)\right\|^2_{\pi} + d_E\left(u^{\delta}(\cdot), \overline{u}(\cdot)\right)^2\right]^{1/2} \\
&\quad \le J^{\delta}(y_0, z(\cdot), c(\cdot,\cdot), u(\cdot)) - J^{\delta}\left(y_0^{\delta}, z^{\delta}(\cdot), c^{\delta}(\cdot,\cdot), u^{\delta}(\cdot)\right), \\
&\quad\quad \forall (y_0, z(\cdot), c(\cdot,\cdot), u(\cdot)) \in \mathbf{R}^m \times L^2_{\mathcal{F},p}([0,T];\mathbf{R}^m) \times F^2_p([0,T];\mathbf{R}^m) \times \mathcal{U}[0,T].
\end{aligned} \tag{3.11}$$

Thus, $(y_0^{\delta}, z^{\delta}(\cdot), c^{\delta}(\cdot,\cdot), u^{\delta}(\cdot))$ is a global minimum point of the following penalized cost functional

$$J^{\delta}(y_0, z(\cdot), c(\cdot,\cdot), u(\cdot)) + \sqrt{\delta}\left[\left\|z^{\delta}(\cdot) - \overline{z}(\cdot)\right\|^2 + \left\|c^{\delta}(\cdot,\cdot) - \overline{c}(\cdot,\cdot)\right\|^2_{\pi} + d_E\left(u^{\delta}(\cdot), \overline{u}(\cdot)\right)^2\right]^{1/2}. \tag{3.12}$$

In other words, if we pose a *penalized optimal control problem* with the state constraint (3.3) and the cost functional (3.12), then $(y_0^{\delta}, z^{\delta}(\cdot), c^{\delta}(\cdot,\cdot), u^{\delta}(\cdot))$ is an optimal 4-tuple of the problem. Note that this problem does not have state constraints, and the optimal 4-tuple $(y_0^{\delta}, z^{\delta}(\cdot), c^{\delta}(\cdot,\cdot), u^{\delta}(\cdot))$ approaches $(\overline{y}_0, \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot))$ as $\delta \to 0$.

Step 3 (*nontriviality of the multiplier*). Note that the state process $(x^\delta(\cdot), y^\delta(\cdot))$ corresponding to the 4-tuple $(y_0^\delta, z^\delta(\cdot), c^\delta(\cdot,\cdot), u^\delta(\cdot))$ satisfies the following SDEJ:

$$\begin{aligned} dx^\delta(t) &= b\Big(t, x^\delta(t), y^\delta(t), z^\delta(t), c^\delta(t,\cdot), u^\delta(t)\Big)dt \\ &\quad + \sigma\Big(t, x^\delta(t), y^\delta(t), z^\delta(t), c^\delta(t,\cdot), u^\delta(t)\Big)dW(t) \\ &\quad + \int_{\mathbf{E}} g\Big(t, x^\delta(t-), y^\delta(t-), z^\delta(t), c^\delta(t,\cdot), u^\delta(t), e\Big)\widetilde{N}(dedt), \\ dy^\delta(t) &= -f\Big(t, x^\delta(t), y^\delta(t), z^\delta(t), c^\delta(t,\cdot), u^\delta(t)\Big)dt + z^\delta(t)dW(t) + \int_{\mathbf{E}} c^\delta(t,e)\widetilde{N}(dedt), \\ x^\delta(0) &= a, \qquad y^\delta(0) = y_0^\delta, \end{aligned} \tag{3.13}$$

with

$$\begin{aligned} &\mathbb{E}\Big|y^\delta(T) - h\Big(x^\delta(T)\Big)\Big|^2 \\ &\quad \le J^\delta\Big(y_0^\delta, z^\delta(\cdot), c^\delta(\cdot,\cdot), u^\delta(\cdot)\Big)^2 \le J^\delta(\overline{y}_0, \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot))^2 = \delta^2 \longrightarrow 0, \quad \text{as } \delta \longrightarrow 0. \end{aligned} \tag{3.14}$$

We now regard $\Theta^\delta(\cdot) \equiv (x^\delta(\cdot), y^\delta(\cdot), z^\delta(\cdot), c^\delta(\cdot,\cdot))$ as the unique solution to the following FBSDEJ:

$$\begin{aligned} dx^\delta(t) &= b\Big(t, x^\delta(t), y^\delta(t), z^\delta(t), c^\delta(t,\cdot), u^\delta(t)\Big)dt \\ &\quad + \sigma\Big(t, x^\delta(t), y^\delta(t), z^\delta(t), c^\delta(t,\cdot), u^\delta(t)\Big)dW(t) \\ &\quad + \int_{\mathbf{E}} g\Big(t, x^\delta(t-), y^\delta(t-), z^\delta(t), c^\delta(t,\cdot), u^\delta(t), e\Big)\widetilde{N}(dedt), \\ -dy^\delta(t) &= f\Big(t, x^\delta(t), y^\delta(t), z^\delta(t), c^\delta(t,\cdot), u^\delta(t)\Big)dt - z^\delta(t)dW(t) - \int_{\mathbf{E}} c^\delta(t,e)\widetilde{N}(dedt), \\ x^\delta(0) &= a, \qquad y^\delta(T) = h\Big(y^\delta(T)\Big) + h_T^\delta, \end{aligned} \tag{3.15}$$

with

$$h_T^\delta \equiv y^\delta(T) - h\Big(x^\delta(T)\Big). \tag{3.16}$$

Next, for any $\eta_T \in L^2_{\mathcal{F}_T}(\Omega; \mathbf{R}^m)$, with

$$\mathbb{E}|\eta_T|^2 \le 1, \tag{3.17}$$

by (H0), the following FBSDEJ

$$\begin{aligned} dx^{\delta,\varepsilon}(t) &= b\Big(t, x^{\delta,\varepsilon}(t), y^{\delta,\varepsilon}(t), z^{\delta,\varepsilon}(t), c^{\delta,\varepsilon}(t,\cdot), u^{\delta}(t)\Big)dt \\ &\quad + \sigma\Big(t, x^{\delta,\varepsilon}(t), y^{\delta,\varepsilon}(t), z^{\delta,\varepsilon}(t), c^{\delta,\varepsilon}(t,\cdot), u^{\delta}(t)\Big)dW(t) \\ &\quad + \int_{\mathbf{E}} g\Big(t, x^{\delta,\varepsilon}(t-), y^{\delta,\varepsilon}(t-), z^{\delta,\varepsilon}(t), c^{\delta,\varepsilon}(t,\cdot), u^{\delta}(t), e\Big)\widetilde{N}(dedt), \\ -dy^{\delta,\varepsilon}(t) &= f\Big(t, x^{\delta,\varepsilon}(t), y^{\delta,\varepsilon}(t), z^{\delta,\varepsilon}(t), c^{\delta,\varepsilon}(t,\cdot), u^{\delta}(t)\Big)dt \\ &\quad - z^{\delta,\varepsilon}(t)dW(t) - \int_{\mathbf{E}} c^{\delta,\varepsilon}(t,e)\widetilde{N}(dedt), \\ x^{\delta,\varepsilon}(0) &= a, \qquad y^{\delta,\varepsilon}(T) = h\Big(y^{\delta,\varepsilon}(T)\Big) + h_T^{\delta} + \varepsilon\eta_T, \end{aligned} \tag{3.18}$$

admits a unique adapted solution $\Theta^{\delta,\varepsilon}(\cdot) \equiv (x^{\delta,\varepsilon}(\cdot), y^{\delta,\varepsilon}(\cdot), z^{\delta,\varepsilon}(\cdot), c^{\delta,\varepsilon}(\cdot,\cdot))$. Note that FBSDEJ (3.18) is nothing but FBSDEJ (3.15) with only h_T^{δ} replaced by $h_T^{\delta} + \varepsilon\eta_T$, and $u^{\delta}(\cdot)$ remains unchanged. Thus, by (1.9), we have

$$\left\|\Theta^{\delta,\varepsilon}(\cdot) - \Theta^{\delta}(\cdot)\right\|_{\mathcal{M}^2[0,T]} \le K\varepsilon^2, \quad \text{as } \varepsilon \longrightarrow 0, \tag{3.19}$$

which implies

$$\lim_{\varepsilon\to 0}\left(\left|y^{\delta,\varepsilon}(0) - y^{\delta}(0)\right|^2 + \mathbb{E}\int_0^T \left|z^{\delta,\varepsilon}(t) - z^{\delta}(t)\right|^2 dt + \mathbb{E}\int_0^T\int_{\mathbf{E}} \left|c^{\delta,\varepsilon}(t,e) - c^{\delta}(t,e)\right|^2 \pi(de)dt\right) = 0. \tag{3.20}$$

Hence, with $y_0^{\delta,\varepsilon} = y^{\delta,\varepsilon}(0)$, by taking $(y_0, z(\cdot), c(\cdot,\cdot), u(\cdot)) = (y_0^{\delta,\varepsilon}, z^{\delta,\varepsilon}(\cdot), c^{\delta,\varepsilon}(\cdot,\cdot), u^{\delta}(\cdot))$ in the last relation in (3.11), we have

$$\begin{aligned} -K\sqrt{\delta}\varepsilon &\le J^{\delta}\Big(y_0^{\delta,\varepsilon}, z^{\delta,\varepsilon}(\cdot), c^{\delta,\varepsilon}(\cdot,\cdot), u^{\delta}(\cdot)\Big) - J^{\delta}\Big(y_0^{\delta}, z^{\delta}(\cdot), c^{\delta}(\cdot,\cdot), u^{\delta}(\cdot)\Big) \\ &= \frac{J^{\delta}\Big(y_0^{\delta,\varepsilon}, z^{\delta,\varepsilon}(\cdot), c^{\delta,\varepsilon}(\cdot,\cdot), u^{\delta}(\cdot)\Big)^2 - J^{\delta}\Big(y_0^{\delta}, z^{\delta}(\cdot), c^{\delta}(\cdot,\cdot), u^{\delta}(\cdot)\Big)^2}{J^{\delta}\Big(y_0^{\delta,\varepsilon}, z^{\delta,\varepsilon}(\cdot), c^{\delta,\varepsilon}(\cdot,\cdot), u^{\delta}(\cdot)\Big) + J^{\delta}\Big(y_0^{\delta}, z^{\delta}(\cdot), c^{\delta}(\cdot,\cdot), u^{\delta}(\cdot)\Big)} \\ &= \frac{\Big[J\Big(y_0^{\delta,\varepsilon}, z^{\delta,\varepsilon}(\cdot), c^{\delta,\varepsilon}(\cdot,\cdot), u^{\delta}(\cdot)\Big) - J(\overline{y}_0, \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot)) + \delta\Big]^2}{J^{\delta}\Big(y_0^{\delta,\varepsilon}, z^{\delta,\varepsilon}(\cdot), c^{\delta,\varepsilon}(\cdot,\cdot), u^{\delta}(\cdot)\Big) + J^{\delta}\Big(y_0^{\delta}, z^{\delta}(\cdot), c^{\delta}(\cdot,\cdot), u^{\delta}(\cdot)\Big)} \end{aligned}$$

$$
\begin{aligned}
&- \frac{\left[J\left(y_0^\delta, z^\delta(\cdot), c^\delta(\cdot,\cdot), u^\delta(\cdot)\right) - J(\overline{y}_0, \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot)) + \delta\right]^2}{J^\delta\left(y_0^{\delta,\varepsilon}, z^{\delta,\varepsilon}(\cdot), c^{\delta,\varepsilon}(\cdot,\cdot), u^\delta(\cdot)\right) + J^\delta\left(y_0^\delta, z^\delta(\cdot), c^\delta(\cdot,\cdot), u^\delta(\cdot)\right)} \\
&+ \frac{\mathbb{E}\left[\left|y^{\delta,\varepsilon}(T) - h(x^{\delta,\varepsilon}(T))\right|^2 - \left|y^\delta(T) - h(x^\delta(T))\right|^2\right]}{J^\delta\left(y_0^{\delta,\varepsilon}, z^{\delta,\varepsilon}(\cdot), c^{\delta,\varepsilon}(\cdot,\cdot), u^\delta(\cdot)\right) + J^\delta\left(y_0^\delta, z^\delta(\cdot), c^\delta(\cdot,\cdot), u^\delta(\cdot)\right)} \\
&\equiv \Phi_0^{\delta,\varepsilon}\left[J\left(y_0^{\delta,\varepsilon}, z^{\delta,\varepsilon}(\cdot), c^{\delta,\varepsilon}(\cdot,\cdot), u^\delta(\cdot)\right) - J\left(y_0^\delta, z^\delta(\cdot), c^\delta(\cdot,\cdot), u^\delta(\cdot)\right)\right] \\
&+ \mathbb{E}\left\langle \Phi_T^{\delta,\varepsilon}, y^{\delta,\varepsilon}(T) - h\left(x^{\delta,\varepsilon}(T)\right) - h_T^\delta \right\rangle \\
&\equiv \left(\Phi_0^\delta + o(1)\right)\left[J\left(y_0^{\delta,\varepsilon}, z^{\delta,\varepsilon}(\cdot), c^{\delta,\varepsilon}(\cdot,\cdot), u^\delta(\cdot)\right) - J\left(y_0^\delta, z^\delta(\cdot), c^\delta(\cdot,\cdot), u^\delta(\cdot)\right)\right] \\
&+ \varepsilon\mathbb{E}\left[\left\langle \Phi_T^\delta + o(1), \eta_T \right\rangle\right],
\end{aligned}
\tag{3.21}
$$

where $o(1)$ stands for certain scalars or vectors that go to 0 as $\varepsilon \to 0$,

$$
\begin{aligned}
\Phi_0^{\delta,\varepsilon} &\doteq \frac{2}{J^\delta\left(y_0^{\delta,\varepsilon}, z^{\delta,\varepsilon}(\cdot), c^{\delta,\varepsilon}(\cdot,\cdot), u^\delta(\cdot)\right) + J^\delta\left(y_0^\delta, z^\delta(\cdot), c^\delta(\cdot,\cdot), u^\delta(\cdot)\right)} \\
&\quad \times \left\{\int_0^1 \left[\beta\left(J\left(y_0^{\delta,\varepsilon}, z^{\delta,\varepsilon}(\cdot), c^{\delta,\varepsilon}(\cdot,\cdot), u^\delta(\cdot)\right) - J(\overline{y}_0, \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot))\right)\right.\right. \\
&\qquad \left.\left. + (1-\beta)\left(J\left(y_0^\delta, z^\delta(\cdot), c^\delta(\cdot,\cdot), u^\delta(\cdot)\right) - J(\overline{y}_0, \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot))\right)\right] d\beta + \delta\right\}, \\
\Phi_T^{\delta,\varepsilon} &\doteq \frac{y^{\delta,\varepsilon}(T) - h(x^{\delta,\varepsilon}(T)) + y^\delta(T) - h(x^\delta(T))}{J^\delta\left(y_0^{\delta,\varepsilon}, z^{\delta,\varepsilon}(\cdot), c^{\delta,\varepsilon}(\cdot,\cdot), u^\delta(\cdot)\right) + J^\delta\left(y_0^\delta, z^\delta(\cdot), c^\delta(\cdot,\cdot), u^\delta(\cdot)\right)},
\end{aligned}
\tag{3.22}
$$

$$
\begin{aligned}
\Phi_0^\delta &\doteq \frac{J\left(y_0^\delta, z^\delta(\cdot), c^\delta(\cdot,\cdot), u^\delta(\cdot)\right) - J(\overline{y}_0, \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot)) + \delta}{J^\delta\left(y_0^\delta, z^\delta(\cdot), c^\delta(\cdot,\cdot), u^\delta(\cdot)\right)} \in [0,1], \\
\Phi_T^\delta &\doteq \frac{y^\delta(T) - h(x^\delta(T))}{J^\delta\left(y_0^\delta, z^\delta(\cdot), c^\delta(\cdot,\cdot), u^\delta(\cdot)\right)} \in L^2_{\mathcal{F}_T}(\Omega; \mathbf{R}^m).
\end{aligned}
\tag{3.23}
$$

We point out that $(\Phi_0^\delta, \Phi_T^\delta)$ is independent of η_T, and

$$
\Phi_0^\delta \geq 0, \quad \left|\Phi_0^\delta\right|^2 + \mathbf{E}\left|\Phi_T^\delta\right|^2 = 1. \tag{3.24}
$$

Then there is a subsequence of $(\Phi_0^\delta, \Phi_T^\delta)$, still denoted by $(\Phi_0^\delta, \Phi_T^\delta)$, such that

$$\lim_{\delta\to 0}\left(\Phi_0^\delta, \Phi_T^\delta\right) = (\Phi_0, \Phi_T), \quad \text{with } |\Phi_0|^2 + \mathbf{E}|\Phi_T|^2 = 1. \tag{3.25}$$

We claim that

$$\Phi_0 \neq 0. \tag{3.26}$$

To show this, let us observe that by (3.19), one has

$$\left(\Phi_0^\delta + o(1)\right)\left[J\left(y_0^{\delta,\varepsilon}, z^{\delta,\varepsilon}(\cdot), c^{\delta,\varepsilon}(\cdot,\cdot), u^\delta(\cdot)\right) - J^\delta\left(y_0^\delta, z^\delta(\cdot), c^\delta(\cdot,\cdot), u^\delta(\cdot)\right)\right] \leq \varepsilon K\left(\Phi_0^\delta + o(1)\right). \tag{3.27}$$

Hence, dividing by ε in (3.21) and then letting $\varepsilon \to 0$ yields

$$-K\sqrt{\delta} \leq K\left|\Phi_0^\delta\right| + \mathbb{E}\left\langle \Phi_T^\delta, \eta_T \right\rangle, \tag{3.28}$$

with the constant K independent of η_T. Now, if $\Phi_0 = 0$, then the above leads to

$$-r^\delta \leq \mathbb{E}\left\langle \Phi_T^\delta, \eta_T \right\rangle, \quad \forall \eta_T \in L^2_{\mathcal{F}_T}(\Omega; \mathbf{R}^m), \qquad \mathbb{E}\left|\eta_T\right|^2 \leq 1, \tag{3.29}$$

with $r^\delta \to 0$, as $\delta \to 0$, uniformly in $\mathbb{E}|\eta_T|^2 \leq 1$. Then,

$$\left(\mathbb{E}\left|\Phi_T^\delta\right|^2\right)^{1/2} = \sup_{\mathbb{E}|\eta_T|^2 \leq 1} \mathbb{E}\left\langle \Phi_T^\delta, \eta_T \right\rangle \leq \left|r^\delta\right| \longrightarrow 0, \quad \text{as } \delta \longrightarrow 0. \tag{3.30}$$

Thus, for $\delta > 0$ small enough, we must have

$$1 = \left|\Phi_0^\delta\right|^2 + \mathbf{E}\left|\Phi_T^\delta\right|^2 < 1. \tag{3.31}$$

It is a contradiction which proves (3.26).

We refer to (Φ_0, Φ_T) as the *Lagrange multiplier* of the corresponding optimal 4-tuple $(\overline{y}_0, \overline{z}(\cdot), \overline{c}(\cdot,\cdot), \overline{u}(\cdot))$. Equation (3.26) shows the nontriviality of the Lagrange multiplier (Φ_0, Φ_T).

Step 4 (*spike variations*). For notational simplicity, we now denote

$$
\begin{gathered}
X = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad v(\cdot) = \begin{pmatrix} z \\ c(\cdot) \\ u \end{pmatrix}, \qquad X_0 = \begin{pmatrix} a \\ y_0 \end{pmatrix}, \qquad X_T = X(T) = \begin{pmatrix} x(T) \\ y(T) \end{pmatrix}, \\
B(t, X, v(\cdot)) = \begin{pmatrix} b(t, x, y, z, c(\cdot), u) \\ -f(t, x, y, z, c(\cdot), u) \end{pmatrix}, \\
\Sigma(t, X, v(\cdot)) = \begin{pmatrix} \sigma(t, x, y, z, c(\cdot), u) \\ z \end{pmatrix}, \\
\Upsilon(t, X, v(\cdot)) = \begin{pmatrix} g(t, x, y, z, c(\cdot), u, \cdot) \\ c(\cdot) \end{pmatrix}, \\
\Gamma(X_0, X_T) = \phi(x_T) + \gamma(y_0), \quad \Pi(X_0, X_T) = \begin{pmatrix} 0 \\ y_T - h(x_T) \end{pmatrix}, \\
\mathbf{R}^l = \mathbf{R}^n \times \mathbf{R}^m, \quad \mathcal{V}[0, T] = L^2_{\mathcal{F},p}([0, T]; \mathbf{R}^m) \times F^2_p([0, T]; \mathbf{R}^m) \times \mathcal{U}[0, T], \\
\mathcal{H} = \mathbf{R}^l \times L^2_{\mathcal{F}_T}\left(\Omega; \mathbf{R}^l\right) \equiv \mathbf{R}^l \times \mathcal{X}^2_l, \quad \mathcal{H}_0 = \mathbf{R}^n \times L^2_{\mathcal{F}_T}(\Omega; \mathbf{R}^m) \equiv \mathbf{R}^n \times \mathcal{X}^2_m.
\end{gathered} \tag{3.32}
$$

Consequently,

$$
\begin{gathered}
J(y_0, z(\cdot), c(\cdot, \cdot), u(\cdot)) = J(y_0, v(\cdot, \cdot)), \\
J^\delta(y_0, z(\cdot), c(\cdot, \cdot), u(\cdot)) = J^\delta(y_0, v(\cdot, \cdot)).
\end{gathered} \tag{3.33}
$$

Note that $\mathcal{H}$ and $\mathcal{H}_0$ are Hilbert spaces. We identify $\mathcal{H}^* = \mathcal{H}$ and $\mathcal{H}_0^* = \mathcal{H}_0$. Also, from the above,

$$
\Gamma : \mathcal{H} \longrightarrow \mathbf{R}, \qquad \Pi : \mathcal{H} \longrightarrow \mathcal{H}_0. \tag{3.34}
$$

We denote the gradient $D\Gamma$ and the Hessian $D^2\Gamma$ of Γ as follows:

$$
\begin{gathered}
D\Gamma(X_0, X_T) \equiv (\Gamma_{X_0}(X_0, X_T), \Gamma_{X_T}(X_0, X_T)) \in \mathcal{L}(\mathcal{H}; \mathbf{R}) \equiv \mathcal{H}^* = \mathcal{H}, \\
D^2\Gamma(X_0, X_T) \equiv \begin{pmatrix} \Gamma_{X_0X_0}(X_0, X_T) & \Gamma_{X_0X_T}(X_0, X_T) \\ \Gamma_{X_TX_0}(X_0, X_T) & \Gamma_{X_TX_T}(X_0, X_T) \end{pmatrix} \in \mathcal{L}_s(\mathcal{H}; \mathcal{H}),
\end{gathered} \tag{3.35}
$$

where $\mathcal{L}(\mathcal{H}_1; \mathcal{H}_2)$ is the set of all linear bounded operators from $\mathcal{H}_1$ to $\mathcal{H}_2$, and $\mathcal{L}_s(\mathcal{H}; \mathcal{H})$ is the set of all linear bounded self-adjoint operators from $\mathcal{H}$ to itself. Clearly,

$$
\begin{gathered}
\Gamma_{X_0}(X_0, X_T) = (0, \gamma_y(y_0))^\top \in \mathbf{R}^l, \\
\Gamma_{X_T}(X_0, X_T) = (\phi_x(x_T), 0)^\top \in \mathcal{X}^2_l,
\end{gathered}
$$

$$\Gamma_{X_0X_0}(X_0, X_T) = \begin{pmatrix} 0 & 0 \\ 0 & \gamma_{yy}(y_0) \end{pmatrix} \in \mathcal{S}^{m+n},$$
$$\Gamma_{X_0X_T}(X_0, X_T) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{L}\left(\mathcal{X}_l^2; \mathbf{R}^l\right),$$
$$\Gamma_{X_TX_0}(X_0, X_T) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{L}\left(\mathbf{R}^l; \mathcal{X}_l^2\right),$$
$$\Gamma_{X_TX_T}(X_0, X_T) = \begin{pmatrix} \phi_{xx}(x_T) & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{L}\left(\mathcal{X}_l^2; \mathcal{X}_l^2\right). \tag{3.36}$$

Also, by the Fréchet differentiability, for $\Pi : \mathscr{H} \to \mathscr{H}_0$, we have

$$\begin{gathered} D\Pi(X_0, X_T) \equiv (\Pi_{X_0}(X_0, X_T), \Pi_{X_T}(X_0, X_T)) \in \mathcal{L}(\mathscr{H}; \mathscr{H}_0), \\ D^2\Pi(X_0, X_T) \equiv \begin{pmatrix} \Pi_{X_0X_0}(X_0, X_T) & \Pi_{X_0X_T}(X_0, X_T) \\ \Pi_{X_TX_0}(X_0, X_T) & \Pi_{X_TX_T}(X_0, X_T) \end{pmatrix} \in \mathcal{L}(\mathscr{H}; \mathcal{L}(\mathscr{H}; \mathscr{H}_0)), \end{gathered} \tag{3.37}$$

To make the above more precise, let us take any $\widehat{\Phi} \equiv (\widehat{\Phi}_0, \widehat{\Phi}_T) \in \mathscr{H}_0$. Then,

$$\left\langle \Pi(X_0, X_T), \widehat{\Phi} \right\rangle = \left\langle y_T - h(x_T), \widehat{\Phi}_T \right\rangle. \tag{3.38}$$

Thus,

$$\begin{aligned} \left[D\Pi(X_0, X_T)\widehat{\Phi}\right] &\equiv D\left[\left\langle \Pi(X_0, X_T), \widehat{\Phi} \right\rangle\right] \\ &= \left(\left\langle \Pi(X_0, X_T), \widehat{\Phi} \right\rangle_{X_0}, \left\langle \Pi(X_0, X_T), \widehat{\Phi} \right\rangle_{X_T} \right), \\ &\equiv \left(\Pi_{X_0}(X_0, X_T)\widehat{\Phi}, \Pi_{X_T}(X_0, X_T)\widehat{\Phi} \right) \in \mathcal{L}(\mathscr{H}; \mathbf{R}), \end{aligned} \tag{3.39}$$

with

$$\Pi_{X_0}(X_0, X_T)\widehat{\Phi} \equiv \left\langle \Pi(X_0, X_T), \widehat{\Phi} \right\rangle_{X_0} = (0, 0),$$
$$\Pi_{X_T}(X_0, X_T)\widehat{\Phi} \equiv \left\langle \Pi(X_0, X_T), \widehat{\Phi} \right\rangle_{X_T} = \left(-h_x(x_T)^\top \widehat{\Phi}_T, \widehat{\Phi}_T \right);$$
$$\left[D^2\Pi(X_0, X_T)\widehat{\Phi}\right] \equiv D^2\left[\left\langle \Pi(X_0, X_T), \widehat{\Phi} \right\rangle\right]$$

$$= \begin{pmatrix} \left\langle \Pi(X_0, X_T), \widehat{\Phi} \right\rangle_{X_0X_0} & \left\langle \Pi(X_0, X_T), \widehat{\Phi} \right\rangle_{X_0X_T} \\ \left\langle \Pi(X_0, X_T), \widehat{\Phi} \right\rangle_{X_TX_0} & \left\langle \Pi(X_0, X_T), \widehat{\Phi} \right\rangle_{X_TX_T} \end{pmatrix}$$
$$\equiv \begin{pmatrix} \Pi_{X_0X_0}(X_0, X_T)\widehat{\Phi} & \Pi_{X_0X_T}(X_0, X_T)\widehat{\Phi} \\ \Pi_{X_TX_0}(X_0, X_T)\widehat{\Phi} & \Pi_{X_TX_T}(X_0, X_T)\widehat{\Phi} \end{pmatrix} \in \mathcal{L}_s(\mathcal{H}; \mathcal{H}), \tag{3.40}$$

with

$$\begin{aligned} \Pi_{X_0X_0}(X_0, X_T)\widehat{\Phi} &\equiv \left\langle \Pi(X_0, X_T), \widehat{\Phi} \right\rangle_{X_0X_0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \\ \Pi_{X_TX_0}(X_0, X_T)\widehat{\Phi} &\equiv \left\langle \Pi(X_0, X_T), \widehat{\Phi} \right\rangle_{X_TX_0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \\ \Pi_{X_0X_T}(X_0, X_T)\widehat{\Phi} &\equiv \left\langle \Pi(X_0, X_T), \widehat{\Phi} \right\rangle_{X_0X_T} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \\ \Pi_{X_TX_T}(X_0, X_T)\widehat{\Phi} &\equiv \left\langle \Pi(X_0, X_T), \widehat{\Phi} \right\rangle_{X_TX_T} = \begin{pmatrix} -h_{xx}(x_T)\widehat{\Phi}_T & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned} \tag{3.41}$$

Next, for fixed $X_0 \in \mathbf{R}^l, v(\cdot) \in \mathcal{V}[0,T]$ and any $\varepsilon \in (0,1)$, let

$$X_0^{\delta,\varepsilon} \doteq X_0^{\delta} + \begin{pmatrix} 0 \\ \sqrt{\varepsilon} y_0 \end{pmatrix}, \tag{3.42}$$

$$v^{\delta,\varepsilon}(t,\cdot) = \begin{cases} v^{\delta}(t,\cdot), & t \in [0,T] \setminus S_\varepsilon, \\ v(t,\cdot), & t \in S_\varepsilon, \end{cases} \tag{3.43}$$

for some measurable set $S_\varepsilon \subseteq [0,T]$ with $|S_\varepsilon| = \varepsilon T$. It is clear that $v^{\delta,\varepsilon}(\cdot,\cdot) \in \mathcal{V}[0,T]$.

Let $X^{\delta,\varepsilon}(\cdot)$ be the state process of (3.1) corresponding to $(X_0^{\delta,\varepsilon}, v^{\delta,\varepsilon}(\cdot,\cdot))$. Further, let $X_1^{\delta,\varepsilon}(\cdot)$ and $X_2^{\delta,\varepsilon}(\cdot)$ be, respectively, the solutions to the following SDEJs:

$$\begin{aligned} dX_1^{\delta,\varepsilon}(t) &= B_X^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t)dt + \left[\Sigma_X^{\delta} + \Delta\Sigma^{\delta}(t,\cdot)I_{S_\varepsilon}(t)\right]dW(t) \\ &\quad + \int_{\mathbf{E}} \left[\Upsilon_X^{\delta}(t,e)X_1^{\delta,\varepsilon}(t-) + \Delta\Upsilon^{\delta}(t,e)I_{S_\varepsilon}(t)\right]\widetilde{N}(dedt), \\ X_1^{\delta,\varepsilon}(0) &= \begin{pmatrix} 0 \\ \sqrt{\varepsilon} y_0 \end{pmatrix}; \\ dX_2^{\delta,\varepsilon}(t) &= \left[B_X^{\delta}(t,\cdot)X_2^{\delta,\varepsilon}(t) + \Delta B^{\delta}(t,\cdot)I_{S_\varepsilon}(t) + \frac{1}{2}B_{XX}^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t)^2\right]dt \\ &\quad + \left[\Sigma_X^{\delta}(t,\cdot)X_2^{\delta,\varepsilon}(t) + \Delta\Sigma_X^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t)I_{S_\varepsilon}(t) + \frac{1}{2}\Sigma_{XX}^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t)^2\right]dW(t) \end{aligned} \tag{3.44}$$

$$+\int_{\mathbf{E}}\left[\Upsilon_X^{\delta}(t,e)X_2^{\delta,\varepsilon}(t-)+\Delta\Upsilon_X^{\delta}(t,e)X_1^{\delta,\varepsilon}(t-)I_{S_\varepsilon}(t)+\frac{1}{2}\Upsilon_{XX}^{\delta}(t,e)X_1^{\delta,\varepsilon}(t-)^2\right]\widetilde{N}(dedt),$$
$$X_2^{\delta,\varepsilon}(0)=0, \tag{3.45}$$

where $I_{S_\varepsilon}(\cdot)$ denotes the indicator function of the set S_ε, and for any $X\in\mathbf{R}^l$,

$$\begin{aligned}
B_X^{\delta}(t,\cdot)&=B_X\Big(t,X^{\delta}(t),v^{\delta}(t,\cdot)\Big),\\
\Sigma_X^{\delta}(t,\cdot)&=\Sigma_X\Big(t,X^{\delta}(t),v^{\delta}(t,\cdot)\Big),\\
\Upsilon_X^{\delta}(t,\cdot)&=\Upsilon_X\Big(t,X^{\delta}(t),v^{\delta}(t,\cdot)\Big),\\
\Delta B^{\delta}(t,\cdot)&=B\Big(t,X^{\delta}(t),v(t,\cdot)\Big)-B\Big(t,X^{\delta}(t),v^{\delta}(t,\cdot)\Big),\\
\Delta\Sigma^{\delta}(t,\cdot)&=\Sigma\Big(t,X^{\delta}(t),v(t,\cdot)\Big)-\Sigma\Big(t,X^{\delta}(t),v^{\delta}(t,\cdot)\Big),\\
\Delta\Upsilon^{\delta}(t,\cdot)&=\Upsilon\Big(t,X^{\delta}(t),v(t,\cdot)\Big)-\Upsilon\Big(t,X^{\delta}(t),v^{\delta}(t,\cdot)\Big),\\
\Delta\Sigma_X^{\delta}(t,\cdot)&=\Sigma_X\Big(t,X^{\delta}(t),v(t,\cdot)\Big)-\Sigma_X\Big(t,X^{\delta}(t),v^{\delta}(t,\cdot)\Big),\\
\Delta\Upsilon_X^{\delta}(t,\cdot)&=\Upsilon_X\Big(t,X^{\delta}(t),v(t,\cdot)\Big)-\Upsilon_X\Big(t,X^{\delta}(t),v^{\delta}(t,\cdot)\Big),\\
B_{XX}^{\delta}(t,\cdot)X^2&=\begin{pmatrix}\left\langle B_{XX}^{1,\delta}(t,\cdot)X,X\right\rangle\\ \left\langle B_{XX}^{2,\delta}(t,\cdot)X,X\right\rangle\\ \vdots\\ \left\langle B_{XX}^{l,\delta}(t,\cdot)X,X\right\rangle\end{pmatrix},\qquad B_{XX}^{i,\delta}(t,\cdot)=B_{XX}^{i}\Big(t,X^{\delta}(t),v^{\delta}(t,\cdot)\Big),\quad 1\le i\le l,\\
\Sigma_{XX}^{\delta}(t,\cdot)X^2&=\begin{pmatrix}\left\langle \Sigma_{XX}^{1,\delta}(t,\cdot)X,X\right\rangle\\ \left\langle \Sigma_{XX}^{2,\delta}(t,\cdot)X,X\right\rangle\\ \vdots\\ \left\langle \Sigma_{XX}^{l,\delta}(t,\cdot)X,X\right\rangle\end{pmatrix},\qquad \Sigma_{XX}^{i,\delta}(t,\cdot)=\Sigma_{XX}^{i}\Big(t,X^{\delta}(t),v^{\delta}(t,\cdot)\Big),\quad 1\le i\le l,\\
\Upsilon_{XX}^{\delta}(t,\cdot)X^2&=\begin{pmatrix}\left\langle \Upsilon_{XX}^{1,\delta}(t,\cdot)X,X\right\rangle\\ \left\langle \Upsilon_{XX}^{2,\delta}(t,\cdot)X,X\right\rangle\\ \vdots\\ \left\langle \Upsilon_{XX}^{l,\delta}(t,\cdot)X,X\right\rangle\end{pmatrix},\qquad \Upsilon_{XX}^{i,\delta}(t,\cdot)=\Upsilon_{XX}^{i}\Big(t,X^{\delta}(t),v^{\delta}(t,\cdot)\Big),\quad 1\le i\le l.
\end{aligned}\tag{3.46}$$

We point out that everything is reduced to the classical optimal control problems for SDEJs. Then by Lemma 2.1 of Tang and Li [24], the following estimates hold for any integers $k \geq 1$:

$$\begin{aligned}
&\sup_{0\leq t\leq T}\mathbb{E}\left|X_1^{\delta,\varepsilon}(t)\right|^{2k} + \sup_{0\leq t\leq T}\mathbb{E}\left|X^{\delta,\varepsilon}(t) - X^{\delta}(t)\right|^{2k} \leq C\varepsilon^{k},\\
&\sup_{0\leq t\leq T}\mathbb{E}\left|X_2^{\delta,\varepsilon}(t)\right|^{2k} + \sup_{0\leq t\leq T}\mathbb{E}\left|X^{\delta,\varepsilon}(t) - X^{\delta}(t) - X_1^{\delta,\varepsilon}(t)\right|^{2k} \leq C\varepsilon^{2k},\\
&\sup_{0\leq t\leq T}\mathbb{E}\left|X^{\delta,\varepsilon}(t) - X^{\delta}(t) - X_1^{\delta,\varepsilon}(t) - X_2^{\delta,\varepsilon}(t)\right|^{2k} = o\left(\varepsilon^{2k}\right).
\end{aligned} \tag{3.47}$$

Now, from the last relation in (3.11), noting (3.19), we obtain

$$\begin{aligned}
&-\sqrt{\delta}\left[\sqrt{\varepsilon}y_0 + \varepsilon(T+K)\right]\\
&\quad\leq J^{\delta}\left(y_0^{\delta,\varepsilon}, v^{\delta,\varepsilon}(\cdot,\cdot)\right) - J^{\delta}\left(y_0^{\delta}, v^{\delta}(\cdot,\cdot)\right)\\
&\quad= \frac{J^{\delta}\left(y_0^{\delta,\varepsilon}, v^{\delta,\varepsilon}(\cdot,\cdot)\right)^2 - J^{\delta}\left(y_0^{\delta}, v^{\delta}(\cdot,\cdot)\right)^2}{J^{\delta}\left(y_0^{\delta,\varepsilon}, v^{\delta,\varepsilon}(\cdot,\cdot)\right) + J^{\delta}\left(y_0^{\delta}, v^{\delta}(\cdot,\cdot)\right)}\\
&\quad= \frac{\left[J\left(y_0^{\delta,\varepsilon}, v^{\delta,\varepsilon}(\cdot,\cdot)\right) - J(\overline{y}_0, \overline{v}(\cdot,\cdot)) + \delta\right]^2 - \left[J\left(y_0^{\delta}, v^{\delta}(\cdot,\cdot)\right) - J(\overline{y}_0, \overline{v}(\cdot,\cdot)) + \delta\right]^2}{J^{\delta}\left(y_0^{\delta,\varepsilon}, v^{\delta,\varepsilon}(\cdot,\cdot)\right) + J^{\delta}\left(y_0^{\delta}, v^{\delta}(\cdot,\cdot)\right)}\\
&\qquad+ \frac{\mathbb{E}\left[\Pi\left(X_0^{\delta,\varepsilon}, X^{\delta,\varepsilon}(T)\right)^2 - \Pi\left(X_0^{\delta}, X^{\delta}(T)\right)^2\right]}{J^{\delta}\left(y_0^{\delta,\varepsilon}, v^{\delta,\varepsilon}(\cdot,\cdot)\right) + J^{\delta}\left(y_0^{\delta}, v^{\delta}(\cdot,\cdot)\right)}\\
&\quad\equiv \Phi_0^{\delta,\varepsilon}\left[J\left(y_0^{\delta,\varepsilon}, v^{\delta,\varepsilon}(\cdot,\cdot)\right) - J\left(y_0^{\delta}, v^{\delta}(\cdot,\cdot)\right)\right]\\
&\qquad+ \mathbb{E}\left\langle \begin{pmatrix} 0 \\ \Phi_T^{\delta,\varepsilon} \end{pmatrix}, \Pi\left(X_0^{\delta,\varepsilon}, X^{\delta,\varepsilon}(T)\right) - \Pi\left(X_0^{\delta}, X^{\delta}(T)\right)\right\rangle\\
&\quad\equiv \left(\Phi_0^{\delta} + o(1)\right)\left[J\left(y_0^{\delta,\varepsilon}, v^{\delta,\varepsilon}(\cdot,\cdot)\right) - J\left(y_0^{\delta}, v^{\delta}(\cdot,\cdot)\right)\right]\\
&\qquad+ \mathbb{E}\left\langle \begin{pmatrix} 0 \\ \Phi_T^{\delta} + o(1) \end{pmatrix}, \Pi\left(X_0^{\delta,\varepsilon}, X^{\delta,\varepsilon}(T)\right) - \Pi\left(X_0^{\delta}, X^{\delta}(T)\right)\right\rangle,
\end{aligned} \tag{3.48}$$

where $(\Phi_0^{\delta}, \Phi_T^{\delta})$ is defined in (3.25) and $(\Phi_0^{\delta,\varepsilon}, \Phi_T^{\delta,\varepsilon})$ is defined similarly to (3.22), with $X_0^{\delta,\varepsilon} = \begin{pmatrix} a \\ y_0^{\delta,\varepsilon} \end{pmatrix}$ replaced by (3.42). We have shown that along a sequence,

$$\lim_{\delta\to 0}\left(\Phi_0^{\delta}, \Phi_T^{\delta}\right) = (\Phi_0, \Phi_T), \quad \text{with } \Phi_0 \neq 0. \tag{3.49}$$

Note that

$$J\left(y_0^{\delta,\varepsilon},v^{\delta,\varepsilon}(\cdot,\cdot)\right)-J\left(y_0^{\delta},v^{\delta}(\cdot,\cdot)\right)=\mathbb{E}\int_0^T\left[l\left(t,X^{\delta,\varepsilon}(t),v^{\delta,\varepsilon}(t,\cdot)\right)-l\left(t,X^{\delta}(t),v^{\delta}(t,\cdot)\right)\right]dt$$
$$+\mathbb{E}\left[\Gamma\left(X_0^{\delta,\varepsilon},X^{\delta,\varepsilon}(T)\right)-\Gamma\left(X_0^{\delta},X^{\delta}(T)\right)\right]\equiv I_1+I_2. \tag{3.50}$$

We first deal with term I_1. We have

$$\begin{aligned}
I_1 &= \mathbb{E}\int_0^T\left[l\left(t,X^{\delta,\varepsilon}(t),v^{\delta,\varepsilon}(t,\cdot)\right)-l\left(t,X^{\delta}(t)+X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t),v^{\delta,\varepsilon}(t,\cdot)\right)\right]dt\\
&\quad+\mathbb{E}\int_0^T\Big[l\left(t,X^{\delta}(t)+X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t),v^{\delta,\varepsilon}(t,\cdot)\right)\\
&\qquad\qquad-l\left(t,X^{\delta}(t)+X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t),v^{\delta}(t,\cdot)\right)\Big]dt\\
&\quad+\mathbb{E}\int_0^T\left[l\left(t,X^{\delta}(t)+X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t),v^{\delta}(t,\cdot)\right)-l\left(t,X^{\delta}(t),v^{\delta}(t,\cdot)\right)\right]dt\\
&=\mathbb{E}\int_0^T\left[\left\langle l_X\left(t,X^{\delta}(t)+X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t),v^{\delta,\varepsilon}(t,\cdot)\right),X^{\delta,\varepsilon}(t)-X^{\delta}(t)-X_1^{\delta,\varepsilon}(t)-X_2^{\delta,\varepsilon}(t)\right\rangle\right]dt\\
&\quad+\frac{1}{2}\mathbb{E}\int_0^T\Big\langle l_{XX}\left(t,X^{\delta}(t)+X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t),v^{\delta,\varepsilon}(t,\cdot)\right)\\
&\qquad\qquad\times\left(X^{\delta,\varepsilon}(t)-X^{\delta}(t)-X_1^{\delta,\varepsilon}(t)-X_2^{\delta,\varepsilon}(t)\right),X^{\delta,\varepsilon}(t)-X^{\delta}(t)-X_1^{\delta,\varepsilon}(t)-X_2^{\delta,\varepsilon}(t)\Big\rangle dt\\
&\quad+\mathbb{E}\int_0^T\Big\langle D^2l^{\delta,\varepsilon,1,2}\left(X^{\delta,\varepsilon}(t)-X^{\delta}(t)-X_1^{\delta,\varepsilon}(t)-X_2^{\delta,\varepsilon}(t)\right),\\
&\qquad\qquad X^{\delta,\varepsilon}(t)-X^{\delta}(t)-X_1^{\delta,\varepsilon}(t)-X_2^{\delta,\varepsilon}(t)\Big\rangle dt\\
&\quad+\mathbb{E}\int_0^T\left[l\left(t,X^{\delta}(t),v^{\delta,\varepsilon}(t,\cdot)\right)-l\left(t,X^{\delta}(t),v^{\delta}(t,\cdot)\right)\right]dt\\
&\quad+\mathbb{E}\int_0^T\left\langle l_X\left(t,X^{\delta}(t),v^{\delta,\varepsilon}(t,\cdot)\right)-l_X\left(t,X^{\delta}(t),v^{\delta}(t,\cdot)\right),X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t)\right\rangle dt\\
&\quad+\frac{1}{2}\mathbb{E}\int_0^T\Big\langle\left[l_{XX}\left(t,X^{\delta}(t),v^{\delta,\varepsilon}(t,\cdot)\right)-l_{XX}\left(t,X^{\delta}(t),v^{\delta}(t,\cdot)\right)\right]\\
&\qquad\qquad\times\left(X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t)\right),X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t)\Big\rangle dt\\
&\quad+\mathbb{E}\int_0^T\left\langle D^2l^{\delta,\varepsilon}\left(X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t)\right),X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t)\right\rangle dt\\
&\quad+\mathbb{E}\int_0^T\left\langle l_X\left(t,X^{\delta}(t),v^{\delta}(t,\cdot)\right),X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t)\right\rangle dt\\
&\quad+\frac{1}{2}\mathbb{E}\int_0^T\left\langle l_{XX}\left(t,X^{\delta}(t),v^{\delta}(t,\cdot)\right)\left(X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t)\right),X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t)\right\rangle dt\\
&\quad+\mathbb{E}\int_0^T\left\langle D^2l^{\delta}\left(X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t)\right),X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t)\right\rangle dt,
\end{aligned} \tag{3.51}$$

with

$$\begin{aligned}
D^2 l^{\delta,\varepsilon,1,2} &\doteq \int_0^1 \beta \Big[l_{XX}\Big(t, \beta\Big(X^{\delta}(t) + X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t)\Big) + (1-\beta) X_1^{\delta,\varepsilon}(t), v^{\delta}(t,\cdot)\Big) \\
&\qquad - l_{XX}\Big(t, X^{\delta}(t) + X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t), v^{\delta}(t,\cdot)\Big)\Big] d\beta, \\
D^2 l^{\delta,\varepsilon} &\doteq \int_0^1 \beta \Big[l_{XX}\Big(t, \beta X^{\delta}(t) + (1-\beta)\Big(X^{\delta}(t) + X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t)\Big), v^{\delta,\varepsilon}(t,\cdot)\Big) \\
&\qquad - l_{XX}\Big(t, \beta X^{\delta}(t) + (1-\beta)\Big(X^{\delta}(t) + X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t)\Big), v^{\delta}(t,\cdot)\Big) \\
&\qquad - l_{XX}\Big(t, X^{\delta}(t), v^{\delta,\varepsilon}(t,\cdot)\Big) + l_{XX}\Big(t, X^{\delta}(t), v^{\delta}(t,\cdot)\Big)\Big] d\beta, \\
D^2 l^{\delta} &\doteq \int_0^1 \beta \Big[l_{XX}\Big(t, \beta X^{\delta}(t) + (1-\beta)\Big(X^{\delta}(t) + X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t)\Big), v^{\delta}(t,\cdot)\Big) \\
&\qquad - l_{XX}\Big(t, X^{\delta}(t), v^{\delta}(t,\cdot)\Big)\Big] d\beta.
\end{aligned} \tag{3.52}$$

Next we deal with term I_2. Similarly, we have

$$\begin{aligned}
I_2 &= \mathbb{E}\Bigg\{ D\Gamma^{\delta} \begin{pmatrix} X_0^{\delta,\varepsilon} - X_0^{\delta} \\ X^{\delta,\varepsilon}(T) - X^{\delta}(T) \end{pmatrix} \\
&\qquad + \frac{1}{2}\left\langle D^2\Gamma^{\delta} \begin{pmatrix} X_0^{\delta,\varepsilon} - X_0^{\delta} \\ X^{\delta,\varepsilon}(T) - X^{\delta}(T) \end{pmatrix}, \begin{pmatrix} X_0^{\delta,\varepsilon} - X_0^{\delta} \\ X^{\delta,\varepsilon}(T) - X^{\delta}(T) \end{pmatrix} \right\rangle \\
&\qquad + \left\langle D^2\Gamma^{\delta,\varepsilon} \begin{pmatrix} X_0^{\delta,\varepsilon} - X_0^{\delta} \\ X^{\delta,\varepsilon}(T) - X^{\delta}(T) \end{pmatrix}, \begin{pmatrix} X_0^{\delta,\varepsilon} - X_0^{\delta} \\ X^{\delta,\varepsilon}(T) - X^{\delta}(T) \end{pmatrix} \right\rangle \Bigg\} \\
&\equiv \mathbb{E}\Bigg\{ \Gamma_{X_0}^{\delta}\Big(X_0^{\delta,\varepsilon} - X_0^{\delta}\Big) + \Gamma_{X_T}^{\delta}\Big(X^{\delta,\varepsilon}(T) - X^{\delta}(T)\Big) \\
&\qquad + \frac{1}{2}\Big\langle \Gamma_{X_0 X_0}^{\delta}\Big(X_0^{\delta,\varepsilon} - X_0^{\delta}\Big), X_0^{\delta,\varepsilon} - X_0^{\delta} \Big\rangle \\
&\qquad + \frac{1}{2}\Big\langle \Gamma_{X_T X_T}^{\delta}\Big(X^{\delta,\varepsilon}(T) - X^{\delta}(T)\Big), X^{\delta,\varepsilon}(T) - X^{\delta}(T) \Big\rangle \\
&\qquad + \left\langle D^2\Gamma^{\delta,\varepsilon} \begin{pmatrix} X_0^{\delta,\varepsilon} - X_0^{\delta} \\ X^{\delta,\varepsilon}(T) - X^{\delta}(T) \end{pmatrix}, \begin{pmatrix} X_0^{\delta,\varepsilon} - X_0^{\delta} \\ X^{\delta,\varepsilon}(T) - X^{\delta}(T) \end{pmatrix} \right\rangle \Bigg\},
\end{aligned} \tag{3.53}$$

with

$$\begin{aligned}
D\Gamma^{\delta} &\doteq \Big(\Gamma_{X_0}^{\delta}, \Gamma_{X_T}^{\delta}\Big) \equiv \Big(\Gamma_{X_0}\Big(X_0^{\delta}, X^{\delta}(T)\Big), \Gamma_{X_T}\Big(X_0^{\delta}, X^{\delta}(T)\Big)\Big), \\
D^2\Gamma^{\delta} &\doteq \begin{pmatrix} \Gamma_{X_0 X_0}^{\delta} & 0 \\ 0 & \Gamma_{X_T X_T}^{\delta} \end{pmatrix} \equiv \begin{pmatrix} \Gamma_{X_0 X_0}\Big(X_0^{\delta}, X^{\delta}(T)\Big) & 0 \\ 0 & \Gamma_{X_T X_T}\Big(X_0^{\delta}, X^{\delta}(T)\Big) \end{pmatrix}, \\
D^2\Gamma^{\delta,\varepsilon} &\doteq \int_0^1 \beta \Big[D^2\Gamma\Big(\beta X_0^{\delta} + (1-\beta) X_0^{\delta,\varepsilon}, \beta X_T^{\delta} + (1-\beta) X_T^{\delta,\varepsilon}\Big) - D^2\Gamma\Big(X_0^{\delta}, X^{\delta}(T)\Big)\Big] d\beta.
\end{aligned} \tag{3.54}$$

Also, we have

$$\begin{aligned}
&\mathbb{E}\left\langle \begin{pmatrix} 0 \\ \Phi_T^{\delta,\varepsilon} \end{pmatrix}, \Pi\left(X_0^{\delta,\varepsilon}, X^{\delta,\varepsilon}(T)\right) - \Pi\left(X_0^{\delta}, X^{\delta}(T)\right) \right\rangle \\
&\quad \equiv \mathbb{E}\left\langle \begin{pmatrix} 0 \\ \Phi_T^{\delta,\varepsilon} \end{pmatrix}, \Pi\left(0, X^{\delta,\varepsilon}(T)\right) - \Pi\left(0, X^{\delta}(T)\right) \right\rangle \\
&\quad = \mathbb{E}\left\{ \left\langle \Pi_{X_T}\left(0, X^{\delta}(T)\right) \begin{pmatrix} 0 \\ \Phi_T^{\delta,\varepsilon} \end{pmatrix}, X^{\delta,\varepsilon}(T) - X^{\delta}(T) \right\rangle \right. \\
&\qquad + \frac{1}{2}\left\langle \Pi_{X_T X_T}\left(0, X^{\delta}(T)\right) \begin{pmatrix} 0 \\ \Phi_T^{\delta,\varepsilon} \end{pmatrix} \left(X^{\delta,\varepsilon}(T) - X^{\delta}(T)\right), X^{\delta,\varepsilon}(T) - X^{\delta}(T) \right\rangle \\
&\qquad \left. + \left\langle D^2\Pi^{\delta,\varepsilon} \begin{pmatrix} 0 \\ \Phi_T^{\delta,\varepsilon} \end{pmatrix} \left(X^{\delta,\varepsilon}(T) - X^{\delta}(T)\right), X^{\delta,\varepsilon}(T) - X^{\delta}(T) \right\rangle \right\},
\end{aligned} \tag{3.55}$$

with

$$D^2\Pi^{\delta,\varepsilon} \doteq \int_0^1 \beta\left[\Pi_{X_T X_T}\left(0, \beta X^{\delta}(T) + (1-\beta) X^{\delta,\varepsilon}(T)\right) - \Pi_{X_T X_T}\left(0, X^{\delta}(T)\right)\right] d\beta. \tag{3.56}$$

Under assumptions (H2), (H3), we have

$$\begin{gathered}
\left|\Phi_0^{\delta,\varepsilon} D^2 l^{\delta,\varepsilon,1,2}\right| \leq K\left|X^{\delta,\varepsilon}(t) - X^{\delta}(t) - X_1^{\delta,\varepsilon}(t) - X_2^{\delta,\varepsilon}(t)\right|, \\
\left|\Phi_0^{\delta,\varepsilon} D^2 l^{\delta,\varepsilon}\right| + \left|\Phi_0^{\delta,\varepsilon} D^2 l^{\delta}\right| \leq K\left|X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t)\right|, \\
\left|\Phi_0^{\delta,\varepsilon} D^2\Gamma^{\delta,\varepsilon}\right| + \left|D^2\Pi^{\delta,\varepsilon}\begin{pmatrix} 0 \\ \Phi_T^{\delta,\varepsilon} \end{pmatrix}\right| \leq K\left(\left|X_0^{\delta,\varepsilon} - X_0^{\delta}\right| + \left|X^{\delta,\varepsilon}(T) - X^{\delta}(T)\right|\right).
\end{gathered} \tag{3.57}$$

Hence, by (3.47), we have

$$\begin{aligned}
\mathbb{E}\Bigg\{ & \left|\left\langle \Phi_0^{\delta,\varepsilon} D^2 l^{\delta,\varepsilon,1,2}\left(X^{\delta,\varepsilon}(t) - X^{\delta}(t) - X_1^{\delta,\varepsilon}(t) - X_2^{\delta,\varepsilon}(t)\right), \right.\right. \\
& \left.\left. X^{\delta,\varepsilon}(t) - X^{\delta}(t) - X_1^{\delta,\varepsilon}(t) - X_2^{\delta,\varepsilon}(t) \right\rangle\right| \\
& + \left|\left\langle \Phi_0^{\delta,\varepsilon} D^2 l^{\delta,\varepsilon}\left(X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(\mathrm{t})\right), X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t) \right\rangle\right| \\
& + \left|\left\langle \Phi_0^{\delta,\varepsilon} D^2 l^{\delta}\left(X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t)\right), X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t) \right\rangle\right| \\
& + \left|\Phi_0^{\delta,\varepsilon}\left\langle D^2\Gamma^{\delta,\varepsilon}\begin{pmatrix} X_0^{\delta,\varepsilon} - X_0^{\delta} \\ X^{\delta,\varepsilon}(T) - X^{\delta}(T) \end{pmatrix}, \begin{pmatrix} X_0^{\delta,\varepsilon} - X_0^{\delta} \\ X^{\delta,\varepsilon}(T) - X^{\delta}(T) \end{pmatrix} \right\rangle\right|
\end{aligned}$$

$$
\begin{aligned}
&+\left|\left\langle D^2\Pi^{\delta,\varepsilon}\begin{pmatrix}0\\ \Phi_T^{\delta,\varepsilon}\end{pmatrix}\left(X^{\delta,\varepsilon}(T)-X^{\delta}(T)\right),X^{\delta,\varepsilon}(T)-X^{\delta}(T)\right\rangle\right|\Bigg\}\\
&\le K\mathbb{E}\bigg(\left|X^{\delta,\varepsilon}(t)-X^{\delta}(t)-X_1^{\delta,\varepsilon}(t)-X_2^{\delta,\varepsilon}(t)\right|^3\\
&\qquad+\left|X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t)\right|^3+\left|X_0^{\delta,\varepsilon}-X_0^{\delta}\right|^3+\left|X^{\delta,\varepsilon}(T)-X^{\delta}(T)\right|^3\bigg)\\
&\le K\varepsilon^{3/2}.
\end{aligned}
\tag{3.58}
$$

Consequently, combining (3.50) with (3.55), from (3.48), we can derive that

$$
\begin{aligned}
&-\sqrt{\delta}\left[\sqrt{\varepsilon}y_0+\varepsilon(T+K)\right]\\
&\le\ \Phi_0^{\delta,\varepsilon}\left[J\left(y_0^{\delta,\varepsilon},v^{\delta,\varepsilon}(\cdot,\cdot)\right)-J\left(y_0^{\delta},v^{\delta}(\cdot,\cdot)\right)\right]\\
&\quad+\mathbb{E}\left\langle\begin{pmatrix}0\\ \Phi_T^{\delta,\varepsilon}\end{pmatrix},\Pi\left(0,X^{\delta,\varepsilon}(T)\right)-\Pi\left(0,X^{\delta}(T)\right)\right\rangle\\
&=\Phi_0^{\delta,\varepsilon}\Bigg\{\mathbb{E}\int_0^T\Big[l\left(t,X^{\delta}(t),v^{\delta,\varepsilon}(t,\cdot)\right)-l\left(t,X^{\delta}(t),v^{\delta}(t,\cdot)\right)\\
&\qquad\qquad+\left\langle l_X\left(t,X^{\delta}(t),v^{\delta}(t,\cdot)\right),X_1^{\delta,\varepsilon}(t)+X_2^{\delta,\varepsilon}(t)\right\rangle\\
&\qquad\qquad+\frac{1}{2}\left\langle l_{XX}\left(t,X^{\delta}(t),v^{\delta}(t,\cdot)\right)X_1^{\delta,\varepsilon}(t),X_1^{\delta,\varepsilon}(t)\right\rangle\Big]dt\\
&\qquad+\mathbb{E}\Big[\Gamma_{X_0}\left(X_0^{\delta},X^{\delta}(T)\right)\left(X_0^{\delta,\varepsilon}-X_0^{\delta}\right)\\
&\qquad\qquad+\frac{1}{2}\left\langle\Gamma_{X_0X_0}\left(X_0^{\delta},X^{\delta}(T)\right)\left(X_0^{\delta,\varepsilon}-X_0^{\delta}\right),X_0^{\delta,\varepsilon}-X_0^{\delta}\right\rangle\\
&\qquad\qquad+\Gamma_{X_T}\left(X_0^{\delta},X^{\delta}(T)\right)\left(X^{\delta,\varepsilon}(T)-X^{\delta}(T)\right)\\
&\qquad\qquad+\frac{1}{2}\left\langle\Gamma_{X_TX_T}\left(X_0^{\delta},X^{\delta}(T)\right)\left(X^{\delta,\varepsilon}(T)-X^{\delta}(T)\right),X^{\delta,\varepsilon}(T)-X^{\delta}(T)\right\rangle\Big]\Bigg\}\\
&\quad+\mathbb{E}\Bigg\{\left\langle\Pi_{X_T}\left(0,X^{\delta}(T)\right)\begin{pmatrix}0\\ \Phi_T^{\delta,\varepsilon}\end{pmatrix},X^{\delta,\varepsilon}(T)-X^{\delta}(T)\right\rangle\\
&\qquad+\frac{1}{2}\left\langle\Pi_{X_TX_T}\left(0,X^{\delta}(T)\right)\begin{pmatrix}0\\ \Phi_T^{\delta,\varepsilon}\end{pmatrix}\left(X^{\delta,\varepsilon}(T)-X^{\delta}(T)\right),X^{\delta,\varepsilon}(T)-X^{\delta}(T)\right\rangle\Bigg\}+O\left(\varepsilon^{3/2}\right)
\end{aligned}
$$

$$
\begin{aligned}
= \mathbb{E}\Bigg\{ \int_0^T \Big\{ & \Phi_0^{\delta,\varepsilon}\Big[l\Big(t, X^\delta(t), v^{\delta,\varepsilon}(t,\cdot)\Big) - l\Big(t, X^\delta(t), v^\delta(t,\cdot)\Big)\Big] \\
& + \Phi_0^{\delta,\varepsilon}\Big\langle l_X\Big(t, X^\delta(t), v^\delta(t,\cdot)\Big), X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t)\Big\rangle \\
& + \frac{1}{2}\Phi_0^{\delta,\varepsilon}\Big\langle l_{XX}\Big(t, X^\delta(t), v^\delta(t,\cdot)\Big)X_1^{\delta,\varepsilon}(t), X_1^{\delta,\varepsilon}(t)\Big\rangle\Big\}dt \\
& + \sqrt{\varepsilon}\left\langle \Phi_0^{\delta,\varepsilon}\Gamma_{X_0}\Big(X_0^\delta, X^\delta(T)\Big), \begin{pmatrix} 0 \\ y_0 \end{pmatrix}\right\rangle \\
& + \frac{\varepsilon}{2}\left\langle \Phi_0^{\delta,\varepsilon}\Gamma_{X_0X_0}\Big(X_0^\delta, X^\delta(T)\Big)\begin{pmatrix} 0 \\ y_0 \end{pmatrix}, \begin{pmatrix} 0 \\ y_0 \end{pmatrix}\right\rangle \\
& + \left\langle \Phi_0^{\delta,\varepsilon}\Gamma_{X_T}\Big(X_0^\delta, X^\delta(T)\Big) + \Pi_{X_T}\Big(0, X^\delta(T)\Big)\begin{pmatrix} 0 \\ \Phi_T^{\delta,\varepsilon} \end{pmatrix}, X_1^{\delta,\varepsilon}(T) + X_2^{\delta,\varepsilon}(T)\right\rangle \\
& + \frac{1}{2}\left\langle \left[\Phi_0^{\delta,\varepsilon}\Gamma_{X_TX_T}\Big(X_0^\delta, X^\delta(T)\Big) + \Pi_{X_TX_T}\Big(0, X^\delta(T)\Big)\begin{pmatrix} 0 \\ \Phi_T^{\delta,\varepsilon} \end{pmatrix}\right]X_1^{\delta,\varepsilon}(T), X_1^{\delta,\varepsilon}(T)\right\rangle\Bigg\} \\
& + O\Big(\varepsilon^{3/2}\Big).
\end{aligned}
\tag{3.59}
$$

Step 5 (duality). Let $(\widetilde{\Phi}^{\delta,\varepsilon}(\cdot), \widetilde{\Psi}^{\delta,\varepsilon}(\cdot), \widetilde{\Xi}^{\delta,\varepsilon}(\cdot,\cdot))$ be the adapted solution to the following BSDEJ:

$$
\begin{aligned}
-d\widetilde{\Phi}^{\delta,\varepsilon}(t) &= \left[B_X^\delta(t,\cdot)^\top\widetilde{\Phi}^{\delta,\varepsilon}(t) + \Sigma_X^\delta(t,\cdot)^\top\widetilde{\Psi}^{\delta,\varepsilon}(t) + \int_{\mathbf{E}}\Upsilon_X^\delta(t,e)^\top\widetilde{\Xi}^{\delta,\varepsilon}(t,e)\pi(de) - \Phi_0^{\delta,\varepsilon}l_X^\delta(t,\cdot)\right]dt \\
&\quad - \widetilde{\Psi}^{\delta,\varepsilon}(t)dW(t) - \int_{\mathbf{E}}\widetilde{\Xi}^{\delta,\varepsilon}(t,e)\widetilde{N}(dedt), \\
d\widetilde{\Phi}^{\delta,\varepsilon}(T) &= -\left[\Phi_0^{\delta,\varepsilon}\Gamma_{X_T}\Big(X_0^\delta, X^\delta(T)\Big) + \Pi_{X_T}\Big(0, X^\delta(T)\Big)\begin{pmatrix} 0 \\ \Phi_T^{\delta,\varepsilon} \end{pmatrix}\right],
\end{aligned}
\tag{3.60}
$$

where $l_X^\delta(t,\cdot) \doteq l_X(t, X^\delta(t), v^\delta(t,\cdot))$. Then, as $(\varepsilon,\delta) \to (0,0)$, $(\widetilde{\Phi}^{\delta,\varepsilon}(\cdot), \widetilde{\Psi}^{\delta,\varepsilon}(\cdot), \widetilde{\Xi}^{\delta,\varepsilon}(\cdot,\cdot))$ goes to $(\widetilde{\Phi}(\cdot), \widetilde{\Psi}(\cdot), \widetilde{\Xi}(\cdot,\cdot))$, which is the adapted solution to the following BSDEJ:

$$
\begin{aligned}
-d\widetilde{\Phi}(t) &= \left[\overline{B}_X(t,\cdot)^\top\widetilde{\Phi}(t) + \overline{\Sigma_X}(t,\cdot)^\top\widetilde{\Psi}(t) + \int_{\mathbf{E}}\overline{\Upsilon}_X(t,e)^\top\widetilde{\Xi}(t,e)\pi(de) - \Phi_0\bar{l}_X(t,\cdot)\right]dt \\
&\quad - \widetilde{\Psi}(t)dW(t) - \int_{\mathbf{E}}\widetilde{\Xi}(t,e)\widetilde{N}(dedt), \\
d\widetilde{\Phi}(T) &= -\left[\Phi_0\Gamma_{X_T}\Big(\overline{X}_0, \overline{X}(T)\Big) + \Pi_{X_T}\Big(0, \overline{X}(T)\Big)\begin{pmatrix} 0 \\ \Phi_T \end{pmatrix}\right],
\end{aligned}
\tag{3.61}
$$

where

$$\overline{B}_X(t,\cdot) \doteq B_X\Big(t,\overline{X}(t),\overline{v}(t,\cdot)\Big), \qquad \overline{\Sigma_X}(t,\cdot) \doteq \Sigma_X\Big(t,\overline{X}(t),\overline{v}(t,\cdot)\Big),$$
$$\overline{\Upsilon}_X(t,\cdot) \doteq B_X\Big(t,\overline{X}(t),\overline{v}(t,\cdot)\Big), \qquad \bar{l}_X(t,\cdot) \doteq l_X\Big(t,\overline{X}(t),\overline{v}(t,\cdot)\Big), \tag{3.62}$$

(Φ_0, Φ_T) is determined by (3.25).

Applying Itô's formula to $\langle \widetilde{\Phi}^{\delta,\varepsilon}(\cdot), X_1^{\delta,\varepsilon}(\cdot) + X_2^{\delta,\varepsilon}(\cdot)\rangle$, we can get

$$\begin{aligned}
&-\mathbb{E}\left\langle \Phi_0^{\delta,\varepsilon}\Gamma_{X_T}\Big(X_0^{\delta}, X^{\delta}(T)\Big) + \Pi_{X_T}\Big(0, X^{\delta}(T)\Big)\begin{pmatrix} 0 \\ \Phi_T^{\delta,\varepsilon}\end{pmatrix}, X_1^{\delta,\varepsilon}(T) + X_2^{\delta,\varepsilon}(T)\right\rangle \\
&= \mathbb{E}\Big\langle \widetilde{\Phi}^{\delta,\varepsilon}(T), X_1^{\delta,\varepsilon}(T) + X_2^{\delta,\varepsilon}(T)\Big\rangle \\
&= \mathbb{E}\left\langle \widetilde{\Phi}^{\delta,\varepsilon}(0), \begin{pmatrix} 0 \\ \sqrt{\varepsilon}y_0\end{pmatrix}\right\rangle \\
&\quad + \mathbb{E}\int_0^T \Bigg\{ \Bigg\langle -B_X^{\delta}(t,\cdot)^{\top}\widetilde{\Phi}^{\delta,\varepsilon}(t) - \Sigma_X^{\delta}(t,\cdot)^{\top}\widetilde{\Psi}^{\delta,\varepsilon}(t) \\
&\qquad\qquad - \int_{\mathbf{E}} \Upsilon_X^{\delta}(t,e)^{\top}\widetilde{\Xi}^{\delta,\varepsilon}(t,e)\pi(de) + \Phi_0^{\delta,\varepsilon} l_X^{\delta}(t,\cdot), X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t)\Bigg\rangle \\
&\qquad + \left\langle \widetilde{\Phi}^{\delta,\varepsilon}(t), B_X^{\delta}(t,\cdot)^{\top}\Big(X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t)\Big) + \Delta B^{\delta}(t,\cdot)I_{S_\varepsilon}(t) + \frac{1}{2}B_{XX}^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t)^2\right\rangle \\
&\qquad + \Bigg\langle \widetilde{\Psi}^{\delta,\varepsilon}(t), \Sigma_X^{\delta}(t,\cdot)^{\top}\Big(X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t)\Big) \\
&\qquad\qquad + \Big(\Delta\Sigma_X^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t) + \Delta\Sigma^{\delta}(t,\cdot)\Big)I_{S_\varepsilon}(t) + \frac{1}{2}\Sigma_{XX}^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t)^2\Bigg\rangle \\
&\qquad + \int_{\mathbf{E}}\Bigg\langle \widetilde{\Xi}^{\delta,\varepsilon}(t,e), \Upsilon_X^{\delta}(t,e)^{\top}\Big(X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t)\Big) \\
&\qquad\qquad + \Big(\Delta\Upsilon_X^{\delta}(t,e)X_1^{\delta,\varepsilon}(t) + \Delta\Upsilon^{\delta}(t,e)\Big)I_{S_\varepsilon}(t) + \frac{1}{2}\Upsilon_{XX}^{\delta}(t,e)X_1^{\delta,\varepsilon}(t)^2\Bigg\rangle\pi(de)\Bigg\}dt \\
&= \mathbb{E}\left\langle \widetilde{\Phi}^{\delta,\varepsilon}(0), \begin{pmatrix} 0 \\ \sqrt{\varepsilon}y_0\end{pmatrix}\right\rangle \\
&\quad + \mathbb{E}\int_0^T \Bigg\{ \left\langle \widetilde{\Phi}^{\delta,\varepsilon}(t), \Delta B^{\delta}(t,\cdot)I_{S_\varepsilon}(t) + \frac{1}{2}B_{XX}^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t)^2\right\rangle \\
&\qquad + \Big\langle \Phi_0^{\delta,\varepsilon} l_X^{\delta}(t,\cdot), X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t)\Big\rangle
\end{aligned}$$

$$
\begin{aligned}
&\quad + \left\langle \widetilde{\Psi}^{\delta,\varepsilon}(t), \Delta\Sigma^{\delta}(t,\cdot) I_{S_\varepsilon}(t) + \frac{1}{2}\Sigma^{\delta}_{XX}(t,\cdot) X_1^{\delta,\varepsilon}(t)^2 \right\rangle \\
&\quad + \int_{\mathbf{E}} \left\langle \widetilde{\Xi}^{\delta,\varepsilon}(t,e), \Delta\Upsilon^{\delta}(t,e) I_{S_\varepsilon}(t) + \frac{1}{2}\Upsilon^{\delta}_{XX}(t,e) X_1^{\delta,\varepsilon}(t)^2 \right\rangle \pi(de) \Big\} dt \\
&+ O\left(\varepsilon^{3/2}\right).
\end{aligned}
\tag{3.63}
$$

Then by (3.59), we have

$$
\begin{aligned}
&-\sqrt{\delta}\left[\sqrt{\varepsilon} y_0 + \varepsilon(T+K)\right] \\
&\leq \mathbb{E}\int_0^T \Big\{ \Phi_0^{\delta,\varepsilon}\left[l\left(t, X^{\delta}(t), v^{\delta,\varepsilon}(t,\cdot)\right) - l\left(t, X^{\delta}(t), v^{\delta}(t,\cdot)\right)\right] \\
&\qquad + \Phi_0^{\delta,\varepsilon}\left\langle l_X^{\delta}(t,\cdot), X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t)\right\rangle + \frac{1}{2}\Phi_0^{\delta,\varepsilon}\left\langle l_{XX}^{\delta}(t,\cdot) X_1^{\delta,\varepsilon}(t), X_1^{\delta,\varepsilon}(t)\right\rangle \Big\} dt \\
&\quad - \mathbb{E}\left\langle \widetilde{\Phi}^{\delta,\varepsilon}(0), \begin{pmatrix} 0 \\ \sqrt{\varepsilon} y_0 \end{pmatrix} \right\rangle \\
&\quad - \mathbb{E}\int_0^T \Big\{ \left\langle \widetilde{\Phi}^{\delta,\varepsilon}(t), \Delta B^{\delta}(t,\cdot) I_{S_\varepsilon}(t) + \frac{1}{2} B^{\delta}_{XX}(t,\cdot) X_1^{\delta,\varepsilon}(t)^2 \right\rangle \\
&\qquad - \left\langle \Phi_0^{\delta,\varepsilon} l_X^{\delta}(t,\cdot), X_1^{\delta,\varepsilon}(t) + X_2^{\delta,\varepsilon}(t)\right\rangle \\
&\qquad + \left\langle \widetilde{\Psi}^{\delta,\varepsilon}(t), \Delta\Sigma^{\delta}(t,\cdot) I_{S_\varepsilon}(t) + \frac{1}{2}\Sigma^{\delta}_{XX}(t,\cdot) X_1^{\delta,\varepsilon}(t)^2 \right\rangle \\
&\qquad + \int_{\mathbf{E}} \left\langle \widetilde{\Xi}^{\delta,\varepsilon}(t,e), \Delta\Upsilon^{\delta}(t,e) I_{S_\varepsilon}(t) + \frac{1}{2}\Upsilon^{\delta}_{XX}(t,e) X_1^{\delta,\varepsilon}(t)^2 \right\rangle \pi(de) \Big\} dt \\
&\quad + \mathbb{E}\Big\{ \sqrt{\varepsilon}\left\langle \Phi_0^{\delta,\varepsilon}\Gamma_{X_0}\left(X_0^{\delta}, X^{\delta}(T)\right), \begin{pmatrix} 0 \\ y_0 \end{pmatrix} \right\rangle \\
&\qquad + \frac{\varepsilon}{2}\left\langle \Phi_0^{\delta,\varepsilon}\Gamma_{X_0X_0}\left(X_0^{\delta}, X^{\delta}(T)\right)\begin{pmatrix} 0 \\ y_0 \end{pmatrix}, \begin{pmatrix} 0 \\ y_0 \end{pmatrix} \right\rangle \\
&\qquad + \frac{1}{2}\Big\langle \Big[\Phi_0^{\delta,\varepsilon}\Gamma_{X_TX_T}\left(X_0^{\delta}, X^{\delta}(T)\right) \\
&\qquad\qquad + \Pi_{X_TX_T}\left(0, X^{\delta}(T)\right)\begin{pmatrix} 0 \\ \Phi_T^{\delta,\varepsilon} \end{pmatrix}\Big] X_1^{\delta,\varepsilon}(T), X_1^{\delta,\varepsilon}(T) \Big\rangle \Big\} + O\left(\varepsilon^{3/2}\right) \\
&= \mathbb{E}\int_0^T \Big\{ \Phi_0^{\delta,\varepsilon}\left[l\left(t, X^{\delta}(t), v^{\delta,\varepsilon}(t,\cdot)\right) - l\left(t, X^{\delta}(t), v^{\delta}(t,\cdot)\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{2}\Phi_0^{\delta,\varepsilon}\left\langle l_{XX}^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t),X_1^{\delta,\varepsilon}(t)\right\rangle\Bigg\}dt\\
&-\mathbb{E}\int_0^T\Bigg\{\left\langle\widetilde{\Phi}^{\delta,\varepsilon}(t),\Delta B^{\delta}(t,\cdot)I_{S_\varepsilon}(t)+\frac{1}{2}B_{XX}^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t)^2\right\rangle\\
&\quad+\left\langle\widetilde{\Psi}^{\delta,\varepsilon}(t),\Delta\Sigma^{\delta}(t,\cdot)I_{S_\varepsilon}(t)+\frac{1}{2}\Sigma_{XX}^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t)^2\right\rangle\\
&\quad+\int_{\mathbf{E}}\left\langle\widetilde{\Xi}^{\delta,\varepsilon}(t,e),\Delta\Upsilon^{\delta}(t,e)I_{S_\varepsilon}(t)+\frac{1}{2}\Upsilon_{XX}^{\delta}(t,e)X_1^{\delta,\varepsilon}(t)^2\right\rangle\pi(de)\Bigg\}dt\\
&+\mathbb{E}\Bigg\{\sqrt{\varepsilon}\left\langle\Phi_0^{\delta,\varepsilon}\Gamma_{X_0}\left(X_0^{\delta},X^{\delta}(T)\right)-\widetilde{\Phi}^{\delta,\varepsilon}(0),\begin{pmatrix}0\\y_0\end{pmatrix}\right\rangle\\
&\quad+\frac{\varepsilon}{2}\left\langle\Phi_0^{\delta,\varepsilon}\Gamma_{X_0X_0}\left(X_0^{\delta},X^{\delta}(T)\right)\begin{pmatrix}0\\y_0\end{pmatrix},\begin{pmatrix}0\\y_0\end{pmatrix}\right\rangle\\
&\quad+\frac{1}{2}\Bigg\langle\Bigg[\Phi_0^{\delta,\varepsilon}\Gamma_{X_TX_T}\left(X_0^{\delta},X^{\delta}(T)\right)\\
&\qquad+\Pi_{X_TX_T}\left(0,X^{\delta}(T)\right)\begin{pmatrix}0\\\Phi_T^{\delta,\varepsilon}\end{pmatrix}\Bigg]X_1^{\delta,\varepsilon}(T),X_1^{\delta,\varepsilon}(T)\Bigg\rangle\Bigg\}+O\left(\varepsilon^{3/2}\right).
\end{aligned}
\tag{3.64}
$$

Note that $Y^{\delta,\varepsilon}(\cdot)\doteq X_1^{\delta,\varepsilon}(\cdot)X_1^{\delta,\varepsilon}(\cdot)^{\top}$ satisfies

$$
\begin{aligned}
&dY^{\delta,\varepsilon}(t)\\
&=\Bigg\{B_X^{\delta}(t,\cdot)Y^{\delta,\varepsilon}(t)+Y^{\delta,\varepsilon}(t)B_X^{\delta}(t,\cdot)^{\top}+\Sigma_X^{\delta}(t,\cdot)Y^{\delta,\varepsilon}(t)\Sigma_X^{\delta}(t,\cdot)^{\top}\\
&\quad+\left[\Delta\Sigma^{\delta}(t,\cdot)\Delta\Sigma^{\delta}(t,\cdot)^{\top}+\Sigma_X^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t)\Delta\Sigma^{\delta}(t,\cdot)^{\top}+\Delta\Sigma^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t)\Sigma_X^{\delta}(t,\cdot)^{\top}\right]I_{S_\varepsilon}(t)\\
&\quad+\int_{\mathbf{E}}\Bigg\{\Upsilon_X^{\delta}(t,e)Y^{\delta,\varepsilon}(t)\Upsilon_X^{\delta}(t,e)^{\top}\\
&\qquad+\Big[\Upsilon_X^{\delta}(t,e)X_1^{\delta,\varepsilon}(t)\Delta\Upsilon^{\delta}(t,e)^{\top}\\
&\qquad\quad+\Delta\Upsilon^{\delta}(t,e)X_1^{\delta,\varepsilon}(t)\Upsilon_X^{\delta}(t,e)^{\top}+\Delta\Upsilon^{\delta}(t,e)\Delta\Upsilon^{\delta}(t,\cdot)^{\top}\Big]I_{S_\varepsilon}(t)\Bigg\}\pi(de)\Bigg\}dt\\
&\quad+\Big\{\Sigma_X^{\delta}(t,\cdot)Y^{\delta,\varepsilon}(t)+Y^{\delta,\varepsilon}(t)\Sigma_X^{\delta}(t,\cdot)^{\top}\\
&\qquad+\left[X_1^{\delta,\varepsilon}(t)\Delta\Sigma^{\delta}(t,\cdot)^{\top}+\Delta\Sigma^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t)^{\top}\right]I_{S_\varepsilon}(t)\Big\}dW(t)\\
&\quad+\int_{\mathbf{E}}\Big\{\Upsilon_X^{\delta}(t,e)Y^{\delta,\varepsilon}(t-)+Y^{\delta,\varepsilon}(t-)\Upsilon_X^{\delta}(t,e)^{\top}+\Upsilon_X^{\delta}(t,e)Y^{\delta,\varepsilon}(t-)\Upsilon_X^{\delta}(t,e)^{\top}\\
&\qquad+\Big[X_1^{\delta,\varepsilon}(t-)\Delta\Upsilon^{\delta}(t,e)^{\top}+\Delta\Upsilon^{\delta}(t,e)X_1^{\delta,\varepsilon}(t-)^{\top}+\Upsilon_X^{\delta}(t,e)X_1^{\delta,\varepsilon}(t-)\Delta\Upsilon^{\delta}(t,e)^{\top}
\end{aligned}
$$

$$+\Delta\Upsilon^{\delta}(t,e)X_1^{\delta,\varepsilon}(t-)\Upsilon_X^{\delta}(t,e)^{\top}+\Delta\Upsilon^{\delta}(t,e)\Delta\Upsilon^{\delta}(t,e)^{\top}\Big]I_{S_\varepsilon}(t)\Big\}\widetilde{N}(dedt),$$
$$Y^{\delta,\varepsilon}(0)=\begin{pmatrix}0 & 0\\ 0 & \varepsilon y_0 y_0^{\top}\end{pmatrix}. \tag{3.65}$$

Now, let $(P^{\delta,\varepsilon}(\cdot),Q^{\delta,\varepsilon}(\cdot),K^{\delta,\varepsilon}(\cdot,\cdot))$ be the adapted solution to the following BSDEJ:

$$\begin{aligned}
-dP^{\delta,\varepsilon}(t)=\Big\{&B_X^{\delta}(t,\cdot)^{\top}P^{\delta,\varepsilon}(t)+P^{\delta,\varepsilon}(t)B_X^{\delta}(t,\cdot)+\Sigma_X^{\delta}(t,\cdot)^{\top}P^{\delta,\varepsilon}(t)\Sigma_X^{\delta}(t,\cdot)\\
&+\Sigma_X^{\delta}(t,\cdot)^{\top}Q^{\delta,\varepsilon}(t)+Q^{\delta,\varepsilon}(t)\Sigma_X^{\delta}(t,\cdot)\\
&+\int_{\mathbf{E}}\Big[\Upsilon_X^{\delta}(t,e)^{\top}P^{\delta,\varepsilon}(t)\Upsilon_X^{\delta}(t,e)+\Upsilon_X^{\delta}(t,e)^{\top}K^{\delta,\varepsilon}(t,e)\Upsilon_X^{\delta}(t,e)\\
&\quad+\Upsilon_X^{\delta}(t,e)^{\top}K^{\delta,\varepsilon}(t,e)+K^{\delta,\varepsilon}(t,e)\Upsilon_X^{\delta}(t,e)\Big]\pi(de)+H_{XX}^{\delta,\varepsilon,\Phi_0^{\delta,\varepsilon}}(t,\cdot)\Big\}dt\\
&-Q^{\delta,\varepsilon}(t)dW(t)-\int_{\mathbf{E}}K^{\delta,\varepsilon}(t,e)\widetilde{N}(dedt),\\
dP^{\delta,\varepsilon}(T)=-\Big[&\Phi_0^{\delta,\varepsilon}\Gamma_{X_TX_T}\Big(X_0^{\delta},X^{\delta}(T)\Big)+\Pi_{X_TX_T}\Big(0,X^{\delta}(T)\Big)\begin{pmatrix}0\\ \Phi_T^{\delta,\varepsilon}\end{pmatrix}\Big],
\end{aligned} \tag{3.66}$$

where

$$H_{XX}^{\delta,\varepsilon,\Phi_0^{\delta,\varepsilon}}(t,\cdot)\doteq H_{XX}\Big(t,\Phi_0^{\delta,\varepsilon},X^{\delta}(t),v^{\delta}(t,\cdot),\widetilde{\Phi}^{\delta,\varepsilon}(t),\widetilde{\Psi}^{\delta,\varepsilon}(t),\widetilde{\Xi}^{\delta,\varepsilon}(t,\cdot)\Big), \tag{3.67}$$

with $H_{XX}(t,\lambda,X,v(\cdot),\widetilde{\Phi},\widetilde{\Psi},\widetilde{\Xi}(\cdot))$ being the Hessian of the following

$$\begin{aligned}
H\Big(t,\lambda,X,v(\cdot),\widetilde{\Phi},\widetilde{\Psi},\widetilde{\Xi}(\cdot)\Big)\doteq{}&\Big\langle\widetilde{\Phi},B(t,X,v(\cdot))\Big\rangle+\Big\langle\widetilde{\Psi},\Sigma(t,X,v(\cdot))\Big\rangle\\
&+\int_{\mathbf{E}}\Big\langle\widetilde{\Xi}(e),\Upsilon(t,X,v(e))\Big\rangle\pi(de)-\lambda l(t,X,v(\cdot)).
\end{aligned} \tag{3.68}$$

Applying Itô's formula to $[P^{\delta,\varepsilon}(\cdot)Y^{\delta,\varepsilon}(\cdot)]$, we have

$$\begin{aligned}
&-\mathbb{E}\left\{\left\langle\left[\Phi_0^{\delta,\varepsilon}\Gamma_{X_TX_T}\left(X_0^{\delta},X^{\delta}(T)\right)+\Pi_{X_TX_T}\left(0,X^{\delta}(T)\right)\begin{pmatrix}0\\ \Phi_T^{\delta,\varepsilon}\end{pmatrix}\right]X_1^{\delta,\varepsilon}(T),X_1^{\delta,\varepsilon}(T)\right\rangle\right.\\
&\quad\left.+\left\langle P^{\delta,\varepsilon}(0)\begin{pmatrix}0\\ \sqrt{\varepsilon}y_0\end{pmatrix},\begin{pmatrix}0\\ \sqrt{\varepsilon}y_0\end{pmatrix}\right\rangle\right\}\\
&=\mathbb{E}\left\{\operatorname{tr}\left[P^{\delta,\varepsilon}(T)Y^{\delta,\varepsilon}(T)\right]-\operatorname{tr}\left[P^{\delta,\varepsilon}(0)Y^{\delta,\varepsilon}(0)\right]\right\}\\
&=\mathbb{E}\int_0^T\operatorname{tr}\left\{P^{\delta,\varepsilon}(t)\left[\Delta\Sigma^{\delta}(t,\cdot)\Delta\Sigma^{\delta}(t,\cdot)^{\top}+\Sigma_X^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t)\Delta\Sigma^{\delta}(t,\cdot)^{\top}\right.\right.\\
&\qquad\qquad\left.+\Delta\Sigma^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t)\Sigma_X^{\delta}(t,\cdot)^{\top}\right]I_{S_\varepsilon}(t)\\
&\qquad+Q^{\delta,\varepsilon}(t)\left[X_1^{\delta,\varepsilon}(t)\Delta\Sigma^{\delta}(t,\cdot)^{\top}+\Delta\Sigma^{\delta}(t,\cdot)X_1^{\delta,\varepsilon}(t)^{\top}\right]I_{S_\varepsilon}(t)-H_{XX}^{\delta,\varepsilon,\Phi_0^{\delta,\varepsilon}}(t,\cdot)Y^{\delta,\varepsilon}(t)\\
&\qquad+\int_{\mathbf{E}}\left\{P^{\delta,\varepsilon}(t)\left[\Upsilon_X^{\delta}(t,e)X_1^{\delta,\varepsilon}(t)\Delta\Upsilon^{\delta}(t,e)^{\top}\right.\right.\\
&\qquad\qquad+\Delta\Upsilon^{\delta}(t,e)X_1^{\delta,\varepsilon}(t)\Upsilon_X^{\delta}(t,e)^{\top}+\left.\Delta\Upsilon^{\delta}(t,e)\Delta\Upsilon^{\delta}(t,e)^{\top}\right]I_{S_\varepsilon}(t)\\
&\qquad\qquad+K^{\delta,\varepsilon}(t,e)\left[X_1^{\delta,\varepsilon}(t-)\Delta\Upsilon^{\delta}(t,e)^{\top}+\Delta\Upsilon^{\delta}(t,e)X_1^{\delta,\varepsilon}(t-)^{\top}\right.\\
&\qquad\qquad+\Upsilon_X^{\delta}(t,e)X_1^{\delta,\varepsilon}(t-)\Delta\Upsilon^{\delta}(t,e)^{\top}+\Delta\Upsilon^{\delta}(t,e)X_1^{\delta,\varepsilon}(t-)\Upsilon_X^{\delta}(t,e)^{\top}\\
&\qquad\qquad\left.\left.\left.+\Delta\Upsilon^{\delta}(t,e)\Delta\Upsilon^{\delta}(t,e)^{\top}\right]I_{S_\varepsilon}(t)\right\}\pi(de)\right\}dt\\
&=\mathbb{E}\int_0^T\operatorname{tr}\left\{P^{\delta,\varepsilon}(t)\left[\Delta\Sigma^{\delta}(t,\cdot)\Delta\Sigma^{\delta}(t,\cdot)^{\top}+\int_{\mathbf{E}}\Delta\Upsilon^{\delta}(t,e)\Delta\Upsilon^{\delta}(t,e)^{\top}\pi(de)\right]I_{S_\varepsilon}(t)\right.\\
&\qquad+\int_{\mathbf{E}}K^{\delta,\varepsilon}(t,e)\Delta\Upsilon^{\delta}(t,e)\Delta\Upsilon^{\delta}(t,e)^{\top}I_{S_\varepsilon}(t)\pi(de)\\
&\qquad\left.-H_{XX}^{\delta,\varepsilon,\Phi_0^{\delta,\varepsilon}}(t,\cdot)Y^{\delta,\varepsilon}(t)\right\}dt+O\left(\varepsilon^{3/2}\right)\\
&=\mathbb{E}\int_0^T\operatorname{tr}\left\{\Delta\Sigma^{\delta}(t,\cdot)^{\top}P^{\delta,\varepsilon}(t)\Delta\Sigma^{\delta}(t,\cdot)I_{S_\varepsilon}(t)\right.\\
&\qquad+\int_{\mathbf{E}}\Delta\Upsilon^{\delta}(t,e)^{\top}\left[P^{\delta,\varepsilon}(t)+K^{\delta,\varepsilon}(t,e)\right]\Delta\Upsilon^{\delta}(t,e)I_{S_\varepsilon}(t)\pi(de)\\
&\qquad\left.-\left\langle H_{XX}^{\delta,\varepsilon,\Phi_0^{\delta,\varepsilon}}(t,\cdot)X_1^{\delta,\varepsilon}(t),X_1^{\delta,\varepsilon}(t)\right\rangle\right\}dt+O\left(\varepsilon^{3/2}\right).
\end{aligned}\tag{3.69}$$

Consequently, from (3.64), we have

$$\begin{aligned}
&-\sqrt{\delta}\left[\sqrt{\varepsilon}y_0+\varepsilon(T+K)\right]\\
&\quad\le \mathbb{E}\int_0^T\Big\{\Phi_0^{\delta,\varepsilon}\Big[l\Big(t,X^\delta(t),v^{\delta,\varepsilon}(t,\cdot)\Big)-l\Big(t,X^\delta(t),v^\delta(t,\cdot)\Big)\Big]\\
&\qquad\qquad-\Big\langle\widetilde{\Phi}^{\delta,\varepsilon}(t),\Delta B^\delta(t,\cdot)\Big\rangle-\Big\langle\widetilde{\Psi}^{\delta,\varepsilon}(t),\Delta\Sigma^\delta(t,\cdot)\Big\rangle\\
&\qquad\qquad-\int_{\mathbf{E}}\Big\langle\widetilde{\Xi}^{\delta,\varepsilon}(t,e),\Delta\Upsilon^\delta(t,e)\Big\rangle\pi(de)\\
&\qquad\qquad-\frac{1}{2}\Delta\Sigma^\delta(t,\cdot)^\top P^{\delta,\varepsilon}(t)\Delta\Sigma^\delta(t,\cdot)\\
&\qquad\qquad-\frac{1}{2}\int_{\mathbf{E}}\Delta\Upsilon^\delta(t,e)^\top\Big[P^{\delta,\varepsilon}(t)+K^{\delta,\varepsilon}(t,e)\Big]\Delta\Upsilon^\delta(t,e)\pi(de)\Big\}I_{S_\varepsilon}(t)dt\\
&\qquad+\mathbb{E}\Bigg\{\sqrt{\varepsilon}\Bigg\langle\Phi_0^{\delta,\varepsilon}\Gamma_{X_0}\Big(X_0^\delta,X^\delta(T)\Big)-\widetilde{\Phi}^{\delta,\varepsilon}(0),\begin{pmatrix}0\\y_0\end{pmatrix}\Bigg\rangle\\
&\qquad\qquad+\frac{\varepsilon}{2}\Bigg\langle\Big(\Phi_0^{\delta,\varepsilon}\Gamma_{X_0X_0}\Big(X_0^\delta,X^\delta(T)\Big)-P^{\delta,\varepsilon}(0)\Big)\begin{pmatrix}0\\y_0\end{pmatrix},\begin{pmatrix}0\\y_0\end{pmatrix}\Bigg\rangle\Bigg\}+O\Big(\varepsilon^{3/2}\Big).
\end{aligned}\tag{3.70}$$

Step 6 (*variational inequality*). In (3.70), dividing $\sqrt{\varepsilon}$ and then sending $\varepsilon\to 0$ followed by sending $\delta\to 0$, we get

$$\begin{aligned}
0&\le\mathbb{E}\Bigg\langle\Phi_0\Gamma_{X_0}(X_0,X(T))-\widetilde{\Phi}(0),\begin{pmatrix}0\\y_0\end{pmatrix}\Bigg\rangle\\
&=\mathbb{E}\Bigg\langle\Phi_0\begin{pmatrix}0\\\gamma_y(y_0)\end{pmatrix}-\widetilde{\Phi}(0),\begin{pmatrix}0\\y_0\end{pmatrix}\Bigg\rangle,\quad\forall y_0\in\mathbf{R}^m,
\end{aligned}\tag{3.71}$$

which implies

$$\widetilde{\Phi}(0)=\Phi_0\begin{pmatrix}0\\\mathbb{E}\gamma_y(y_0)\end{pmatrix}.\tag{3.72}$$

Using a standard argument of Tang and Li [24] in (3.70), we have the following variational inequality:

$$\begin{aligned}
&\Phi_0\Big[l\Big(t,\overline{X}(t),v(t,\cdot)\Big)-l\Big(t,\overline{X}(t),\overline{v}(t,\cdot)\Big)\Big]-\Big\langle\widetilde{\Phi}(t),B\Big(t,\overline{X}(t),v(t,\cdot)\Big)-B\Big(t,\overline{X}(t),\overline{v}(t,\cdot)\Big)\Big\rangle\\
&\quad-\Big\langle\widetilde{\Psi}(t),\Sigma\Big(t,\overline{X}(t),v(t,\cdot)\Big)-\Sigma\Big(t,\overline{X}(t),\overline{v}(t,\cdot)\Big)\Big\rangle
\end{aligned}$$

$$
\begin{aligned}
&-\int_{\mathbf{E}}\left\langle\tilde{\Xi}(t,e),\Upsilon\Big(t,\overline{X}(t),v(t,e)\Big)-\Upsilon\Big(t,\overline{X}(t),\overline{v}(t,e)\Big)\right\rangle\pi(de)\\
&-\frac{1}{2}\left[\Sigma\Big(t,\overline{X}(t),v(t,\cdot)\Big)-\Sigma\Big(t,\overline{X}(t),\overline{v}(t,\cdot)\Big)\right]^{\top}P(t)\left[\Sigma\Big(t,\overline{X}(t),v(t,\cdot)\Big)-\Sigma\Big(t,\overline{X}(t),\overline{v}(t,\cdot)\Big)\right]\\
&-\frac{1}{2}\int_{\mathbf{E}}\left[\Upsilon\Big(t,\overline{X}(t),v(t,e)\Big)-\Upsilon\Big(t,\overline{X}(t),\overline{v}(t,e)\Big)\right]^{\top}[P(t)+K(t,e)]\\
&\times\left[\Upsilon\Big(t,\overline{X}(t),v(t,e)\Big)-\Upsilon\Big(t,\overline{X}(t),\overline{v}(t,e)\Big)\right]\pi(de)\geq 0,\quad\forall v(\cdot),\ \text{a.e., a.s.}
\end{aligned}
\tag{3.73}
$$

Step 7 (finalizing the proof). Since $\Phi_0\neq 0$, by rescaling, we let $\Phi_0=1$. Then the first-order adjoint equation (3.61) reduces to

$$
\begin{aligned}
-d\tilde{\Phi}(t)&=\left[\overline{B}_X(t,\cdot)^{\top}\tilde{\Phi}(t)+\overline{\Sigma_X}(t,\cdot)^{\top}\tilde{\Psi}(t)+\int_{\mathbf{E}}\overline{\Upsilon}_X(t,e)^{\top}\tilde{\Xi}(t,e)\pi(de)-\bar{l}_X(t,\cdot)\right]dt\\
&\quad-\tilde{\Psi}(t)dW(t)-\int_{\mathbf{E}}\tilde{\Xi}(t,e)\widetilde{N}(dedt),\\
\tilde{\Phi}(T)&=-\left[\Gamma_{X_T}\Big(\overline{X}_0,\overline{X}(T)\Big)+\Pi_{X_T}\Big(0,\overline{X}(T)\Big)\begin{pmatrix}0\\ \Phi_T\end{pmatrix}\right].
\end{aligned}
\tag{3.74}
$$

Let $(\delta,\varepsilon)\to 0$ in (3.66), then $(P^{\delta,\varepsilon}(\cdot),Q^{\delta,\varepsilon}(\cdot),K^{\delta,\varepsilon}(\cdot,\cdot))\to(P(\cdot),Q(\cdot),K(\cdot,\cdot))$ which is the adapted solution to the following second-order adjoint equation:

$$
\begin{aligned}
-dP(t)&=\Bigg\{\overline{B}_X(t,\cdot)^{\top}P(t)+P(t)\overline{B}_X(t,\cdot)+\overline{\Sigma_X}(t,\cdot)^{\top}P(t)\overline{\Sigma_X}(t,\cdot)+\overline{\Sigma_X}(t,\cdot)^{\top}Q(t)+Q(t)\overline{\Sigma_X}(t,\cdot)\\
&\qquad+\int_{\mathbf{E}}\Big[\overline{\Upsilon}_X(t,e)^{\top}P(t)\overline{\Upsilon}_X(t,e)+\overline{\Upsilon}_X(t,e)^{\top}K(t,e)\overline{\Upsilon}_X(t,e)\\
&\qquad\qquad+\overline{\Upsilon}_X(t,e)^{\top}K(t,e)+K(t,e)\overline{\Upsilon}_X(t,e)\Big]\pi(de)+\overline{H}_{XX}(t,\cdot)\Bigg\}dt\\
&\qquad-Q(t)dW(t)-\int_{\mathbf{E}}K(t,e)\widetilde{N}(dedt),\\
P(T)&=-\left[\Gamma_{X_TX_T}\Big(\overline{X}_0,\overline{X}(T)\Big)+\Pi_{X_TX_T}\Big(0,\overline{X}(T)\Big)\begin{pmatrix}0\\ \Phi_T\end{pmatrix}\right],
\end{aligned}
\tag{3.75}
$$

where

$$
\overline{H}_{XX}(t,\cdot)\doteq\overline{H}_{XX}\Big(t,\overline{X}(t),\overline{v}(t,\cdot),\tilde{\Phi}(t),\tilde{\Psi}(t),\tilde{\Xi}(t,\cdot)\Big),
\tag{3.76}
$$

with $\overline{H}_{XX}(t,X,v(\cdot),\widetilde{\Phi},\widetilde{\Psi},\widetilde{\Xi}(\cdot))$ being the Hessian of the following

$$\begin{aligned}\overline{H}\Big(t,X,v(\cdot),\widetilde{\Phi},\widetilde{\Psi},\widetilde{\Xi}(\cdot)\Big) &\doteq \Big\langle \widetilde{\Phi}, B(t,X,v(\cdot))\Big\rangle + \Big\langle \widetilde{\Psi}, \Sigma(t,X,v(\cdot))\Big\rangle \\ &\quad + \int_{\mathbf{E}} \Big\langle \widetilde{\Xi}(e), \Upsilon(t,X,v(e))\Big\rangle \pi(de) - l(t,X,v(\cdot)).\end{aligned} \tag{3.77}$$

Also, from (3.72), we have

$$\widetilde{\Phi}(0) = \begin{pmatrix} 0 \\ \mathbb{E}\gamma_y(y_0) \end{pmatrix}. \tag{3.78}$$

Note that

$$\Gamma_{X_T}\Big(\overline{X}_0,\overline{X}(T)\Big) = \begin{pmatrix} \phi_x(\overline{x}(T)) \\ 0 \end{pmatrix}, \qquad \Gamma_{X_T X_T}\Big(\overline{X}_0,\overline{X}(T)\Big) = \begin{pmatrix} \phi_{xx}(\overline{x}(T)) & 0 \\ 0 & 0 \end{pmatrix}. \tag{3.79}$$

On the other hand, since

$$\left\langle \Pi\Big(0,\overline{X}(T)\Big), \begin{pmatrix} 0 \\ \Phi_T \end{pmatrix}\right\rangle = \Big(\overline{y}(T) - h\Big(\overline{X}(T)\Big)\Big)\Phi_T, \tag{3.80}$$

we have

$$\begin{aligned}\Pi_{X_T}\Big(0,\overline{X}(T)\Big)\begin{pmatrix} 0 \\ \Phi_T \end{pmatrix} &= \begin{pmatrix} -h_x(\overline{x}(T))^\top \Phi_T \\ \Phi_T \end{pmatrix}, \\ \Pi_{X_T X_T}\Big(0,\overline{X}(T)\Big)\begin{pmatrix} 0 \\ \Phi_T \end{pmatrix} &= \begin{pmatrix} -h_{xx}(\overline{x}(T))^\top \Phi_T & 0 \\ 0 & 0 \end{pmatrix}.\end{aligned} \tag{3.81}$$

Let

$$\widetilde{\Phi}(\cdot) \equiv \begin{pmatrix} p(\cdot) \\ q(\cdot) \end{pmatrix}, \qquad \widetilde{\Psi}(\cdot) \equiv \begin{pmatrix} k(\cdot) \\ \widetilde{k}(\cdot) \end{pmatrix}, \qquad \widetilde{\Xi}(\cdot,\cdot) \equiv \begin{pmatrix} r(\cdot,\cdot) \\ \widetilde{r}(\cdot,\cdot) \end{pmatrix}. \tag{3.82}$$

Then it follows from (3.74), (3.78), and (3.82) that

$$\begin{pmatrix} p(0) \\ q(0) \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbb{E}\gamma_y(y_0) \end{pmatrix}, \qquad \begin{pmatrix} p(T) \\ q(T) \end{pmatrix} = \begin{pmatrix} -\phi_x(\overline{x}(T)) + h_x(\overline{x}(T))^\top \Phi_T \\ -\Phi_T \end{pmatrix}. \tag{3.83}$$

By (3.82), we can rewrite the equations in (3.74) as

$$-d\begin{pmatrix} p(t) \\ q(t) \end{pmatrix} = \left\{ \begin{pmatrix} \overline{b}_x(t,\cdot)^\top & -\overline{f}_x(t,\cdot)^\top \\ \overline{b}_y(t,\cdot)^\top & -\overline{f}_y(t,\cdot)^\top \end{pmatrix} \begin{pmatrix} p(t) \\ q(t) \end{pmatrix} + \begin{pmatrix} \overline{\sigma}_x(t,\cdot)^\top & 0 \\ \overline{\sigma}_y(t,\cdot)^\top & 0 \end{pmatrix} \begin{pmatrix} k(t) \\ \widetilde{k}(t) \end{pmatrix} \right.$$
$$\left. + \int_{\mathbf{E}} \begin{pmatrix} \overline{g}_x(t,e)^\top & 0 \\ \overline{g}_y(t,e)^\top & 0 \end{pmatrix} \begin{pmatrix} r(t,e) \\ \widetilde{r}(t,e) \end{pmatrix} \pi(de) - \begin{pmatrix} \overline{l}_x(t,\cdot) \\ \overline{l}_y(t,\cdot) \end{pmatrix} \right\} dt \tag{3.84}$$
$$- \begin{pmatrix} k(t) \\ \widetilde{k}(t) \end{pmatrix} dW(t) - \int_{\mathbf{E}} \begin{pmatrix} r(t,e) \\ \widetilde{r}(t,e) \end{pmatrix} \widetilde{N}(dedt),$$

and the variational inequality (3.73) now takes the form:

$$\begin{aligned}
& l(t,\overline{x}(t),\overline{y}(t),z,c(\cdot),u) - l(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) \\
& - \langle p(t), b(t,\overline{x}(t),\overline{y}(t),z,c(\cdot),u) - b(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) \rangle \\
& + \langle q(t), f(t,\overline{x}(t),\overline{y}(t),z,c(\cdot),u) - f(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) \rangle \\
& - \langle k(t), \sigma(t,\overline{x}(t),\overline{y}(t),z,c(\cdot),u) - \sigma(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) \rangle \\
& - \int_{\mathbf{E}} \langle r(t,e), g(t,\overline{x}(t),\overline{y}(t),z,c(e),u) - g(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,e),\overline{u}(t)) \rangle \pi(de) \\
& - \left\langle \widetilde{k}(t), z-\overline{z}(t) \right\rangle - \int_{\mathbf{E}} \langle \widetilde{r}(t,e), c(e)-\overline{c}(t,e) \rangle \pi(de) \\
& - \frac{1}{2} \begin{pmatrix} \sigma(t,\overline{x}(t),\overline{y}(t),z,c(\cdot),u) - \sigma(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) \\ z-\overline{z}(t) \end{pmatrix}^\top P(t) \\
& \times \begin{pmatrix} \sigma(t,\overline{x}(t),\overline{y}(t),z,c(\cdot),u) - \sigma(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) \\ z-\overline{z}(t) \end{pmatrix} \\
& - \frac{1}{2} \int_{\mathbf{E}} \begin{pmatrix} g(t,\overline{x}(t),\overline{y}(t),z,c(e),u,e) - g(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,e),\overline{u}(t),e) \\ c(e)-\overline{c}(t,e) \end{pmatrix}^\top \\
& \times [P(t)+K(t,e)] \\
& \times \begin{pmatrix} g(t,\overline{x}(t),\overline{y}(t),z,c(e),u,e) - g(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,e),\overline{u}(t),e) \\ c(e)-\overline{c}(t,e) \end{pmatrix} \pi(de) \\
& \geq 0, \quad \forall z, c(\cdot), u, \ \text{a.e., a.s.}
\end{aligned} \tag{3.85}$$

Thus, taking $u = \overline{u}(\cdot)$, $z = \overline{z}(t) + \varepsilon z_0$ and $c(\cdot) = \overline{c}(t,\cdot) + \varepsilon c_0(\cdot)$, then dividing by ε and sending $\varepsilon \to 0$, we get

$$\begin{aligned}
0 \leq & \langle l_z(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)), z_0 \rangle - \langle p(t), b_z(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) z_0 \rangle \\
& + \langle q(t), f_z(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) z_0 \rangle - \langle k(t), \sigma_z(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) z_0 \rangle
\end{aligned}$$

$$
\begin{aligned}
&- \int_{\mathbf{E}} \langle r(t,e), g_z(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,e),\overline{u}(t)) z_0 \rangle - \left\langle \widetilde{k}(t), z_0 \right\rangle \\
&+ \langle l_c(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)), c_0(\cdot) \rangle - \langle p(t), b_c(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) c_0(\cdot) \rangle \\
&+ \langle q(t), f_c(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) c_0(\cdot) \rangle - \langle k(t), \sigma_c(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) c_0(\cdot) \rangle \\
&- \int_{\mathbf{E}} \langle r(t,e), g_c(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,e),\overline{u}(t)) c_0(e) \rangle \\
&- \int_{\mathbf{E}} \langle \widetilde{r}(t,e), c_0(e) \rangle \pi(de), \quad \forall z_0 \in \mathbf{R}^m, \ c_0(\cdot) \in \mathcal{L}^2(\mathbf{E}, \mathcal{B}(\mathbf{E}), \pi; \mathbf{R}^m).
\end{aligned}
\tag{3.86}
$$

Hence,

$$
\begin{aligned}
\widetilde{k}(t) = {} & l_z(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) - b_z(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t))^\top p(t) \\
& + f_z(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t))^\top q(t) - \sigma_z(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t))^\top k(t) \\
& - \int_{\mathbf{E}} g_z(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,e),\overline{u}(t))^\top r(t,e) \pi(de), \\
\widetilde{r}(t,\cdot) = {} & l_c(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) - b_c(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t))^\top p(t) \\
& + f_c(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t))^\top q(t) - \sigma_c(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t))^\top k(t) \\
& - \int_{\mathbf{E}} g_c(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,e),\overline{u}(t))^\top r(t,e) \pi(de).
\end{aligned}
\tag{3.87}
$$

Combining (3.83), (3.84) with (3.87), we arrive at (2.10).

Next, we have

$$
\begin{aligned}
P(T) &= -\left[\Gamma_{X_T X_T}\left(\overline{X}_0, \overline{X}(T)\right) + \Pi_{X_T X_T}\left(0, \overline{X}(T)\right) \begin{pmatrix} 0 \\ \Phi_T \end{pmatrix} \right] \\
&= \begin{pmatrix} -\phi_{xx}(\overline{x}(T)) + h_{xx}(\overline{x}(T))^\top \Phi_T & 0 \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} -\phi_{xx}(\overline{x}(T)) - h_{xx}(\overline{x}(T))^\top q(T) & 0 \\ 0 & 0 \end{pmatrix}.
\end{aligned}
\tag{3.88}
$$

Then (2.11) follows from (3.75).

Further, by taking $z = \overline{z}(\cdot), c(\cdot) = \overline{c}(\cdot,\cdot)$ in the variational inequality (3.85), we obtain

$$
\begin{aligned}
& l(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),u) - l(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) \\
& \quad - \langle p(t), b(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),u) - b(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) \rangle \\
& \quad + \langle q(t), f(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),u) - f(t,\overline{x}(t),\overline{y}(t),\overline{z}(t),\overline{c}(t,\cdot),\overline{u}(t)) \rangle
\end{aligned}
$$

$$
\begin{aligned}
&- \left\langle k(t), \sigma\big(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{c}(t,\cdot), u\big) - \sigma\big(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{c}(t,\cdot), \bar{u}(t)\big) \right\rangle \\
&- \int_{\mathbf{E}} \left\langle r(t,e), g\big(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{c}(t,e), u, e\big) - g\big(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{c}(t,e), \bar{u}(t), e\big) \right\rangle \pi(de) \\
&- \frac{1}{2} \begin{pmatrix} \sigma\big(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{c}(t,\cdot), u\big) - \sigma\big(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{c}(t,\cdot), \bar{u}(t)\big) \\ 0 \end{pmatrix}^{\top} P(t) \\
&\times \begin{pmatrix} \sigma\big(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{c}(t,\cdot), u\big) - \sigma\big(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{c}(t,\cdot), \bar{u}(t)\big) \\ 0 \end{pmatrix} \\
&- \frac{1}{2} \int_{\mathbf{E}} \begin{pmatrix} g\big(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{c}(t,\cdot), u, e\big) - g\big(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{c}(t,e), \bar{u}(t), e\big) \\ 0 \end{pmatrix}^{\top} \\
&\times [P(t) + K(t,e)] \\
&\times \begin{pmatrix} g\big(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{c}(t,\cdot), u, e\big) - g\big(t, \bar{x}(t), \bar{y}(t), \bar{z}(t), \bar{c}(t,e), \bar{u}(t), e\big) \\ 0 \end{pmatrix} \pi(de) \\
\geq{}& 0, \quad \forall u \in \mathbf{U}, \ \text{a.e., a.s.},
\end{aligned}
\tag{3.89}
$$

which gives (2.13). This completes the proof of our main result—Theorem 2.1.

4. A Linear Quadratic Example

In this section, we briefly discuss a linear quadratic case, which serves as an illustrating example of our main result. Consider the following linear controlled FBSDEJ:

$$
\begin{aligned}
dx(t) &= \big[b_1 x(t) + b_2 y(t) + b_3 z(t) + b_4 c(t,\cdot) + b_5 u(t)\big] dt \\
&\quad + \big[\sigma_1 x(t) + \sigma_2 y(t) + \sigma_3 z(t) + \sigma_4 c(t,\cdot) + \sigma_5 u(t)\big] dW(t) \\
&\quad + \int_{\mathbf{E}} \big[g_1 x(t-) + g_2 y(t-) + g_3 z(t) + g_4 c(t,e) + g_5 u(t)\big] \widetilde{N}(dedt), \\
-dy(t) &= \big[f_1 x(t) + f_2 y(t) + f_3 z(t) + f_4 c(t,\cdot) + f_5 u(t)\big] dt \\
&\quad - z(t) dW(t) - \int_{\mathbf{E}} c(t,e) \widetilde{N}(dedt), \\
x(0) &= a, \qquad y(T) = h_1 x(T),
\end{aligned}
\tag{4.1}
$$

with the quadratic cost functional given by

$$
J(u(\cdot)) = \frac{1}{2} \mathbb{E} \left\{ \int_0^T \Big[l_1 x^2(t) + l_2 y^2(t) + l_3 z^2(t) + l_4 c^2(t,\cdot) + l_5 u^2(t) \Big] dt + \phi_1 x^2(T) + \gamma_1 y^2(0) \right\}.
\tag{4.2}
$$

In the above, $b_i, \sigma_i, g_i, f_i, l_i$, $i = 1, \ldots, 5$ and h_1, ϕ_1, γ_1 are all real constants of $\mathbf{R}$. The control domain $\mathbf{U} \subseteq \mathbf{R}$ could be very arbitrary; in particular, $\mathbf{U}$ does not have to be a convex set. We let

$\mathcal{U}[0,T]$ be the set of all $\mathcal{F}_t$-predictable processes $u:[0,T]\times\Omega \to \mathbf{U}$ such that $\sup_{0\le t\le T}\mathbb{E}|u(t)|^i < +\infty$, $\forall i = 1,2,\ldots$.

For any $u(\cdot) \in \mathcal{U}[0,T]$, we denote

$$\Xi = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad \mathcal{A}(t,\Xi,c(\cdot),u) = \begin{pmatrix} -f_1x - f_2y - f_3z - f_4c(\cdot) - f_5u \\ b_1x + b_2y + b_3z + b_4c(\cdot) + b_5u \\ \sigma_1x + \sigma_2y + \sigma_3z + \sigma_4c(\cdot) + \sigma_5u \end{pmatrix}, \tag{4.3}$$

and assume that (H4.1)

$$\begin{aligned} &\langle \mathcal{A}(t,\Xi,c(\cdot),u) - \mathcal{A}(t,\Xi',c'(\cdot),u), \Xi - \Xi' \rangle \\ &\quad + \int_{\mathbf{E}} \langle g_1(x-x') + g_2(y-y') + g_3(z-z') + g_4(c(e)-c'(e)), c(e)-c'(e)\rangle \pi(de) \\ &\le -\beta_1|x-x'|^2 - \beta_2\left(|y-y'|^2 + |z-z'|^2 + \int_{\mathbf{E}} |c(e)-c'(e)|^2\pi(de)\right), \\ &h_1 \ge \mu_1 > 0, \end{aligned} \tag{4.4}$$

where β_1,β_2, and μ_1 are given constants with $\beta_1 > 0$, $\mu_1 > 0$, $\beta_2 \ge 0$. Then by Wu [1], there exists a unique adapted solution $(x(\cdot),y(\cdot),z(\cdot),c(\cdot,\cdot))$ to (4.1).

Now, suppose that the corresponding stochastic control problem admits an optimal 5-tuple $(\overline{x}(\cdot),\overline{y}(\cdot),\overline{z}(\cdot),\overline{c}(\cdot,\cdot),\overline{u}(\cdot))$. Now let us look at Theorem 2.1. The first-order adjoint equation (2.10) reads

$$\begin{aligned} -dp(t) &= \left[b_1p(t) - f_1q(t) + \sigma_1k(t) + \int_{\mathbf{E}} g_1r(t,e)\pi(de) - l_1\overline{x}(t)\right]dt \\ &\quad - k(t)dW(t) - \int_{\mathbf{E}} r(t,e)\widetilde{N}(dedt), \\ dq(t) &= \left[-b_2p(t) + f_2q(t) - \sigma_2k(t) - \int_{\mathbf{E}} g_2r(t,e)\pi(de) + l_2\overline{y}(t)\right]dt \\ &\quad + \left[-b_3p(t) + f_3q(t) - \sigma_3k(t) - \int_{\mathbf{E}} g_3r(t,e)\pi(de) + l_3\overline{z}(t)\right]dW(t) \\ &\quad + \int_{\mathbf{E}} [-b_4p(t-) + f_4q(t-) - \sigma_4k(t) - g_4r(t,e) + l_4\overline{c}(t,e)]\widetilde{N}(dedt), \\ p(T) &= -\phi_1\overline{x}(T) - h_1q(T), \qquad q(0) = \gamma_1\overline{y}_0. \end{aligned} \tag{4.5}$$

Similarly, we denote

$$\vartheta = \begin{pmatrix} q \\ p \\ k \end{pmatrix}, \qquad \mathcal{B}(t,\vartheta,r(\cdot)) = \begin{pmatrix} -b_1p + f_1q - \sigma_1k - g_1r(\cdot) + l_1\overline{x} \\ -b_2p + f_2q - \sigma_2k - g_2r(\cdot) + l_2\overline{y} \\ -b_2p + f_3q - \sigma_3k - g_3r(\cdot) + l_3\overline{z} \end{pmatrix}. \tag{4.6}$$

Since controlled FBSDEJ satisfies (H4.1), we can easily check that the above FBSDEJ (4.4) satisfies the following condition (H4.2):

$$
\begin{aligned}
&\langle \mathcal{B}(t,\vartheta,r(\cdot),u) - \mathcal{B}(t,\vartheta',r'(\cdot),u), \vartheta - \vartheta' \rangle \\
&\quad + \int_{\mathbf{E}} \langle -b_4(p-p') + f_4(q-q') - \sigma_4(k-k') \\
&\quad - g_4(r(e)-r'(e)) + l_4\bar{c}(e), r(e)-r'(e) \rangle \pi(de) \\
&\geq \beta_1 |q-q'|^2 + \beta_2 \left(|p-q'|^2 + |k-k'|^2 + \int_{\mathbf{E}} |r(e)-r'(e)|^2 \pi(de) \right), \\
&\quad - h_1 \leq -\mu_1,
\end{aligned}
\tag{4.7}
$$

where β_1, β_2, and μ_1 are the same as in (H4.1). Then by Wu [38], FBSDEJ (4.4) admits a unique adapted solution $(q(\cdot), p(\cdot), k(\cdot), r(\cdot,\cdot))$.

The 2×2 matrix-valued second-order adjoint equation (2.11) becomes

$$
\begin{aligned}
-dP(t) &= \Big\{ B^{\top}P(t) + P(t)B + \Sigma^{\top}P(t)\Sigma + \Sigma^{\top}Q(t) + Q(t)\Sigma + \Upsilon^{\top}P(t)\Upsilon \\
&\qquad + \int_{\mathbf{E}} \left[\Upsilon^{\top}K(t,e)\Upsilon + \Upsilon^{\top}K(t,e) + K(t,e)\Upsilon \right] \pi(de) - L \Big\} dt \\
&\qquad - Q(t)dW(t) - \int_{\mathbf{E}} K(t,e)\widetilde{N}(dedt), \\
P(T) &= \begin{pmatrix} -\phi_1 & 0 \\ 0 & 0 \end{pmatrix},
\end{aligned}
\tag{4.8}
$$

where

$$
B = \begin{pmatrix} b_1 & b_2 \\ -f_1 & -f_2 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \sigma_1 & \sigma_2 \\ 0 & 0 \end{pmatrix}, \qquad \Upsilon = \begin{pmatrix} g_1 & g_2 \\ 0 & 0 \end{pmatrix}, \qquad L = \begin{pmatrix} l_1 & 0 \\ 0 & l_2 \end{pmatrix}. \tag{4.9}
$$

As before, (4.5) can be split into the following three scalar BSDEJs:

$$
\begin{aligned}
-dP_1(t) = \Big\{ & 2b_1P_1(t) + \sigma_1^2P_1(t) + 2\sigma_1Q_1(t) - 2f_1P_2(t) + g_1^2P_1(t) \\
& + \int_{\mathbf{E}} \left[g_1^2K_1(t,e) + 2g_1K_1(t,e) \right] \pi(de) - l_1 \Big\} dt
\end{aligned}
$$

$$
\begin{aligned}
&\qquad - Q_1(t)dW(t) - \int_{\mathbf{E}} K_1(t,e)\widetilde{N}(dedt),\\
&P_1(T) = -\phi_1,
\end{aligned}
\tag{4.10}
$$

$$
\begin{aligned}
-dP_2(t) = &\Big\{ b_2P_1(t) + b_1P_2(t) + \sigma_1\sigma_2P_1(t) + \sigma_1Q_2(t) + \sigma_2Q_1(t) - f_1P_3(t) - f_2P_2(t)\\
&+ g_1g_2P_1(t) + \int_{\mathbf{E}} \big[g_1g_2K_1(t,e) + g_1K_2(t,e) + g_2K_1(t,e)\big]\pi(de)\Big\}dt\\
&- Q_2(t)dW(t) - \int_{\mathbf{E}} K_2(t,e)\widetilde{N}(dedt),\\
P_2(T) = 0,&
\end{aligned}
\tag{4.11}
$$

$$
\begin{aligned}
-dP_3(t) = &\Big\{ 2b_2P_2(t) + \sigma_2^2P_1(t) + 2\sigma_2Q_2(t) - 2f_2P_3(t) + g_2^2P_1(t)\\
&+ \int_{\mathbf{E}} \Big[g_2^2K_1(t,e) + 2g_2K_2(t,e)\Big]\pi(de) - l_2\Big\}dt\\
&- Q_3(t)dW(t) - \int_{\mathbf{E}} K_3(t,e)\widetilde{N}(dedt),\\
P_3(T) = 0.&
\end{aligned}
\tag{4.12}
$$

Note that since all the coefficients in (4.5) are constants and the terminal condition is deterministic, we must have

$$
Q(\cdot) = 0, \qquad K(\cdot,\cdot) = 0, \tag{4.13}
$$

and $P(\cdot)$ is deterministic. Consequently,

$$
\dot{P}(t) = -\big[B^\top P(t) + P(t)B + \Sigma^\top P(t)\Sigma + \Upsilon^\top P(t)\Upsilon - L\big], \qquad P(T) = \begin{pmatrix} -\phi_1 & 0\\ 0 & 0\end{pmatrix}. \tag{4.14}
$$

And at the same time (4.7)–(4.10) reduce to

$$
\begin{aligned}
\dot{P}_1(t) &= -\Big(2b_1 + \sigma_1^2 + g_1^2\Big)P_1(t) + 2f_2P_2(t) + l_1, \qquad P_1(T) = -\phi_1,\\
\dot{P}_2(t) &= -(b_2 + \sigma_1\sigma_2 - g_1g_2)P_1(t) + (b_1 - f_2)P_2(t) - f_1P_3(t), \qquad P_2(T) = 0,\\
\dot{P}_3(t) &= -(\sigma_2^2 + g_2^2)P_1(t) - \sigma_2^2P_2(t) + 2f_2P_3(t) - l_2, \qquad P_3(T) = 0,
\end{aligned}
\tag{4.15}
$$

and $Q_i(\cdot) = 0$, $K_i(\cdot,\cdot) = 0$, $i = 1,2,3$. Clearly, linear ODE (4.11) admits a unique explicit solution and then ODEs (4.12) do.

Since the Hamiltonian function (2.12) reads

$$
\begin{aligned}
&H(x,y,z,c(\cdot),u,p,q,k,r(\cdot)) \\
&\quad = (b_1x + b_2y + b_3z + b_4c(\cdot) + b_5u)p - (f_1x + f_2y + f_3z + f_4c(\cdot) + f_5u)q \\
&\qquad + (\sigma_1x + \sigma_2y + \sigma_3z + \sigma_4c(\cdot) + \sigma_5u)k \\
&\qquad + \int_{\mathbf{E}} (g_1x + g_2y + g_3z + g_4c(e) + g_5u)r(e)\pi(de) \\
&\qquad - \frac{1}{2}\left(l_1x^2 + l_2y^2 + l_3z^2 + l_4c^2(\cdot) + l_5u^2\right),
\end{aligned}
\tag{4.16}
$$

the maximum condition (2.13) now takes the form (noting that $K_1(\cdot,\cdot) = 0$)

$$
\begin{aligned}
&\left[b_5p(t) - f_5q(t) + \sigma_5k(t) + \int_{\mathbf{E}} g_5r(t,e)\pi(de)\right][\overline{u}(t) - u] \\
&\quad - \frac{1}{2}l_5\left[\overline{u}^2(t) - u^2\right] - \frac{1}{2}\left[\sigma_5^2 + g_5^2\right][\overline{u}(t) - u]^2P_1(t) \geq 0, \quad \forall u \in \mathbf{U},\ t \in [0,T].
\end{aligned}
\tag{4.17}
$$

This implies that

$$
\begin{aligned}
&b_5p(t) - f_5q(t) + \sigma_5k(t) + \int_{\mathbf{E}} g_5r(t,e)\pi(de) \\
&\quad \geq \frac{1}{2}l_5[\overline{u}(t) + u] + \frac{1}{2}\left[\sigma_5^2 + g_5^2\right][\overline{u}(t) - u]P_1(t), \quad \forall u \in \mathbf{U},\ u > \overline{u}(t),\ t \in [0,T]; \\
&b_5p(t) - f_5q(t) + \sigma_5k(t) + \int_{\mathbf{E}} g_5r(t,e)\pi(de) \\
&\quad \leq \frac{1}{2}l_5[\overline{u}(t) + u] + \frac{1}{2}\left[\sigma_5^2 + g_5^2\right][\overline{u}(t) - u]P_1(t), \quad \forall u \in \mathbf{U},\ u \leq \overline{u}(t),\ t \in [0,T].
\end{aligned}
\tag{4.18}
$$

Now, let us look at an interesting special case. Noting that some related problems without random jumps were discussed by Shi and Wu [39] by a purely completion-of-squares technique.

Suppose that

$$
\begin{gathered}
f_1 > 0, \qquad b_2 \leq 0, \qquad \sigma_3 \leq 0, \qquad g_4 \leq 0, \\
b_1 = f_2, \qquad \sigma_1 = f_3, \qquad g_1 = f_4, \qquad \sigma_2 = -b_3, \qquad g_2 = -b_4, \qquad g_3 = -\sigma_4.
\end{gathered}
\tag{4.19}
$$

It is very easy to check that the assumptions (H4.1) and (H4.2) hold. Now the first-order adjoint equation (4.4) reads

$$
\begin{aligned}
-dp(t) = &\left[b_1p(t) - f_1q(t) + \sigma_1k(t) + \int_{\mathbf{E}} g_1r(t,e)\pi(de) - l_1\overline{x}(t)\right]dt \\
&- k(t)\ dW(t) - \int_{\mathbf{E}} r(t,e)\widetilde{N}(dedt),
\end{aligned}
$$

$$\begin{aligned}
dq(t) &= \left[-b_2 p(t) + b_1 q(t) - \sigma_2 k(t) - \int_{\mathbf{E}} g_2 r(t,e)\pi(de) + l_2 \overline{y}(t)\right] dt \\
&\quad + \left[\sigma_2 p(t) + \sigma_1 q(t) - \sigma_3 k(t) + \int_{\mathbf{E}} \sigma_4 r(t,e)\pi(de) + l_3 \overline{z}(t)\right] dW(t) \\
&\quad + \int_{\mathbf{E}} \left[g_2 p(t-) + g_1 q(t-) - \sigma_4 k(t) - g_4 r(t,e) + l_4 \overline{c}(t,e)\right] \widetilde{N}(dedt), \\
p(T) &= -\phi_1 \overline{x}(T) - h_1 q(T), \qquad q(0) = \gamma_1 \overline{y}_0,
\end{aligned} \tag{4.20}$$

and the second-order adjoint ODEs (4.12) reduce to

$$\begin{aligned}
\dot{P}_1(t) &= -\left(2b_1 + \sigma_1^2 + g_1^2\right) P_1(t) + 2b_1 P_2(t) + l_1, \qquad P_1(T) = -\phi_1, \\
\dot{P}_2(t) &= -\left(b_2 + \sigma_1\sigma_2 - g_1 g_2\right) P_1(t) - f_1 P_3(t), \qquad P_2(T) = 0, \\
\dot{P}_3(t) &= -\left(\sigma_2^2 + g_2^2\right) P_1(t) - \sigma_2^2 P_2(t) + 2b_1 P_3(t) - l_2, \qquad P_3(T) = 0.
\end{aligned} \tag{4.21}$$

If, in addition, $b_1 = \sigma_2 = 0$ (noting that this does not mean that the state equation (4.1) is decoupled), then (4.21) further reduce to

$$\dot{P}_1(t) = -\left(\sigma_1^2 + g_1^2\right) P_1(t) + l_1, \qquad P_1(T) = -\phi_1, \tag{4.22}$$

$$\dot{P}_2(t) = -\left(b_2 - g_1 g_2\right) P_1(t) - f_1 P_3(t), \qquad P_2(T) = 0, \tag{4.23}$$

$$\dot{P}_3(t) = -g_2^2 P_1(t) - l_2, \qquad P_3(T) = 0. \tag{4.24}$$

We can first solve $P_1(\cdot)$ explicitly from (4.22):

$$P_1(t) = -e^{(\sigma_1^2+g_1^2)(T-t)} \left[\int_t^T l_1 e^{-(\sigma_1^2+g_1^2)(T-s)} ds + \phi_1\right], \tag{4.25}$$

and then solve $P_3(\cdot)$ explicitly from (4.24). Finally, explicit $P_2(\cdot)$ can be obtained from (4.23). The maximum condition (4.18) remains the same.

However, we cannot obtain the explicit optimal control from the maximum condition (4.18) in general, since this depends heavily on the adjoint processes $(p(\cdot), q(\cdot), k(\cdot), r(\cdot,\cdot))$. The explicit, adapted solution to fully coupled FBSDEJ such as (4.20) is an interesting and challenging open problem. And the explicit optimal control in its state feedback form is rather difficult to obtain even in some very simple cases when there are no random jumps. We will consider some relevant issues in the future work.

5. Concluding Remarks

In this paper, we have discussed a general optimal control problem for fully coupled forward-backward stochastic differential equations with random jumps (FBSDEJs). The control domain is not assumed to be convex, and the control variable appears in both diffusion and jump coefficients of the forward equation. Enlightened by Tang and Li [24], Wu [13], Yong

[20], necessary conditions of Pontraygin's type for the optimal controls are derived by means of spike variation technique and Ekeland variational principle. And the general maximum principle for forward stochastic control systems with random jumps [24] and the maximum principles for forward-backward stochastic control systems with random jumps [25, 26] are recovered in this paper. A linear quadratic stochastic optimal control problem is discussed as an illustrating example.

Acknowledgment

This work was supported by China Postdoctoral Science Foundation Funded Project (no. 20100481278), Postdoctoral Innovation Foundation Funded Project of Shandong Province (no. 201002026), National Natural Sciences Foundations of China (no. 11126209) and Shandong Province (no. ZR2011AQ012), and Independent Innovation Foundation of Shandong University (IIFSDU, no. 2010TS060).

References

[1] Z. Wu, "Forward-backward stochastic differential equations with Brownian motion and Poisson process," *Acta Mathematicae Applicatae Sinica*, vol. 15, no. 4, pp. 433–443, 1999.

[2] S. G. Peng, "A general stochastic maximum principle for optimal control problems," *SIAM Journal on Control and Optimization*, vol. 28, no. 4, pp. 966–979, 1990.

[3] J. M. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, vol. 43, Springer, New York, NY, USA, 1999.

[4] N. El Karoui, S. G. Peng, and M. C. Quenez, "Backward stochastic differential equations in finance," *Mathematical Finance*, vol. 7, no. 1, pp. 1–71, 1997.

[5] N. El Karoui, S. G. Peng, and M. C. Quenez, "A dynamic maximum principle for the optimization of recursive utilities under constraints," *The Annals of Applied Probability*, vol. 11, no. 3, pp. 664–693, 2001.

[6] M. Schroder and C. Skiadas, "Optimal consumption and portfolio selection with stochastic differential utility," *Journal of Economic Theory*, vol. 89, no. 1, pp. 68–126, 1999.

[7] F. Antonelli, E. Barucci, and M. E. Mancino, "Asset pricing with a forward-backward stochastic differential utility," *Economics Letters*, vol. 72, no. 2, pp. 151–157, 2001.

[8] S. L. Ji and X. Y. Zhou, "A maximum principle for stochastic optimal control with terminal state constraints, and its applications," *Communications in Information and Systems*, vol. 6, no. 4, pp. 321–337, 2006.

[9] N. Williams, "On dynamic principal-agent problems in continuous time," Working paper, 2008, http://www.ssc.wisc.edu/~nwilliam/dynamic-pa1.pdf.

[10] J. Cvitanić, X. H Wan, and J. F. Zhang, "Optimal contracts in continuous-time models," *Journal of Applied Mathematics and Stochastic Analysis*, vol. 2006, Article ID 95203, 27 pages, 2006.

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

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

[13] Z. Wu, "A general maximum principle for optimal control problems of forward-backward stochastic control systems," *Working paper*, Shandong University, China, 2009.

[14] J. Cvitanić and J. Ma, "Hedging options for a large investor and forward-backward SDE's," *The Annals of Applied Probability*, vol. 6, no. 2, pp. 370–398, 1996.

[15] D. Cuoco and J. Cvitanić, "Optimal consumption choices for a "large" investor," *Journal of Economic Dynamics & Control*, vol. 22, no. 3, pp. 401–436, 1998.

[16] R. Buckdahn and Y. Hu, "Hedging contingent claims for a large investor in an incomplete market," *Advances in Applied Probability*, vol. 30, no. 1, pp. 239–255, 1998.

[17] 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.

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

[19] 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.

[20] J. M. 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.

[21] R. C. Merton, "Option pricing when underlying stock returns are discontinuous," *Journal of Financial Economics*, vol. 3, no. 1-2, pp. 125–144, 1976.

[22] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*, Chapman & Hall/CRC, Boca Raton, Fla, USA, 2004.

[23] R. Situ, "A maximum principle for optimal controls of stochastic systems with random jumps," in *Proceedings of the National Conference on Control Theory and Applications*, Qingdao, China, 1991.

[24] S. J. Tang and X. J. Li, "Necessary conditions for optimal control of stochastic systems with random jumps," *SIAM Journal on Control and Optimization*, vol. 32, no. 5, pp. 1447–1475, 1994.

[25] B. Øksendal and A. Sulem, "Maximum principles for optimal control of forward-backward stochastic differential equations with jumps," *SIAM Journal on Control and Optimization*, vol. 48, no. 5, pp. 2945–2976, 2009.

[26] J. T. Shi and Z. Wu, "Maximum principle for forward-backward stochastic control system with random jumps and applications to finance," *Journal of Systems Science & Complexity*, vol. 23, no. 2, pp. 219–231, 2010.

[27] J. T. Shi, "Global maximum principle for the forward-backward stochastic optimal control problem with Poisson jumps," *Early View in Asian Journal of Control*, vol. 16, no. 1, pp. 1–11, 2014.

[28] J. T. Shi and Z. Wu, "Maximum principle for fully coupled forward-backward stochastic control system with random jumps," in *Proceedings of the 26th Chinese Control Conference*, pp. 375–380, Zhangjiajie, China, 2007.

[29] Q. X. Meng and Y. Z. Sun, "Optimal control problem of fully coupled forward-backward stochastic systems with Poisson jumps under partial information," in *Proceedings of the 30th Chinese Control Conference (CCC '11)*, pp. 1317–1322, Yantai, China, 2011.

[30] M. Kohlmann and X. Y. Zhou, "Relationship between backward stochastic differential equations and stochastic controls: a linear-quadratic approach," *SIAM Journal on Control and Optimization*, vol. 38, no. 5, pp. 1392–1407, 2000.

[31] J. Ma and J. M. Yong, *Forward-Backward Stochastic Differential Equations and Their Applications*, vol. 1702 of *Lecture Notes in Mathematics*, Springer, Berlin, Germany, 1999.

[32] A. Bensoussan, "Lectures on stochastic control," in *Nonlinear Filtering and Stochastic Control*, vol. 972 of *Lecture Notes in Mathematics*, pp. 1–62, Springer, Berlin, Germany, 1982.

[33] N. Dokuchaev and X. Y. Zhou, "Stochastic controls with terminal contingent conditions," *Journal of Mathematical Analysis and Applications*, vol. 238, no. 1, pp. 143–165, 1999.

[34] G. C. 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.

[35] 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.

[36] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, vol. 24, North-Holland, Amsterdam, The Netherlands, 2nd edition, 1989.

[37] I. Ekeland, "Nonconvex minimization problems," *Bulletin of The American Mathematical Society*, vol. 1, no. 3, pp. 443–474, 1979.

[38] Z. Wu, "Fully coupled FBSDE with Brownian motion and Poisson process in stopping time duration," *Journal of the Australian Mathematical Society*, vol. 74, no. 2, pp. 249–266, 2003.

[39] J. T. Shi and Z. Wu, "The LQ problem for fully coupled forward-backward stochastic control system," *Journal of Shandong University*, vol. 40, no. 1, pp. 11–17, 2005.