Research Article

Maximum Principle for Mean-Field FBSDEs with Mixed Initial-Terminal Conditions

Ruijing Li

School of Statistics and Mathematics, Guangdong University of Finance and Economics, Guangzhou 510320, Guangdong, China

Correspondence should be addressed to Ruijing Li; li209981@sina.com

Received 9 February 2022; Accepted 18 April 2022; Published 10 May 2022

Academic Editor: Binchang Wang

Copyright © 2022 Ruijing Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The present paper concerns with optimal control problems allowing for time inconsistent utility functions of mean-field FBSDEs with mixed initial-terminal conditions. Moreover, the control variable enters the diffusion coefficient, and the control domain is with nonconvexity. Via extended Ekeland’s variational principle as well as the reduction method, a general stochastic maximum principle is established in the framework of mean-field theory. Finally, a linear-quadratic example is worked out to illustrate the application of the results.

1. Introduction

It is well known that the key tools in solving optimal control problems are based upon Pontryagin’s maximum principle and the Bellman dynamic programming principle. Recently, the mean-field approaches have been widely applied in various fields, such as in Finance, Statistical Mechanics, and Game theory. The systems these fields involve are mainly of mean-field type, where the performance functionals, drifts, and diffusion coefficients depend not only on the state but also on the probability distribution of the state. In this case, the mean-variance utility functions are of time inconstistency, for which the Bellman dynamic programming principle does not hold. Therefore, problems involving such utilities cannot be approached by the classical Hamilton–Jacobi–Bellman equation. In this paper, we study a mean-variance setting of the Principle-Agent problem by means of Pontryagin’s maximum principle.

The pioneering paper in which nonlinear mean-field backward stochastic differential equations (BSDEs) first appeared is in [1]. Taking advantage of a BSDE approach, the paper [2] considered the mean-field BSDE and gave a probabilistic interpretation to related nonlocal partial differential equations. Since then, the theory of mean-field forward-backward stochastic differential equations (FBSDEs) has been intensively investigated, and rich results have been obtained for the system of this kind of McKean–Vlasov type (see [3]) in the literature. For instance, by using the penalization method to construct approximation solution, the authors in [4] proved the existence of solutions to first-order mean-field games arising in optimal switching. In [5], by translating the splitting method to BSDEs, the author proved the existence and the uniqueness of the solution of the split equations; moreover, under appropriate regularity assumptions on the coefficients, the value function turned out to be the unique classical solution of the related nonlocal quasi-linear integral-partial differential equation of mean-field type. On the other hand, with the development of mean-field FBSDEs, research on stochastic control problems in mean-field context got rapid growth. For instance, by utilizing a spike variation of the optimal control, the authors in [6] established Peng’s-type stochastic maximum principle. When the controlled system contains the Markov jump parameters, the authors in [7] established the corresponding maximum principle for mean-field stochastic linear-quadratic optimal control problems. Furthermore, with the help of maximum principle, the authors in [8] presented necessary as well as sufficient solvability condition of the finite horizon mean-field optimal control problem in an explicit expression form.
and the author in [9] established necessary and sufficient optimality conditions for the mean-field jump-diffusion system with moving-average as well as pointwise delays. Further, by using a backward separation method with a dominated growth rate as well as an approximate technique, the authors in [10] presented a maximum principle for optimality of mean-field stochastic differential equations, where the state is partially observed via a noisy process. For some related topics, we refer the readers to [4, 11–13], [14], [7, 15, 16], [17], and the references therein.

However, to the best of our knowledge, when the control domain is not assumed convexity, for the stochastic optimal control problems with mixed initial-terminal constraints, their corresponding stochastic maximum principle is all established on the basis of requirements of lower semi-continuity of the penalty functional. So, how to reduce this requirement is a new problem. This paper is trying to remedy this issue and establish stochastic maximum principle for mean-field FBSDEs under general control domains. With the help of extended Ekeland’s variational principle and the reduction method, a general stochastic maximum principle will be acquired. It is necessary to point out that, due to the application of extended Ekeland’s variational principle, the reduction method we adopt is different from those given in [18]; meantime, our results extend those of [19] essentially to the framework of mean-field theory.

The rest of this paper is organized as follows. Preliminaries are in Section 2. In Section 3, we give some prior estimates and establish the general stochastic maximum principle in the framework of mean-field theory. In Section 4, a linear-quadratic stochastic control problem is worked out to illustrate the theoretical results. Finally, some concluding remarks are given in Section 5.

2. Preliminaries

Let \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)\) be a filtered probability space satisfying the usual condition, on which a one-dimensional Brownian motion \((B_t)_{t \geq 0}\) is defined, and \(\mathcal{F} = \{\mathcal{F}_s, \ 0 \leq s \leq T\}\) be the natural filtration generated by \((B_t)_{t \geq 0}\) and augmented by all \(P\)-null sets, i.e.,

\[
\mathcal{F}_s = \sigma[B_r, \ r \leq s] \vee \mathcal{N}_p, \ s \in [0, T],
\]

where \(\mathcal{N}_p\) is the set of all \(P\)-null subsets. We denote by \(\delta_{\mathcal{F}}^2(0, T; R)\) the set of \(\mathcal{F}\)-adapted and continuous stochastic processes \(\psi\) such that \(E[\sup_{t \in [0, T]}|\psi|^2] < \infty\), by \(\mathcal{F}^2(0, T; R)\) the set of \(\mathcal{F}\)-adapted stochastic processes \(\psi\) such that \(E[\int_0^T |\psi|^2 \, dt] < \infty\), and \(M^2[0, T] = \mathcal{F}^2(0, T; R) \times \delta_{\mathcal{F}}^2(0, T; R)\).

Let \((\bar{\Omega}, \bar{\mathcal{F}}, \bar{P}) = (\Omega \times \Omega, \bar{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}, \bar{P} = P \otimes P)\) be the (non-completed) product of \((\Omega, \mathcal{F}, P)\) with itself, and this product space is endowed with the filtration \(\bar{\mathcal{F}} = \{\mathcal{F}_t = \mathcal{F} \otimes \mathcal{F}_t, \ 0 \leq t \leq T\}\). Then, any random variable \(\chi \in L^2(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P})\) originally defined on \(\Omega\) can be extended canonically to \(\bar{\Omega}: \chi'(w', w) = \chi(w'), \ (w', w) \in \bar{\Omega} = \Omega \times \Omega\). For any \(\rho \in L^1(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P})\), the variable \(\rho(\cdot, w): \Omega \longrightarrow R\) belongs to \(L^1(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P}), P(\bar{P})\)-a.s., and its expectation is denoted by

\[
E' [\rho(\cdot, w)] = \int_{\bar{\Omega}} \rho(w', w)P(\bar{P}dw').
\]

Notice that \(E'[\rho] = E'[\rho(\cdot, w)]\), and \(\bar{E}[\rho] = \int_{\bar{\Omega}} \rho(\cdot, w)P(\bar{P}dw) = E[E'[\rho]]\).

Consider a fully coupled mean-field stochastic differential equation with mixed initial-terminal conditions as follows:

\[
\begin{aligned}
\text{d}X_t &= E' \left[ b(t, X_t, Y_t, Z_t, X_0, Y_0, Z_0) \right] \text{d}t + E' \left[ \sigma(t, X_t, Y_t, Z_t, X_0, Y_0, Z_0) \right] \text{d}B_t, \\
-\text{d}Y_t &= E' \left[ f(t, X_t, Y_t, Z_t, X_0, Y_0, Z_0) \right] \text{d}t - Z_t \text{d}B_t, \\
X_0 &= \gamma(X_T, Y_T), \quad Y_T = h(X_T, Y_T),
\end{aligned}
\]

where \(b, \sigma, f: \ [0, T] \times \mathbb{R}^6 \times U \longrightarrow \mathbb{R} \); \(\gamma, h: \mathbb{R}^2 \longrightarrow \mathbb{R} \); \(U\) is a nonempty closed and nonconvex subset of \(\mathbb{R}\).

\begin{remark}
The drift terms \(b\) and \(f\) and the diffusion term \(\sigma\) of (3) have to be interpreted as follows:
\end{remark}
The cost functional associated with (3) is of the form as follows:

$$J(u) = E\left\{ \int_0^T E'[f(t, X'_t, Y'_t, Z'_t, X_t, Y_t, Z_t, u_t)] \, dt \right\} + E[X(T, Y_0)],$$

where $I: [0, T] \times \mathbb{R}^6 \times U \to \mathbb{R}$; $\varphi: \mathbb{R}^2 \to \mathbb{R}$. The admissible control set is defined as

$$\mathcal{U} = \{ u_t \in L^2_T(0, T, U) | u_t(w, t) : [0, T] \times \Omega \to U \}.$$  \hspace{1cm} (6)

Then, the optimal control problem we are concerned with is as follows.

**Problem 1.** For the system in (3), we seek a control $u \in \mathcal{U}$ such that

$$J(u) = \inf_{u_0 \in \mathcal{U}} J(u).$$

If $u$ attains the infimum, then it is called optimal; the solution $(X, Y, Z)$ corresponding to $u$ is called an optimal trajectory. In what follows, some notations and assumptions shall be assumed to ensure the well-posedness of stochastic system (3). We denote the scalar product by $\langle \cdot, \cdot \rangle$ and the norm by $|\cdot|$ of an Euclidean space. For $\Lambda = (x, y, z, \bar{x}, \bar{y})$, $\Xi = (X_T, Y_0)$ and define $F(t, \Lambda, u) = (-f(t, \Lambda, u), b(t, \Lambda, u), \sigma(t, \Lambda, u))$, $Y(\Xi) = (y(\Xi), h(\Xi))$.

(H1) The functions $b, \sigma, f, I, \varphi, h, \gamma$ are twice continuously differentiable with respect to $(x, y, \bar{x}, \bar{y})$. Moreover, all their derivatives with respect to $(x, y, \bar{x}, \bar{y})$ up to second-order are continuous in $(x, y, \bar{x}, \bar{y}, u)$ and bounded.

(H2) The functions $b, \sigma, f, I, \varphi, h, \gamma$ grow linearly with respect to $(\bar{x}, \bar{y}, x, y, z, u)$ and are continuous in $(u, t)$, respectively.

(H3) For any $\Lambda$, $u \in \mathcal{U}$, $F(t, \Lambda, u) \in H^2_\mathbb{R} (0, T; \mathbb{R}^3)$, there exists a constant $C > 0$ such that

$$|F(t, \Lambda_1, u) - F(t, \Lambda_2, u)| \leq C|\Lambda_1 - \Lambda_2|,$$

where $\beta_1 > 0$. \hspace{1cm} (8)

Remark 2. Under the assumptions (H1) - (H4) or (H1) - (H4)', via Theorem 1 in [20], the mean-field system (3) admits a unique adapted solution $(X_t, Y_t, Z_t)_{t \in [0, T]} \in \delta^2_\mathbb{R} (0, T; R) \times \delta^2_\mathbb{R} (0, T; R) \times \mathcal{H}_\mathbb{R}^2 (0, T; R)$.

In fact, due to the mixed initial-terminal conditions in the state equation, even if the well-posedness of the state equation is ensured via the Lyapunov operator introduced in [21], the well-posedness of the first-order adjoint equation seems to be not guaranteed. On the other hand, since both $Z$ and $u$ appear in the coefficients of the backward equation (3), the regularity of process $Z$ (as a part of the state process) is not enough to obtain a second-order adjoint equation. Thus,
the classical method cannot be directly used. To overcome this difficulty, we introduce a reduction method inspired by the study of optimality variational principle for controlled FBSDEs [18]. First, a variation of problem \( \mathcal{A} \) is formulated as follows.

**Problem 2.** \( \mathcal{B} \) Minimize

\[
\begin{aligned}
\tilde{J}(X_0, Y_0, Z, u) &= E \left[ \int_0^T E' \left[ l(t, X'_t, Y'_t, Z'_t, X_t, Y_t, Z_t, u_t) \right] dt \right] \\
&+ E \left[ \varphi(X_T, Y_0) \right],
\end{aligned}
\]

subject to

\[
\begin{aligned}
dX_t &= E' \left[ b(t, X'_t, Y'_t, Z'_t, X_t, Y_t, Z_t, u_t) \right] dt + E' \left[ \sigma(t, X'_t, Y'_t, Z'_t, X_t, Y_t, Z_t, u_t) \right] dB_t, \\
-dY_t &= E' \left[ f(t, X'_t, Y'_t, Z'_t, X_t, Y_t, Z_t, u_t) \right] dt - Z_t dB_t, \\
X(0) &= X_0, \quad Y(0) = Y_0.
\end{aligned}
\]

Over \( (X_0, Y_0, Z, u) \in \mathcal{B} := \mathbb{R} \times \mathbb{R} \times \mathcal{H}_2^1 (0, T; \mathbb{R}) \times \mathcal{U} \) with the mixed initial-terminal state constraints,

\[
X_0 = \gamma(X_T, Y_0), \quad Y_T = h(X_T, Y_0).
\]

**Remark 3.** Via Theorem 2 in [22], under the assumptions \((H_1) - (H_2)\), the mean-field forward stochastic system (10) admits a unique solution \( (X_t, Y_t, Z_t)_{t \in [0, T]} \in \delta_{\mathcal{H}_2} (0, T; \mathcal{R}) \times \delta_{\mathcal{H}_2} (0, T; \mathcal{R}) \times \mathcal{H}_2^1 (0, T; \mathcal{R}). \)

It is remarkable that problem \( \mathcal{A} \) is embedded into problem \( \mathcal{B} \). Hence, if \( (\overline{X}_0, \overline{Y}_0, \overline{Z}, \overline{u}) \) is the optimal control of \( \mathcal{B} \), then \( \overline{u} \) is optimal for \( \mathcal{A} \). In the following section, the classical second-order variational technique will be adopted to solve problem \( \mathcal{B} \).

Before presenting the main results, we introduce the extended Ekeland’s variational principle [23]. Let \( K \) be a nonempty closed subset of a complete metric space \( (S, d) \) and \( G: K \times K \rightarrow \mathbb{R} \) be a bifunction. Assume that \( \varepsilon > 0 \) and the following assumptions are satisfied:

\[
J^\varepsilon(X_0, Y_0, Z, u) = \left\{ \left[ J(X_0, Y_0, Z, u) + \varepsilon \right]^2 + E \left[ \left| X_0 - \gamma(X_T, Y_0) \right|^2 + \left| Y_T - h(X_T, X_0) \right|^2 \right] \right\}^{1/2}.
\]

Endow the set \( \mathcal{U} \) and \( \mathcal{H}_2^1 (0, T; \mathcal{R}) \) with the metric \( d \),

\[
d(v, \overline{v}) = E \left[ \left| \int_0^T |v_t - \overline{v}_t|^2 \right| \right], \quad \forall v, \overline{v} \in \mathcal{U} \text{ or } \mathcal{H}_2^1 (0, T; \mathcal{R}).
\]

Clearly, (13) subject to (10) is a forward stochastic control problem without the state constraints. However, we have to notice that the unboundedness of \( \mathcal{R} \) cannot assure lower semicontinuity of \( J^\varepsilon(X_0, Y_0, Z, u) \) in \( (\mathcal{R}, \overline{d}) \). So,
Ekeland’s variational principle cannot be directly used. Fortunately, extended Ekeland’s variational principle can reduce the requirement of lower semicontinuity of \( J^\delta (X_0, Y_0, Z, u) \). Hence, Theorem 1 can be adopted to discuss this issue. First, we consider the case that \( X_0, Y_0, \) and \( Z \) take values in \( \mathcal{M} \subset \mathbb{R} \), \( \mathbb{M} \subset \mathcal{M} \), and \( \mathcal{N} \subset \mathbb{R} \). Moreover, \( \mathcal{M} \) is convex. Then, there exists \( (X_0^\delta, Y_0^\delta, Z^\delta, u^\delta) \in \mathcal{R}_1 \) such that \( F(\cdot, \cdot) = J^\delta(\cdot) - J^\delta(\cdot) \) with \( \cdot = (X_0, Y_0, Z, u) \in \mathcal{R}_1 \) and \( \tau = (X_0, Y_0, Z, u) \in \mathcal{R}_1 \). Then, \( F(\cdot, \cdot) \) satisfies the conditions of Theorem 1, the demonstration process is similar as the appendix in [25], and thus the detail is omitted. Hence, \( (X_0^\delta, Y_0^\delta, Z^\delta, u^\delta) \in \mathcal{R}_1 \) such that \( \delta \) further, we can get
\[
\begin{aligned}
J^\delta(X_0, Y_0, Z, u) &\leq \delta,
\delta \leq \bar{d}(X_0^\delta, Y_0^\delta, Z^\delta, u^\delta), (X_0, Y_0, Z, u) \in \mathcal{R}_1,
\end{aligned}
\]

Then, (17) means that the control process \( (X_0^\delta, Y_0^\delta, Z^\delta, u^\delta) \) is a global minimum point of the following penalized cost functional:
\[
J^\delta(X_0, Y_0, Z, u) + \sqrt{\delta} \mathcal{D}(X_0^\delta, Y_0^\delta, Z^\delta, u^\delta), (X_0, Y_0, Z, u).
\]

Since \( Z \) and \( u \) are dynamic, spike technique for \( Z^\delta \) and \( u^\delta \) is employed. For sufficiently small \( \varepsilon > 0 \), define
\[
\mu^\varepsilon = \begin{cases} 
\mu^\varepsilon, & t \in [0, T] \setminus S_e, \\
\mu_t, & t \in S_e.
\end{cases}
\]

These notations also work for their corresponding derivatives.

Suppose \( X_0^\varepsilon, \) and \( X^\delta \) are solutions of (10) corresponding to \( (X_0^\varepsilon, \delta) \) and \( (X_0^\delta, \varepsilon) \), respectively. We introduce the first-order and second-order variational equations.

\[
\begin{aligned}
X = \left( X \right), \quad v = \left( \begin{array}{c}
Z' \\
u
\end{array} \right), \quad x_0 = \left( \begin{array}{c}
X_0 \\
Y_0
\end{array} \right), \quad x_T = \left( \begin{array}{c}
X_T \\
Y_T
\end{array} \right),
B(t, X', X, v) = \left( \begin{array}{c}
b_0, (X_0)' , (Y_0)', (Z_0)', X_0, Y_0, Z_0, u_t
\end{array} \right),
\Sigma(t, X', X, v) = \left( \begin{array}{c}
\sigma(t, (X_0)', (Y_0)', (Z_0)', X_0, Y_0, Z_0, u_t)
\end{array} \right),
\Gamma(X_0, X_T) = \left( \begin{array}{c}
X_0 - y(X_T, Y_0) \\
Y_T - h((X_T, Y_0))
\end{array} \right), \quad F(\cdot, \cdot) = \varphi(X_T, Y_0).
\end{aligned}
\]
where $I_{S_t}$ denotes the indicator function of a set $S_t$. It is easy to verify that (21) and (22) admit unique solutions, and these solutions satisfy the following prior estimates.

\[ dX_t = E' \left[ B_{\chi} \left( X_{t-}, X_T \right) \right] + B_X X_t + dB_t, \]

\[ dY_t = \left[ \sum_{i=1}^{n} \left( \frac{\left( Y_t \right)^i}{\gamma} \right) \right] dX_t, \]

\[ d\tilde{Z}_t = \left[ \sum_{i=1}^{n} \left( \frac{\left( \tilde{Z}_t \right)^i}{\gamma} \right) \right] dX_t, \]

\[ d\tilde{B}_t = \left[ \sum_{i=1}^{n} \left( \frac{\left( \tilde{B}_t \right)^i}{\gamma} \right) \right] dX_t, \]

\[ \xi_T = \theta \varphi X_T - \tilde{\theta} Y X_T - \tilde{\theta} h X_T, \quad \eta_T = \tilde{\theta} T, \]

such that

\[ E[\sup_{t \in [0, T]} |X_t - X_T|^{2k}] + E[\sup_{t \in [0, T]} |\xi_t - \xi_T|^{2k}] \leq C e^k, \]

\[ E[\sup_{t \in [0, T]} |\tilde{X}_t - \tilde{X}_T|^{2k}] + E[\sup_{t \in [0, T]} |\tilde{\xi}_t - \tilde{\xi}_T|^{2k}] \leq C e^k, \]

\[ E[\sup_{t \in [0, T]} |\tilde{Y}_t - \tilde{Y}_T|^{2k}] + E[\sup_{t \in [0, T]} |\tilde{\eta}_t - \tilde{\eta}_T|^{2k}] \leq C e^k, \]
Proof 2. This state constrained stochastic maximum principle will be proved into two steps. □

\begin{align}
P_0 \geq \begin{pmatrix} 0, 0, & 0 \\ 0, & \theta_0 y + \theta_0 v - \theta_0 \varphi y - \theta_0 \varphi v \\ & \theta_0 \end{pmatrix}, \quad \text{and} \\
E[H(t, x, \bar{x}, v, \Phi, \Psi, \theta)] - H(t, x, \bar{x}, v, \Phi, \Psi, \theta)] + \frac{1}{2} (\Delta \Sigma)^T \Delta \Sigma \geq 0. \quad \forall X_0 \in \mathcal{M}, Y_0 \in \mathcal{M}, Z \in \mathcal{N}, u \in U, a.e., a.s. \end{align}

where $H$ is the Hamiltonian function defined by

\begin{align}
H(t, x, \bar{x}, v, \Phi, \Psi, \theta) = \langle \Phi, B(t, x, \bar{x}, v) \rangle + \langle \Psi, \Sigma(t, x, \bar{x}, v) \rangle + \langle \theta, l(t, x, \bar{x}, v) \rangle.
\end{align}

And $(P, Q) \in \mathcal{S}_T^2 (0, T; \mathbb{R}^{2 \times 2}) \times \mathcal{F}_T^2 (0, T; \mathbb{R}^{2 \times 2})$ is the unique adapted solution of the following BSDE:

\begin{align}
dP_t = -E \left[ B^T \Sigma \Sigma B + B^T \Sigma \Sigma + \Sigma^T \Sigma \right] dt + Q dB_t.
\end{align}

Proof 2. This state constrained stochastic maximum principle will be proved into two steps. □

Step 1. Let $X_0, Y_0, Z$ take values in $\mathcal{M} \subset R, \mathcal{M} \subset R$, and $\mathcal{N} \subset R$, respectively, and $\mathcal{M}$ be convex. Moreover, $\mathcal{M}$ and $\mathcal{N}$ are all closed.

From (17) and the definition of $\beta$, we have

\begin{align}
\delta f \left( x_0, y_0, z_0, u_0 \right) - \delta f \left( x_0, y_0, z_0, u_0 \right) + \sqrt{\delta f} \left( x_0, y_0, z_0, u_0 \right) = \delta f \left( x_0, y_0, z_0, u_0 \right) - \delta f \left( x_0, y_0, z_0, u_0 \right) + \sqrt{\delta f} \left( x_0, y_0, z_0, u_0 \right) \geq 0.
\end{align}

By using Taylor expansion with addition of Lemma 1, we deduce
\[
\frac{j_0^\delta(x_0^\delta, v^\delta) - j_0^\delta(x_0^\delta, v^\delta)}{\int j_0^\delta(x_0^\delta, v^\delta) + j_1^\delta(x_0^\delta, v^\delta)} = \theta_0^\delta \left( J(x_0^\delta, v^\delta) - J(x_0^\delta, v^\delta) \right) + \mathcal{E} \left( \vartheta_0^\delta, \Gamma(x_0^\delta, x_1^\delta) - \Gamma(x_0^\delta, x_1^\delta) \right)
\]

\[
\cdot \left[ \theta_0^\delta + o(1) \left[ J(x_0^\delta, v^\delta) - J(x_0^\delta, v^\delta) \right] + \mathcal{E} \left( \vartheta_0^\delta + o(1), \Gamma(x_0^\delta, x_1^\delta) - \Gamma(x_0^\delta, x_1^\delta) \right) \right].
\]

where \( o(1) \) stands for certain vectors that tend to 0 as \( \varepsilon \rightarrow 0 \) and \( \cdot \) stands for inner product of vectors. Here,
\[
\vartheta_0^\delta = (\vartheta_0^\delta, \vartheta_0^\delta, \vartheta_0^\delta), \quad \vartheta_0^\delta = (\vartheta_0^\delta, \vartheta_0^\delta), \quad \text{and}
\]

\[
\begin{align*}
\vartheta_0^\delta &= \frac{2 \int_0^1 \lambda J \left( \frac{x_0^\delta, v^\delta}{x_0^\delta, v^\delta} + (1 - \lambda) J(x_0^\delta, v^\delta) + \delta \right) d\lambda}{\int j_0^\delta(x_0^\delta, v^\delta) + j_1^\delta(x_0^\delta, v^\delta)} , \\
\vartheta_0^\delta &= \frac{x_0^\delta - \gamma(x_1^\delta, v^\delta)}{j_0^\delta(x_0^\delta, v^\delta) + j_1^\delta(x_0^\delta, v^\delta)} , \\
\vartheta_0^\delta &= \frac{y^\delta - h(x_1^\delta, v^\delta)}{j_0^\delta(x_0^\delta, v^\delta) + j_1^\delta(x_0^\delta, v^\delta)}.
\end{align*}
\]

Then, \( \theta_0^\delta \geq 0 \) and \( |\theta_0^\delta|^2 + E|\vartheta_0^\delta|^2 + E|\vartheta_0^\delta|^2 = 1 \). Thus, there exists a subsequence still denoted by \((\theta_0^\delta, \vartheta_0^\delta, \vartheta_0^\delta)\) convergent, i.e.,

\[
\lim_{\delta \to 0} \left( \theta_0^\delta, \vartheta_0^\delta, \vartheta_0^\delta \right) = \left( \theta_0^\delta, \vartheta_0^\delta, \vartheta_0^\delta \right).
\]

We claim that \( \theta_0^\delta \neq 0 \). The detailed illustration of this point refers to [19]. Here, \( (\theta_0^\delta, \vartheta_0^\delta, \vartheta_0^\delta) \) is called the Lagrange multiplier of the corresponding optimal 4-tuple \((X_0, Y_0, Z, \bar{Z})\). While, by Taylor expansion,
\[ E \int_0^T E^t \left\{ \left( f^i_{X_t} \left( X^\delta_{l,t} + X^\delta_{z,t} \right) \right) + \left( I_{X_t}^0 X^\delta_{l,t} + I_{X_t}^0 X^\delta_{z,t} \right) \right\} \, dt + E \left\{ \left( F^0_{X_t} X^\delta_{l,t} - X^\delta_{0,t} \right) + \left( F^0_{X_t} X^\delta_{z,t} - X^\delta_{0,t} \right) \right\} \] 

\[ + \frac{1}{2} \left[ \left( I_{X_t}^0 X^\delta_{l,t} - X^\delta_{0,t} \right), X^\delta_{l,t} - X^\delta_{0,t} \right] + \left( I_{X_t}^0 X^\delta_{z,t} - X^\delta_{0,t} \right), X^\delta_{z,t} - X^\delta_{0,t} \right) \] 

\[ + 2 \left( F^0_{X_t} X^\delta_{l,t} - X^\delta_{0,t}, X^\delta_{l,t} - X^\delta_{0,t} \right) \right] + o(\varepsilon) \right\} \] 

where

\[
\begin{align*}
\mathbf{l}^\varepsilon_i &= I_i(t, X^\varepsilon, \left( \Gamma_x^\varepsilon \right), v^\varepsilon_i) \\
\mathbf{v}^\varepsilon_i &= I_{ij}(t, X^\varepsilon, \left( \Gamma_x^\varepsilon \right), v^\varepsilon_i) \\
F^\delta_j &= F_j(X^\delta_0, X^\delta_T) \\
F^\delta_{ij} &= F_{ij}(X^\delta_0, X^\delta_T), i, j = X, X_\varepsilon.
\end{align*}
\]

Similarly,

\[ E \langle \vartheta^\delta, \Gamma(X^\delta_0, X^\delta_T) - \Gamma(X^\delta_0, X^\delta_T) \rangle = E \left\{ \left( I_{X_0}^0 \vartheta^\delta, X_0^\delta - X_0^\delta \right) + \left( I_{X_0}^0 \vartheta^\delta, X_0^\delta - X_0^\delta \right) \right\} \] 

\[ + \frac{1}{2} \left[ \left( I_{X_0}^0 \vartheta^\delta (X^\delta_0 - X^\delta_0), X^\delta_0 - X^\delta_0 \right) + \left( I_{X_0}^0 \vartheta^\delta (X^\delta_0 - X^\delta_0), X^\delta_0 - X^\delta_0 \right) \right] + \frac{1}{2} \left[ \left( I_{X_0}^0 \vartheta^\delta (X^\delta_0 - X^\delta_0), X^\delta_0 - X^\delta_0 \right) + \left( I_{X_0}^0 \vartheta^\delta (X^\delta_0 - X^\delta_0), X^\delta_0 - X^\delta_0 \right) \right] \] 

\[ + 2 \left( F^0_{X_0} X^\delta_0 - X^\delta_0, X^\delta_0 - X^\delta_0 \right) \right] + o(\varepsilon) \right\} \] 

where \( \Gamma^\delta_j = \Gamma_j(X^\delta_0, X^\delta_T) \) and \( \Gamma^\delta_{ij} = \Gamma_{ij}(X^\delta_0, X^\delta_T) \), \( i, j = X, X_\varepsilon, X_T \).

Consequently, from (29) – (35), the variational inequality is derived as
\[-\varepsilon \sqrt{\delta} \sqrt{2t^2 + |X_0|^2 + |Y_0|^2} \leq \Theta^\varepsilon_0 E \left[ J\left( X^\varepsilon_0, Y^\varepsilon_0 \right) - J\left( X^d_0, Y^d_0 \right) \right] \]

\[+ E\langle \tilde{\theta}^\varepsilon_0, \Gamma\left( X^\varepsilon_0, X^\varepsilon_T \right) - \Gamma\left( X^d_0, X^d_T \right) \rangle \]

\[= \Theta^\varepsilon_0 E \int_0^T E' \left\{ \left( \rho^\delta_0 \left( X^\varepsilon_{1,T} + X^\varepsilon_{2,T} \right) \right) + \left( \rho^\delta_0 \left( X^\varepsilon_{1,T} + X^\varepsilon_{2,T} \right) \right) + \left( \rho^\delta_0 \left( X^\varepsilon_{1,T} + X^\varepsilon_{2,T} \right) \right) \right\} dt \]

\[+ \frac{1}{2} \left[ \left( \rho^\delta_0 \left( X^\varepsilon_{1,T} \right) \right) + \left( \rho^\delta_0 \left( X^\varepsilon_{1,T} \right) \right) + \left( \rho^\delta_0 \left( X^\varepsilon_{1,T} \right) \right) + \left( \rho^\delta_0 \left( X^\varepsilon_{1,T} \right) \right) \right] \]

To get rid of $X^\delta_0$ in the above inequality, we introduce the first-order adjoint equations as

\begin{align*}
\frac{d}{dt} \Phi^\delta_t &= -E' \left[ \Phi^\delta_t B^\delta + \langle \Phi^\delta_t, R^\delta \rangle B^\delta + \Psi^\delta_t \Sigma^\delta \right] dt + \psi^\delta_t dB_t, \\
\Phi^\delta_T &= \Theta^\varepsilon_0 F^\delta_{X_t} + \left[ \Gamma^\delta_{X_t} \tilde{\theta}^\varepsilon_0 \right].
\end{align*}

Applying Itô’s formula to $\langle \Phi^\delta_t, X^\varepsilon_{1,T} + X^\varepsilon_{2,T} \rangle$ fulfills

\begin{align*}
E\left( \Phi^\delta_0 F^\delta_{X_t} + \left[ \Gamma^\delta_{X_t} \tilde{\theta}^\varepsilon_0 \right], X^\varepsilon_{1,T} + X^\varepsilon_{2,T} \right) &= E\left( \Phi^\delta_0, X^\varepsilon_T \right) + \int_0^T E' \left[ \langle \Phi^\delta_t, \left( \Delta B^\delta_{X_t} \right) + \Delta B^\delta_{X_t} \rangle \right] I_{\lambda_j}, \\
&+ \frac{1}{2} B^\delta_{X_t} \left( \langle X^\varepsilon_{1,T} \rangle \right) \right] + \frac{1}{2} B^\delta_{X_t} \left( \langle X^\varepsilon_{1,T} \rangle \right) \right] \]

+ \psi^\delta_t \left( \Delta \Sigma^\delta_{X_t} \left( \langle X^\varepsilon_{1,T} \rangle \right) \right) + \Delta \Sigma^\delta_{X_t} \left( \langle X^\varepsilon_{1,T} \rangle \right) + \Delta \Sigma \right\} I_{\lambda_j}, \\
&+ \frac{1}{2} \Sigma^\delta_{X_t} \left( \langle X^\varepsilon_{1,T} \rangle \right) \right] + \frac{1}{2} \Sigma^\delta_{X_t} \left( \langle X^\varepsilon_{1,T} \rangle \right) \right] \right\} dt.
\end{align*}

Combining (36) and (38), again with the help of Lemma 1, the variational inequality (36) can be changed as
with $H_{ii}^{\delta x}$ being the Hessian of

$$H(t, x, \bar{x}, v, \Phi, \Psi, \theta) = \langle \Phi, B(t, x, \bar{x}, v) \rangle + \langle \Psi, \Sigma(t, x, \bar{x}, v) \rangle + \langle \theta, I(t, x, \bar{x}, v) \rangle.$$  

We see that in (39), there left some second-order terms in $x_i^{\delta x}$. Hence, we want to further get rid of them. Firstly, the equations satisfied by $\Theta^{\delta x} = x_i^{\delta x} (x_i^{\delta x})^T$ are presented by appealing to Itô’s formula.

$$
\begin{align*}
\frac{d}{dt} \Theta^{\delta x}_t &= E\left[ B^\delta \Theta^{\delta x} + \Theta^{\delta x} (B^\delta)^T + B^\delta \left( \Theta^{\delta x} \right)^T + \left( \Theta^{\delta x} \right)^T \left( B^\delta \right)^T \right] \\
&+ \Sigma^\delta \Theta^{\delta x} (\Sigma^\delta)^T + \Sigma^\delta \left( \Theta^{\delta x} \right)^T (\Sigma^\delta)^T + \Lambda^{\delta x}_t \\
&+ \left[ \Delta \Sigma^\delta (\Delta \Sigma^\delta)^T + \Lambda^{\delta x}_t \right] I_{S_{\delta t}} \\
&+ E\left[ \Sigma^\delta_0 \Theta^{\delta x} + \Theta^{\delta x} (\Sigma^\delta_0)^T + \Sigma^\delta_0 \left( \Theta^{\delta x} \right)^T + \left( \Theta^{\delta x} \right)^T (\Sigma^\delta_0)^T + \Lambda^{\delta x}_0 \right] dB_t,
\end{align*}
$$

where

$$
\begin{align*}
\Lambda^{\delta x}_{1t} &= \Sigma^\delta \Sigma^\delta \left[ (x_i^{\delta x})^T \right] (\Sigma^\delta_x)^T + \Sigma^\delta \left( x_i^{\delta x} \right)^T (x_i^{\delta x})^T, \\
\Lambda^{\delta x}_{2t} &= \Sigma^\delta \Sigma^\delta (\Delta \Sigma^\delta)^T + \Delta \Sigma^\delta \left( x_i^{\delta x} \right)^T (\Delta \Sigma^\delta)^T + x_i^{\delta x} (\Delta B^\delta)^T, \\
\Lambda^{\delta x}_{3t} &= \Sigma^\delta \Sigma^\delta \left( x_i^{\delta x} \right)^T (\Delta \Sigma^\delta)^T + \Delta \Sigma^\delta \left( x_i^{\delta x} \right)^T (\Delta \Sigma^\delta)^T + x_i^{\delta x} (\Delta B^\delta)^T, \\
\Lambda^{\delta x}_{4t} &= \Sigma^\delta \Sigma^\delta \left( x_i^{\delta x} \right)^T (\Delta \Sigma^\delta)^T + \Delta \Sigma^\delta \left( x_i^{\delta x} \right)^T (\Delta \Sigma^\delta)^T + x_i^{\delta x} (\Delta B^\delta)^T.
\end{align*}
$$

Secondly, we introduce the second-order adjoint equations corresponding to $u^{\delta x}$, respectively, to remove $x_i^{\delta x} (x_i^{\delta x})^T$ in (39). Now, let $(P^{\delta x}, Q^{\delta x})$ be the adapted solution of the following BSDE:
Inserting (44) into (39), we can derive

\[
\begin{align*}
\left\{ \begin{array}{l}
dP_{t}^{0,\delta} = -E'\left[ (B_{X}^{0})^{T} P_{t}^{0,\delta} + P_{t}^{0,\delta} B_{X}^{0} + (\Sigma_{X}^{0})^{T} (P_{t}^{0,\delta})' + (P_{t}^{0,\delta})' B_{X}^{0} \\
+ (\Sigma_{X}^{0})^{T} (P_{t}^{0,\delta}) + (\Sigma_{X}^{0})^{T} (P_{t}^{0,\delta})' + (\Sigma_{X}^{0})^{T} Q_{t}^{0,\delta} + Q_{t}^{0,\delta} (\Sigma_{X}^{0})
+ (\Sigma_{X}^{0})^{T} (Q_{t}^{0,\delta})' + (Q_{t}^{0,\delta})' (\Sigma_{X}^{0}) + H_{X,X}^{0,\delta} + H_{X}^{0,\delta} \right] dt + Q_{t}^{0,\delta} dB_{t},
\end{array} \right. \\
P_{t}^{0,\delta} = \delta_{0}^{0,\delta} F_{X_{0},X_{0}} + \left[ \Gamma_{X_{0},X_{0}}^{0,\delta} \right].
\end{align*}
\]

Then, by Lemma 1 and applying Itô's formula to \( \langle P_{t}^{0,\delta}, \Theta_{0}^{0,\delta} \rangle \) yield

\[
E\left[ \langle \left( \delta_{0}^{0,\delta} F_{X_{0},X_{0}}^{0} + \left[ \Gamma_{X_{0},X_{0}}^{0,\delta} \right] \right) X_{1,1,1}, X_{1,1,1} \rangle \right] - \varepsilon^{2} \langle P_{0}^{0,\delta} X_{0}, X_{0} \rangle
\]
\[= E \int_{0}^{T} E' \left[ \left( \Delta \Sigma_{0}^{0,\delta} \right)^{T} P_{t}^{0,\delta} \Delta \Sigma_{0,1} I_{S_{1}} \right] + o(\varepsilon) \]  

Inserting (44) into (39), we can derive

\[
-\varepsilon \sqrt{2T} \left| X_{0} \right| + \varepsilon \left| Y_{0} \right| \leq E \int_{0}^{T} 
\left\{ \Delta X_{0}^{0,\delta} + \Gamma_{X_{0},X_{0}}^{0,\delta} \delta_{0}^{0,\delta} + \Phi_{0}^{0,\delta} \right\} dt
\]
\[+ \frac{1}{2} \left( \Delta \Sigma_{0}^{0,\delta} \right)^{T} P_{t}^{0,\delta} \Delta \Sigma_{0,1} I_{S_{1}} \right] + o(\varepsilon),
\]

where

\[
\begin{align*}
F_{X_{0}} &= (0, \varphi_{Y_{0}}), \\
F_{X_{1}} &= (\varphi_{X_{1}}, 0), \\
F_{X_{0},X_{0}} &= (0, \varphi_{Y_{0}}, 0), \\
F_{X_{0},X_{0},X_{0}} &= (0, \varphi_{Y_{0},X_{0}}).
\end{align*}
\]

By the Frechet differentiability, for \( \Gamma \), we have

\[
\langle \Gamma(X_{0},X_{0}), \tilde{\gamma} \rangle = (X_{0} - \gamma(X_{0},Y_{0})) \tilde{\gamma}_{0} + (Y_{0} - h(X_{0},Y_{0})) \tilde{\gamma}_{0}.
\]
Hence,

\[
\begin{align*}
\Gamma_x \bar{\theta} &= \langle \Gamma(X_0, X_T), \bar{\theta} \rangle_{X_0} = (\bar{\theta}_0, -\bar{\theta}_0 Y_{X_0} - \bar{\theta}_T H_{X_0}), \\
\Gamma_x \bar{\theta} &= \langle \Gamma(X_0, X_T), \bar{\theta} \rangle_{X_1} = (-\bar{\theta}_0 Y_{X_1} - \bar{\theta}_T H_{X_1}, \bar{\theta}_T), \\
\Gamma_{x_0} \bar{\theta} &= \langle \Gamma(X_0, X_T), \bar{\theta} \rangle_{X_0} = \left(0, 0 \right), \\
\Gamma_{x_2} \bar{\theta} &= \langle \Gamma(X_0, X_T), \bar{\theta} \rangle_{X_2} = \left(0, 0 \right), \\
\Gamma_{x_1} \bar{\theta} &= \langle \Gamma(X_0, X_T), \bar{\theta} \rangle_{X_1} = \left(0, 0 \right), \\
\Gamma_{x_3} \bar{\theta} &= \langle \Gamma(X_0, X_T), \bar{\theta} \rangle_{X_3} = \left(0, 0 \right), \\
\Gamma_{x_4} \bar{\theta} &= \langle \Gamma(X_0, X_T), \bar{\theta} \rangle_{X_4} = \left(0, 0 \right),
\end{align*}
\]

Note that \(\lim_{x \to 0} \|\bar{\theta}_x^x\|^2 + E|\bar{\theta}_x^x|^2 = 1\); there exists a subsequence of \((\bar{\theta}_x^x, \bar{\theta}_x^x)\) still denoted by \((\bar{\theta}_x^x, \bar{\theta}_x^x)\) such that \(\lim_{x \to 0} \|\bar{\theta}_x^x\|^2 + E|\bar{\theta}_x^x|^2 = 1\). Consequently, by the continuous dependence of the solution of mean-field BSDEs on parameters, we have the following conclusions, as \(x \to 0\), \((\Phi^x, \Psi^x) \to (\Phi^0, \Psi^0)\), \((p^x, q^x) \to (p^0, q^0)\) in \(H_{x^2}(0, T; R \times R)\) where \((\Phi^x, \Psi^x)\) and \((p^x, q^x)\) are solutions of first-order and second-order adjoint equations corresponding to \(\mu^x\), respectively. Hence, let (3.17) \(\times 1/\epsilon\) and send \(\epsilon \to 0\), \(\delta \to 0\), we see that

\[
E(\theta_0 F_X + \Gamma_x \bar{\theta} + \Phi_0, X_0) \geq 0, \quad \forall X_0 \in R^2,
\]

which implies

\[
\Phi_0 = -\theta_0 F_X - \Gamma_x \bar{\theta}. \tag{50}
\]

Also, dividing \(\epsilon^2 |X_0|^2\) in (45) and sending \(|X_0| \to \infty\), \(\epsilon \to 0\), \(\delta \to 0\), we get

\[
E(\theta_0 F_X, X_0 + \Gamma_x \bar{\theta} + P_0) \geq 0, \tag{51}
\]

twhich implies that the necessary condition (25) holds. Finally, by taking \(X_0 = 0\) in (45) and sending \(\epsilon \to 0\), \(\delta \to 0\), then using a standard argument, the variational inequality (26) follows.

Finally, let

\[
\Phi = (\xi, \eta), \quad \Psi = (\bar{\zeta}, \bar{\bar{\zeta}}). \tag{52}
\]

Then, it follows from (37) and (50) that

\[
\begin{align*}
\xi_0 &= -\bar{\theta}_0, \\
\eta_0 &= -\theta_0 \phi_{x_0} + \bar{\theta}_0 Y_{X_0} - \bar{\theta}_T H_{X_0}, \\
\xi_T &= \bar{\theta}_0 \phi_{x_T} - \bar{\theta}_0 Y_{X_T} - \bar{\theta}_T H_{X_T}, \\
\eta_T &= \bar{\theta}_T.
\end{align*}
\]

By use of (52), we may rewrite the first-order adjoint (37), further giving the one-dimensional first-order adjoint (24).

\[\text{Step 2. The general case of control domains.}\]

First, we set

\[
\begin{align*}
\mathcal{M}^n &= \{X_0 \in R^2; |X_0| \leq |X_0| + n\}, \\
\mathcal{N}^n &= \{Z_t \in R; |Z_t| \leq |Z_t| + n\}, \\
\mathcal{H}_{x^2}(0, T; \mathcal{N}^n) &= \{Z_t \in \mathcal{H}_{x^2}(0, T, R); Z_t \in \mathcal{N}^n\}, \\
\mathcal{H}_u &= \{u \in \mathcal{H}; \sup_{t \in [0, T]} |u| \leq \sup_{t \in [0, T]} |u| + n\}, \quad n = 1, 2, 3, \ldots
\end{align*}
\]

It is easy to see that \(\mathcal{M}^n\) is convex and

\[
\begin{align*}
\mathcal{X}_0 \in \mathcal{M}^n \subseteq \mathcal{M}^{n+1}, \\
R^2 = U_{n=1}^{\infty} \mathcal{M}^n, \quad \mathcal{U} \subseteq \mathcal{U}^{n+1}, \quad \mathcal{U} = U_{n=1}^{\infty} \mathcal{U}^n, \\
Z_t \in \mathcal{H}_{x^2}(0, T; \mathcal{N}^n) \subseteq \mathcal{H}_{x^2}(0, T, \mathcal{N}^{n+1}), \quad \mathcal{H}_{x^2}(0, T, R) = U_{n=1}^{\infty} \mathcal{H}_{x^2}(0, T, \mathcal{N}^n). \tag{55}
\end{align*}
\]

Then, by using the similar methods in [25], the desired conclusion (26) can be drawn. Since \((X_0, Y_0, Z, u)\) is arbitrary, we draw the desired conclusion and summarize it as follows.
Remark 4. When the coefficients $b, \sigma, \text{and } f$ do not depend on $w'$, system (3) reduces to the standard one, then Theorem 3 reduces to the necessary conditions for classical fully coupled FBSDEs, and the corresponding results were obtained by [19]. So our result is nontrivially more general of [19].

Remark 5. Compared with reference [19], the extended Ekeland’s variational principle is firstly introduced to reduce the requirement of lower semicontinuity of $f^t$. Secondly, a convergence technique is used to establish the stochastic maximum principle in general control domains. These methods are all different from the ones in [19].

Remark 6. For tackling the mixed initial-terminal conditions, a reduction method is adopted to transform the original mixed initial-terminal conditions as mixed initial-terminal state constraints. This treatment not only ensures the well-posedness of the first-order adjoint equations but also simplifies the derivation of the second-order adjoint equations.

4. Application

Consider the optimal control problem.

Problem 3. \( \mathcal{D} \) Minimize \( J(u) = 1/2E\left[ \int_0^T (l_tZ_t^2 + l_tu_t^2)dt + \phi_1X_T^2 + \phi_2Y_T^2 \right] \) over \( \mathcal{U} \), subject to

\[
\begin{align*}
\frac{dX_t}{dt} &= \left[ b_1EX_t + b_2EY_t + b_3EZ_t + b_4u_t \right]dt + \sigma_1\sigma_2EY_t + \sigma_3EZ_t + \sigma_4u_t dB_t, \\
\frac{dY_t}{dt} &= \left[ f_1EX_t + f_2EY_t + f_3EZ_t + f_4u_t \right]dt - Z_t dB_t, \\
X_0 &= \gamma_1X_T + \gamma_2Y_T, \quad Y_T = h_1X_T + h_2Y_T,
\end{align*}
\]

where \( b_i, \sigma_j, f_j, i \in \{1, \ldots, 4\} \) and \( \phi_j, \gamma_j, h_j, j \in \{1, 2\} \) are real constants of \( R \). Moreover, \( l_1 > 0, l_2 > 0, \phi_1, \phi_2, 0, \) and \( \phi_2 > 0 \). In practice, the range of control variable is limited for some reasons such as \( U = (-\infty, -1] \cup [1, +\infty) \) (see [18]). Assume that

\[
\begin{pmatrix}
-2f_1 & b_1 - f_2 & \sigma_1 - f_3 \\
 b_1 - f_2 & 2b_2 & \sigma_2 + b_3 \\
\sigma_1 - f_3 & \sigma_2 + b_3 & 2\sigma_3
\end{pmatrix}
= \begin{pmatrix}
-f_1 & -f_2 & -f_3 \\
 b_1 & b_2 & b_3 \\
-\phi_1 & -\phi_2 & -\phi_3
\end{pmatrix} + \begin{pmatrix}
-f_1 & b_1 & \phi_1 \\
 b_1 & b_2 & \phi_2 \\
-\phi_1 & -\phi_2 & -\phi_3
\end{pmatrix},
\]

is negative definite matrix, and

\[
\begin{pmatrix}
2h_1 & h_2 - \gamma_1 \\
h_2 - \gamma_1 & -2\gamma_2
\end{pmatrix}
\]

is positive definite matrix. Then, for any \( u \in \mathcal{U}[0, T] \), there exists a unique state process \((X, Y, Z)\). Now, suppose \( \mathcal{T} \) is optimal for problem \( \mathcal{D} \), then the adjoint equations are read as follows:

\[
\begin{align*}
d\xi_t &= -(b_1EX_t + \sigma_1E\xi_t - f_1E\eta_t)dt + \xi_t dB_t, \\
-d\eta_t &= (b_2E\xi_t + \sigma_2E\xi_t - f_2E\eta_t)dt - \eta_t dB_t, \\
\xi_T &= \theta_0\phi_1X_T - \theta_0\gamma_1 - \theta_1h_1, \quad \eta_T = \theta_0h_1.
\end{align*}
\]

Clearly, (59) admits unique solution \((\xi, \eta, \zeta)\) by assumptions \((H'_1) - (H'_4)\). And the BSDE for \((P, Q)\) becomes

\[
\begin{align*}
dP_t &= -E[B^TP_t + P_tB + \Sigma^TP_t\Sigma]dt + Q_t dB_t, \\
P_T &= \begin{pmatrix}
2\theta_0\phi_1 & 0 \\
0 & 0
\end{pmatrix},
\end{align*}
\]

where \( B = \begin{pmatrix} b_1 & b_2 \\ b_1 & b_2 \end{pmatrix} \) and \( \Sigma = \begin{pmatrix} \sigma_1 & \sigma_2 \\ \sigma_1 & \sigma_2 \end{pmatrix} \). According to the existence and uniqueness of solutions,

\[
\begin{align*}
\frac{1}{2}\theta_2u_t^2 + (b_4\xi_t - f_4\eta_t + \sigma_4\zeta_t)u_t + \frac{1}{2}\sigma_4^2P_{1t}u_t^2 - \sigma_4^2P_{1t}u_t \\
\geq \frac{1}{2}\theta_2u_t^2 + (b_4\xi_t - f_4\eta_t + \sigma_4\zeta_t)u_t - \frac{1}{2}\sigma_4^2P_{1t}u_t^2, \quad \forall u \in U.
\end{align*}
\]

Then, we have

\[
\begin{cases}
\mu_t, & \mu_t \in (-\infty, -1] \cup [1, +\infty), \\
1, & \mu_t \in [0, 1), \\
-1, & \mu_t \in (-1, 0),
\end{cases}
\]

with \( \mu_t = -(b_4\xi_t - f_4\eta_t + \sigma_4\zeta_t)/\theta_2 \).
Remark 7. From the above proof process, we note that when the coefficients of system (56) only depend on the expected values of the solution states, the second-order adjoint equation is mean-field independent and solvable.

Proposition 1. The control \( \pi \) defined by (63) is optimal, together with the corresponding trajectory \((X, Y, Z)\) which is an optimal solution for problem \( \mathcal{P} \).

Proof 3. Suppose \((X, Y, Z)\) is the trajectory of system (56) controlled by \( u \in \mathcal{U} \). By the convexity of a function, we have
\[
\frac{1}{2} \varphi_1 X_i^2 + \frac{1}{2} \varphi_1 X_i^2 \geq \varphi_1 X_i^2 (X_i - X_i T),
\] (64)
and
\[
\frac{1}{2} \varphi_2 Y_i^2 + \frac{1}{2} \varphi_2 Y_i^2 \geq \varphi_2 Y_i^2 (Y_i - Y_i T).
\] (65)

Applying Itô’s formula to \( \xi_t (X_i - X_i) + \eta_t (Y_i - Y_i) \), one has
\[
E[\theta_0 \varphi_1 X_i T (X_i - X_i T) + \theta_0 \varphi_2 Y_i T (Y_i - Y_i T)]
= E \int_0^T \left\{ (b_i \xi_t - f_i \eta_t + \sigma_i \xi_t) (u_t - \bar{u}_t) + \tilde{\xi}_t (Z_t - Z_t) \right\} dt.
\] (66)

Taking notice of (64)–(66),
\[
J(u) - J(\bar{u}) \geq \frac{1}{2} E \int_0^T \left\{ 2 \left( Z_i^2 - Z_i^2 \right) + 2 \left( u_t^2 - \bar{u}_t^2 \right) \right\} dt
+ E\left[ \varphi_1 X_i T (X_i - X_i T) + \varphi_2 Y_i T (Y_i - Y_i T) \right]
\]
\[
= \frac{1}{2} E \int_0^T \left\{ 2 \left( Z_i^2 - Z_i^2 \right) + 2 \left( u_t^2 - \bar{u}_t^2 \right) \right\} dt
+ E \int_0^T \left\{ b_i \xi_t - f_i \eta_t + \sigma_i \xi_t \left( u_t - \bar{u}_t \right) + \tilde{\xi}_t (Z_t - Z_t) \right\} dt
\]
\[
= \frac{1}{2} E \int_0^T \left\{ 2 \left( Z_i^2 - Z_i^2 \right) + 2 \left( u_t^2 - \bar{u}_t^2 \right) \right\} dt
+ E \int_0^T \left\{ \frac{b_i \xi_t - f_i \eta_t + \sigma_i \xi_t}{\theta_0} \left( u_t - \bar{u}_t \right) \right\} dt.
\] (67)

From (63), we know that
\[
\frac{1}{2} \varphi_1 X_i^2 + \frac{1}{2} \varphi_1 X_i^2 \geq \varphi_1 X_i^2 (X_i - X_i T),
\] holds for all \( u \in \mathcal{U} \). This means that \( J(u) - J(\bar{u}) \geq 0, \forall u \in \mathcal{U} \), which yields the optimality of \( \bar{u} \).

5. Conclusion

This paper discussed optimal control problems of mean-field FBSDEs with mixed initial-terminal conditions. Firstly, we initially introduce extended Ekeland’s variational principle to reduce requirements of lower semicontinuity of the penalty functional. Secondly, the reduction method is adopted to guarantee the well-posedness of the first-order and second-order adjoint equations also with mixed initial-terminal conditions. Via spike variational technique, a general stochastic maximum principle is established in the framework of mean-field theory. Finally, a linear-quadratic example is worked out to illustrate the application of the results.

Data Availability

Data sharing is not applicable to this article as no datasets were generated during the current study.

Conflicts of Interest

The author declares that she has no conflicts of interest regarding the publication of this paper.
Acknowledgments
This research was partially supported by National Natural Science Foundation of China (Grant no. 11901112).

References


