

Research Article

Maximum Principle for Near-Optimality of Mean-Field FBSDEs

Ruijing Li¹ and Chaozhu Hu² 

¹*School of Statistics and Mathematics, Guangdong University of Finance and Economics, Guangzhou 510320, Guangdong, China*

²*School of Science, Hubei University of Technology, Wuhan 430068, Hubei, China*

Correspondence should be addressed to Chaozhu Hu; huchaozhu@126.com

Received 13 April 2020; Revised 16 May 2020; Accepted 18 May 2020; Published 8 June 2020

Guest Editor: Rongwei Guo

Copyright © 2020 Ruijing Li and Chaozhu Hu. 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 a near-optimal control problem for systems governed by mean-field forward-backward stochastic differential equations (FBSDEs) with mixed initial-terminal conditions. Utilizing Ekeland's variational principle as well as the reduction method, the necessary and sufficient near-optimality conditions are established in the form of Pontryagin's type. The results are obtained under restriction on the convexity of the control domain. As an application, a linear-quadratic stochastic control problem is solved explicitly.

1. Introduction

Near-optimal control problems have attracted more attentions in recent years due to its distinct advantages, such as existence under minimal assumptions, availability in most practical cases, and convenience for implementation both analytically and numerically. The study of this theory can be traced back to Ekeland [1] and later greatly developed by Zhou [2–4] for deterministic and stochastic cases. Since then, many works have been devoted to the near-optimality of various stochastic control systems. Without being exhaustive, let us refer to [5–13] and the references therein.

In 2015, Zhang et al. [14] investigated the near-optimality necessary conditions for classical linear FBSDEs, where the control domain was with nonconvexity. Via convergence technique as well as reduction method, they established the near-optimal maximum principle. Soon afterwards, under the same assumptions, Zhang [15] presented the near-optimal sufficient conditions for such classical linear FBSDEs. Especially, in 2018, by defining viscosity solution with perturbation factor to dispense the illusory differentiability condition of value function, Zhang and Zhou [16] established the necessary near-optimality conditions for stochastic recursive systems by virtue of dynamic programming principle. Another noteworthy thing is that,

for recent years, some authors started research studies on near-optimal control problems for delay systems. For example, Zhang [17] first studied near-optimal control problems for linear stochastic delay systems. By anticipated backward stochastic differential equations method as well as maximum principle, necessary condition and sufficient verification theorem were provided. Then, also under restriction on convexity control domain, Wang and Wu [18] investigated near-optimal control problem for nonlinear stochastic delay systems. By Ekeland's variational principle and corresponding moment estimations, they presented the sufficient as well as necessary near-optimality conditions. For more details, refer to [19, 20] and the references therein.

However, to the best of our knowledge, few papers can be found in the literature on the near-optimality of mean-field backward stochastic differential equations (BSDEs). This new kind of mean-field BSDEs was first introduced by Buckdahn et al. [21], which were derived as a limit of some highly dimensional system of FBSDEs, corresponding to a large number of particles. It has been shown in Buckdahn et al. [22] that, such a mean-field BSDE described the viscosity solution of the associated nonlocal partial differential equations. Henceforth, many authors take into account of this system of McKean-Vlasov type (Lasry and Lions [23]) adapted for different frameworks, for example, Xu and Wu

[24] presented a maximum principle for optimal control problems governed by backward stochastic partial differential equations of mean-field type, and for other related works, refer to [25–28].

As we can see that all the above literature studies are about mean-field problems involving expectations as mean-field terms. In fact, there is another line dealing with mean-field problems, which involve large-population as mean-field terms to describe the impact of the population's collective behaviors on all agents (Huang et al. [29]) such as the work of Huang [30] and Xu and Shi [31] as well as the work of Xu and Zhang [32] all concerned with general mean-field linear-quadratic-Gaussian (LQG) games of stochastic large-population systems; through the consistency condition, they derived the decentralized strategies and further verified the asymptotic near-optimality property (namely, ε -Nash equilibrium) of decentralized strategies for the LQ games. On the contrary, a relevant work of Hafayed and Abbas [8] dealing with near-optimal control problems has established necessary and sufficient conditions for mean-field singular stochastic systems in the case of controlled diffusion coefficient. Particularly, in the concluding section, it is pointed out that the establishment of necessary and sufficient near-optimal conditions for mean-field FBSDEs also remains an open problem. Motivated by this fact with the addition of above described mean-field theory application background in economics and finance, this paper is to discuss near-optimal control problems for mean-field FBSDEs, where the controlled state systems are with mixed initial-terminal conditions.

The main contribution of this paper lies in the initial introduction of three first-order adjoint equations to eliminate the corresponding variational processes during dual analysis; another is rooted in the usage of reduction method to guarantee the well-posedness of the first-order adjoint equations with mixed initial-terminal conditions. Via classical convex variational technique and Ekeland's variational principle, a necessary condition of Pontryagin's type is derived. Then, under some additional assumptions, we prove that the near-maximum condition on the Hamiltonian function is a sufficient condition for near-optimality. It is remarkable that our results extend those of [5] essentially to the framework of mean-field theory.

The rest of this paper is organized as follows. In Section 2, we state some preliminaries and basic definitions. In Sections 3 and 4, we establish the main theorems and provide its detailed proof. In Section 5, an example of a linear-quadratic control problem is worked out to illustrate the theoretical applications. Finally, some concluding remarks are given in Section 6.

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 standard Brownian motion $(W_t)_{t \geq 0}$ is defined, $\mathcal{F} = \{\mathcal{F}_s, 0 \leq s \leq T\}$ be the natural filtration generated by $(W_t)_{t \geq 0}$ and augmented by all P -null sets, i.e.,

$$\mathcal{F}_s = \sigma\{W_r, r \leq s\} \vee \mathcal{N}_P, \quad s \in [0, T], \quad (1)$$

where \mathcal{N}_P is the set of all P -null subsets. We now introduce some spaces of random variables and stochastic processes.

$$\begin{aligned} L^2_{\mathcal{F}}(\Omega; R) &= \{\mathcal{F} - \text{measurable random variable} \\ X: E|X|^2 < \infty\} \\ \mathcal{S}^2_{\mathcal{F}}(0, T; R) &= \{\mathcal{F} - \text{adapted and continuous process} \\ \psi: E[\sup_{t \in [0, T]} |\psi_t|^2] < \infty\} \\ \mathcal{H}^2_{\mathcal{F}}(0, T; R) &= \{\mathcal{F} - \text{adapted process} \\ \psi: E[\int_0^T |\psi_t|^2 dt] < \infty\} \\ M^2[0, T] &:= \mathcal{S}^2_{\mathcal{F}}(0, T; R) \times \mathcal{S}^2_{\mathcal{F}}(0, T; R) \times \mathcal{H}^2_{\mathcal{F}}(0, T; R) \end{aligned}$$

Clearly, $M^2[0, T]$ is a Banach space. Any process in $M^2[0, T]$ is defined by $\Theta := (x, y, z)$ with the norm

$$\|\Theta\|_{M^2[0, T]} := \left\{ E \left[\sup_{t \in [0, T]} |x_t|^2 + \sup_{t \in [0, T]} |y_t|^2 + \int_0^T |z_t|^2 dt \right] \right\}^{1/2}. \quad (2)$$

We study the near-optimal control problem of the following controlled mean-field FBSDEs having mixed initial-terminal conditions:

$$\begin{cases} dx_t = b(t, x_t, Ex_t, u_t)dt + \sigma(t, x_t, Ex_t, u_t)dW_t, \\ -dy_t = f(t, x_t, y_t, z_t, Ex_t, Ey_t, Ez_t, u_t)dt - z_t dW_t, \\ x_0 = \gamma(x_T, y_0), \\ y_T = h(x_T, y_0), \end{cases} \quad (3)$$

where $b, \sigma: [0, T] \times R^2 \times U \rightarrow R$; $f: [0, T] \times R^6 \times U \rightarrow R$; $h, \gamma: R \times R \rightarrow R$; and U is a given convex closed set of R . The cost functional to be minimized over the space $\mathcal{U} = L^2_{\mathcal{F}}(0, T; U)$ of admissible controls takes the form

$$J(u) = E \left\{ \int_0^T l(t, x_t, y_t, z_t, Ex_t, Ey_t, Ez_t, u_t)dt + \varphi(x_T, y_0) \right\}, \quad (4)$$

with $l: [0, T] \times R^6 \times U \rightarrow R$; $\varphi: R \times R \rightarrow R$.

Definition 1 (see [4]). Both a family of admissible pairs $\{(x^\varepsilon, y^\varepsilon, z^\varepsilon, u^\varepsilon)\}$ parameterized by $\varepsilon > 0$ and any element $(x^\varepsilon, y^\varepsilon, z^\varepsilon, u^\varepsilon)$ in the family are called near-optimal if

$$\left| J(u^\varepsilon) - \inf_{u \in \mathcal{U}} J(u) \right| \leq r(\varepsilon), \quad (5)$$

holds for sufficiently small ε , where r is a function of ε satisfying $r(\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow 0$. The estimate $r(\varepsilon)$ is called an error bound. If $r(\varepsilon) = C\varepsilon^\delta$ for some $\delta > 0$ independent of the constant C , then u^ε is called near-optimal with order ε^δ . Particularly, when $r(\varepsilon) = \varepsilon$, u^ε is called ε -optimal. The near-optimal control problem under consideration in this paper is as follows.

Problem \mathcal{A} . Find $u^\varepsilon \in \mathcal{U}$ such that

$$J(u^\varepsilon) = \inf_{v \in \mathcal{U}} J(v) + \varepsilon. \quad (6)$$

Some notations and assumptions are presented before giving the well-posedness of system (3). We denote the norm by $|\cdot|$ of an Euclidean space.

(A₁) The functions b , σ , f , and l are \mathcal{F} -progressively measurable in u , continuously differentiable in $x, y, z, \bar{x}, \bar{y}$, and \bar{z} , and the derivatives of b , σ , f , and l with respect to $x, y, z, \bar{x}, \bar{y}$, and \bar{z} are bounded. Moreover, for some constant $C > 0$,

$$\begin{aligned} & (1 + |x| + |\bar{x}|)^{-1} |b(t, x, \bar{x}, u)| + (1 + |x| + |\bar{x}|)^{-1} |\sigma(t, x, \bar{x}, u)| \\ & + (1 + |x| + |y| + |z| + |\bar{x}| + |\bar{y}| + |\bar{z}|)^{-1} |f(t, x, y, z, \bar{x}, \bar{y}, \bar{z}, u)| \\ & + (1 + |x| + |y| + |z| + |\bar{x}| + |\bar{y}| + |\bar{z}|)^{-1} |l(t, x, y, z, \bar{x}, \bar{y}, \bar{z}, u)| \leq C. \end{aligned} \quad (7)$$

(A₂) h , γ , and φ are continuously differentiable in x and y , and the derivatives of h , γ , and φ with respect to x and y are bounded. Moreover, for some constant $C > 0$, $\rho = h, \gamma, \varphi$

$$(1 + |x| + |y|)^{-1} |\rho(x, y)| \leq C. \quad (8)$$

(A₃) There is a constant $C > 0$ and $\beta \in [0, 1]$ such that

$$\begin{aligned} & \left| b_x(t, x, \bar{x}, u) - b_x(t, x_l, \bar{x}', u) \right| + \left| b_x^-(t, x, \bar{x}, u) - b_x^-(t, x_l, \bar{x}', u) \right| \\ & + \left| \sigma_x(t, x, \bar{x}, u) - \sigma_x(t, x_l, \bar{x}', u) \right| + \left| \sigma_x^-(t, x, \bar{x}, u) - \sigma_x^-(t, x_l, \bar{x}', u) \right| \\ & \leq C \left(|x - x_l|^\beta + |\bar{x} - \bar{x}'|^\beta \right), \\ & \left| f_i(t, x, y, z, \bar{x}, \bar{y}, \bar{z}, u) - f_i(t, x_l, y_l, z_l, \bar{x}', \bar{y}', \bar{z}', u) \right| \\ & + \left| l_i(t, x, y, z, \bar{x}, \bar{y}, \bar{z}, u) - l_i(t, x_l, y_l, z_l, \bar{x}', \bar{y}', \bar{z}', u) \right| \\ & \leq C \left(|x - x_l|^\beta + |y - y_l|^\beta + |z - z_l|^\beta + |\bar{x} - \bar{x}'|^\beta + |\bar{y} - \bar{y}'|^\beta + |\bar{z} - \bar{z}'|^\beta \right). \end{aligned} \quad (9)$$

Furthermore,

$$\begin{aligned} & \left| \rho_x(x, y) - \rho_x(x_l, y_l) \right| + \left| \rho_y(x, y) - \rho_y(x_l, y_l) \right| \\ & \leq C \left(|x - x_l|^\beta + |y - y_l|^\beta \right), \end{aligned} \quad (10)$$

where $i = x, y, z, \bar{x}, \bar{y}, \bar{z}$ and $\rho = h, \gamma, \varphi$.

Remark 1. Under assumptions (A₁) – (A₃) via Theorem 2 in [33], the mean-field stochastic system (3) admits a unique adapted solution $(x, y, z) \in M^2[0, T]$.

In fact, due to the mixed initial-terminal conditions in the state equation, even if we have the well-posedness of the state equation via the Lyapunov operator introduced in [34], the well-posedness of the first-order adjoint equation seems to be not guaranteed. To overcome this difficulty, we introduce a reduction method inspired by the study of optimality variational principle for controlled FBSDEs with mixed initial-terminal conditions [35]. First, we pose the following problem.

Problem \mathcal{B} . Find $(x_0^\varepsilon, y_0^\varepsilon, u^\varepsilon) \in \mathcal{R} := \mathbb{R} \times \mathbb{R} \times \mathcal{U}$ such that

$$J(x_0^\varepsilon, y_0^\varepsilon, u^\varepsilon) = \inf_{(x_0, y_0, u) \in \mathcal{R}} J(x_0, y_0, u) + \varepsilon, \quad (11)$$

where (x_0, y_0, u) is subject to the forward control system:

$$\begin{cases} dx_t = b(t, x_t, Ex_t, u_t)dt + \sigma(t, x_t, Ex_t, u_t)dW_t, \\ -dy_t = f(t, x_t, y_t, z_t, Ex_t, Ey_t, Ez_t, u_t)dt - z_t dW_t, \\ x(0) = x_0, \\ y(0) = y_0, \end{cases} \quad (12)$$

with the mixed initial-terminal state constraints:

$$\begin{aligned} x_0 &= \gamma(x_T, y_0), \\ y_T &= h(x_T, y_0). \end{aligned} \quad (13)$$

It is remarkable that, for Problem \mathcal{A} , the mean-field system (3) has a unique solution (x, y, z) under (A₁) – (A₃), which implies that $y(0)$ is unique and completely determinate. While, for Problem \mathcal{B} , $y(0)$ is arbitrary and viewed as a control variable. It just needs to satisfy the near-optimal state constraints at time T . So, Problem \mathcal{A} is embedded into Problem \mathcal{B} . Hence, if the triple $(x_0^\varepsilon, y_0^\varepsilon, u^\varepsilon)$ is the near-optimal control of \mathcal{B} , then u^ε is near-optimal for Problem \mathcal{A} . In the following section, we will adopt the classical convex variational technique to solve Problem \mathcal{B} .

3. Necessary Condition of Near-Optimality

This section is devoted to the study of the main theorem. For simplicity, we denote

$$\left\{ \begin{array}{l} b_j := b_j(t, x_t, Ex_t, u_t), \\ \sigma_j := \sigma_j(t, x_t, Ex_t, u_t), \\ f_j := f_j(t, \Gamma_t, u_t), \\ l_j := l_j(t, \Gamma_t, u_t), \\ \rho_{x_T} = \rho_{x_T}(x_T, y_0), \\ \rho_{y_0} = \rho_{y_0}(x_T, y_0), \\ j \in \{x, y, z, \tilde{x}, \tilde{y}, \tilde{z}\}, \\ \Gamma_t = (x_t, y_t, z_t, Ex_t, Ey_t, Ez_t), \\ \rho = h, \gamma, \varphi. \end{array} \right. \quad (14)$$

For any $u \in \mathcal{U}$ and the corresponding state processes (x, y, z) , we define the first-order adjoint equation as

$$\left\{ \begin{array}{l} d\xi_t = \left\{ f_y \xi_t + l_y + E \left[f_{\tilde{y}} \xi_t + l_{\tilde{y}} \right] \right\} dt \\ \quad + \left\{ f_z \xi_t + l_z + E \left[f_{\tilde{z}} \xi_t + l_{\tilde{z}} \right] \right\} dW_t, \\ -d\eta_t = \left\{ -f_x \xi_t + b_x \eta_t + \sigma_x \zeta_t - l_x \right\} dt \\ \quad + \left\{ E \left[-f_{\tilde{x}} \xi_t + b_{\tilde{x}} \eta_t + \sigma_{\tilde{x}} \zeta_t - l_{\tilde{x}} \right] \right\} dt - \zeta_t dW_t, \\ \xi_0 = -\theta_0 \varphi_{y_0} - \eta_0 \gamma_{y_0} + \xi_T h_{y_0}, \\ \eta_T = -\xi_T h_{x_T} + \eta_0 \gamma_{x_T} + \theta_0 \varphi_{x_T}. \end{array} \right. \quad (15)$$

Remark 2. Under assumptions $(A_1) - (A_3)$, the adjoint equation (15) admits a unique adapted solution $(\xi, \eta, \zeta) \in M^2[0, T]$. The well-posedness of the corresponding adjoint system will be provided in the derivation process of Theorem 1.

Define a metric on \mathcal{U} by

$$d(u, u') = \left[E \int_0^T |u - u'|^2 dt \right]^{1/2}, \quad \forall u, u' \in \mathcal{U}. \quad (16)$$

Since \mathcal{U} is closed, it can be shown that (\mathcal{U}, d) is a complete metric space. Next, we will present some

$$\begin{aligned} -C\varepsilon^k \theta_0^\varepsilon \leq E \int_0^T & \left(\eta_t^\varepsilon [b(t, x_t^\varepsilon, Ex_t^\varepsilon, u_t) - b(t, x_t^\varepsilon, Ex_t^\varepsilon, u_t^\varepsilon)] + \zeta_t^\varepsilon [\sigma(t, x_t^\varepsilon, Ex_t^\varepsilon, u_t) - \sigma(t, x_t^\varepsilon, Ex_t^\varepsilon, u_t^\varepsilon)] \right) \\ & - \xi_t^\varepsilon [f(t, \Gamma_t^\varepsilon, u_t) - f(t, \Gamma_t^\varepsilon, u_t^\varepsilon)] - [l(t, \Gamma_t^\varepsilon, u_t) - l(t, \Gamma_t^\varepsilon, u_t^\varepsilon)] dt, \quad \forall u \in \mathcal{U}, \end{aligned} \quad (19)$$

where $(\xi^\varepsilon, \eta^\varepsilon, \zeta^\varepsilon) \in M^2[0, T]$ is the solution of the first-order adjoint equation (15) corresponding to u^ε .

continuity of the state processes and adjoint processes with respect to the metric d .

Lemma 1. For any $0 < \alpha < 1$ and $0 < p \leq 2$, there is a constant $C = C(\alpha, p) > 0$ such that, for any $u, \bar{u} \in \mathcal{U}$ along with the corresponding trajectories (x, y, z) and $(\bar{x}, \bar{y}, \bar{z})$, it follows that

$$E \left(\sup_{t \in [0, T]} |x_t - \bar{x}_t|^p \right) \leq C d(u, \bar{u})^{\alpha p/2}, \quad (17)$$

$$\sup_{t \in [0, T]} E |y_t - \bar{y}_t|^p + E \int_0^T |z_t - \bar{z}_t|^p dt \leq C d(u, \bar{u})^{\alpha p/2}.$$

Proof. Applying the classical methods as Lemma 4 in [5] for dealing with mean-field FBSDEs, together with Burkholder–Davis–Gundy inequality and Gronwall’s inequality, we can logically obtain the estimates. \square

Lemma 2. Let $(A_1) - (A_3)$ hold, for any $0 < \alpha < 1$ and $1 < p < 2$ satisfying $(1 + \alpha\beta)p < 2$, and there is a constant $C = C(\alpha, \beta, p) > 0$ such that, for any $u, \bar{u} \in \mathcal{U}$, along with the corresponding trajectories (x, y, z) and $(\bar{x}, \bar{y}, \bar{z})$ and the solutions (ξ, η, ζ) and $(\bar{\xi}, \bar{\eta}, \bar{\zeta})$ of the corresponding adjoint equation (15), it holds that

$$E \int_0^T |\xi_t - \bar{\xi}_t|^p dt \leq C d(u, \bar{u})^{\alpha\beta p/2}, \quad (18)$$

$$E \int_0^T \left(|\eta_t - \bar{\eta}_t|^p + |\zeta_t - \bar{\zeta}_t|^p \right) dt \leq C d(u, \bar{u})^{\alpha\beta p/2}.$$

Proof. Applying the classical methods as Lemma 5 in [5] for dealing with mean-field FBSDEs, we can naturally obtain the estimates. \square

Theorem 1. Let $(A_1) - (A_3)$ hold, for any $\varepsilon > 0$, u^ε is an ε -optimal control of problem \mathcal{B} . Then, for any $\kappa \in [0, (1/3))$, there exist three parameters θ_0^ε , $\bar{\theta}_0^\varepsilon$, and $\bar{\theta}_T^\varepsilon$ with $|\theta_0^\varepsilon|^2 + E|\bar{\theta}_0^\varepsilon|^2 + E|\bar{\theta}_T^\varepsilon|^2 = 1$, and $\theta_0^\varepsilon \geq 0$ holds that

Proof. Under the assumption (A_2) , it is easy to check that $J(x_0, y_0, u)$ is lower semicontinuous on $\mathcal{R} := R \times R \times \mathcal{U}$, which is a complete metric space under the following metric:

$$d_{\mathcal{R}}(\Theta, \bar{\Theta}) = \left[|x_0 - \bar{x}_0|^2 + |y_0 - \bar{y}_0|^2 + d(u, \bar{u})^2 \right]^{1/2}, \quad \forall \Theta = (x_0, y_0, u), \bar{\Theta} = (\bar{x}_0, \bar{y}_0, \bar{u}) \in \mathcal{R}. \quad (20)$$

By Ekeland's variational principle [1], there exists an admissible control $(\tilde{x}_0^\varepsilon, \tilde{y}_0^\varepsilon, \tilde{u}^\varepsilon) \in \mathcal{R}$ such that

$$\begin{cases} d_{\mathcal{R}}((x_0^\varepsilon, y_0^\varepsilon, u^\varepsilon), (\tilde{x}_0^\varepsilon, \tilde{y}_0^\varepsilon, \tilde{u}^\varepsilon)) \leq \varepsilon^{2/3}, \\ \tilde{J}^\varepsilon((\tilde{x}_0^\varepsilon, \tilde{y}_0^\varepsilon, \tilde{u}^\varepsilon)) \leq \tilde{J}^\varepsilon(x_0, y_0, u), \quad \forall (x_0, y_0, u) \in \mathcal{R}, \end{cases} \quad (21)$$

where

$$\tilde{J}^\varepsilon(x_0, y_0, u) = J(x_0, y_0, u) + \varepsilon^{1/3} d_{\mathcal{R}}((x_0, y_0, u), (\tilde{x}_0^\varepsilon, \tilde{y}_0^\varepsilon, \tilde{u}^\varepsilon)). \quad (22)$$

It means that $(\tilde{x}_0^\varepsilon, \tilde{y}_0^\varepsilon, \tilde{u}^\varepsilon)$ is optimal for system (12) with the new cost functional $\tilde{J}^\varepsilon(x_0, y_0, u)$. On the contrary, due to the mixed initial-terminal endpoint constraints in problem \mathcal{B} , we need to introduce the penalty functional to transform the original problem with endpoint constraints to the penalized optimal control problem with no endpoint constraints.

$$\begin{cases} J^{\delta, \varepsilon}(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon}) \leq J^{\delta, \varepsilon}(\tilde{x}_0^\varepsilon, \tilde{y}_0^\varepsilon, \tilde{u}^\varepsilon) = \delta, \\ |x_0^{\delta, \varepsilon} - \tilde{x}_0^\varepsilon|^2 + |y_0^{\delta, \varepsilon} - \tilde{y}_0^\varepsilon|^2 + d|u^{\delta, \varepsilon}, \tilde{u}^\varepsilon|^2 \leq \delta, \\ -\sqrt{\delta} \left[|x_0^{\delta, \varepsilon} - x_0|^2 + |y_0^{\delta, \varepsilon} - y_0|^2 + d(u^{\delta, \varepsilon}, u)^2 \right]^{1/2} \leq J^{\delta, \varepsilon}(x_0, y_0, u) - J^{\delta, \varepsilon}(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon}), \quad \forall (x_0, y_0, u) \in \mathcal{R}. \end{cases} \quad (25)$$

Therefore, $(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})$ is optimal for system (13) with the new cost functional:

$$J^{\delta, \varepsilon}(x_0, y_0, u) + \sqrt{\delta} \left[|x_0^{\delta, \varepsilon} - x_0|^2 + |y_0^{\delta, \varepsilon} - y_0|^2 + d(u^{\delta, \varepsilon}, u)^2 \right]^{1/2}. \quad (26)$$

So far, we have transformed the original problem with endpoint constraints to the penalized optimal control problem with no endpoint constraints, and the optimal 3-tuple $(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})$ approaches $(\tilde{x}_0^\varepsilon, \tilde{y}_0^\varepsilon, \tilde{u}^\varepsilon)$ as $\delta \rightarrow 0$. In the following, a convex perturbation is employed to obtain a

Let $(\tilde{x}_0^\varepsilon, \tilde{y}_0^\varepsilon, \tilde{u}^\varepsilon)$ be an optimal control of problem \mathcal{B} , with the corresponding optimal state process $(\tilde{x}^\varepsilon, \tilde{y}^\varepsilon, \tilde{z}^\varepsilon)$. Without loss of generality, we assume that $\tilde{J}^\varepsilon(\tilde{x}_0^\varepsilon, \tilde{y}_0^\varepsilon, \tilde{u}^\varepsilon) = 0$. For any $\delta > 0$ and $(x_0, y_0, u) \in \mathcal{R}$, we define the penalty functional:

$$J^{\delta, \varepsilon}(x_0, y_0, u) = \left\{ [\tilde{J}^\varepsilon(x_0, y_0, u) + \delta]^2 + E[|x_0 - \gamma(x_T, y_0)|^2 + |y_T - h(x_T, y_0)|^2] \right\}^{1/2}. \quad (23)$$

Obviously,

$$\begin{aligned} J^{\delta, \varepsilon}(x_0, y_0, u) &> 0, \\ J^{\delta, \varepsilon}(\tilde{x}_0^\varepsilon, \tilde{y}_0^\varepsilon, \tilde{u}^\varepsilon) &= \delta \leq \inf_{(x_0, y_0, u) \in \mathcal{R}} J^{\delta, \varepsilon}(x_0, y_0, u) + \delta. \end{aligned} \quad (24)$$

By Ekeland's variational principle, there exists a 3-tuple $(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon}) \in \mathcal{R}$ such that

maximum principle for $(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})$. To this end, let $\Theta = (x_0, y_0, u) \in \mathcal{R}$ such that $\Theta^{\delta, \varepsilon} + \Theta = (x_0^{\delta, \varepsilon} + x_0, y_0^{\delta, \varepsilon} + y_0, u^{\delta, \varepsilon} + u) \in \mathcal{R}$; then, for any $\varepsilon > 0$,

$$\Theta_t^{\varepsilon, \delta, \varepsilon} = \Theta_t^{\delta, \varepsilon} + \varepsilon \Theta_t \in \mathcal{R}, \quad t \in [0, T]. \quad (27)$$

Let $(x^{\varepsilon, \delta, \varepsilon}, y^{\varepsilon, \delta, \varepsilon}, z^{\varepsilon, \delta, \varepsilon})$ be the state processes corresponding to $(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon})$, and the processes $(x_1^{\varepsilon, \delta, \varepsilon}, y_1^{\varepsilon, \delta, \varepsilon}, z_1^{\varepsilon, \delta, \varepsilon})$ be the solution of the following variational equations:

$$\begin{cases} dx_{1,t}^{\varepsilon, \delta, \varepsilon} = \left\{ b_x^{\delta, \varepsilon} x_{1,t}^{\varepsilon, \delta, \varepsilon} + b_x^{\delta, \varepsilon} E[x_{1,t}^{\varepsilon, \delta, \varepsilon}] + \Delta b^{\delta, \varepsilon} I_{S_{ct}} \right\} dt + \left\{ \sigma_x^{\delta, \varepsilon} x_{1,t}^{\varepsilon, \delta, \varepsilon} + \sigma_x^{\delta, \varepsilon} E[x_{1,t}^{\varepsilon, \delta, \varepsilon}] + \Delta \sigma^{\delta, \varepsilon} I_{S_{ct}} \right\} dW_t, \\ -dy_{1,t}^{\varepsilon, \delta, \varepsilon} = \left\{ f_x^{\delta, \varepsilon} x_{1,t}^{\varepsilon, \delta, \varepsilon} + f_y^{\delta, \varepsilon} y_{1,t}^{\varepsilon, \delta, \varepsilon} + f_z^{\delta, \varepsilon} z_{1,t}^{\varepsilon, \delta, \varepsilon} + f_x^{\delta, \varepsilon} E[x_{1,t}^{\varepsilon, \delta, \varepsilon}] + f_y^{\delta, \varepsilon} E[y_{1,t}^{\varepsilon, \delta, \varepsilon}] + f_z^{\delta, \varepsilon} E[z_{1,t}^{\varepsilon, \delta, \varepsilon}] + \Delta f^{\delta, \varepsilon} I_{S_{ct}} \right\} dt - z_{1,t}^{\varepsilon, \delta, \varepsilon} dW_t, \\ x_{1,0}^{\varepsilon, \delta, \varepsilon} = \sqrt{\varepsilon} x_0, \\ y_{1,0}^{\varepsilon, \delta, \varepsilon} = \sqrt{\varepsilon} y_0, \end{cases} \quad (28)$$

where for simplicity of notations, we still use $\rho_j^{\delta, \varepsilon}$ corresponding to $(x^{\delta, \varepsilon}, y^{\delta, \varepsilon}, z^{\delta, \varepsilon}, u^{\delta, \varepsilon})$, $\rho = b, \sigma, f$. Then, we have

the following estimates, whose proofs are similar to those given in [27]

$$\left\{ \begin{array}{l}
E\left(\sup_{t \in [0, T]} |x_{1,t}^{\varepsilon, \delta, \varepsilon}|^2\right) + E\left(\sup_{t \in [0, T]} |x_t^{\varepsilon, \delta, \varepsilon} - x_t^{\delta, \varepsilon}|^2\right) \leq C\varepsilon, \\
E\left(\sup_{t \in [0, T]} |x_t^{\varepsilon, \delta, \varepsilon} - x_t^{\delta, \varepsilon} - x_{1,t}^{\varepsilon, \delta, \varepsilon}|^2\right) \leq C\varepsilon^2, \\
E\left(\sup_{t \in [0, T]} |y_{1,t}^{\varepsilon, \delta, \varepsilon}|^2\right) + E\left(\sup_{t \in [0, T]} |y_t^{\varepsilon, \delta, \varepsilon} - y_t^{\delta, \varepsilon}|^2\right) \leq C\varepsilon, \\
E\left(\sup_{t \in [0, T]} |y_t^{\varepsilon, \delta, \varepsilon} - y_t^{\delta, \varepsilon} - y_{1,t}^{\varepsilon, \delta, \varepsilon}|^2\right) \leq C\varepsilon^2, \\
E \int_0^T |z_{1,t}^{\varepsilon, \delta, \varepsilon}|^2 dt + E \int_0^T |z_t^{\varepsilon, \delta, \varepsilon} - z_t^{\delta, \varepsilon}|^2 dt \leq C\varepsilon, \\
E \int_0^T |z_t^{\varepsilon, \delta, \varepsilon} - z_t^{\delta, \varepsilon} - z_{1,t}^{\varepsilon, \delta, \varepsilon}|^2 dt \leq C\varepsilon^2.
\end{array} \right. \quad \text{Noting that } d(u_t^{\varepsilon, \delta, \varepsilon}, u_t^{\delta, \varepsilon}) \leq C\varepsilon, \text{ from the last relation in} \\
(25), \text{ we derive}$$

(29)

$$\begin{aligned}
& -\varepsilon\sqrt{\delta}\sqrt{|x_0|^2 + |y_0|^2} + C \leq J^{\delta, \varepsilon}(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) - J^{\delta, \varepsilon}(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon}) \\
& = \frac{J^{\delta, \varepsilon}(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon})^2 - J^{\delta, \varepsilon}(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})^2}{J^{\delta, \varepsilon}(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) + J^{\delta, \varepsilon}(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})} \\
& = \frac{[\bar{J}^\varepsilon(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) + \delta]^2 - [\bar{J}^\varepsilon(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon}) + \delta]^2}{J^{\delta, \varepsilon}(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) + J^{\delta, \varepsilon}(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})} \\
& + \frac{E\left[|x_0^{\varepsilon, \delta, \varepsilon} - \gamma(x_T^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon})|^2 - |x_0^{\delta, \varepsilon} - \gamma(x_T^{\delta, \varepsilon}, y_0^{\delta, \varepsilon})|^2\right]}{J^{\delta, \varepsilon}(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) + J^{\delta, \varepsilon}(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})} \\
& + \frac{E\left[|y_T^{\varepsilon, \delta, \varepsilon} - h(x_T^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon})|^2 - |y_T^{\delta, \varepsilon} - h(x_T^{\delta, \varepsilon}, y_0^{\delta, \varepsilon})|^2\right]}{J^{\delta, \varepsilon}(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) + J^{\delta, \varepsilon}(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})} \\
& = \theta_0^{\varepsilon, \delta, \varepsilon} [\bar{J}^\varepsilon(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) - \bar{J}^\varepsilon(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})] \\
& + E\left\{\bar{\theta}_0^{\varepsilon, \delta, \varepsilon} \left(x_0^{\varepsilon, \delta, \varepsilon} - \gamma(x_T^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}) - [x_0^{\delta, \varepsilon} - \gamma(x_T^{\delta, \varepsilon}, y_0^{\delta, \varepsilon})]\right)\right\} \\
& + E\left\{\bar{\theta}_T^{\varepsilon, \delta, \varepsilon} \left(y_T^{\varepsilon, \delta, \varepsilon} - h(x_T^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}) - [y_T^{\delta, \varepsilon} - h(x_T^{\delta, \varepsilon}, y_0^{\delta, \varepsilon})]\right)\right\} \\
& = (\theta_0^{\delta, \varepsilon} + o(1)) [\bar{J}^\varepsilon(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) - \bar{J}^\varepsilon(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})] \\
& + E\left[\left(\bar{\theta}_0^{\delta, \varepsilon} + o(1)\right) \left(x_0^{\varepsilon, \delta, \varepsilon} - \gamma(x_T^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}) - [x_0^{\delta, \varepsilon} - \gamma(x_T^{\delta, \varepsilon}, y_0^{\delta, \varepsilon})]\right)\right] \\
& + E\left[\left(\bar{\theta}_T^{\delta, \varepsilon} + o(1)\right) \left(y_T^{\varepsilon, \delta, \varepsilon} - h(x_T^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}) - [y_T^{\delta, \varepsilon} - h(x_T^{\delta, \varepsilon}, y_0^{\delta, \varepsilon})]\right)\right],
\end{aligned} \tag{30}$$

with

$$\left\{ \begin{aligned} \theta_0^{\varepsilon, \delta, \varepsilon} &= \frac{2 \int_0^1 [\lambda \bar{J}^\varepsilon(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) + (1 - \lambda) \bar{J}^\varepsilon(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon}) + \delta] d\lambda}{J^{\delta, \varepsilon}(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) + J^{\delta, \varepsilon}(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})}, \\ \bar{\theta}_0^{\varepsilon, \delta, \varepsilon} &= \frac{x_0^{\varepsilon, \delta, \varepsilon} - \gamma(x_T^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}) + x_0^{\delta, \varepsilon} - \gamma(x_T^{\delta, \varepsilon}, y_0^{\delta, \varepsilon})}{J^{\delta, \varepsilon}(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) + J^{\delta, \varepsilon}(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})}, \\ \theta_T^{\varepsilon, \delta, \varepsilon} &= \frac{y_T^{\varepsilon, \delta, \varepsilon} - h(x_T^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}) + y_T^{\delta, \varepsilon} - h(x_T^{\delta, \varepsilon}, y_0^{\delta, \varepsilon})}{J^{\delta, \varepsilon}(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) + J^{\delta, \varepsilon}(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})}, \end{aligned} \right. \quad (31)$$

and

$$\left\{ \begin{aligned} \theta_0^{\delta, \varepsilon} &= \frac{\bar{J}^\varepsilon(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon}) + \delta}{J^{\delta, \varepsilon}(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})}, \\ \bar{\theta}_0^{\delta, \varepsilon} &= \frac{x_0^{\delta, \varepsilon} - \gamma(x_T^{\delta, \varepsilon}, y_0^{\delta, \varepsilon})}{J^{\delta, \varepsilon}(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})}, \\ \theta_T^{\delta, \varepsilon} &= \frac{y_T^{\delta, \varepsilon} - h(x_T^{\delta, \varepsilon}, y_0^{\delta, \varepsilon})}{J^{\delta, \varepsilon}(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})}. \end{aligned} \right. \quad (32)$$

It is necessary to point out that $\theta_0^{\delta, \varepsilon} \geq 0$, and $|\theta_0^{\delta, \varepsilon}|^2 + E|\bar{\theta}_0^{\delta, \varepsilon}|^2 + E|\theta_T^{\delta, \varepsilon}|^2 = 1$. Thus, there exists a subsequence still denoted by $(\theta_0^{\delta, \varepsilon}, \bar{\theta}_0^{\delta, \varepsilon}, \theta_T^{\delta, \varepsilon})$ convergent, i.e.,

$$\lim_{\delta \rightarrow 0} (\theta_0^{\delta, \varepsilon}, \bar{\theta}_0^{\delta, \varepsilon}, \theta_T^{\delta, \varepsilon}) = (\theta_0^\varepsilon, \bar{\theta}_0^\varepsilon, \theta_T^\varepsilon). \quad (33)$$

We claim that $\theta_0^\varepsilon \neq 0$. The detailed illustration of this point refers to [35]. Here, $(\theta_0^\varepsilon, \bar{\theta}_0^\varepsilon, \theta_T^\varepsilon)$ is called the Lagrange multiplier of the corresponding optimal 3-tuple $(\bar{x}_0^\varepsilon, \bar{y}_0^\varepsilon, \bar{u}^\varepsilon)$. On the contrary,

$$\begin{aligned} & \theta_0^{\varepsilon, \delta, \varepsilon} [\bar{J}^\varepsilon(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) - \bar{J}^\varepsilon(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})] \\ &= \theta_0^{\varepsilon, \delta, \varepsilon} [J(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) + \varepsilon^{1/3} d_{\mathcal{R}}((x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}), (\bar{x}_0^\varepsilon, \bar{y}_0^\varepsilon, \bar{u}^\varepsilon)) - J(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon}) \\ &\quad - \varepsilon^{1/3} d_{\mathcal{R}}((x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon}), (\bar{x}_0^\varepsilon, \bar{y}_0^\varepsilon, \bar{u}^\varepsilon))] \\ &= \theta_0^{\varepsilon, \delta, \varepsilon} [J(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) - J(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})] \\ &\quad + \varepsilon^{1/3} \theta_0^{\varepsilon, \delta, \varepsilon} [d_{\mathcal{R}}((x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}), (\bar{x}_0^\varepsilon, \bar{y}_0^\varepsilon, \bar{u}^\varepsilon)) - d_{\mathcal{R}}((x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon}), (\bar{x}_0^\varepsilon, \bar{y}_0^\varepsilon, \bar{u}^\varepsilon))] \\ &\leq \theta_0^{\varepsilon, \delta, \varepsilon} [J(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) - J(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})] + \varepsilon^{1/3} \theta_0^{\varepsilon, \delta, \varepsilon} d_{\mathcal{R}}((x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}), (x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})) \\ &\leq \theta_0^{\varepsilon, \delta, \varepsilon} [J(x_0^{\varepsilon, \delta, \varepsilon}, y_0^{\varepsilon, \delta, \varepsilon}, u^{\varepsilon, \delta, \varepsilon}) - J(x_0^{\delta, \varepsilon}, y_0^{\delta, \varepsilon}, u^{\delta, \varepsilon})] + \varepsilon \varepsilon^{1/3} \theta_0^{\varepsilon, \delta, \varepsilon} \sqrt{|x_0|^2 + |y_0|^2} + C. \end{aligned} \quad (34)$$

From (A₃) and (29), we can deduce

$$\begin{aligned}
& J(x_0^{\varepsilon,\delta,\varepsilon}, y_0^{\varepsilon,\delta,\varepsilon}, u^{\varepsilon,\delta,\varepsilon}) - J(x_0^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}, u^{\delta,\varepsilon}) \\
&= E \int_0^T \{l(t, \Gamma_t^{\varepsilon,\delta,\varepsilon}, u_t^{\varepsilon,\delta,\varepsilon}) - l(t, \Gamma_t^{\delta,\varepsilon}, u_t^{\delta,\varepsilon})\} dt + E \{ \varphi(x_T^{\varepsilon,\delta,\varepsilon}, y_0^{\varepsilon,\delta,\varepsilon}) - \varphi(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) \} \\
&= E \int_0^T \{l(t, \Gamma_t^{\delta,\varepsilon}, u_t^{\varepsilon,\delta,\varepsilon}) - l(t, \Gamma_t^{\delta,\varepsilon}, u_t^{\delta,\varepsilon})\} dt \\
&\quad + E \int_0^T \int_0^1 \{l_x^{\varepsilon,\delta,\varepsilon}(x_t^{\varepsilon,\delta,\varepsilon} - x_t^{\delta,\varepsilon}) + l_y^{\varepsilon,\delta,\varepsilon}(y_t^{\varepsilon,\delta,\varepsilon} - y_t^{\delta,\varepsilon}) + l_z^{\varepsilon,\delta,\varepsilon}(z_t^{\varepsilon,\delta,\varepsilon} - z_t^{\delta,\varepsilon})\} d\lambda dt \\
&\quad + E \int_0^T \int_0^1 \{l_x^{\varepsilon,\delta,\varepsilon}(Ex_t^{\varepsilon,\delta,\varepsilon} - Ex_t^{\delta,\varepsilon}) + l_y^{\varepsilon,\delta,\varepsilon}(Ey_t^{\varepsilon,\delta,\varepsilon} - Ey_t^{\delta,\varepsilon}) + l_z^{\varepsilon,\delta,\varepsilon}(Ez_t^{\varepsilon,\delta,\varepsilon} - Ez_t^{\delta,\varepsilon})\} d\lambda dt \\
&\quad + E \{ \varphi(x_T^{\varepsilon,\delta,\varepsilon}, y_0^{\varepsilon,\delta,\varepsilon}) - \varphi(x_T^{\delta,\varepsilon}, y_0^{\varepsilon,\delta,\varepsilon}) + \varphi(x_T^{\delta,\varepsilon}, y_0^{\varepsilon,\delta,\varepsilon}) - \varphi(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) \} \\
&= E \int_0^T \{l(t, \Gamma_t^{\delta,\varepsilon}, u_t^{\varepsilon,\delta,\varepsilon}) - l(t, \Gamma_t^{\delta,\varepsilon}, u_t^{\delta,\varepsilon}) + l_x^{\delta,\varepsilon} x_{1,t}^{\varepsilon,\delta,\varepsilon} + l_x^{\delta,\varepsilon} E[x_{1,t}^{\varepsilon,\delta,\varepsilon}] + l_y^{\delta,\varepsilon} y_{1,t}^{\varepsilon,\delta,\varepsilon} + l_y^{\delta,\varepsilon} E[y_{1,t}^{\varepsilon,\delta,\varepsilon}] + l_z^{\delta,\varepsilon} z_{1,t}^{\varepsilon,\delta,\varepsilon} + l_z^{\delta,\varepsilon} E[z_{1,t}^{\varepsilon,\delta,\varepsilon}]\} dt \\
&\quad + E \int_0^T \int_0^1 \{ (l_x^{\varepsilon,\delta,\varepsilon} - l_x^{\delta,\varepsilon}) x_{1,t}^{\varepsilon,\delta,\varepsilon} + (l_y^{\varepsilon,\delta,\varepsilon} - l_y^{\delta,\varepsilon}) y_{1,t}^{\varepsilon,\delta,\varepsilon} + (l_z^{\varepsilon,\delta,\varepsilon} - l_z^{\delta,\varepsilon}) z_{1,t}^{\varepsilon,\delta,\varepsilon} + (l_x^{\varepsilon,\delta,\varepsilon} - l_x^{\delta,\varepsilon}) E[x_{1,t}^{\varepsilon,\delta,\varepsilon}] + (l_y^{\varepsilon,\delta,\varepsilon} - l_y^{\delta,\varepsilon}) E[y_{1,t}^{\varepsilon,\delta,\varepsilon}] \\
&\quad + (l_z^{\varepsilon,\delta,\varepsilon} - l_z^{\delta,\varepsilon}) E[z_{1,t}^{\varepsilon,\delta,\varepsilon}] \} d\lambda dt \tag{35} \\
&\quad + E \int_0^1 \varphi_{x_T}(x_T^{\delta,\varepsilon} + \lambda(x_T^{\varepsilon,\delta,\varepsilon} - x_T^{\delta,\varepsilon}), y_0^{\varepsilon,\delta,\varepsilon})(x_T^{\varepsilon,\delta,\varepsilon} - x_T^{\delta,\varepsilon}) d\lambda \\
&\quad + E \int_0^1 \varphi_{y_0}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon} + \lambda(y_0^{\varepsilon,\delta,\varepsilon} - y_0^{\delta,\varepsilon}))(y_0^{\varepsilon,\delta,\varepsilon} - y_0^{\delta,\varepsilon}) d\lambda \\
&= E \int_0^T \{l(t, \Gamma_t^{\delta,\varepsilon}, u_t^{\varepsilon,\delta,\varepsilon}) - l(t, \Gamma_t^{\delta,\varepsilon}, u_t^{\delta,\varepsilon}) + l_x^{\delta,\varepsilon} x_{1,t}^{\varepsilon,\delta,\varepsilon} + l_x^{\delta,\varepsilon} E[x_{1,t}^{\varepsilon,\delta,\varepsilon}] + l_y^{\delta,\varepsilon} y_{1,t}^{\varepsilon,\delta,\varepsilon} + l_y^{\delta,\varepsilon} E[y_{1,t}^{\varepsilon,\delta,\varepsilon}] + l_z^{\delta,\varepsilon} z_{1,t}^{\varepsilon,\delta,\varepsilon} + l_z^{\delta,\varepsilon} E[z_{1,t}^{\varepsilon,\delta,\varepsilon}]\} dt \\
&\quad + E \{ \varphi_{x_T}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) x_{1,T}^{\varepsilon,\delta,\varepsilon} + \varphi_{y_0}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) y_{1,0}^{\varepsilon,\delta,\varepsilon} \} \\
&\quad + E \int_0^1 \{ [\varphi_{x_T}(x_T^{\delta,\varepsilon} + \lambda(x_T^{\varepsilon,\delta,\varepsilon} - x_T^{\delta,\varepsilon}), y_0^{\varepsilon,\delta,\varepsilon}) - \varphi_{x_T}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon})] x_{1,T}^{\varepsilon,\delta,\varepsilon} \} d\lambda \\
&\quad + E \int_0^1 \{ [\varphi_{y_0}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon} + \lambda(y_0^{\varepsilon,\delta,\varepsilon} - y_0^{\delta,\varepsilon})) - \varphi_{y_0}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon})] y_{1,0}^{\varepsilon,\delta,\varepsilon} \} d\lambda + o(\varepsilon) \\
&= E \int_0^T \{l(t, \Gamma_t^{\delta,\varepsilon}, u_t^{\varepsilon,\delta,\varepsilon}) - l(t, \Gamma_t^{\delta,\varepsilon}, u_t^{\delta,\varepsilon}) + l_x^{\delta,\varepsilon} x_{1,t}^{\varepsilon,\delta,\varepsilon} + l_x^{\delta,\varepsilon} E[x_{1,t}^{\varepsilon,\delta,\varepsilon}] + l_y^{\delta,\varepsilon} y_{1,t}^{\varepsilon,\delta,\varepsilon} + l_y^{\delta,\varepsilon} E[y_{1,t}^{\varepsilon,\delta,\varepsilon}] + l_z^{\delta,\varepsilon} z_{1,t}^{\varepsilon,\delta,\varepsilon} + l_z^{\delta,\varepsilon} E[z_{1,t}^{\varepsilon,\delta,\varepsilon}]\} dt \\
&\quad + E \{ \varphi_{x_T}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) x_{1,T}^{\varepsilon,\delta,\varepsilon} + \varphi_{y_0}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) y_{1,0}^{\varepsilon,\delta,\varepsilon} \} + o(\varepsilon),
\end{aligned}$$

where $l_j^{\varepsilon,\delta,\varepsilon} = l_j(t, \Gamma_t^{\delta,\varepsilon} + \varepsilon(\Gamma_t^{\varepsilon,\delta,\varepsilon} - \Gamma_t^{\delta,\varepsilon}), u_t^{\varepsilon,\delta,\varepsilon})$, $j = x, y, z$, $\tilde{x}, \tilde{y}, \tilde{z}$. Similarly,

$$\begin{aligned}
& E \left\{ \bar{\theta}_0^{\varepsilon,\delta,\varepsilon} (x_0^{\varepsilon,\delta,\varepsilon} - \gamma(x_T^{\varepsilon,\delta,\varepsilon}, y_0^{\varepsilon,\delta,\varepsilon}) - [x_0^{\delta,\varepsilon} - \gamma(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon})]) \right\} \\
&= E \left\{ \bar{\theta}_0^{\varepsilon,\delta,\varepsilon} (x_{1,0}^{\varepsilon,\delta,\varepsilon} - \gamma_{x_T}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) x_{1,T}^{\varepsilon,\delta,\varepsilon} - \gamma_{y_0}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) y_{1,0}^{\varepsilon,\delta,\varepsilon}) \right\} + o(\varepsilon), \\
& E \left\{ \bar{\theta}_T^{\varepsilon,\delta,\varepsilon} (y_T^{\varepsilon,\delta,\varepsilon} - h(x_T^{\varepsilon,\delta,\varepsilon}, y_0^{\varepsilon,\delta,\varepsilon}) - [y_T^{\delta,\varepsilon} - h(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon})]) \right\} \\
&= E \left\{ \bar{\theta}_T^{\varepsilon,\delta,\varepsilon} (y_{1,T}^{\varepsilon,\delta,\varepsilon} - h_{x_T}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) x_{1,T}^{\varepsilon,\delta,\varepsilon} - h_{y_0}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) y_{1,0}^{\varepsilon,\delta,\varepsilon}) \right\} + o(\varepsilon).
\end{aligned} \tag{36}$$

Then, taking notice of (30), we can further obtain

$$\begin{aligned}
& -\varepsilon\sqrt{|x_0|^2 + |y_0|^2} + C(\sqrt{\delta} + \varepsilon^{1/3}\theta_0^{\varepsilon,\delta,\varepsilon}) \\
& \leq \theta_0^{\varepsilon,\delta,\varepsilon} E \int_0^T \left\{ l(t, \Gamma_t^{\delta,\varepsilon}, u_t^{\varepsilon,\delta,\varepsilon}) - l(t, \Gamma_t^{\delta,\varepsilon}, u_t^{\delta,\varepsilon}) + l_x^{\delta,\varepsilon} x_{1,t}^{\varepsilon,\delta,\varepsilon} + l_x^{\delta,\varepsilon} E[x_{1,t}^{\varepsilon,\delta,\varepsilon}] + l_y^{\delta,\varepsilon} y_{1,t}^{\varepsilon,\delta,\varepsilon} + l_y^{\delta,\varepsilon} E[y_{1,t}^{\varepsilon,\delta,\varepsilon}] + l_z^{\delta,\varepsilon} z_{1,t}^{\varepsilon,\delta,\varepsilon} + l_z^{\delta,\varepsilon} E[z_{1,t}^{\varepsilon,\delta,\varepsilon}] \right\} dt \\
& + \theta_0^{\varepsilon,\delta,\varepsilon} E \left\{ \varphi_{x_T}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) x_{1,T}^{\varepsilon,\delta,\varepsilon} + \varphi_{y_0}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) y_{1,0}^{\varepsilon,\delta,\varepsilon} \right\} \\
& + E \left\{ \bar{\theta}_0^{\varepsilon,\delta,\varepsilon} (x_{1,0}^{\varepsilon,\delta,\varepsilon} - \gamma_{x_T}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) x_{1,T}^{\varepsilon,\delta,\varepsilon} - \gamma_{y_0}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) y_{1,0}^{\varepsilon,\delta,\varepsilon}) \right\} \\
& + E \left\{ \bar{\theta}_T^{\varepsilon,\delta,\varepsilon} (y_{1,T}^{\varepsilon,\delta,\varepsilon} - h_{x_T}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) x_{1,T}^{\varepsilon,\delta,\varepsilon} - h_{y_0}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon}) y_{1,0}^{\varepsilon,\delta,\varepsilon}) \right\} + o(\varepsilon).
\end{aligned} \tag{37}$$

Let us introduce the following first-order BSDEs:

$$\begin{cases}
d\bar{\zeta}_t^{\varepsilon,\delta,\varepsilon} = \left\{ f_y^{\delta,\varepsilon} \bar{\zeta}_t^{\varepsilon,\delta,\varepsilon} + l_y^{\delta,\varepsilon} + E \left[f_y^{\delta,\varepsilon} \bar{\zeta}_t^{\varepsilon,\delta,\varepsilon} + l_y^{\delta,\varepsilon} \right] \right\} dt + \left\{ f_z^{\delta,\varepsilon} \bar{\zeta}_t^{\varepsilon,\delta,\varepsilon} + l_z^{\delta,\varepsilon} + E \left[f_z^{\delta,\varepsilon} \bar{\zeta}_t^{\varepsilon,\delta,\varepsilon} + l_z^{\delta,\varepsilon} \right] \right\} dW_t, \\
-d\bar{\eta}_t^{\varepsilon,\delta,\varepsilon} = \left\{ -f_x^{\delta,\varepsilon} \bar{\zeta}_t^{\varepsilon,\delta,\varepsilon} + b_x^{\delta,\varepsilon} \bar{\eta}_t^{\varepsilon,\delta,\varepsilon} + \sigma_x^{\delta,\varepsilon} \bar{\zeta}_t^{\varepsilon,\delta,\varepsilon} - l_x^{\delta,\varepsilon} + E \left[-f_x^{\delta,\varepsilon} \bar{\zeta}_t^{\varepsilon,\delta,\varepsilon} + b_x^{\delta,\varepsilon} \bar{\eta}_t^{\varepsilon,\delta,\varepsilon} + \sigma_x^{\delta,\varepsilon} \bar{\zeta}_t^{\varepsilon,\delta,\varepsilon} - l_x^{\delta,\varepsilon} \right] \right\} dt - \bar{\zeta}_t^{\varepsilon,\delta,\varepsilon} dW_t, \\
\bar{\zeta}_T^{\varepsilon,\delta,\varepsilon} = \bar{\theta}_T^{\varepsilon,\delta,\varepsilon}, \\
\bar{\eta}_T^{\varepsilon,\delta,\varepsilon} = -\bar{\theta}_T^{\varepsilon,\delta,\varepsilon} h_{x_T}^{\delta,\varepsilon} - \bar{\theta}_0^{\varepsilon,\delta,\varepsilon} \gamma_{x_T}^{\delta,\varepsilon} + \theta_0^{\varepsilon,\delta,\varepsilon} \varphi_{x_T}^{\delta,\varepsilon},
\end{cases} \tag{38}$$

where $\rho_{x_T}^{\delta,\varepsilon} = \rho_{x_T}(x_T^{\delta,\varepsilon}, y_0^{\delta,\varepsilon})$, $\rho = h, \gamma, \varphi$. Applying Itô's formula to $\bar{\zeta}_t^{\varepsilon,\delta,\varepsilon} y_{1,t}^{\varepsilon,\delta,\varepsilon} + \bar{\eta}_t^{\varepsilon,\delta,\varepsilon} x_{1,t}^{\varepsilon,\delta,\varepsilon}$ fulfills

$$\begin{aligned}
& E \left\{ \bar{\theta}_T^{\varepsilon,\delta,\varepsilon} y_{1,T}^{\varepsilon,\delta,\varepsilon} - \left(\bar{\theta}_T^{\varepsilon,\delta,\varepsilon} h_{x_T}^{\delta,\varepsilon} + \bar{\theta}_0^{\varepsilon,\delta,\varepsilon} \gamma_{x_T}^{\delta,\varepsilon} - \theta_0^{\varepsilon,\delta,\varepsilon} \varphi_{x_T}^{\delta,\varepsilon} \right) x_{1,T}^{\varepsilon,\delta,\varepsilon} \right\} - \sqrt{\varepsilon} E \left\{ \bar{\zeta}_0^{\varepsilon,\delta,\varepsilon} y_0 + \bar{\eta}_0^{\varepsilon,\delta,\varepsilon} x_0 \right\} \\
& = E \int_0^T \left\{ \left(l_x^{\delta,\varepsilon} + E l_x^{\delta,\varepsilon} \right) x_{1,t}^{\varepsilon,\delta,\varepsilon} + \left(l_y^{\delta,\varepsilon} + E l_y^{\delta,\varepsilon} \right) y_{1,t}^{\varepsilon,\delta,\varepsilon} + \left(l_z^{\delta,\varepsilon} + E l_z^{\delta,\varepsilon} \right) z_{1,t}^{\varepsilon,\delta,\varepsilon} + \bar{\eta}_t^{\varepsilon,\delta,\varepsilon} \Delta b^{\delta,\varepsilon} + \bar{\zeta}_t^{\varepsilon,\delta,\varepsilon} \Delta \sigma^{\delta,\varepsilon} - \bar{\zeta}_t^{\varepsilon,\delta,\varepsilon} \Delta f^{\delta,\varepsilon} \right\} I_{S_{ct}} dt.
\end{aligned} \tag{39}$$

Combining (37) and (39), yields

$$\begin{aligned}
& -\varepsilon\sqrt{|x_0|^2 + |y_0|^2} + C(\sqrt{\delta} + \varepsilon^{1/3}\theta_0^{\varepsilon,\delta,\varepsilon}) \\
& \leq \sqrt{\varepsilon} E \left\{ \left(\bar{\theta}_0^{\varepsilon,\delta,\varepsilon} + \bar{\eta}_0^{\varepsilon,\delta,\varepsilon} \right) x_0 + \left(\theta_0^{\varepsilon,\delta,\varepsilon} \varphi_{y_0}^{\delta,\varepsilon} - \bar{\theta}_0^{\varepsilon,\delta,\varepsilon} \gamma_{y_0}^{\delta,\varepsilon} - \bar{\theta}_T^{\varepsilon,\delta,\varepsilon} h_{y_0}^{\delta,\varepsilon} + \bar{\zeta}_0^{\varepsilon,\delta,\varepsilon} \right) y_0 \right\} \\
& + E \int_0^T \left(\bar{\eta}_t^{\varepsilon,\delta,\varepsilon} \Delta b^{\delta,\varepsilon} + \bar{\zeta}_t^{\varepsilon,\delta,\varepsilon} \Delta \sigma^{\delta,\varepsilon} - \bar{\zeta}_t^{\varepsilon,\delta,\varepsilon} \Delta f^{\delta,\varepsilon} - \Delta l^{\delta,\varepsilon} \right) I_{S_{ct}} dt + o(\varepsilon).
\end{aligned} \tag{40}$$

To derive the first-order adjoint equation with mixed initial-terminal conditions, divide $\sqrt{\varepsilon}$ in (40) and then send $\varepsilon \rightarrow 0$, $\delta \rightarrow 0$, and we see that

$$E \left\{ \left(\bar{\theta}_0^\varepsilon + \bar{\eta}_0^\varepsilon \right) x_0 + \left(\theta_0^\varepsilon \varphi_{y_0}^\varepsilon - \bar{\theta}_0^\varepsilon \gamma_{y_0}^\varepsilon - \bar{\theta}_T^\varepsilon h_{y_0}^\varepsilon + \bar{\zeta}_0^\varepsilon \right) y_0 \right\} \geq 0, \quad \forall (x_0, y_0) \in R \times R, \tag{41}$$

which implies

$$\begin{aligned}\tilde{\eta}_0^\varepsilon &= -\bar{\theta}_0^\varepsilon, \\ \tilde{\xi}_0^\varepsilon &= -\theta_0^\varepsilon \varphi_{y_0}^\varepsilon - \tilde{\eta}_0^\varepsilon \gamma_{y_0}^\varepsilon + \tilde{\xi}_T^\varepsilon h_{y_0}^\varepsilon.\end{aligned}\quad (42)$$

Meanwhile, by taking $(x_0, y_0) = (0, 0)$ in (40), dividing this inequality by ε , and then sending $\varepsilon \rightarrow 0$, $\delta \rightarrow 0$, the variational inequality follows:

$$\begin{aligned}-C\varepsilon^{1/3}\theta_0^\varepsilon \leq E \int_0^T & (\tilde{\eta}_t^\varepsilon [b(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, u_t) - b(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, \tilde{u}_t^\varepsilon)] + \tilde{\zeta}_t^\varepsilon [\sigma(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, u_t) - \sigma(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, \tilde{u}_t^\varepsilon)] \\ & - \tilde{\xi}_t^\varepsilon [f(t, \tilde{\Gamma}_t^\varepsilon, u_t) - f(t, \tilde{\Gamma}_t^\varepsilon, \tilde{u}_t^\varepsilon)] - [l(t, \tilde{\Gamma}_t^\varepsilon, u_t) - l(t, \tilde{\Gamma}_t^\varepsilon, \tilde{u}_t^\varepsilon)] dt,\end{aligned}\quad (43)$$

where $\tilde{\Gamma}_t^\varepsilon = (\tilde{x}_t^\varepsilon, \tilde{y}_t^\varepsilon, \tilde{z}_t^\varepsilon, E\tilde{x}_t^\varepsilon, E\tilde{y}_t^\varepsilon, E\tilde{z}_t^\varepsilon)$. On the contrary, from (38) and (42), we can present the adjoint equation with mixed initial-terminal conditions as follows:

$$\begin{cases} d\tilde{\xi}_t^\varepsilon = \left\{ f_y^\varepsilon \tilde{\xi}_t^\varepsilon + l_y^\varepsilon + E \left[f_y^\varepsilon \tilde{\xi}_t^\varepsilon + l_y^\varepsilon \right] \right\} dt + \left\{ f_z^\varepsilon \tilde{\xi}_t^\varepsilon + l_z^\varepsilon + E \left[f_z^\varepsilon \tilde{\xi}_t^\varepsilon + l_z^\varepsilon \right] \right\} dW_t, \\ -d\tilde{\eta}_t^\varepsilon = \left\{ -f_x^\varepsilon \tilde{\xi}_t^\varepsilon + b_x^\varepsilon \tilde{\eta}_t^\varepsilon + \sigma_x^\varepsilon \tilde{\xi}_t^\varepsilon - l_x^\varepsilon + E \left[-f_x^\varepsilon \tilde{\xi}_t^\varepsilon + b_x^\varepsilon \tilde{\eta}_t^\varepsilon + \sigma_x^\varepsilon \tilde{\xi}_t^\varepsilon - l_x^\varepsilon \right] \right\} dt - \tilde{\zeta}_t^\varepsilon dW_t, \\ \tilde{\xi}_0^\varepsilon = -\theta_0^\varepsilon \varphi_{y_0}^\varepsilon - \tilde{\eta}_0^\varepsilon \gamma_{y_0}^\varepsilon + \tilde{\xi}_T^\varepsilon h_{y_0}^\varepsilon, \\ \tilde{\eta}_T^\varepsilon = -\tilde{\xi}_T^\varepsilon h_{x_T}^\varepsilon + \tilde{\eta}_0^\varepsilon \gamma_{x_T}^\varepsilon + \theta_0^\varepsilon \varphi_{x_T}^\varepsilon,\end{cases}\quad (44)$$

whose well-posedness can be deduced directly virus the above derivative process (Remark 2). Now, consider (43) again but with $(\tilde{x}^\varepsilon, \tilde{y}^\varepsilon, \tilde{z}^\varepsilon, \tilde{u}^\varepsilon)$, etc., replaced by $(x^\varepsilon, y^\varepsilon, z^\varepsilon, u^\varepsilon)$,

etc. We are about to derive an estimate for the term similar to the right side of (43) with respect to $(x^\varepsilon, y^\varepsilon, z^\varepsilon, u^\varepsilon)$, etc. To this end, we first estimate the following difference:

$$\begin{aligned}E \int_0^T & \{ \tilde{\eta}_t^\varepsilon [b(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, u_t) - b(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, \tilde{u}_t^\varepsilon)] - \eta_t^\varepsilon [b(t, x_t^\varepsilon, Ex_t^\varepsilon, u_t) - b(t, x_t^\varepsilon, Ex_t^\varepsilon, u_t^\varepsilon)] \} dt \\ & = E \int_0^T (\tilde{\eta}_t^\varepsilon - \eta_t^\varepsilon) [b(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, u_t) - b(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, \tilde{u}_t^\varepsilon)] dt \\ & \quad + E \int_0^T \eta_t^\varepsilon [b(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, u_t) - b(t, x_t^\varepsilon, Ex_t^\varepsilon, u_t)] dt - E \int_0^T \eta_t^\varepsilon [b(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, \tilde{u}_t^\varepsilon) - b(t, x_t^\varepsilon, Ex_t^\varepsilon, u_t^\varepsilon)] dt \\ & = \Xi_1 + \Xi_2 + \Xi_3.\end{aligned}\quad (45)$$

Due to Lemma 2, for any $\kappa \in [0, (1/3))$, by using the similar arguments as developed in [7] the proof of Theorem 1, we can also prove that

$$\begin{aligned}\Xi_1 &= E \int_0^T (\tilde{\eta}_t^\varepsilon - \eta_t^\varepsilon) [b(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, u_t) - b(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, \tilde{u}_t^\varepsilon)] dt \leq C\varepsilon^\kappa, \\ \Xi_2 &= E \int_0^T \eta_t^\varepsilon [b(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, u_t) - b(t, x_t^\varepsilon, Ex_t^\varepsilon, u_t)] dt \leq C\varepsilon^\kappa, \\ \Xi_3 &= E \int_0^T \eta_t^\varepsilon [b(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, \tilde{u}_t^\varepsilon) - b(t, x_t^\varepsilon, Ex_t^\varepsilon, u_t^\varepsilon)] dt \leq C\varepsilon^\kappa,\end{aligned}\quad (46)$$

i.e.,

$$E \int_0^T \{ \tilde{\eta}_t^\varepsilon [b(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, u_t) - b(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, \tilde{u}_t^\varepsilon)] - \eta_t^\varepsilon [b(t, x_t^\varepsilon, Ex_t^\varepsilon, u_t) - b(t, x_t^\varepsilon, Ex_t^\varepsilon, u_t^\varepsilon)] \} dt \leq C\varepsilon^\kappa. \quad (47)$$

Similarly, via Lemma 2, we also have

$$\begin{aligned} E \int_0^T & \{ \tilde{\zeta}_t^\varepsilon [\sigma(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, u_t) - \sigma(t, \tilde{x}_t^\varepsilon, E\tilde{x}_t^\varepsilon, \tilde{u}_t^\varepsilon)] - \zeta_t^\varepsilon [\sigma(t, x_t^\varepsilon, Ex_t^\varepsilon, u_t) - \sigma(t, x_t^\varepsilon, Ex_t^\varepsilon, u_t^\varepsilon)] \\ & - \tilde{\xi}_t^\varepsilon [f(t, \tilde{\Gamma}_t^\varepsilon, u_t) - f(t, \tilde{\Gamma}_t^\varepsilon, \tilde{u}_t^\varepsilon)] + \xi_t^\varepsilon [f(t, \Gamma_t^\varepsilon, u_t) - f(t, \Gamma_t^\varepsilon, u_t^\varepsilon)] \\ & - [l(t, \tilde{\Gamma}_t^\varepsilon, u_t) - l(t, \tilde{\Gamma}_t^\varepsilon, \tilde{u}_t^\varepsilon)] + [l(t, \Gamma_t^\varepsilon, u_t) - l(t, \Gamma_t^\varepsilon, u_t^\varepsilon)] \} dt \leq C\varepsilon^\kappa. \end{aligned} \quad (48)$$

Therefore, the desired result (20) follows immediately by combining (43)–(48).

Since (x_0, y_0, u) is arbitrary, we draw the desired conclusion and summarize it as follows. \square

Theorem 2. *Suppose $(A_1) - (A_3)$ hold. For any $\varepsilon > 0$, u^ε is an ε -optimal control of problem \mathcal{A} . Then, for any $\kappa \in [0, (1/3))$, there exist three nonnegative parameters $\theta_0^\varepsilon, \tilde{\theta}_0^\varepsilon$, and $\tilde{\theta}_T^\varepsilon$ with $|\theta_0^\varepsilon|^2 + E|\tilde{\theta}_0^\varepsilon|^2 + E|\tilde{\theta}_T^\varepsilon|^2 = 1$ and $\theta_0^\varepsilon \geq 0$ such that, for any $x_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}$, and $u \in \mathcal{U}$, the necessary condition (20) holds a.e. a.s., where $(\xi^\varepsilon, \eta^\varepsilon, \zeta^\varepsilon)$ is the solution of (15) corresponding to u^ε .*

Define the Hamiltonian $H(t, x, y, z, u, \xi, \eta, \zeta)$ by

$$\begin{aligned} H(t, x, y, z, u, \xi, \eta, \zeta) &= -\xi f(t, \Gamma, u) + \eta b(t, x, Ex, u) \\ &\quad + \zeta \sigma(t, x, Ex, u) - l(t, \Gamma, u), \end{aligned} \quad (49)$$

then we have the following form of necessity conditions.

Corollary 1. *Under the assumptions of Theorem 2, it holds that*

$$\begin{aligned} E \int_0^T & H(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, u_t^\varepsilon, \xi_t^\varepsilon, \eta_t^\varepsilon, \zeta_t^\varepsilon) dt \\ & \geq \sup_{u \in \mathcal{U}} E \int_0^T H(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, u_t, \xi_t^\varepsilon, \eta_t^\varepsilon, \zeta_t^\varepsilon) dt - C\theta_0^\varepsilon \varepsilon^\kappa. \end{aligned} \quad (50)$$

Proof. According to the definition of the control $u^{\varepsilon, \delta, \varepsilon}$, the point $u \in U$ can be replaced by any admissible control $u \in \mathcal{U}$, and the subsequent arguments still go through. Therefore, the conclusion in Theorem 2 holds for any $u \in \mathcal{U}$, which is an easy variant of our corollary. \square

Remark 3. If the coefficients of system (3) do not depend on the expected values of the states, Theorem 2 reduces to the near-maximum condition for the classical system under convex control domain.

Remark 4. For exact optimality, the integral form and the pointwise form of the maximum condition is equivalent; however, it is not the case for near-optimality. We can only deduce the near maximum condition in an integral form.

Remark 5. If $\varepsilon = 0$, we can obtain a stochastic maximum principle for controlled mean-field FBSDEs with the control domain of convexity assumption.

4. Sufficient Condition of Near-Optimality

In this section, we will prove that the near-maximum condition of the Hamiltonian H in the integral form is sufficient for near-optimality under some additional assumptions.

(A_4) Let b, σ, f , and l be differentiable in u , and there exists a constant $C > 0$ such that

$$\begin{cases} |b(t, x, \tilde{x}, u_1) - b(t, x, \tilde{x}, u_2)| + |b_u(t, x, \tilde{x}, u_1) - b_u(t, x, \tilde{x}, u_2)| + |\sigma(t, x, \tilde{x}, u_1) - \sigma(t, x, \tilde{x}, u_2)| + |\sigma_u(t, x, \tilde{x}, u_1) - \sigma_u(t, x, \tilde{x}, u_2)| \leq C|u_1 - u_2|, \\ |f(t, \Gamma, u_1) - f(t, \Gamma, u_2)| + |f_u(t, \Gamma, u_1) - f_u(t, \Gamma, u_2)| + |l(t, \Gamma, u_1) - l(t, \Gamma, u_2)| + |l_u(t, \Gamma, u_1) - l_u(t, \Gamma, u_2)| \leq C|u_1 - u_2|. \end{cases} \quad (51)$$

Theorem 3. *Let $(A_1) - (A_4)$ hold, and assume that the Hamiltonian $H(t, \cdot, \cdot, \cdot, \cdot, \xi_t^\varepsilon, \eta_t^\varepsilon, \zeta_t^\varepsilon)$ is concave for a.e. $t \in [0, T]$, P-a.s., h is concave, and φ and γ are convex. Let*

$\Lambda_t^\varepsilon = (\xi^\varepsilon, \eta^\varepsilon, \zeta^\varepsilon)$ be the solution of the adjoint equation (15) associated with $(x^\varepsilon, y^\varepsilon, z^\varepsilon, u^\varepsilon)$. If, for any $u \in \mathcal{U}$ and some $\varepsilon > 0$,

$$\begin{aligned} & \sup_{u \in \mathcal{U}} E \int_0^T H(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, u_t, \Lambda_t^\varepsilon) dt \\ & \leq E \int_0^T H(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, u_t^\varepsilon, \Lambda_t^\varepsilon) dt + \varepsilon, \end{aligned} \quad (52)$$

holds, then we have

$$J(u^\varepsilon) \leq \inf_{u \in \mathcal{U}} J(u) + C\varepsilon^{1/2}, \quad (53)$$

where $C > 0$ is a constant independent of ε .

Proof. Fix $\varepsilon > 0$, for any $u, v \in \mathcal{U}$, and define a new metric \bar{d} on \mathcal{U} as follows:

$$\bar{d}(u, v) = E \int_0^T \nu_t^\varepsilon |u_t - v_t| dt, \quad (54)$$

with $\nu_t^\varepsilon = 1 + |\xi_t^\varepsilon| + |\eta_t^\varepsilon| + |\zeta_t^\varepsilon|$. Obviously, \bar{d} is a complete metric on \mathcal{U} as a weighted L^1 norm. Define a functional \bar{J} on \mathcal{U} by

$$\bar{J}(u) = E \int_0^T H(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, u_t, \Lambda_t^\varepsilon) dt. \quad (55)$$

A simple calculation shows that

$$|\bar{J}(u) - \bar{J}(v)| \leq CE \int_0^T \nu_t^\varepsilon |u_t - v_t| dt. \quad (56)$$

Therefore, \bar{J} is continuous on \mathcal{U} with respect to \bar{d} . Then, by using (52) and Ekeland's variational principle, there exists a $\tilde{u}^\varepsilon \in \mathcal{U}$ such that

$$\bar{d}(\tilde{u}^\varepsilon, u^\varepsilon) \leq \varepsilon^{1/2}, \quad (57)$$

and

$$E \int_0^T \tilde{H}(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, \tilde{u}_t^\varepsilon) dt = \max_{u \in \mathcal{U}} E \int_0^T \tilde{H}(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, u_t) dt, \quad (58)$$

where

$$\tilde{H}(t, x, y, z, u) = H(t, x, y, z, u, \Lambda) - \varepsilon^{1/2} \nu_t^\varepsilon |u - \tilde{u}^\varepsilon|. \quad (59)$$

The integral-form maximum condition (58) implies a pointwise maximum condition, that is, for a.e. $t \in [0, T]$ and P -a.s.,

$$\tilde{H}(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, \tilde{u}_t^\varepsilon) = \max_{u \in \mathcal{U}} \tilde{H}(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, u_t). \quad (60)$$

Then, by Lemma 2.3 of Yong and Zhou [36], we have $0 \in \partial_u \tilde{H}(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, \tilde{u}_t^\varepsilon)$. By using (59) and the fact that the

generalized gradient of the sum of two functions is contained in the sum of the generalized gradients of the two functions, we deduce

$$\begin{aligned} \partial_u \tilde{H}(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, \tilde{u}_t^\varepsilon) &= \partial_u H(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, \tilde{u}_t^\varepsilon, \Lambda_t^\varepsilon) \\ &+ [-\varepsilon^{1/2} \nu_t^\varepsilon, \varepsilon^{1/2} \nu_t^\varepsilon]. \end{aligned} \quad (61)$$

Furthermore, since H is differentiable in u , there exists a $\vartheta_t^\varepsilon \in [-\varepsilon^{1/2} \nu_t^\varepsilon, \varepsilon^{1/2} \nu_t^\varepsilon]$, such that

$$H_u(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, \tilde{u}_t^\varepsilon, \Lambda_t^\varepsilon) = -\vartheta_t^\varepsilon. \quad (62)$$

Consequently, by (A₄), we can prove that

$$\begin{aligned} & |H_u(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, u_t^\varepsilon, \Lambda_t^\varepsilon)| \\ & \leq |H_u(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, u_t^\varepsilon, \Lambda_t^\varepsilon) - H_u(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, \tilde{u}_t^\varepsilon, \Lambda_t^\varepsilon)| \\ & + |H_u(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, \tilde{u}_t^\varepsilon, \Lambda_t^\varepsilon)| \\ & \leq C \nu_t^\varepsilon |u_t^\varepsilon - \tilde{u}_t^\varepsilon| + \varepsilon^{1/2} \nu_t^\varepsilon. \end{aligned} \quad (63)$$

By the concavity of $H(t, \cdot, \cdot, \cdot, \cdot, \Lambda_t^\varepsilon)$, we have

$$\begin{aligned} & H(t, x_t, y_t, z_t, u_t, \Lambda_t^\varepsilon) - H(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, u_t^\varepsilon, \Lambda_t^\varepsilon) \\ & \leq H_x^\varepsilon(E x_t - E x_t^\varepsilon) + H_y^\varepsilon(E y_t - E y_t^\varepsilon) + H_z^\varepsilon(E z_t - E z_t^\varepsilon) \\ & + H_x^\varepsilon(x_t - x_t^\varepsilon) + H_y^\varepsilon(y_t - y_t^\varepsilon) + H_z^\varepsilon(z_t - z_t^\varepsilon) + H_u^\varepsilon(u_t - u_t^\varepsilon), \end{aligned} \quad (64)$$

for any (x, y, z, u) , where $H_j^\varepsilon = H_j(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, u_t^\varepsilon, \Lambda_t^\varepsilon)$, $j = x, y, z, \tilde{x}, \tilde{y}, \tilde{z}$.

Taking integrations, from (57) and (63), follows

$$\begin{aligned} & E \int_0^T [H(t, x_t, y_t, z_t, u_t, \Lambda_t^\varepsilon) - H(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, u_t^\varepsilon, \Lambda_t^\varepsilon)] dt \\ & \leq E \int_0^T H_x^\varepsilon(E x_t - E x_t^\varepsilon) dt + E \int_0^T H_y^\varepsilon(E y_t - E y_t^\varepsilon) dt \\ & + E \int_0^T H_z^\varepsilon(E z_t - E z_t^\varepsilon) dt + E \int_0^T H_x^\varepsilon(x_t - x_t^\varepsilon) dt \\ & + E \int_0^T H_y^\varepsilon(y_t - y_t^\varepsilon) dt + E \int_0^T H_z^\varepsilon(z_t - z_t^\varepsilon) dt + C\varepsilon^{1/2}. \end{aligned} \quad (65)$$

Applying Itô's formula to $\xi_t^\varepsilon(y_t - y_t^\varepsilon) + \eta_t^\varepsilon(x_t - x_t^\varepsilon)$, we obtain

$$\begin{aligned}
 & E[\xi_T^\varepsilon (h(x_T, y_0) - h(x_T^\varepsilon, y_0^\varepsilon))] + E[[-\xi_T^\varepsilon h_{x_T}^\varepsilon + \eta_0^\varepsilon \gamma_{x_T}^\varepsilon + \theta_0^\varepsilon \varphi_{x_T}^\varepsilon](x_T - x_T^\varepsilon)] \\
 & - E[[-\theta_0^\varepsilon \varphi_{y_0}^\varepsilon - \eta_0^\varepsilon \gamma_{y_0}^\varepsilon + \xi_T^\varepsilon h_{y_0}^\varepsilon](y_0 - y_0^\varepsilon)] - E[\eta_0^\varepsilon (\gamma(x_T, y_0) - \gamma(x_T^\varepsilon, y_0^\varepsilon))] \\
 & = E \int_0^T [H(t, x_t, y_t, z_t, u_t, \Lambda_t^\varepsilon) - H(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, u_t^\varepsilon, \Lambda_t^\varepsilon) + l(t, \Gamma_t, u_t) - l(t, \Gamma_t^\varepsilon, u_t^\varepsilon)] dt \\
 & - E \int_0^T [EH_{x_t}^\varepsilon](x_t - x_t^\varepsilon) dt - E \int_0^T [EH_{y_t}^\varepsilon](y_t - y_t^\varepsilon) dt \\
 & - E \int_0^T [EH_{z_t}^\varepsilon](z_t - z_t^\varepsilon) dt - E \int_0^T H_{x_t}^\varepsilon (x_t - x_t^\varepsilon) dt \\
 & - E \int_0^T H_{y_t}^\varepsilon (y_t - y_t^\varepsilon) dt - E \int_0^T H_{z_t}^\varepsilon (z_t - z_t^\varepsilon) dt.
 \end{aligned} \tag{66}$$

Since h is concave, we have

$$\begin{aligned}
 & E\{\xi_T^\varepsilon (h(x_T, y_0) - h(x_T^\varepsilon, y_0^\varepsilon)) - \xi_T^\varepsilon h_{x_T}^\varepsilon (x_T - x_T^\varepsilon) \\
 & - \xi_T^\varepsilon h_{y_0}^\varepsilon (y_0 - y_0^\varepsilon)\} \geq 0.
 \end{aligned} \tag{67}$$

Similarly, by the convexity of φ and γ yielding

$$E\{\varphi_{x_T}^\varepsilon (x_T - x_T^\varepsilon) + \varphi_{y_0}^\varepsilon (y_0 - y_0^\varepsilon)\} \leq E[\varphi(x_T, y_0) - \varphi(x_T^\varepsilon, y_0^\varepsilon)], \tag{68}$$

$$E\{\gamma_{x_T}^\varepsilon (x_T - x_T^\varepsilon) + \gamma_{y_0}^\varepsilon (y_0 - y_0^\varepsilon)\} \leq E[\gamma(x_T, y_0) - \gamma(x_T^\varepsilon, y_0^\varepsilon)]. \tag{69}$$

Combining (65) and (67)–(69), we can finally get the following result:

$$J(u^\varepsilon) \leq J(u) + C\varepsilon^{1/2}. \tag{70}$$

Since u is arbitrary, we can draw the conclusion. \square

Corollary 2. Under the assumptions of Theorem 3, a sufficient condition for an admissible control u^ε to be ε -optimal is

$$\begin{aligned}
 & \sup_{u \in \mathcal{U}} E \int_0^T H(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, u, \Lambda_t^\varepsilon) dt \\
 & \leq E \int_0^T H(t, x_t^\varepsilon, y_t^\varepsilon, z_t^\varepsilon, u_t^\varepsilon, \Lambda_t^\varepsilon) dt + \left(\frac{\varepsilon}{C}\right)^2.
 \end{aligned} \tag{71}$$

5. A Linear-Quadratic Problem

Consider the near-optimal control problem.

Problem \mathcal{P}^ε . Minimize $J^\varepsilon(u) = E[\int_0^1 ((\sqrt{2}/2)u_t^2 + (\sqrt{2}\varepsilon/2)u_t^2)dt + (\sqrt{2}/2)x_1^2 + (\sqrt{2}/2)y_0^2]$, subject to

$$\begin{cases} dx_t = \{x_t + Ex_t + u_t\}dt + \{x_t + Ex_t + u_t\}dW_t, \\ -dy_t = \{x_t + y_t + z_t + Ex_t + Ey_t + Ez_t + u_t\}dt - z_t dW_t, \\ x_0 = x_1 - \frac{1}{2}y_0, \\ y_1 = \frac{1}{2}x_1 + \frac{1}{2}y_0, \end{cases} \tag{72}$$

where $\varepsilon > 0$ is a small parameter, and the control domain is limited as $\mathcal{U} = [-1, 1]$.

In view of the practical point, it is difficult to find a near-optimal control for \mathcal{P}^ε directly. So, we adopt the approach, in which \mathcal{P}^ε is firstly approximated by a simpler one (Problem \mathcal{P}), and then an optimal control u^* for \mathcal{P} is obtained. Finally, it is proved that u^* is near-optimal for \mathcal{P}^ε . This method has been applied to the hierarchical controls of stochastic manufacturing systems [37].

Problem \mathcal{P} . Minimize $J(u) = E[\int_0^1 ((\sqrt{2}/2)u_t^2 dt + (\sqrt{2}/2)x_1^2 + (\sqrt{2}/2)y_0^2)]$, subject to

$$\begin{cases} dx_t = \{x_t + Ex_t + u_t\}dt + \{x_t + Ex_t + u_t\}dW_t, \\ -dy_t = \{x_t + y_t + z_t + Ex_t + Ey_t + Ez_t + u_t\}dt - z_t dW_t, \\ x_0 = x_1 - \frac{1}{2}y_0, \\ y_1 = \frac{1}{2}x_1 + \frac{1}{2}y_0. \end{cases} \tag{73}$$

Let u be an optimal admissible control of \mathcal{P} , and the corresponding optimal trajectory is denoted by (x, y, z) . Set $\theta_0 = (\sqrt{2}/2)$, for a given admissible triple (x, y, z, u) , the corresponding first-order adjoint equation is presented as

$$\left\{ \begin{array}{l} d\xi_t = \{\xi_t + E\xi_t\}dt + \{\xi_t + E\xi_t\}dW_t, \\ -d\eta_t = \{-\xi_t + \eta_t + \zeta_t + E[-\xi_t + \eta_t + \zeta_t]\}dt - \zeta_t dW_t, \\ \xi_0 = -\gamma_0 + \frac{1}{2}\eta_0 + \frac{1}{2}\xi_1, \\ \eta_1 = -\frac{1}{2}\xi_1 + \eta_0 + x_1. \end{array} \right. \quad (74)$$

From Remark 5 and Theorem 2, the candidate optimal control u should satisfy

$$\frac{\sqrt{2}}{2}v_t^2 - (\eta_t + \zeta_t - \xi_t)v_t \leq \frac{\sqrt{2}}{2}u_t^2 - (\eta_t + \zeta_t - \xi_t)u_t, \quad \forall v \in \mathcal{U}. \quad (75)$$

Then, we have

$$u_t = \begin{cases} \mu_t, & \mu_t \in [-1, 1], \\ 1, & \mu_t \in (-\infty, -1], \\ -1, & \mu_t \in [1, +\infty), \end{cases} \quad (76)$$

with $\mu_t = (\sqrt{2}/2)(\eta_t + \zeta_t - \xi_t)$.

We claim that the control u defined by (76) is optimal for Problem \mathcal{P} , which will be illustrated in the following proposition. Now, we are about to show that the same optimal control is near-optimal for \mathcal{P}^ε when ε is sufficiently small. Denote by (x^*, y^*, z^*, u^*) the optimal state and optimal control under (76) and (ξ^*, η^*, ζ^*) the

corresponding solution of (74). Then, the Hamiltonian function for \mathcal{P} is

$$\begin{aligned} H(t, x_t, y_t, z_t, u_t, \Lambda_t) = & -\frac{\sqrt{2}}{2}u_t^2 - (\xi_t - \eta_t - \zeta_t)u_t \\ & + (\eta_t + \zeta_t)(x_t + Ex_t) \\ & - \xi_t(x_t + y_t + z_t + Ex_t + Ey_t + Ez_t). \end{aligned} \quad (77)$$

Since u^* is optimal, it necessarily maximizes the Hamiltonian function a.s., namely,

$$u_t^* + \frac{\sqrt{2}}{2}(\xi_t^* - \eta_t^* - \zeta_t^*) = 0, \quad P - \text{a.s. a.e.} \quad (78)$$

However, the Hamiltonian function for \mathcal{P}^ε is

$$\begin{aligned} H_\varepsilon(t, x_t, y_t, z_t, u_t, \Lambda_t) = & -\frac{\sqrt{2}}{2}(1 + \varepsilon)u_t^2 - (\xi_t - \eta_t - \zeta_t)u_t \\ & + (\eta_t + \zeta_t)(x_t + Ex_t) \\ & - \xi_t(x_t + y_t + z_t + Ex_t + Ey_t + Ez_t). \end{aligned} \quad (79)$$

Obviously, it is concave. Moreover, it is maximized at u^ε , which satisfies

$$u_t^\varepsilon = \frac{u_t^*}{1 + \varepsilon}. \quad (80)$$

Hence,

$$\begin{aligned} & \sup_{u \in \mathcal{U}} E \int_0^1 H_\varepsilon(t, x_t^*, y_t^*, z_t^*, u_t, \Lambda_t^*) dt - E \int_0^1 H_\varepsilon(t, x_t^*, y_t^*, z_t^*, u_t^*, \Lambda_t^*) dt \\ & \leq E \int_0^1 \sup_{u \in \mathcal{U}} H_\varepsilon(t, x_t^*, y_t^*, z_t^*, u_t, \Lambda_t^*) dt - E \int_0^1 H_\varepsilon(t, x_t^*, y_t^*, z_t^*, u_t^*, \Lambda_t^*) dt \\ & = E \int_0^1 \{H_\varepsilon(t, x_t^*, y_t^*, z_t^*, u_t^\varepsilon, \Lambda_t^*) - H_\varepsilon(t, x_t^*, y_t^*, z_t^*, u_t^*, \Lambda_t^*)\} dt \\ & = E \int_0^1 \frac{\varepsilon^2}{2(1 + \varepsilon)} (u_t^*)^2 dt \\ & \leq C\varepsilon. \end{aligned} \quad (81)$$

According to Theorem 3, u^* is near-optimal for \mathcal{P}^ε with an error order of ε when ε is sufficiently small.

Proposition 1. *The control u^* defined by (76) together with the corresponding trajectory (x^*, y^*, z^*) is an optimal solution for Problem \mathcal{P} .*

Proof. Suppose (x, y, z) is the trajectory of system (73) controlled by $u \in \mathcal{U}$. By the convexity of a function, we have

$$\begin{aligned} \frac{\sqrt{2}}{2}x_1^2 - \frac{\sqrt{2}}{2}(x_1^*)^2 & \geq \sqrt{2}x_1^*(x_1 - x_1^*), \\ \frac{\sqrt{2}}{2}y_0^2 - \frac{\sqrt{2}}{2}(y_0^*)^2 & \geq \sqrt{2}y_0^*(y_0 - y_0^*). \end{aligned} \quad (82)$$

Applying Itô's formula to $\xi_t^*(y_t - y_t^*) + \eta_t^*(x_t - x_t^*)$, we obtain

$$E[x_1^*(x_1 - x_1^*) + y_0^*(y_0 - y_0^*)] = E \int_0^1 (\xi_t^* - \eta_t^* - \zeta_t^*)(u_t - u_t^*) dt. \quad (83)$$

Then,

$$\begin{aligned} J(u) - J(u^*) &\geq E \left\{ \int_0^1 \left[\frac{\sqrt{2}}{2} u_t^2 - \frac{\sqrt{2}}{2} (u_t^*)^2 \right] dt + \sqrt{2} x_1^* (x_1 - x_1^*) + \sqrt{2} y_0^* (y_0 - y_0^*) \right\} \\ &= E \int_0^1 \left[\frac{\sqrt{2}}{2} u_t^2 - \frac{\sqrt{2}}{2} (u_t^*)^2 + \sqrt{2} (\xi_t^* - \eta_t^* - \zeta_t^*) (u_t - u_t^*) \right] dt. \end{aligned} \quad (84)$$

From (76), we have

$$\begin{aligned} \frac{\sqrt{2}}{2} u_t^2 + \sqrt{2} (\xi_t^* - \eta_t^* - \zeta_t^*) u_t &\geq \frac{\sqrt{2}}{2} (u_t^*)^2 \\ + \sqrt{2} (\xi_t^* - \eta_t^* - \zeta_t^*) u_t^*, \quad \forall u \in \mathcal{U}. \end{aligned} \quad (85)$$

The inequality above implies that

$$J(u) - J(u^*) \geq 0, \quad \forall u \in \mathcal{U}. \quad (86)$$

Therefore, u^* is the optimal control for \mathcal{P} , and (x^*, y^*, z^*) is the optimal trajectory. \square

6. Conclusion

This paper discussed near-optimal control problems for mean-field FBSDEs with mixed initial-terminal conditions. Firstly, we initially introduce three first-order adjoint equations to operate dual analysis with corresponding variational processes. Secondly, the reduction method is adopted to guarantee the well-posedness of the first-order adjoint equations also with mixed initial-terminal conditions. Furthermore, by introducing the penalty functional, the original problem with endpoint constraints is transformed as penalized optimal control problem with no endpoint constraints. Via convex variational technique as well as Ekeland's variational principle, the necessary condition of Pontryagin's type is established. Finally, to illustrate the application of our theoretical results, a linear-quadratic problem is worked out. In our future work, we will develop the research on near-optimality to solve both theoretical and practical problems.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant no. 11901112), and National Natural Science Foundation of China (Grant no. 11701161), and High level introduction of talent research start-up fund of Hubei University of Technology (Grant no. BSQD14067).

References

- [1] I. Ekeland, "Non convex minimization problems," *Bulletin of the American Mathematical Society*, vol. 1, no. 3, pp. 443–475, 1979.
- [2] X. Y. Zhou, "Deterministic near-optimal control, part I: necessary and sufficient conditions for near-optimality," *Journal of Optimization Theory and Applications*, vol. 85, no. 2, pp. 473–488, 1995.
- [3] X. Y. Zhou, "Deterministic near-optimal controls. Part II: dynamic programming and viscosity solution approach," *Mathematics of Operations Research*, vol. 21, no. 3, pp. 655–674, 1996.
- [4] X. Y. Zhou, "Stochastic near-optimal controls: necessary and sufficient conditions for near-optimality," *SIAM Journal on Control and Optimization*, vol. 36, no. 3, pp. 929–947, 1998.
- [5] K. Bahlali, N. Khelfallah, and B. Mezerdi, "Necessary and sufficient conditions for near-optimality in stochastic control of FBSDEs," *Systems & Control Letters*, vol. 58, no. 12, pp. 857–864, 2009.
- [6] R. Guo, "Projective synchronization of a class of chaotic systems by dynamic feedback control method," *Nonlinear Dynamics*, vol. 90, no. 1, pp. 53–64, 2017.
- [7] M. Hafayed, A. Abba, and S. Boukaf, "On Zhou's maximum principle for near-optimal control of mean-field forward-backward stochastic systems with jumps and its applications," *International Journal of Modelling, Identification and Control*, vol. 25, no. 1, pp. 1–16, 2016.
- [8] M. Hafayed and S. Abbas, "On near-optimal mean-field stochastic singular controls: necessary and sufficient conditions for near-optimality," *Journal of Optimization Theory and Applications*, vol. 160, no. 3, pp. 778–808, 2014.
- [9] J. Huang, X. Li, and G. Wang, "Near-optimal control problems for linear forward-backward stochastic systems," *Automatica*, vol. 46, no. 2, pp. 397–404, 2010.
- [10] E. Hui, J. Huang, X. Li, and G. Wang, "Near-optimal control for stochastic recursive problems," *Systems & Control Letters*, vol. 60, no. 3, pp. 161–168, 2011.
- [11] M. Tang, "Stochastic maximum principle of near-optimal control of fully coupled forward-backward stochastic differential equation," *Abstract and Applied Analysis*, vol. 2014, Article ID 361259, 12 pages, 2014.
- [12] S. Yang, "The necessary and sufficient conditions for stochastic differential systems with multi-time state cost functional," *Systems & Control Letters*, vol. 114, pp. 11–18, 2018.
- [13] X. Yi, R. Guo, and Y. Qi, "Stabilization of chaotic systems with both uncertainty and disturbance by the UDE-based control method," *IEEE Access*, vol. 8, no. 1, pp. 62471–62477, 2020.
- [14] L. Zhang, J. Huang, and X. Li, "Necessary condition for near optimal control of linear forward-backward stochastic differential equations," *International Journal of Control*, vol. 88, no. 8, pp. 1594–1608, 2015.

- [15] L. Zhang, "Sufficient condition for near-optimal control of general controlled linear forward-backward stochastic differential equations," *International Journal of Dynamics and Control*, vol. 5, no. 2, pp. 306–313, 2017.
- [16] L. Zhang and Q. Zhou, "Near-optimal control of stochastic recursive systems via viscosity solution," *Journal of Optimization Theory and Applications*, vol. 178, no. 2, pp. 363–382, 2018.
- [17] F. Zhang, "Maximum principle for near-optimality of stochastic delay control problem," *Advances in Difference Equations*, vol. 2017, no. 1, p. 98, 2017.
- [18] Y. Wang and Z. Wu, "Necessary and sufficient conditions for near-optimality of stochastic delay systems," *International Journal of Control*, vol. 91, no. 8, pp. 1–24, 2017.
- [19] G. Ren and B. Liu, "Near-optimal control for a singularly perturbed linear stochastic singular system with Markovian jumping parameters," *European Journal of Control*, vol. 50, pp. 88–95, 2019.
- [20] Y. Zhang, S. Li, and L. Liao, "Near-optimal control of nonlinear dynamical systems: a brief survey," *Annual Reviews in Control*, vol. 47, pp. 71–80, 2019.
- [21] R. Buckdahn, B. Djehiche, J. Li, and S. Peng, "Mean-field backward stochastic differential equations: a limit approach," *The Annals of Probability*, vol. 37, no. 4, pp. 1524–1565, 2009.
- [22] R. Buckdahn, J. Li, and S. Peng, "Mean-field backward stochastic differential equations and related partial differential equations," *Stochastic Processes and Their Applications*, vol. 119, no. 10, pp. 3133–3154, 2009.
- [23] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229–260, 2007.
- [24] R. M. Xu and T. T. Wu, "Mean-field backward stochastic evolution equations in Hilbert spaces and optimal control for BSPDEs," *Mathematical Problems in Engineering*, vol. 2014, Article ID 718948, 18 pages, 2014.
- [25] D. Andersson and B. Djehiche, "A maximum principle for SDEs of mean-field type," *Applied Mathematics & Optimization*, vol. 63, no. 3, pp. 341–356, 2011.
- [26] R. Buckdahn, B. Djehiche, and J. Li, "A general stochastic maximum principle for SDEs of mean-field type," *Applied Mathematics & Optimization*, vol. 64, no. 2, pp. 197–216, 2011.
- [27] J. Li, "Stochastic maximum principle in the mean-field controls," *Automatica*, vol. 48, no. 2, pp. 366–373, 2012.
- [28] Y. Shen and T. K. Siu, "The maximum principle for a jump-diffusion mean-field model and its application to the mean-variance problem," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 86, pp. 58–73, 2013.
- [29] J. Huang, S. Wang, and Z. Wu, "Backward mean-field linear-quadratic-Gaussian (LQG) games: full and partial information," *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 3784–3796, 2016.
- [30] M. Huang, "Large-population LQG games involving a major player: the Nash certainty equivalence principle," *SIAM Journal on Control and Optimization*, vol. 48, no. 5, pp. 3318–3353, 2010.
- [31] R. Xu and J. Shi, " ϵ -Nash mean-field games for linear-quadratic systems with random jumps and applications," *International Journal of Control*, pp. 1–11, 2019.
- [32] R. M. Xu and F. Zhang, " ϵ -Nash mean-field games for general linear-quadratic systems with applications," *Automatica*, vol. 114, Article ID 108835, 2020.
- [33] A. Bensoussan, S. C. P. Yam, and Z. Zhang, "Well-posedness of mean-field type forward-backward stochastic differential equations," *Stochastic Processes and Their Applications*, vol. 125, no. 9, pp. 3327–3354, 2015.
- [34] J. M. Yong, "Forward-backward stochastic differential equations with mixed initial-terminal conditions," *Transactions of the American Mathematical Society*, vol. 362, no. 2, pp. 1047–1096, 2010.
- [35] J. Yong, "Optimality variational principle for controlled forward-backward stochastic differential equations with mixed initial-terminal conditions," *SIAM Journal on Control and Optimization*, vol. 48, no. 6, pp. 4119–4156, 2010.
- [36] J. M. Yong and X. Y. Zhou, *Stochastic Controls, Hamiltonian Systems and HJB Equations*, Springer-Verlag, Berlin, Germany, 1999.
- [37] X. Y. Zhou and S. P. Sethi, "A sufficient condition for near-optimal stochastic controls and its application to manufacturing systems," *Applied Mathematics & Optimization*, vol. 29, no. 1, pp. 67–92, 1994.