1-1-2013

Linear-Quadratic Optimal Control Problems for Mean-Field Stochastic Differential Equations

Jiongmin Yong
University of Central Florida

Recommended Citation

Showcase of Text, Archives, Research & Scholarship

LINEAR-QUADRATIC OPTIMAL CONTROL PROBLEMS FOR MEAN-FIELD STOCHASTIC DIFFERENTIAL EQUATIONS*

JIONGMIN YONG†

Abstract. Linear-quadratic optimal control problems are considered for mean-field stochastic differential equations with deterministic coefficients. By a variational method, the optimality system is derived, which is a linear mean-field forward-backward stochastic differential equation. Using a decoupling technique, two Riccati differential equations are obtained which are uniquely solvable under certain conditions. Then a feedback representation is obtained for the optimal control.

Key words. mean-field stochastic differential equation, linear-quadratic optimal control, Riccati differential equation, feedback representation

AMS subject classifications. 49N10, 49N35, 93E20

DOI. 10.1137/120892477

1. Introduction. Let $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ be a complete filtered probability space on which a one-dimensional standard Brownian motion $W(\cdot)$ is defined with $\mathbb{F} \equiv \{\mathcal{F}_t\}_{t \geq 0}$ being its natural filtration augmented by all the $\mathbb{P}$-null sets. Consider the following controlled linear stochastic differential equation (SDE):

$$(1.1) \quad \begin{cases} dX(s) = \big\{A(s)X(s) + \bar{A}(s)\mathbb{E}[X(s)] + B(s)u(s) + \bar{B}(s)\mathbb{E}[u(s)]\big\}ds \\ \qquad\quad + \big\{C(s)X(s) + \bar{C}(s)\mathbb{E}[X(s)] + D(s)u(s) + \bar{D}(s)\mathbb{E}[u(s)]\big\}dW(s), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad s \in [0, T], \\ X(0) = x, \end{cases}$$

where $A(\cdot), \bar{A}(\cdot), B(\cdot), \bar{B}(\cdot), C(\cdot), \bar{C}(\cdot), D(\cdot), \bar{D}(\cdot)$ are given matrix valued deterministic functions. In the above, $X(\cdot)$, valued in $\mathbb{R}^n$, is the *state process*, and $u(\cdot)$, valued in $\mathbb{R}^m$, is the *control process*.

We note that $\mathbb{E}[X(\cdot)]$ and $\mathbb{E}[u(\cdot)]$ appear in the state equation. Such an equation is referred to as a mean-field (forward) SDE (MF-FSDE). MF-FSDEs can be used to describe particle systems at the mesoscopic level, which is of great importance in applications. Historically, what would later be called the McKean–Vlasov SDE, a kind of MF-FSDE, was suggested by Kac [24] in 1956 as a stochastic toy model for the Vlasov kinetic equation of plasma, and a study of this was initiated by McKean [29] in 1966. Since then, many authors have made contributions to McKean–Vlasov type SDEs and applications; see, for example, Dawson [15], Dawson and Gärtner [16], Gärtner [20], Scheutzow [34], Graham [21], Chan [12], Chiang [13], and Ahmed and Ding [1]. In recent years, related topics and problems have attracted more researchers' attention; see, for example, Veretennikov [35], Huang, Malhamé, and Caines [23], Mahmudov and McKibben [28], Buckdahn et al. [10], Buckdahn, Li, and Peng [11], Borkar and Kumar [8], Crisan and Xiong [14], and Kotelenez and Kurtz [25], to mention a few. More interestingly, control problems of the McKean–Vlasov equation or MF-FSDEs

†Department of Mathematics, University of Central Florida, Orlando, FL 32816 (jiongmin. yong@ucf.edu).

were investigated by Ahmed and Ding [2], Ahmed [3], Buckdahn, Djehiche, and Li [9], Park, Balasubramaniam, and Kang [33], Andersson and Djehiche [4], Meyer-Brandis, Oksendal, and Zhou [30], and so on. This paper can be regarded as an addition to the study of optimal control for MF-FSDEs.

For the state equation (1.1), we introduce the following set:

$$
\begin{aligned}
\mathcal{U}[0,T] = L_{\mathbb{F}}^2(0,T;\mathbb{R}^m) \\
\triangleq \Big\{ u : [0,T] \times \Omega \to \mathbb{R}^m \ \big| \ u(\cdot) \text{ is } \mathbb{F}\text{-progressively measurable,} \\
\mathbb{E} \int_0^T |u(s)|^2 ds < \infty \Big\}.
\end{aligned}
$$

Any $u(\cdot) \in \mathcal{U}[0,T]$ is called an *admissible control*. Under mild conditions, one can show that (see below) for any $(x, u(\cdot)) \in \mathbb{R}^n \times \mathcal{U}[0,T]$, (1.1) admits a unique solution $X(\cdot) = X(\cdot\,; x, u(\cdot))$. We introduce the following cost functional:

$$
\begin{aligned}
(1.2) \quad J(x; u(\cdot)) = \mathbb{E}\Big\{ &\int_0^T \Big[\langle Q(s)X(s), X(s) \rangle + \langle \bar{Q}(s)\mathbb{E}[X(s)], \mathbb{E}[X(s)] \rangle \\
&+ \langle R(s)u(s), u(s) \rangle + \langle \bar{R}(s)\mathbb{E}[u(s)], \mathbb{E}[u(s)] \rangle \Big] ds \\
&+ \langle GX(T), X(T) \rangle + \langle \bar{G}\mathbb{E}[X(T)], \mathbb{E}[X(T)] \rangle \Big\}
\end{aligned}
$$

with $Q(\cdot), \bar{Q}(\cdot), R(\cdot), \bar{R}(\cdot)$ being suitable symmetric matrix-valued functions and $G, \bar{G}$ being symmetric matrices. Our optimal control problem can be stated as follows.

Problem (MF-LQ). For given $x \in \mathbb{R}^n$, find a $u^*(\cdot) \in \mathcal{U}[0,T]$ such that

$$
(1.3) \quad J(x; u^*(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}[0,T]} J(x; u(\cdot)).
$$

Any $u^*(\cdot) \in \mathcal{U}[0,T]$ satisfying the above is called an *optimal control* and the corresponding state process $X^*(\cdot) \equiv X(\cdot\,; x, u^*(\cdot))$ is called an *optimal state process*; the pair $(X^*(\cdot), u^*(\cdot))$ is called an *optimal pair*.

From the above-listed literature, one has some motivation for the inclusion of $\mathbb{E}[X(\cdot)]$ and $\mathbb{E}[u(\cdot)]$ in the state equation. We now briefly explain a motivation of including $\mathbb{E}[X(\cdot)]$ and $\mathbb{E}[u(\cdot)]$ in the cost functional. We recall that for a classical LQ problem with the state equation

$$
\begin{aligned}
(1.4) \\
\begin{cases}
dX(s) = \Big[A(s)X(s) + B(s)u(s) \Big] ds + \Big[C(s)X(s) + D(s)u(s) \Big] dW(s), s \in [0,T], \\
X(0) = x,
\end{cases}
\end{aligned}
$$

one has the following cost functional:

$$
\begin{aligned}
(1.5) \quad J_0(x; u(\cdot)) = \mathbb{E}\Big\{ &\int_0^T \Big[\langle Q_0(s)X(s), X(s) \rangle + \langle R_0(s)u(s), u(s) \rangle \Big] ds \\
&+ \langle G_0 X(T), X(T) \rangle \Big\}.
\end{aligned}
$$

For such a corresponding optimal control problem, it is natural to hope that the optimal state process and/or control process could be not too sensitive with respect

to the possible variation of the random events. One way to achieve this is trying to keep the variations $\text{var}\,[X(\cdot)]$ and $\text{var}\,[u(\cdot)]$ small. Therefore, one could include $\text{var}\,[X(\cdot)]$ and $\text{var}\,[u(\cdot)]$ in the cost functional. A natural form of the cost functional, replacing (1.5), could be the following:

(1.6)
$$\widehat{J}_0(x;u(\cdot)) = \mathbb{E}\Bigg\{\int_0^T \Big[\,\langle\, Q_0(s)X(s), X(s)\,\rangle + q(s)\text{var}\,[X(s)] + \langle\, R_0(s)u(s), u(s)\,\rangle$$
$$+\rho(s)\text{var}\,[u(s)]\Big]ds + \langle\, G_0 X(T), X(T)\,\rangle + g\text{var}\,[X(T)]\Bigg\}$$

for some (positive) weighting factors $q(\cdot)$, $\rho(\cdot)$, and g. Since

$$\text{var}\,[X(s)] = \mathbb{E}|X(s)|^2 - \Big(\mathbb{E}[X(s)]\Big)^2,$$

and similar things hold for $\text{var}\,[X(T)]$ and $\text{var}\,[u(s)]$, we see that

$$\widehat{J}_0(x;u(\cdot)) = \mathbb{E}\Bigg\{\int_0^T \Big[\,\langle[Q_0(s)+q(s)I]X(s), X(s)\,\rangle -q(s)\Big(\mathbb{E}[X(s)]\Big)^2$$
$$+\langle[R_0(s)+\rho(s)I]u(s), u(s)\,\rangle -\rho(s)\Big(\mathbb{E}[u(s)]\Big)^2\Big]ds$$
$$+\langle[G_0+gI]X(T), X(T)\,\rangle -g\Big(\mathbb{E}[X(T)]\Big)^2\Bigg\}.$$

Clearly, the above is a special case of (1.2) with

$$Q(\cdot) = Q_0(\cdot) + q(\cdot)I, \quad R(\cdot) = R_0(\cdot) + \rho(\cdot)I, \quad G = G_0 + gI,$$
$$\bar{Q}(\cdot) = -q(\cdot)I, \quad \bar{R}(\cdot) = -\rho(\cdot)I, \quad \bar{G} = -gI.$$

We point out that $\bar{Q}(\cdot)$, $\bar{R}(\cdot)$, and $\bar{G}$ in the above are not positive semidefinite. Because of this, we will not assume the positive semidefiniteness of $\bar{Q}(\cdot)$, $\bar{R}(\cdot)$, and $\bar{G}$ in (1.2).

Next, let us make an observation. By (1.1), we have

(1.7)
$$\begin{cases} d\mathbb{E}[X(t)] = \Big\{\big[A(t) + \bar{A}(t)\big]\mathbb{E}[X(t)] + \big[B(t) + \bar{B}(t)\big]\mathbb{E}[u(t)]\Big\}dt, \\ \mathbb{E}[X(0)] = x. \end{cases}$$

Then

$$d\Big(X(t) - \mathbb{E}[X(t)]\Big) = \Big\{A(t)\Big(X(t) - \mathbb{E}[X(t)]\Big) + B(t)\Big(u(t) - \mathbb{E}[u(t)]\Big)\Big\}dt$$
$$+\Big\{C(t)\Big(X(t) - \mathbb{E}[X(t)]\Big) + \Big(C(t) + \bar{C}(t)\Big)\mathbb{E}[X(t)]$$
$$+D(t)\Big(u(t) - \mathbb{E}[u(t)]\Big) + \Big(D(t) + \bar{D}(t)\Big)\mathbb{E}[u(t)]\Big\}dW(t).$$

Let

$$\mathbf{X}(t) = \begin{pmatrix} X(t) - \mathbb{E}[X(t)] \\ \mathbb{E}[X(t)] \end{pmatrix}, \qquad \mathbf{u}(t) = \begin{pmatrix} u(t) - \mathbb{E}[u(t)] \\ \mathbb{E}[u(t)] \end{pmatrix}, \qquad \mathbf{X}_0 = \begin{pmatrix} 0 \\ x \end{pmatrix},$$

and

$$\begin{cases} \mathcal{A}(t) = \begin{pmatrix} A(t) & 0 \\ 0 & A(t) + \bar{A}(t) \end{pmatrix}, & \mathcal{B}(t) = \begin{pmatrix} B(t) & 0 \\ 0 & B(t) + \bar{B}(t) \end{pmatrix}, \\ \mathcal{C}(t) = \begin{pmatrix} C(t) & C(t) + \bar{C}(t) \\ 0 & 0 \end{pmatrix}, & \mathcal{D}(t) = \begin{pmatrix} D(t) & D(t) + \bar{D}(t) \\ 0 & 0 \end{pmatrix}. \end{cases}$$

Then (1.1) and (1.7) can be put together as

$$(1.8) \qquad \begin{cases} d\mathbf{X}(t) = \big\{ \mathcal{A}(t)\mathbf{X}(t) + \mathcal{B}(t)\mathbf{u}(t) \big\} dt + \big\{ \mathcal{C}(t)\mathbf{X}(t) + \mathcal{D}(t)\mathbf{u}(t) \big\} dW(t), \\ \mathbf{X}(0) = \mathbf{X}_0. \end{cases}$$

On the other hand, we note that

$$\begin{aligned} \mathbb{E}\{ \langle\, Q(t)X(t), X(t) \,\rangle &+ \langle\, \bar{Q}(t)\mathbb{E}[X(t)], \mathbb{E}[X(t)] \,\rangle \} \\ &= \mathbb{E}\{ \langle\, Q(t)(X(t) - \mathbb{E}[X(t)]), X(t) - \mathbb{E}[X(t)] \,\rangle \\ &\qquad + \langle\, (Q(t) + \bar{Q}(t))\mathbb{E}[X(t)], \mathbb{E}[X(t)] \,\rangle \} \\ &= \mathbb{E}\left\langle \begin{pmatrix} Q(t) & 0 \\ 0 & Q(t) + \bar{Q}(t) \end{pmatrix} \mathbf{X}(t), \mathbf{X}(t) \right\rangle. \end{aligned}$$

Similarly,

$$\begin{aligned} \mathbb{E}\Big\{ \langle\, R(t)u(t), u(t) \,\rangle &+ \langle\, \bar{R}(t)\mathbb{E}[u(t)], \mathbb{E}[u(t)] \,\rangle \Big\} \\ &= \mathbb{E}\left\langle \begin{pmatrix} R(t) & 0 \\ 0 & R(t) + \bar{R}(t) \end{pmatrix} \mathbf{u}(t), \mathbf{u}(t) \right\rangle \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}\Big\{ \langle\, GX(T), X(T) \,\rangle &+ \langle\, \bar{G}\mathbb{E}[X(T)], \mathbb{E}[X(T)] \,\rangle \Big\} \\ &= \mathbb{E}\left\langle \begin{pmatrix} G & 0 \\ 0 & G + \bar{G} \end{pmatrix} \mathbf{X}(T), \mathbf{X}(T) \right\rangle. \end{aligned}$$

Hence, by letting

$$\mathcal{Q}(t) = \begin{pmatrix} Q(t) & 0 \\ 0 & Q(t) + \bar{Q}(t) \end{pmatrix}, \quad \mathcal{R}(t) = \begin{pmatrix} R(t) & 0 \\ 0 & R(t) + \bar{R}(t) \end{pmatrix}, \quad \mathcal{G} = \begin{pmatrix} G & 0 \\ 0 & G + \bar{G} \end{pmatrix},$$

we have
(1.9)

$$\begin{aligned} J(x; u(\cdot)) &= \mathbb{E}\bigg\{ \int_0^T \Big[\langle\, Q(s)X(s), X(s) \,\rangle + \langle\, \bar{Q}(s)\mathbb{E}[X(s)], \mathbb{E}[X(s)] \,\rangle \\ &\qquad + \langle\, R(s)u(s), u(s) \,\rangle + \langle\, \bar{R}(s)\mathbb{E}[u(s)], \mathbb{E}[u(s)] \,\rangle \Big] ds \\ &\qquad + \langle\, GX(T), X(T) \,\rangle + \langle\, \bar{G}\mathbb{E}[X(T)], \mathbb{E}[X(T)] \,\rangle \bigg\} \\ &= \mathbb{E}\bigg\{ \int_0^T \Big[\langle\, \mathcal{Q}(t)\mathbf{X}(t), \mathbf{X}(t) \,\rangle + \langle\, \mathcal{R}(t)\mathbf{u}(t), \mathbf{u}(t) \,\rangle \Big] dt + \langle\, \mathcal{G}\mathbf{X}(T), \mathbf{X}(T) \,\rangle \bigg\}. \end{aligned}$$

Note that state equation (1.8) with cost functional (1.9) forms a standard-looking stochastic LQ problem. However, the control $\mathbf{u}(\cdot)$ has to be of the form

$$\mathbf{u}(\cdot) = \begin{pmatrix} u(\cdot) - \mathbb{E}[u(\cdot)] \\ \mathbb{E}[u(\cdot)] \end{pmatrix}.$$

The collection of all such processes is not $\mathcal{U}[0,T] \times \mathcal{U}[0,T]$, which should be the set of all admissible controls if (1.8)–(1.9) is regarded as a standard stochastic LQ problem. Hence, the above reduction is just formal and does not lead to a direct application of standard stochastic LQ theory. This also shows that problem (MF-LQ) cannot be regarded as a standard LQ problem. On the other hand, the above observation motivates studies of LQ problems with a (linearly) constrained control set. We will explore some relevant results along this line in our future publications.

The purpose of this paper is to study problem (MF-LQ). We will begin with the well-posedness of the state equation and the solvability of problem (MF-LQ) in section 2. Then, in section 3, we will establish necessary and sufficient conditions for optimal pairs. It turns out that the optimality system is a coupled mean-field type forward-backward SDE (MF-FBSDE). Inspired by the invariant embedding [5] and the four-step scheme for FBSDEs [27], we derive two Riccati differential equations in section 4, so that the optimal control can be represented as a state feedback form. We also present a verification of optimality for the state feedback control. In section 5, we looked carefully at the one-dimensional case. Some interesting comparisons are presented.

2. Preliminaries. First, for any Euclidean space $H = \mathbb{R}^n, \mathbb{R}^{n \times m}, \mathcal{S}^n$ (with $\mathcal{S}^n$ being the set of all $(n \times n)$ symmetric matrices), we let $L^p(0,t;H)$ be the set of all H-valued functions that are L^p-integrable on $[0,t]$, $p \in [1,\infty]$. Next, we introduce the following spaces:

$$\mathcal{X}_t \equiv L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^n) = \left\{ \xi : \Omega \to \mathbb{R}^n \mid \xi \text{ is } \mathcal{F}_t\text{-measurable}, \mathbb{E}|\xi|^2 < \infty \right\},$$

$$\mathcal{U}_t \equiv L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^m) = \left\{ \eta : \Omega \to \mathbb{R}^m \mid \eta \text{ is } \mathcal{F}_t\text{-measurable}, \mathbb{E}|\eta|^2 < \infty \right\},$$

$$L^2_{\mathbb{F}}(0,t;\mathbb{R}^n) = \left\{ X : [0,t] \times \Omega \to \mathbb{R}^n \mid X(\cdot) \text{ is } \mathbb{F}\text{-adapted}, \mathbb{E}\int_0^t |X(s)|^2 ds < \infty \right\},$$

$$\mathcal{X}[0,t] \equiv C_{\mathbb{F}}([0,t];\mathbb{R}^n) = \left\{ X : [0,t] \times \Omega \to \mathbb{R}^n \mid X(\cdot) \text{ is } \mathbb{F}\text{-adapted}, L^2\text{-countinuous} \right.$$
$$\left. \sup_{s \in [0,t]} \mathbb{E}|X(s)|^2 < \infty \right\},$$

$$\widehat{\mathcal{X}}[0,t] \equiv L^2_{\mathbb{F}}(\Omega;C([0,t];\mathbb{R}^n)) = \left\{ X : [0,t] \times \Omega \to \mathbb{R}^n \mid X(\cdot) \text{ is } \mathbb{F}\text{-adapted}, \right.$$
$$\left. X(\cdot) \text{ has continuous paths}, \mathbb{E}\left[\sup_{s \in [0,t]} |X(s)|^2 \right] < \infty \right\}.$$

Note that the L^2-continuity of $X(s)$ means that as a map from $[0,t]$ to $\mathcal{X}_t$, $s \mapsto X(s)$ is continuous. $X(\cdot)$ with continuous paths means that almost surely $\omega \in \Omega$, $s \mapsto X(s,\omega)$ is continuous. It is known that

$$\widehat{\mathcal{X}}[0,t] \subseteq \mathcal{X}[0,t] \subseteq L^2_{\mathbb{F}}(0,t;\mathbb{R}^n), \qquad \widehat{\mathcal{X}}[0,t] \neq \mathcal{X}[0,t] \neq L^2_{\mathbb{F}}(0,t;\mathbb{R}^n).$$

We now introduce the following assumption for the coefficients of the state equation.

(H1) The following hold:

$$(2.1) \qquad \begin{cases} A(\cdot), \bar{A}(\cdot), C(\cdot), \bar{C}(\cdot) \in L^\infty(0, T; \mathbb{R}^{n \times n}), \\ B(\cdot), \bar{B}(\cdot), D(\cdot), \bar{D}(\cdot) \in L^\infty(0, T; \mathbb{R}^{n \times m}). \end{cases}$$

For the weighting matrices in the cost functional, we introduce the following assumption.

(H2) The following hold:

$$(2.2) \qquad Q(\cdot), \bar{Q}(\cdot) \in L^\infty(0, T; \mathcal{S}^n), \quad R(\cdot), \bar{R}(\cdot) \in L^\infty(0, T; \mathcal{S}^m), \quad G, \bar{G} \in \mathcal{S}^n,$$

and for some $\delta > 0$,

$$(2.3) \qquad \begin{cases} Q(s), \ Q(s) + \bar{Q}(s) \geq 0, \qquad R(s), \ R(s) + \bar{R}(s) \geq \delta I, \qquad s \in [0, T], \\ G, \ G + \bar{G} \geq 0. \end{cases}$$

We point out that $\bar{Q}(\cdot)$, $\bar{R}(\cdot)$, and $\bar{G}$ are not assumed to be positive semidefinite. Now, for any $X(\cdot) \in \mathcal{X}[0, T]$ and any $u(\cdot) \in \mathcal{U}[0, T]$, we define
(2.4)

$$\begin{cases} [\mathcal{A}X(\cdot)](t) = \int_0^t \Big(A(s)X(s) + \bar{A}(s)\mathbb{E}[X(s)] \Big) ds + \int_0^t \Big(C(s)X(s) + \bar{C}(s)\mathbb{E}[X(s)] \Big) dW(s), \\ [\mathcal{B}u(\cdot)](t) = \int_0^t \Big(B(s)u(s) + \bar{B}(s)\mathbb{E}[u(s)] \Big) ds + \int_0^t \Big(D(s)u(s) + \bar{D}(s)\mathbb{E}[u(s)] \Big) dW(s), \\ \hspace{9cm} t \in [0, T]. \end{cases}$$

The following result is concerned with operators $\mathcal{A}$ and $\mathcal{B}$.

LEMMA 2.1. *The following estimates hold as long as the involved norms on the right-hand sides are meaningful: For any $t \in [0, T]$,*
(2.5)

$$\|\mathcal{A}X(\cdot)\|_{\widehat{\mathcal{X}}[0,t]}^2 \leq K \Big[\|A(\cdot)\|_{L^2(0,t;\mathbb{R}^{n \times n})}^2 + \|\bar{A}(\cdot)\|_{L^2(0,t;\mathbb{R}^{n \times n})}^2$$
$$+ \|C(\cdot)\|_{L^\infty(0,t;\mathbb{R}^{n \times n})}^2 + \|\bar{C}(\cdot)\|_{L^\infty(0,t;\mathbb{R}^{n \times n})}^2 \Big] \|X(\cdot)\|_{L^2_{\mathbb{F}}(0,t;\mathbb{R}^n)}^2,$$

$$(2.6) \qquad \|\mathcal{A}X(\cdot)\|_{\widehat{\mathcal{X}}[0,t]}^2 \leq K \Big[\|A(\cdot)\|_{L^1(0,t;\mathbb{R}^{n \times n})}^2 + \|\bar{A}(\cdot)\|_{L^1(0,t;\mathbb{R}^{n \times n})}^2$$
$$+ \|C(\cdot)\|_{L^2(0,t;\mathbb{R}^{n \times n})}^2 + \|\bar{C}(\cdot)\|_{L^2(0,t;\mathbb{R}^{n \times n})}^2 \Big] \|X(\cdot)\|_{\mathcal{X}[0,t]}^2,$$

and

$$(2.7) \qquad \|\mathcal{B}u(\cdot)\|_{\widehat{\mathcal{X}}[0,t]}^2 \leq K \Big[\|B(\cdot)\|_{L^2(0,t;\mathbb{R}^{n \times n})}^2 + \|\bar{B}(\cdot)\|_{L^2(0,t;\mathbb{R}^{n \times n})}^2$$
$$+ \|D(\cdot)\|_{L^\infty(0,t;\mathbb{R}^{n \times n})}^2 + \|\bar{D}(\cdot)\|_{L^\infty(0,t;\mathbb{R}^{n \times n})}^2 \Big] \|u(\cdot)\|_{\mathcal{U}[0,t]}^2.$$

Hereafter, $K > 0$ represents a generic constant which can be different from line to line.

Proof. For any $t \in (0, T]$ and any $X(\cdot) \in L_{\mathbb{F}}^2(0, t; \mathbb{R}^n)$,

$$
\mathbb{E}\left(\sup_{s \in [0,t]} |[\mathcal{A}X(\cdot)](s)| \right) \leq K \left\{ \mathbb{E}\left(\int_0^t |A(s)||X(s)|ds \right)^2 + \left(\int_0^t |\bar{A}(s)||\mathbb{E}[X(s)]|ds \right)^2 \right.
$$
$$
\left. + \mathbb{E}\int_0^t |C(s)|^2 |X(s)|^2 ds + \int_0^t |\bar{C}(s)|^2 |\mathbb{E}[X(s)]|^2 ds \right\}
$$
$$
\leq K \left\{ \left(\int_0^t |A(s)|^2 ds \right) + \left(\int_0^t |\bar{A}(s)|^2 ds \right) \right.
$$
$$
\left. + \sup_{s \in [0,t]} |C(s)|^2 + \sup_{s \in [0,t]} |\bar{C}(s)|^2 \right\} \int_0^t \mathbb{E}|X(s)|^2 ds.
$$

Thus, estimate (2.5) holds. Next, for any $X(\cdot) \in \mathcal{X}[0, t]$,

$$
\mathbb{E}\left(\sup_{s \in [0,t]} |[\mathcal{A}X(\cdot)](s)|^2 \right)
$$
$$
\leq K \left\{ \mathbb{E}\left(\int_0^t |A(s)||X(s)|ds \right)^2 + \mathbb{E}\left(\int_0^t |\bar{A}(s)||\mathbb{E}[X(s)]|ds \right)^2 \right.
$$
$$
\left. + \mathbb{E}\int_0^t |C(s)|^2 |X(s)|^2 ds + \mathbb{E}\int_0^t |\bar{C}(s)|^2 |\mathbb{E}[X(s)]|^2 ds \right\}
$$
$$
\leq K \left[\left(\int_0^t |A(s)|ds \right)^2 + \left(\int_0^t |\bar{A}(s)|ds \right)^2 + \int_0^t |C(s)|^2 ds + \int_0^t |\bar{C}(s)|^2 ds \right] \left(\sup_{s \in [0,t]} \mathbb{E}|X(s)|^2 \right).
$$

Hence, (2.6) follows. Similar to the proof of (2.5), we can prove (2.7). □

The above lemma leads to the following corollary.

COROLLARY 2.2. *If* (H1) *holds, then* $\mathcal{A} : \mathcal{X}[0, T] \to \widehat{\mathcal{X}}[0, T]$ *and* $\mathcal{B} : \mathcal{U}[0, T] \to \widehat{\mathcal{X}}[0, T]$ *are bounded,* $\mathcal{A} : L_{\mathbb{F}}^2(0, T; \mathbb{R}^n) \to \widehat{\mathcal{X}}[0, T]$ *is also bounded.*

Next, we define

$$
\begin{cases}
I_T X(\cdot) = X(T), \\
\mathcal{A}_T X(\cdot) = I_T \mathcal{A}X(\cdot) \equiv [\mathcal{A}X(\cdot)](T) = \int_0^T \left(A(s)X(s) + \widehat{A}(s)\mathbb{E}[X(s)] \right) ds \\
\qquad\qquad + \int_0^T \left(A_1(s)X(s) + \widehat{A}_1(s)\mathbb{E}[X(s)] \right) dW(s), \\
\mathcal{B}_T u(\cdot) = I_T \mathcal{B}u(\cdot) \equiv [\mathcal{B}u(\cdot)](T) = \int_0^T \left(B(s)u(s) + \widehat{B}(s)\mathbb{E}[u(s)] \right) ds \\
\qquad\qquad + \int_0^T \left(B_1(s)u(s) + \widehat{B}_1(s)\mathbb{E}[u(s)] \right) dW(s).
\end{cases}
$$

It is easy to see that

$$
I_T : \mathcal{X}[0, T] \to \mathcal{X}_T
$$

is bounded. According to Lemma 2.1*, we have the following result.*

COROLLARY 2.3. *If* (H1) *holds, then* $\mathcal{A}_T : \mathcal{X}[0, T] \to \mathcal{X}_T$ *and* $\mathcal{B}_T : \mathcal{U}[0, T] \to \mathcal{X}_T$ *are bounded, and* $\mathcal{A}_T : L_{\mathbb{F}}^2(0, T; \mathbb{R}^n) \to \mathcal{X}_T$ *is also bounded.*

Recall that if $\eta \in \mathbb{D}^{1,2} \cap \mathcal{X}_T$, then there exists a unique $\zeta(\cdot) \in L_{\mathbb{F}}^2(0, T; \mathbb{R}^n)$ such that

$$\eta = \mathbb{E}\eta + \int_0^T \zeta(s)dW(s).$$

According to [32],

$$\zeta(s) = \mathbb{E}[D_s\eta \mid \mathcal{F}_s] \equiv \mathcal{D}_s\eta, \qquad s \in [0, T],$$

where $D_s\eta$ is called the *Malliavin derivative* of η. Next, we have the following results which give representation of the adjoint operators of $\mathcal{A}$, $\mathcal{B}$, $\mathcal{A}_T$, and $\mathcal{B}_T$.

PROPOSITION 2.4. *For any $\eta \in \mathbb{D}^{1,2} \cap \mathcal{X}_T$ and $Y(\cdot) \in L^2(0, T; \mathbb{D}^{1,2})$, the following hold:*

$$\begin{cases}
(\mathcal{A}^*Y)(s) = \int_s^T \Big(A(s)^T Y(t) + \bar{A}(s)^T \mathbb{E}[Y(t)] + C(s)^T \mathcal{D}_s Y(t) + \bar{C}(s)^T \mathbb{E}[\mathcal{D}_s Y(t)]\Big)dt, \\
(\mathcal{B}^*Y)(s) = \int_s^T \Big(B(s)^T Y(t) + \bar{B}(s)^T \mathbb{E}[Y(t)] + D(s)^T \mathcal{D}_s Y(t) + \bar{D}(s)^T \mathbb{E}[\mathcal{D}_s Y(t)]\Big)dt, \\
(\mathcal{A}_T^*\eta)(s) = A(s)^T\eta + \bar{A}(s)^T\mathbb{E}\eta + C(s)^T\mathcal{D}_s\eta + \bar{C}(s)^T\mathbb{E}[\mathcal{D}_s\eta], \\
(\mathcal{B}_T^*\eta)(s) = B(s)^T\eta + \bar{B}(s)^T\mathbb{E}\eta + D(s)^T\mathcal{D}_s\eta + \bar{D}(s)^T\mathbb{E}[\mathcal{D}_s\eta].
\end{cases}$$

Proof. For any $Y(\cdot) \in L^2(0, T; \mathbb{D}^{1,2})$,

$$\langle X, \mathcal{A}^*Y \rangle = \langle \mathcal{A}X, Y \rangle = \mathbb{E}\int_0^T \langle [\mathcal{A}X](t), Y(t) \rangle dt$$

$$= \mathbb{E}\int_0^T \Big\langle \int_0^t \Big(A(s)X(s) + \bar{A}(s)\mathbb{E}[X(s)]\Big)ds + \int_0^t \Big(C(s)X(s) + \bar{C}(s)\mathbb{E}[X(s)]\Big)dW(s), Y(t)\Big\rangle dt$$

$$= \mathbb{E}\int_0^T \int_s^T \langle A(s)X(s) + \bar{A}(s)\mathbb{E}[X(s)], Y(t) \rangle dt ds$$

$$+ \mathbb{E}\int_0^T \int_s^T \langle C(s)X(s) + \bar{C}(s)\mathbb{E}[X(s)], \mathcal{D}_s Y(t) \rangle dt ds$$

$$= \mathbb{E}\int_0^T \Big\langle X(s), \int_s^T \Big(A(s)^T Y(t) + \bar{A}(s)^T \mathbb{E}[Y(t)] + C(s)^T \mathcal{D}_s Y(t) + \bar{C}(s)^T \mathbb{E}[\mathcal{D}_s Y(t)]\Big)dt \Big\rangle ds.$$

Thus, the representation of $\mathcal{A}^*$ follows. We can obtain the representation of $\mathcal{B}^*$ similarly. Next, for any $\eta \in \mathbb{D}^{1,2} \cap \mathcal{X}_T$,

$$\langle X, \mathcal{A}_T^*\eta \rangle = \langle \mathcal{A}_T X, \eta \rangle$$

$$= \mathbb{E}\Big\langle \int_0^T \Big(A(s)X(s) + \bar{A}(s)\mathbb{E}[X(s)]\Big)ds, \eta \Big\rangle$$

$$+ \mathbb{E}\Big\langle \int_0^T \Big(C(s)X(s) + \bar{C}(s)\mathbb{E}[X(s)]\Big)dW(s), \eta \Big\rangle$$

$$= \mathbb{E}\int_0^T \langle X(s), A(s)^T\eta + \bar{A}(s)^T\mathbb{E}\eta + C(s)^T\mathcal{D}_s\eta + \bar{C}(s)^T\mathbb{E}[\mathcal{D}_s\eta] \rangle ds.$$

Therefore, the representation of $\mathcal{A}_T^*$ follows. Similarly, we can obtain the representation of $\mathcal{B}_T^*$. □

For completeness, let us also prove the following result.

PROPOSITION 2.5. *It holds that*

$$(2.8) \qquad I_T^* \xi = \xi \delta_{\{T\}} \qquad \forall \eta \in \mathcal{X}_T,$$

where $\delta_{\{T\}}$ is the Dirac measure at T, and

$$(2.9) \qquad \mathbb{E}^* x = x^T \mathbb{E} \qquad \forall x \in \mathbb{R}^n.$$

Proof. First, since $I_T : \mathcal{X}[0,T] \to \mathcal{X}_T$ is bounded, we have $I_T^* : \mathcal{X}_T^* \equiv \mathcal{X}_T \to \mathcal{X}[0,T]^*$. For any $\xi \in \mathcal{X}_T$ and any $Y(\cdot) \in \mathcal{X}[0,T]$, we have

$$\langle I_T^* \xi, Y(\cdot) \rangle = \langle \xi, I_T Y(\cdot) \rangle = \mathbb{E} \langle \xi, Y(T) \rangle = \mathbb{E} \int_0^T \langle Y(s), \xi \rangle \delta_{\{T\}}(ds).$$

This gives (2.8). Next, since $\mathbb{E} : \mathcal{X}_T \to \mathbb{R}$, we have $\mathbb{E}^* : \mathbb{R} \to \mathcal{X}_T$. For any $\xi \in \mathcal{X}_T$ and $x \in \mathbb{R}$,

$$\langle \mathbb{E}^* x, \xi \rangle = \langle x, \mathbb{E}\xi \rangle = \mathbb{E} \langle x, \xi \rangle = x^T \mathbb{E}\xi.$$

This completes the proof. □

With operators $\mathcal{A}$ and $\mathcal{B}$, one can write the state equation (1.1) as follows:

$$(2.10) \qquad X = x + \mathcal{A}X + \mathcal{B}u.$$

We now have the following well-posedness of the state equation.

PROPOSITION 2.6. *Let* (H1) *hold. Then for any $(x, u(\cdot)) \in \mathbb{R}^n \times \mathcal{U}[0,T]$, state equation* (1.1) *admits a unique solution $X(\cdot) \equiv X(\cdot; x, u(\cdot)) \in \widehat{\mathcal{X}}[0,T]$.*

Proof. For any $X(\cdot) \in \mathcal{X}[0,T]$ and $u(\cdot) \in \mathcal{U}[0,T]$, by (2.6), we have

$$\mathbb{E} \left[\sup_{s \in [0,t]} |(\mathcal{A}X)(s)|^2 \right] \le \alpha(t) \mathbb{E} \left[\sup_{s \in [0,t]} |X(s)|^2 \right]$$

with $\alpha(t) \in (0,1)$ when $t > 0$ is small. Hence, by the contraction mapping theorem, we obtain the well-posedness of the state equation on $[0,t]$. Then by a usual continuation argument, we obtain the well-posedness of the state equation on $[0,T]$. □

Proposition 2.6 tells us that under (H1), the operator $I - \mathcal{A} : \widehat{\mathcal{X}}[0,T] \to \widehat{\mathcal{X}}[0,T]$ is invertible and the solution X to the state equation corresponding to $(x, u(\cdot)) \in \mathbb{R}^n \times \mathcal{U}[0,T]$ is given by

$$X = (I - \mathcal{A})^{-1} x + (I - \mathcal{A})^{-1} \mathcal{B}u.$$

Note that

$$I_T \left[(I - \mathcal{A})^{-1} x + (I - \mathcal{A})^{-1} \mathcal{B}u \right] = I_T X = X(T) = x + \mathcal{A}_T X + \mathcal{B}_T u$$
$$= \left[I + \mathcal{A}_T (I - \mathcal{A})^{-1} \right] x + \left[\mathcal{A}_T (I - \mathcal{A})^{-1} \mathcal{B} + \mathcal{B}_T \right] u.$$

Therefore,

$$I_T (I - \mathcal{A})^{-1} = I + \mathcal{A}_T (I - \mathcal{A})^{-1}, \qquad I_T (I - \mathcal{A})^{-1} \mathcal{B} = \mathcal{A}_T (I - \mathcal{A})^{-1} \mathcal{B} + \mathcal{B}_T.$$

Now, let

$$\begin{cases} [\mathcal{Q}X(\cdot)](s) = Q(s)X(s), & s \in [0,T], \quad \forall X(\cdot) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^n), \\ [\bar{\mathcal{Q}}\varphi(\cdot)](s) = \bar{Q}(s)\varphi(s), & s \in [0,T], \quad \forall \varphi(\cdot) \in L^2(0,T;\mathbb{R}^n), \\ [\mathcal{R}u(\cdot)](s) = R(s)u(s), & s \in [0,T], \quad \forall u(\cdot) \in \mathcal{U}[0,T], \\ [\bar{\mathcal{R}}\varphi(\cdot)](s) = \bar{R}(s)\varphi(s), & s \in [0,T], \quad \forall \varphi(\cdot) \in L^2(0,T;\mathbb{R}^m), \\ \mathcal{G}\xi = G\xi \quad \forall \xi \in \mathcal{X}_T, \qquad \bar{\mathcal{G}}x = \bar{G}x \quad \forall x \in \mathbb{R}^n. \end{cases}$$

Then the cost functional can be written as

$$J(x; u(\cdot)) = \langle \mathcal{Q}X, X \rangle + \langle \bar{\mathcal{Q}}\mathbb{E}X, \mathbb{E}X \rangle + \langle \mathcal{R}u, u \rangle + \langle \bar{\mathcal{R}}\mathbb{E}u, \mathbb{E}u \rangle$$
$$+ \langle \mathcal{G}X(T), X(T) \rangle + \langle \bar{\mathcal{G}}\mathbb{E}X(T), \mathbb{E}X(T) \rangle$$
$$\equiv \langle \Theta_2 u, u \rangle + 2 \langle \Theta_1 x, u \rangle + \langle \Theta_0 x, x \rangle,$$

where

$$\begin{cases} \Theta_2 = \mathcal{R} + \mathbb{E}^* \bar{\mathcal{R}} \mathbb{E} + \mathcal{B}^* (I - \mathcal{A}^*)^{-1} (\mathcal{Q} + \mathbb{E}^* \bar{\mathcal{Q}} \mathbb{E})(I - \mathcal{A})^{-1} \mathcal{B} \\ \qquad + [\mathcal{B}^* (I - \mathcal{A}^*)^{-1} \mathcal{A}_T^* + \mathcal{B}_T^*](\mathcal{G} + \mathbb{E}^* \bar{\mathcal{G}} \mathbb{E})[\mathcal{A}_T (I - \mathcal{A})^{-1} \mathcal{B} + \mathcal{B}_T], \\ \Theta_1 = \mathcal{B}^* (I - \mathcal{A}^*)^{-1} (\mathcal{Q} + \mathbb{E}^* \bar{\mathcal{Q}} \mathbb{E})(I - \mathcal{A})^{-1} \\ \qquad + [\mathcal{B}^* (I - \mathcal{A}^*)^{-1} \mathcal{A}_T^* + \mathcal{B}_T^*](\mathcal{G} + \mathbb{E}^* \bar{\mathcal{G}} \mathbb{E})[I + \mathcal{A}_T (I - \mathcal{A})^{-1}], \\ \Theta_0 = (I - \mathcal{A}^*)^{-1} (\mathcal{Q} + \mathbb{E}^* \bar{\mathcal{Q}} \mathbb{E})(I - \mathcal{A})^{-1} \\ \qquad + [I + (I - \mathcal{A}^*)^{-1} \mathcal{A}_T^*](\mathcal{G} + \mathbb{E}^* \bar{\mathcal{G}} \mathbb{E})[I + \mathcal{A}_T (I - \mathcal{A})^{-1}]. \end{cases}$$

Consequently, for any $u(\cdot), v(\cdot) \in \mathcal{U}[0, T]$, and $x \in \mathbb{R}^n$,

$$J(x; v(\cdot)) = J(x; u(\cdot) + [v(\cdot) - u(\cdot)])$$
$$= \langle \Theta_2 [u + (v - u)], u + (v - u) \rangle + 2 \langle \Theta_1 x, u + (v - u) \rangle + \langle \Theta_0 x, x \rangle$$
$$= \langle \Theta_2 u, u \rangle + 2 \langle \Theta_1 x, u \rangle + \langle \Theta_0 x, x \rangle + 2 \langle \Theta_2 u + \Theta_1 x, v - u \rangle + \langle \Theta_2 (v - u), v - u \rangle$$
$$= J(x; u(\cdot)) + 2 \langle \Theta_2 u + \Theta_1 x, v - u \rangle + \langle \Theta_2 (v - u), v - u \rangle.$$

We now present the following result, whose proof is standard, making use of the above (see [31]).

PROPOSITION 2.7. *If $u(\cdot) \mapsto J(x; u(\cdot))$ admits a minimum, then*

$$(2.11) \qquad\qquad\qquad\qquad \Theta_2 \geq 0.$$

Conversely, if in addition to (2.11), one has

$$\Theta_1 x \in \Theta_2 \Big(\mathcal{U}[0, T] \Big),$$

then $u(\cdot) \mapsto J(x; u(\cdot))$ admits a minimum $u^*(\cdot) \in \mathcal{U}[0, T]$. Further, if

$$(2.12) \qquad\qquad\qquad\qquad \Theta_2 \geq \delta I$$

for some $\delta > 0$, then for any given $x \in \mathbb{R}^n$, $u(\cdot) \mapsto J(x; u(\cdot))$ admits a unique minimum.

By the definition of Θ_2, we see that (2.11) is implied by

$$(2.13) \qquad \mathcal{R} + \mathbb{E}^* \bar{\mathcal{R}} \mathbb{E} \geq 0, \quad \mathcal{Q} + \mathbb{E}^* \bar{\mathcal{Q}} \mathbb{E} \geq 0, \quad \mathcal{G} + \mathbb{E}^* \bar{\mathcal{G}} \mathbb{E} \geq 0$$

and (2.12) is implied by

$$(2.14) \qquad \mathcal{R} + \mathbb{E}^* \bar{\mathcal{R}} \mathbb{E} \geq \delta I, \quad \mathcal{Q} + \mathbb{E}^* \bar{\mathcal{Q}} \mathbb{E} \geq 0, \quad \mathcal{G} + \mathbb{E}^* \bar{\mathcal{G}} \mathbb{E} \geq 0$$

for some $\delta > 0$. Now, we would like to present more direct conditions under which (2.12) holds.

PROPOSITION 2.8. *Let (H1)–(H2) hold. Then (2.12) holds and problem (MF-LQ) admits a unique solution.*

Proof. For any $\xi \in \mathcal{X}_T$,

$$\mathbb{E}\Big[\langle G\xi, \xi\rangle + \langle \bar{G}\mathbb{E}[\xi], \mathbb{E}[\xi]\rangle\Big]$$
$$= \mathbb{E}\Big[\langle G(\xi - \mathbb{E}[\xi]), \xi - \mathbb{E}[\xi]\rangle + \langle (G + \bar{G})\mathbb{E}[\xi], \mathbb{E}[\xi]\rangle\Big] \geq 0,$$
$$\mathbb{E}\Big[\langle Q(s)\xi, \xi\rangle + \langle \bar{Q}(s)\mathbb{E}[\xi], \mathbb{E}[\xi]\rangle\Big]$$
$$= \mathbb{E}\Big[\langle Q(s)(\xi - \mathbb{E}[\xi]), \xi - \mathbb{E}[\xi]\rangle + \langle [Q(s) + \bar{Q}(s)]\mathbb{E}[\xi], \mathbb{E}[\xi]\rangle\Big] \geq 0,$$

and for any $\eta \in \mathcal{U}_T$,

$$\mathbb{E}\Big[\langle R(s)\eta, \eta\rangle + \langle \bar{R}(s)\mathbb{E}[\eta], \mathbb{E}[\eta]\rangle\Big]$$
$$= \mathbb{E}\Big[\langle R(s)(\eta - \mathbb{E}[\eta]), \eta - \mathbb{E}[\eta]\rangle + \langle [R(s) + \bar{R}(s)]\mathbb{E}[\eta], \mathbb{E}[\eta]\rangle\Big]$$
$$\geq \delta\mathbb{E}\Big[|\eta - \mathbb{E}[\eta]|^2 + |\mathbb{E}[\eta]|^2\Big] = \delta\mathbb{E}|\eta|^2.$$

Hence, (2.12) holds. $\quad\square$

3. Optimality conditions. In this section, we first derive a necessary condition for an optimal pair of problem (MF-LQ).

THEOREM 3.1. *Let* (H1)–(H2) *hold. Let* $(X^*(\cdot), u^*(\cdot))$ *be the optimal pair of problem* (MF-LQ). *Then the following mean-field backward SDE (MF-BSDE) admits a unique adapted solution* $(Y(\cdot), Z(\cdot))$:

$$(3.1) \quad \begin{cases} dY(s) = -\Big(A(s)^T Y(s) + \bar{A}(s)^T \mathbb{E}[Y(s)] + C(s)^T Z(s) + \bar{C}(s)^T \mathbb{E}[Z(s)] \\ \qquad\qquad + Q(s)X^*(s) + \bar{Q}(s)\mathbb{E}[X^*(s)]\Big)ds + Z(s)dW(s), \qquad s \in [0, T], \\ Y(T) = GX^*(T) + \bar{G}\mathbb{E}[X^*(T)] \end{cases}$$

such that

$$(3.2) \quad \begin{aligned} R(s)u^*(s) + \bar{R}(s)\mathbb{E}[u^*(s)] + B(s)^T Y(s) + \bar{B}(s)^T \mathbb{E}[Y(s)] \\ + D(s)^T Z(s) + \bar{D}(s)^T \mathbb{E}[Z(s)] = 0, \qquad s \in [0, T], \quad \text{a.s.} \end{aligned}$$

Proof. Let $(X^*(\cdot), u^*(\cdot))$ be an optimal pair of problem (MF-LQ). For any $u(\cdot) \in \mathcal{U}[0, T]$, let $X(\cdot)$ be the state process corresponding to the zero initial condition and control $u(\cdot)$. Then we have

$$0 = \mathbb{E}\Bigg\{\int_0^T \Big[\langle Q(s)X^*(s), X(s)\rangle + \langle \bar{Q}(s)\mathbb{E}[X^*(s)], \mathbb{E}[X(s)]\rangle$$
$$+ \langle R(s)u^*(s), u(s)\rangle + \langle \bar{R}(s)\mathbb{E}[u^*(s)], \mathbb{E}[u(s)]\rangle\Big]ds$$
$$+ \langle GX^*(T), X(T)\rangle + \langle \bar{G}\mathbb{E}[X^*(T)], \mathbb{E}[X(T)]\rangle\Bigg\}$$
$$= \mathbb{E}\Bigg\{\int_0^T \Big[\langle Q(s)X^*(s) + \bar{Q}(s)\mathbb{E}[X^*(s)], X(s)\rangle$$
$$+ \langle R(s)u^*(s) + \bar{R}(s)\mathbb{E}[u^*(s)], u(s)\rangle\Big]ds + \langle GX^*(T) + \bar{G}\mathbb{E}[X^*(T)], X(T)\rangle\Bigg\}.$$

Next, by [10] and [11], MF-BSDE (3.1) admits a unique adapted solution $(Y(\cdot), Z(\cdot))$. Then

$$\mathbb{E}\langle X(T), Y(T)\rangle$$
$$= \mathbb{E}\int_0^T \Big(\langle A(s)X(s) + \bar{A}(s)\mathbb{E}[X(s)] + B(s)u(s) + \bar{B}(s)\mathbb{E}[u(s)], Y(s)\rangle$$
$$- \langle X(s), A(s)^T Y(s) + \bar{A}(s)^T\mathbb{E}[Y(s)] + C(s)^T Z(s)$$
$$+ \bar{C}(s)^T\mathbb{E}[Z(s)] + Q(s)X^*(s) + \bar{Q}(s)\mathbb{E}[X^*(s)]\rangle$$
$$+ \langle C(s)X(s) + \bar{C}(s)\mathbb{E}[X(s)] + D(s)u(s) + \bar{D}(s)\mathbb{E}[u(s)], Z(s)\rangle\Big)ds$$
$$= \mathbb{E}\int_0^T \Big(-\langle Q(s)X^*(s), X(s)\rangle - \langle\bar{Q}(s)\mathbb{E}[X^*(s)], \mathbb{E}[X(s)]\rangle$$
$$+ \langle u(s), B(s)^T Y(s) + \bar{B}(s)^T\mathbb{E}[Y(s)] + D(s)^T Z(s) + \bar{D}(s)^T\mathbb{E}[Z(s)]\rangle\Big)ds.$$

Hence,

$$0 = \mathbb{E}\left\{\int_0^T \Big[\langle Q(s)X^*(s) + \bar{Q}(s)\mathbb{E}[X^*(s)], X(s)\rangle + \langle R(s)u^*(s) + \bar{R}(s)\mathbb{E}[u^*(s)], u(s)\rangle\Big]ds\right.$$
$$\left. + \langle GX^*(T) + \bar{G}\mathbb{E}[X^*(T)], X(T)\rangle\right\}$$
$$= \mathbb{E}\int_0^T \langle u(s), R(s)u^*(s) + \bar{R}(s)\mathbb{E}[u^*(s)] + B(s)^T Y(s) + \bar{B}(s)^T\mathbb{E}[Y(s)]$$
$$+ D(s)^T Z(s) + \bar{D}(s)^T\mathbb{E}[Z(s)]\rangle\, ds.$$

Therefore, the optimal control satisfies

$$R(s)u^*(s) + \mathbb{E}[\bar{R}(s)]\mathbb{E}[u^*(s)] + B(s)^T Y(s) + \mathbb{E}[\bar{B}(s)^T Y(s)]$$
$$+ D(s)^T Z(s) + \mathbb{E}[\bar{D}(s)^T Z(s)] = 0.$$

This completes the proof.   □

From the above, we end up with the following optimality system (with s suppressed):

$$(3.3)\quad\begin{cases} dX^* = \Big(AX^* + \bar{A}\mathbb{E}[X^*] + Bu^* + \bar{B}\mathbb{E}[u^*]\Big)ds \\ \qquad\quad + \Big(CX^* + \bar{C}\mathbb{E}[X^*] + Du^* + \bar{D}\mathbb{E}[u^*]\Big)dW(s), \\ dY = -\Big(A^T Y + \bar{A}^T\mathbb{E}[Y] + C^T Z + \bar{C}^T\mathbb{E}[Z] \\ \qquad\quad + QX^* + \bar{Q}\mathbb{E}[X^*]\Big)ds + ZdW(s), \\ X^*(0) = x, \qquad Y(T) = GX^*(T) + \bar{G}\mathbb{E}[X^*(T)], \\ Ru^* + \bar{R}\mathbb{E}[u^*] + B^T Y + \bar{B}^T\mathbb{E}[Y] + D^T Z + \bar{D}^T\mathbb{E}[Z] = 0. \end{cases}$$

This is the MF-FBSDE. Note that the coupling comes from the last relation (which is essentially the maximum condition in the usual Pontryagin type maximum principle). The 4-tuple $(X^*(\cdot), u^*(\cdot), Y(\cdot), Z(\cdot))$ of $\mathbb{F}$-adapted processes satisfying the above is called an *adapted solution* of (3.3). We now look at the sufficiency of the above result.

THEOREM 3.2. *Let* (H1)–(H2) *hold. Suppose* $(X^*(\cdot), u^*(\cdot), Y(\cdot), Z(\cdot))$ *is an adapted solution to the MF-FBSDE* (3.3). *Then* $(X^*(\cdot), u^*(\cdot))$ *is the optimal pair.*

Proof. First, we know that under (H1)–(H2), one has (2.12). Let $(X^*(\cdot), u^*(\cdot), Y(\cdot), Z(\cdot))$ be an adapted solution to the MF-FBSDE. For any $u(\cdot) \in \mathcal{U}[0, T]$, let

$$X_1(\cdot) \equiv X(\cdot\,; 0, u(\cdot) - u^*(\cdot)).$$

Then

$$X(s; x, u(\cdot)) = X^*(s) + X_1(s), \qquad s \in [0, T].$$

Hence (suppressing s),

$$
\begin{aligned}
&J(x; u(\cdot)) - J(x; u^*(\cdot)) \\
&\quad = 2\mathbb{E}\bigg\{ \int_0^T \Big[\langle QX^*, X_1 \rangle + \langle \bar{Q}\mathbb{E}[X^*], \mathbb{E}[X_1] \rangle \\
&\qquad\qquad + \langle Ru^*, u - u^* \rangle + \langle \bar{R}\mathbb{E}[u^*], \mathbb{E}[u - u^*] \rangle \Big] ds \\
&\qquad\quad + \langle GX^*(T), X_1(T) \rangle + \langle \bar{G}\mathbb{E}[X^*(T)], \mathbb{E}[X_1(T)] \rangle \bigg\} \\
&\qquad + \mathbb{E}\bigg\{ \int_0^T \Big[\langle QX_1, X_1 \rangle + \langle \bar{Q}\mathbb{E}[X_1], \mathbb{E}[X_1] \rangle \\
&\qquad\qquad + \langle R(u - u^*), u - u^* \rangle + \langle \bar{R}\mathbb{E}[u - u^*], \mathbb{E}[u - u^*] \rangle \Big] ds \\
&\qquad\quad + \langle GX_1(T), X_1(T) \rangle + \langle \bar{G}\mathbb{E}[X_1(T)], \mathbb{E}[X_1(T)] \rangle \bigg\} \\
&\quad = 2\mathbb{E}\bigg\{ \int_0^T \Big[\langle X_1, QX^* + \bar{Q}\mathbb{E}[X^*] \rangle + \langle u - u^*, Ru^* + \bar{R}\mathbb{E}[u^*] \rangle \Big] ds \\
&\qquad\quad + \langle X_1(T), GX^*(T) + \bar{G}\mathbb{E}[X^*(T)] \rangle \bigg\} + J(0; u(\cdot) - u^*(\cdot)).
\end{aligned}
$$

Note that

$$
\begin{aligned}
&\mathbb{E}\langle X_1(T), GX^*(T) + \bar{G}\mathbb{E}[X^*(T)] \rangle = \mathbb{E}\langle X_1(T), Y(T) \rangle \\
&\quad = \mathbb{E}\int_0^T (\langle AX_1 + B(u - u^*) + \bar{A}\mathbb{E}[X_1] + \bar{B}\mathbb{E}[u - u^*], Y \rangle \\
&\qquad\quad - \langle X_1, A^T Y + C^T Z + \bar{A}^T \mathbb{E}[Y] + \bar{C}^T \mathbb{E}[Z] + QX^* + \bar{Q}\mathbb{E}[X^*] \rangle \\
&\qquad\quad + \langle CX_1 + D(u - u^*) + \bar{C}\mathbb{E}[X_1] + \bar{D}\mathbb{E}[u - u^*], Z \rangle) ds \\
&\quad = \mathbb{E}\int_0^T (- \langle X_1, QX^* + \bar{Q}\mathbb{E}[X^*] \rangle \\
&\qquad\quad + \langle u - u^*, B^T Y + D^T Z + \bar{B}^T \mathbb{E}[Y] + \bar{D}^T \mathbb{E}[Z] \rangle) ds.
\end{aligned}
$$

Thus, by (2.12),

$$
\begin{aligned}
&J(x; u(\cdot)) - J(x; u^*(\cdot)) \\
&\quad = 2\mathbb{E}\int_0^T \langle u - u^*, Ru^* + \bar{R}\mathbb{E}[u^*] + B^T Y + D^T Z + \bar{B}^T \mathbb{E}[Y] + \bar{D}^T \mathbb{E}[Z] \rangle ds \\
&\qquad + J(0; u(\cdot) - u^*(\cdot)) \\
&\quad = J(0; u(\cdot) - u^*(\cdot)) = \langle \Theta_2(u - u^*), u - u^* \rangle \geq 0.
\end{aligned}
$$

Hence, $(X^*(\cdot), u^*(\cdot))$ is optimal. □

We have the following corollary.

COROLLARY 3.3. *Let* (H1)–(H2) *hold. Then MF-FBSDE* (3.3) *admits a unique adapted solution* $(X^*(\cdot), u^*(\cdot), Y(\cdot), Z(\cdot))$ *of which* $(X^*(\cdot), u^*(\cdot))$ *is the unique optimal pair of problem* (MF-LQ).

Proof. We know from Proposition 2.8 that under (H1) and (H2), problem (MF-LQ) admits a unique optimal pair $(X^*(\cdot), u^*(\cdot))$. Then by Theorem 3.1, for some $(Y(\cdot), Z(\cdot))$, the 4-tuple $(X^*(\cdot), u^*(\cdot), Y(\cdot), Z(\cdot))$ is an adapted solution to MF-FBSDE (3.3). Next, if (3.3) has another adapted solution $(\widetilde{X}(\cdot), \widetilde{u}(\cdot), \widetilde{Y}(\cdot), \widetilde{Z}(\cdot))$, then by Theorem 3.2, $(\widetilde{X}(\cdot), \widetilde{u}(\cdot))$ must be an optimal pair of problem (MF-LQ). Hence, by the uniqueness of the optimal pair of problem (MF-LQ), we must have

$$\widetilde{X}(\cdot) = X^*(\cdot), \qquad \widetilde{u}(\cdot) = u^*(\cdot).$$

Then by the uniqueness of MF-BSDE (3.1), we must have

$$\widetilde{Y}(\cdot) = Y(\cdot), \qquad \widetilde{Z}(\cdot) = Z(\cdot),$$

proving the corollary. $\square$

4. Decoupling the MF-FBSDE and Riccati equations. From Corollary 3.3, we see that under (H1) and (H2), MF-FBSDE (3.3) admits a unique adapted solution $(X^*(\cdot), u^*(\cdot), Y(\cdot), Z(\cdot))$ and $(X^*(\cdot), u^*(\cdot))$ is the optimal pair that we are looking for. Note that $u^*(\cdot)$ is represented in terms of $X^*(\cdot)$ and $(Y(\cdot), Z(\cdot))$ at the moment. We now would like to obtain a representation of $u^*(\cdot)$ only in terms of $X^*(\cdot)$. To this end, we use the idea of decoupling inspired by the four-step scheme introduced in [26, 27] for general FBSDEs. This will lead to a derivation of our Riccati equations. We now carry out the detailed derivation.

Let $(X^*(\cdot), u^*(\cdot), Y(\cdot), Z(\cdot))$ be the adapted solution to MF-FBSDE (3.3). Note that (suppressing s)

$$d\mathbb{E}[X^*] = \Big((A + \bar{A})\mathbb{E}[X^*] + (B + \bar{B})\mathbb{E}[u^*]\Big)ds.$$

Hence,

$$\begin{aligned} d\Big(X^* - \mathbb{E}[X^*]\Big) &= \Big(A(X^* - \mathbb{E}[X^*]) + B(u^* - \mathbb{E}[u^*])\Big)ds \\ &\quad + \Big(C(X^* - \mathbb{E}[X^*]) + (C + \bar{C})\mathbb{E}[X^*] + D(u^* - \mathbb{E}[u^*]) \\ &\quad + (D + \bar{D})\mathbb{E}[u^*]\Big)dW(s). \end{aligned}$$

Now, we assume that

$$(4.1) \qquad Y(s) = P(s)\Big(X^*(s) - \mathbb{E}[X^*(s)]\Big) + \Pi(s)\mathbb{E}[X^*(s)], \qquad s \in [0, T],$$

for some deterministic and differentiable functions $P(\cdot)$ and $\Pi(\cdot)$, taking values in $\mathcal{S}^n$, such that

$$P(T) = G, \qquad \Pi(T) = G + \bar{G}.$$

Then

$$\mathbb{E}[Y(s)] = \Pi(s)\mathbb{E}[X^*(s)]$$

and

$$Y(s) - \mathbb{E}[Y(s)] = P(s)\Big(X^*(s) - \mathbb{E}[X^*(s)]\Big).$$

Therefore,

$$
\begin{aligned}
\Big(& -A^T(Y - \mathbb{E}[Y]) - (A + \bar{A})^T \mathbb{E}[Y] - C^T(Z - \mathbb{E}[Z]) - (C + \bar{C})^T \mathbb{E}[Z] \\
& -Q(X^* - \mathbb{E}[X^*]) - (Q + \bar{Q})\mathbb{E}[X^*] \Big) ds + Z dW(s) = dY \\
& = d\Big[P(X^* - \mathbb{E}[X^*]) + \Pi \mathbb{E}[X^*] \Big] \\
& = \Big[\dot{P}(X^* - \mathbb{E}[X^*]) + P\Big(A(X^* - \mathbb{E}[X^*]) + B(u^* - \mathbb{E}[u^*]) \Big) \\
& \quad + \dot{\Pi}\mathbb{E}[X^*] + \Pi\Big((A + \bar{A})\mathbb{E}[X^*] + (B + \bar{B})\mathbb{E}[u^*] \Big) \Big] ds \\
& \quad + P\Big(C(X^* - \mathbb{E}[X^*]) + (C + \bar{C})\mathbb{E}[X^*] \\
& \quad + D(u^* - \mathbb{E}[u^*]) + (D + \bar{D})\mathbb{E}[u^*] \Big) dW(s) \\
& = \Big[(\dot{P} + PA)(X^* - \mathbb{E}[X^*]) + PB(u^* - \mathbb{E}[u^*]) \\
& \quad + \Big(\dot{\Pi} + \Pi(A + \bar{A}) \Big) \mathbb{E}[X^*] + \Pi(B + \bar{B})\mathbb{E}[u^*] \Big] ds \\
& \quad + P\Big(C(X^* - \mathbb{E}[X^*]) + (C + \bar{C})\mathbb{E}[X^*] \\
& \quad + D(u^* - \mathbb{E}[u^*]) + (D + \bar{D})\mathbb{E}[u^*] \Big) dW(s).
\end{aligned}
\tag{4.2}
$$

Comparing the diffusion terms, we should have

$$
\begin{aligned}
Z = P\Big(& C(X^* - \mathbb{E}[X^*]) + (C + \bar{C})\mathbb{E}[X^*] + D(u^* - \mathbb{E}[u^*]) \\
& + (D + \bar{D})\mathbb{E}[u^*] \Big).
\end{aligned}
\tag{4.3}
$$

Then

$$
\mathbb{E}[Z] = P\Big((C + \bar{C})\mathbb{E}[X^*] + (D + \bar{D})\mathbb{E}[u^*] \Big)
$$

and

$$
Z - \mathbb{E}[Z] = P\Big(C(X^* - \mathbb{E}[X^*]) + D(u^* - \mathbb{E}[u^*]) \Big).
$$

These yield, taking into account (3.2),

$$
\begin{aligned}
0 & = R(u^* - \mathbb{E}[u^*]) + (R + \bar{R})\mathbb{E}[u^*] + B^T(Y - \mathbb{E}[Y]) + (B + \bar{B})^T \mathbb{E}[Y] \\
& \quad + D^T(Z - \mathbb{E}[Z]) + (D + \bar{D})^T \mathbb{E}[Z] \\
& = R(u^* - \mathbb{E}[u^*]) + (R + \bar{R})\mathbb{E}[u^*] + B^T P(X^* - \mathbb{E}[X^*]) + (B + \bar{B})^T \Pi \mathbb{E}[X^*] \\
& \quad + D^T P\Big(C(X^* - \mathbb{E}[X^*]) + D(u^* - \mathbb{E}[u^*]) \Big) \\
& \quad + (D + \bar{D})^T P\Big((C + \bar{C})\mathbb{E}[X^*] + (D + \bar{D})\mathbb{E}[u^*] \Big) \\
& = (R + D^T PD)(u^* - \mathbb{E}[u^*]) + (R + \bar{R} + (D + \bar{D})^T P(D + \bar{D})\Big) \mathbb{E}[u^*] \\
& \quad + (B^T P + D^T PC)(X^* - \mathbb{E}[X^*]) \\
& \quad + \Big((B + \bar{B})^T \Pi + (D + \bar{D})^T P(C + \bar{C}) \Big) \mathbb{E}[X^*] \\
& \equiv \Sigma_0(u^* - \mathbb{E}[u^*]) + \Sigma_1 \mathbb{E}[u^*] + (B^T P + D^T PC)(X^* - \mathbb{E}[X^*]) \\
& \quad + \Big((B + \bar{B})^T \Pi + (D + \bar{D})^T P(C + \bar{C}) \Big) \mathbb{E}[X^*],
\end{aligned}
$$

where we denote

$$\Sigma_0 = R + D^T P D, \qquad \Sigma_1 = R + \bar{R} + (D + \bar{D})^T P(D + \bar{D}).$$

Applying $\mathbb{E}$ to the above, we obtain

$$0 = \Sigma_1 \mathbb{E}[u^*] + \Big[(B + \bar{B})^T \Pi + (D + \bar{D})^T P(C + \bar{C})\Big] \mathbb{E}[X^*],$$

which implies

$$0 = \Sigma_0(u^* - \mathbb{E}[u^*]) + (B^T P + D^T P C)(X^* - \mathbb{E}[X^*]).$$

Assuming Σ_0 and Σ_1 to be invertible, one gets

$$\mathbb{E}[u^*] = -\Sigma_1^{-1}\Big[(B + \bar{B})^T \Pi + (D + \bar{D})^T P(C + \bar{C})\Big] \mathbb{E}[X^*]$$

and

$$
\begin{aligned}
(4.4) \qquad u^* &= \mathbb{E}[u^*] - \Sigma_0^{-1}(B^T P + D^T P C)(X^* - \mathbb{E}[X^*]) \\
&= -\Sigma_0^{-1}(B^T P + D^T P C)(X^* - \mathbb{E}[X^*]) \\
&\quad - \Sigma_1^{-1}\Big[(B + \bar{B})^T \Pi + (D + \bar{D})^T P(C + \bar{C})\Big] \mathbb{E}[X^*].
\end{aligned}
$$

Next, comparing the drift terms in (4.2), we have

$$
\begin{aligned}
0 &= (\dot{P} + PA)(X^* - \mathbb{E}[X^*]) + PB(u^* - \mathbb{E}[u^*]) \\
&\quad + \Big(\dot{\Pi} + \Pi(A + \bar{A})\Big)\mathbb{E}[X^*] + \Pi(B + \bar{B})\mathbb{E}[u^*] \\
&\quad + A^T(Y - \mathbb{E}[Y]) + (A + \bar{A})^T \mathbb{E}[Y] + C^T(Z - \mathbb{E}[Z]) \\
&\quad + (C + \bar{C})^T \mathbb{E}[Z] + Q(X^* - \mathbb{E}[X^*]) + (Q + \bar{Q})\mathbb{E}[X^*] \\
&= (\dot{P} + PA)(X^* - \mathbb{E}[X^*]) + PB(u^* - \mathbb{E}[u^*]) \\
&\quad + \Big(\dot{\Pi} + \Pi(A + \bar{A})\Big)\mathbb{E}[X^*] + \Pi(B + \bar{B})\mathbb{E}[u^*] \\
&\quad + A^T P(X^* - \mathbb{E}[X^*]) + (A + \bar{A})^T \Pi \mathbb{E}[X^*] \\
&\quad + C^T P\Big(C(X^* - \mathbb{E}[X^*]) + D(u^* - \mathbb{E}[u^*])\Big) \\
&\quad + (C + \bar{C})^T P\Big((C + \bar{C})\mathbb{E}[X^*] + (D + \bar{D})\mathbb{E}[u^*]\Big) \\
&\quad + Q(X^* - \mathbb{E}[X^*]) + (Q + \bar{Q})\mathbb{E}[X^*] \\
&= \Big[\dot{P} + PA + A^T P + C^T P C + Q\Big](X^* - \mathbb{E}[X^*]) \\
&\quad + (PB + C^T P D)(u^* - \mathbb{E}[u^*]) \\
&\quad + \Big[\dot{\Pi} + \Pi(A + \bar{A}) + (A + \bar{A})^T \Pi + (C + \bar{C})^T P(C + \bar{C}) + Q + \bar{Q}\Big]\mathbb{E}[X^*] \\
&\quad + \Big[\Pi(B + \bar{B}) + (C + \bar{C})^T P(D + \bar{D})\Big]\mathbb{E}[u^*] \\
&= \Big[\dot{P} + PA + A^T P + C^T P C + Q\Big](X^* - \mathbb{E}[X^*]) \\
&\quad - (PB + C^T P D)\Sigma_0^{-1}(B^T P + D^T P C)(X^* - \mathbb{E}[X^*]) \\
&\quad + \Big[\dot{\Pi} + \Pi(A + \bar{A}) + (A + \bar{A})^T \Pi + (C + \bar{C})^T P(C + \bar{C}) + Q + \bar{Q}\Big]\mathbb{E}[X^*] \\
&\quad - \Big[\Pi(B + \bar{B}) + (C + \bar{C})^T P(D + \bar{D})\Big]\Sigma_1^{-1} \\
&\qquad \cdot \Big[(B + \bar{B})^T \Pi + (D + \bar{D})^T P(C + \bar{C})\Big]\mathbb{E}[X^*].
\end{aligned}
$$

$$
\begin{aligned}
= &\Big[\dot{P} + PA + A^T P + C^T PC + Q \\
&\quad - (PB + C^T PD)\Sigma_0^{-1}(B^T P + D^T PC)\Big](X^* - \mathbb{E}[X^*]) \\
&+ \Big\{\dot{\Pi} + \Pi(A + \bar{A}) + (A + \bar{A})^T \Pi + (C + \bar{C})^T P(C + \bar{C}) + Q + \bar{Q} \\
&\quad - \Big[\Pi(B + \bar{B}) + (C + \bar{C})^T P(D + \bar{D})\Big]\Sigma_1^{-1} \\
&\qquad \cdot \Big[(B + \bar{B})^T \Pi + (D + \bar{D})^T P(C + \bar{C})\Big]\Big\}\mathbb{E}[X^*].
\end{aligned}
$$

Therefore, we should let $P(\cdot)$ and $\Pi(\cdot)$ be the solutions to the following Riccati equations, respectively:

$$
(4.5) \qquad
\begin{cases}
\dot{P} + PA + A^T P + C^T PC + Q \\
\qquad - (PB + C^T PD)\Sigma_0^{-1}(B^T P + D^T PC) = 0, \\
P(T) = G,
\end{cases}
$$

and

$$
(4.6) \qquad
\begin{cases}
\dot{\Pi} + \Pi(A + \bar{A}) + (A + \bar{A})^T \Pi + (C + \bar{C})^T P(C + \bar{C}) + Q + \bar{Q} \\
\qquad - \Big[\Pi(B + \bar{B}) + (C + \bar{C})^T P(D + \bar{D})\Big]\Sigma_1^{-1} \\
\qquad\qquad \cdot \Big[(B + \bar{B})^T \Pi + (D + \bar{D})^T P(C + \bar{C})\Big] = 0, \\
\Pi(T) = G + \bar{G}.
\end{cases}
$$

It is standard [38] that Riccati equation (4.5) admits a unique solution $P(\cdot)$ which is positive definite since (2.3) holds. For the solvability of (4.6), we first rewrite it as the following:

$$
(4.7) \qquad
\begin{cases}
\dot{\Pi} + \Pi\Big[A + \bar{A} - (B + \bar{B})\Sigma_1^{-1}(D + \bar{D})^T P(C + \bar{C})\Big] \\
\qquad + \Big[(A + \bar{A})^T + (C + \bar{C})^T P(D + \bar{D})\Sigma_1^{-1}(B + \bar{B})^T\Big]\Pi \\
\qquad + (C + \bar{C})^T[P - P(D + \bar{D})\Sigma_1^{-1}(D + \bar{D})^T P](C + \bar{C}) \\
\qquad + Q + \bar{Q} - \Pi(B + \bar{B})\Sigma_1^{-1}(B + \bar{B})^T \Pi = 0, \\
\Pi(T) = G + \bar{G}.
\end{cases}
$$

Note that

$$
\begin{aligned}
&P - P(D + \bar{D})\Sigma_1^{-1}(D + \bar{D})^T P \\
&= P - P(D + \bar{D})\Big[R + \bar{R} + (D + \bar{D})^T P(D + \bar{D})\Big]^{-1}(D + \bar{D})^T P \\
&\equiv P - P\widetilde{D}(\widetilde{R} + \widetilde{D}^T P\widetilde{D})^{-1}\widetilde{D}^T P \\
&= P^{\frac{1}{2}}\Big[I - P^{\frac{1}{2}}\widetilde{D}\widetilde{R}^{-\frac{1}{2}}\Big(I + \widetilde{R}^{-\frac{1}{2}}\widetilde{D}^T P^{\frac{1}{2}}P^{\frac{1}{2}}\widetilde{D}\widetilde{R}^{-\frac{1}{2}}\Big)^{-1}\widetilde{R}^{-\frac{1}{2}}\widetilde{D}^T P^{\frac{1}{2}}\Big]P^{\frac{1}{2}} \\
&\equiv P^{\frac{1}{2}}\Big[I - \Gamma(I + \Gamma^T \Gamma)^{-1}\Gamma^T\Big]P^{\frac{1}{2}} \\
&= P^{\frac{1}{2}}(I + \Gamma\Gamma^T)^{-1}P^{\frac{1}{2}} \equiv P^{\frac{1}{2}}(I + P^{\frac{1}{2}}\widetilde{D}\widetilde{R}^{-1}\widetilde{D}^T P^{\frac{1}{2}})^{-1}P^{\frac{1}{2}} \\
&\equiv P^{\frac{1}{2}}[I + P^{\frac{1}{2}}(D + \bar{D})(R + \bar{R})^{-1}(D + \bar{D})^T P^{\frac{1}{2}}]^{-1}P^{\frac{1}{2}} \geq 0.
\end{aligned}
$$

In the above, we denoted

$$\widetilde{D} = D + \bar{D}, \quad \widetilde{R} = R + \bar{R}, \quad \Gamma = P^{\frac{1}{2}}\widetilde{D}\widetilde{R}^{-\frac{1}{2}}$$

and used the fact

$$I - \Gamma(I + \Gamma^T\Gamma)^{-1}\Gamma^T = (I + \Gamma\Gamma^T)^{-1}.$$

Consequently, by the conditions (recall (2.3))

$$\begin{cases} (C + \bar{C})^T\Big[P - P(D + \bar{D})\Sigma_1^{-1}(D + \bar{D})^TP\Big](C + \bar{C}) + Q + \bar{Q} \geq 0, \\ \Sigma_1 \equiv R + \bar{R} + (D + \bar{D})^TP(D + \bar{D}) \geq \delta I, \quad G + \bar{G} \geq 0, \end{cases}$$

we see that the Riccati equation (4.7) admits a unique solution $\Pi(\cdot)$ which is positive definite, and it is also the solution to the Riccati equation (4.6).

Now, from (4.3), it follows that

$$\begin{aligned}
Z &= P\Big(C(X^* - \mathbb{E}[X^*]) + (C + \bar{C})\mathbb{E}[X^*] + D(u^* - \mathbb{E}[u^*]) + (C + \bar{C})\mathbb{E}[X^*]\Big) \\
&= P\Big\{C(X^* - \mathbb{E}[X^*]) + (C + \bar{C})\mathbb{E}[X^*] - D\Sigma_0^{-1}(B^TP + D^TPC)(X^* - \mathbb{E}[X^*]) \\
&\quad - (C + \bar{C})\Sigma_1^{-1}\Big[(B + \bar{B})^T\Pi + (D + \bar{D})^TP(C + \bar{C})\Big]\mathbb{E}[X^*]\Big\} \\
&= \Big[PC - PD\Sigma_0^{-1}(B^TP + D^TPC)\Big]\Big(X^* - \mathbb{E}[X^*]\Big) \\
&\quad + \Big[P(C + \bar{C}) - P(D + \bar{D})\Sigma_1^{-1}\Big((B + \bar{B})^T\Pi + (D + \bar{D})^TP(C + \bar{C})\Big)\Big]\mathbb{E}[X^*].
\end{aligned}$$

Next, we plug u^* of form (4.4) into the state equation to get the following closed-loop system:

$$\begin{aligned}
dX^* &= \Big\{AX^* - B\Big[\Sigma_0^{-1}(B^TP + D^TPC)\Big(X^* - \mathbb{E}[X^*]\Big) \\
&\quad - \Sigma_1^{-1}\Big((B + \bar{B})^T\Pi + (D + \bar{D})^TP(C + \bar{C})\Big)\mathbb{E}[X^*]\Big] \\
&\quad + \bar{A}\mathbb{E}[X^*] - \bar{B}\Sigma_1^{-1}\Big[(B + \bar{B})^T\Pi + (D + \bar{D})^TP(C + \bar{C})\Big]\mathbb{E}[X^*]\Big\}dt \\
&\quad + \Big\{CX^* - D\Big[\Sigma_0^{-1}(B^TP + D^TPC)\Big(X^* - \mathbb{E}[X^*]\Big) \\
&\quad - \Sigma_1^{-1}\Big((B + \bar{B})^T\Pi + (D + \bar{D})^TP(C + \bar{C})\Big)\mathbb{E}[X^*]\Big] \\
&\quad + \bar{C}\mathbb{E}[X^*] - \bar{D}\Sigma_1^{-1}\Big[(B + \bar{B})^T\Pi + (D + \bar{D})^TP(C + \bar{C})\Big]\mathbb{E}[X^*]\Big\}dW(t) \\
&= \Big\{\Big[A - B\Sigma_0^{-1}(B^TP + D^TPC)\Big]\Big(X^* - \mathbb{E}[X^*]\Big) \\
&\quad + \Big[(A + \bar{A}) - (B + \bar{B})\Sigma_1^{-1}\Big((B + \bar{B})^T\Pi + (D + \bar{D})^TP(C + \bar{C})\Big)\Big]\mathbb{E}[X^*]\Big\}dt \\
&\quad + \Big\{\Big[C - D\Sigma_0^{-1}(B^TP + D^TPC)\Big]\Big(X^* - \mathbb{E}[X^*]\Big) \\
&\quad + \Big[(C + \bar{C}) - (D + \bar{D})\Sigma_1^{-1}\Big((B + \bar{B})^T\Pi + (D + \bar{D})^TP(C + \bar{C})\Big)\Big]\mathbb{E}[X^*]\Big\}dW(t).
\end{aligned}$$

To summarize the above, we have the following result.

THEOREM 4.1. *Let* (H1)–(H2) *hold. Then the following Riccati equations admit unique solutions* $P(\cdot)$ *and* $\Pi(\cdot)$, *respectively:*

$$(4.8) \quad \begin{cases} \dot{P} + PA + A^T P + C^T PC + Q \\ \quad - (PB + C^T PD)\Sigma_0^{-1}(B^T P + D^T PC) = 0, \quad s \in [0,T], \\ P(T) = G, \end{cases}$$

$$(4.9) \quad \begin{cases} \dot{\Pi} + \Pi(A + \bar{A}) + (A + \bar{A})^T \Pi + (C + \bar{C})^T P(C + \bar{C}) + Q + \bar{Q} \\ \quad - \Big[\Pi(B + \bar{B}) + (C + \bar{C})^T P(D + \bar{D})\Big]\Sigma_1^{-1} \\ \qquad \cdot \Big[(B + \bar{B})^T \Pi + (D + \bar{D})^T P(C + \bar{C})\Big] = 0, \\ \Pi(T) = G + \bar{G}, \end{cases}$$

where

$$\Sigma_0 = R + D^T PD, \qquad \Sigma_1 = R + \bar{R} + (D + \bar{D})^T P(D + \bar{D}).$$

Further, the following closed-loop system admits a unique solution $X^*(\cdot)$:

$$\begin{cases} dX^* = \Big\{\Big[A - B\Sigma_0^{-1}(B^T P + D^T PC)\Big]\Big(X^* - \mathbb{E}[X^*]\Big) \\ \qquad + \Big[(A+\bar{A}) - (B+\bar{B})\Sigma_1^{-1}\Big((B+\bar{B})^T\Pi + (D+\bar{D})^T P(C+\bar{C})\Big)\Big]\mathbb{E}[X^*]\Big\}dt \\ \qquad + \Big\{\Big[C - D\Sigma_0^{-1}(B^T P + D^T PC)\Big]\Big(X^* - \mathbb{E}[X^*]\Big) \\ \qquad + \Big[(C+\bar{C}) - (D+\bar{D})\Sigma_1^{-1}\Big((B+\bar{B})^T\Pi + (D+\bar{D})^T P(C+\bar{C})\Big)\Big]\mathbb{E}[X^*]\Big\}dW(t), \\ X^*(0) = x, \end{cases}$$

and by defining

$$\begin{cases} u^* = -\Sigma_0^{-1}(B^T P + D^T PC)\Big(X^* - \mathbb{E}[X^*]\Big) \\ \qquad - \Sigma_1^{-1}\Big[(B + \bar{B})^T \Pi + (D + \bar{D})^T P(C + \bar{C})\Big]\mathbb{E}[X^*], \\ Y = P\Big(X^* - \mathbb{E}[X^*]\Big) + \Pi\mathbb{E}[X^*], \\ Z = \Big[PC - PD\Sigma_0^{-1}(B^T P + D^T PC)\Big]\Big(X^* - \mathbb{E}[X^*]\Big) \\ \qquad + \Big[P(C+\bar{C}) - P(D+\bar{D})\Sigma_1^{-1}\Big((B+\bar{B})^T\Pi + (D+\bar{D})^T P(C+\bar{C})\Big)\Big]\mathbb{E}[X^*], \end{cases}$$

the 4-tuple $(X^*(\cdot), u^*(\cdot), Y(\cdot), Z(\cdot))$ *is the adapted solution to the MF-FBSDE* (3.3), *and* $(X^*(\cdot), u^*(\cdot))$ *is the optimal pair of problem* (MF-LQ). *Moreover,*

$$(4.10) \qquad \inf_{u(\cdot) \in \mathcal{U}[0,T]} J(x; u(\cdot)) = \langle \Pi(0)x, x \rangle \qquad \forall x \in \mathbb{R}^n.$$

Proof. We need only to prove (4.10), which is a verification theorem. Let $P(\cdot)$ and $\Pi(\cdot)$ be the solutions to the Riccati equations (4.8) and (4.9). We observe

$$J(x; u(\cdot)) - \langle \Pi(0)x, x \rangle$$
$$= J(x; u(\cdot)) - \langle P(0)(X(0) - \mathbb{E}[X(0)]), X(0) - \mathbb{E}[X(0)] \rangle$$
$$- \langle \Pi(0)\mathbb{E}[X(0)], \mathbb{E}[X(0)] \rangle$$
$$= \mathbb{E} \int_0^T \Big\{ \langle Q(X - \mathbb{E}[X]), X - \mathbb{E}[X] \rangle + \langle (Q + \bar{Q})\mathbb{E}[X], \mathbb{E}[X] \rangle$$
$$+ \langle R(u - \mathbb{E}[u]), u - \mathbb{E}[u] \rangle + \langle (R + \bar{R})\mathbb{E}[u], \mathbb{E}[u] \rangle \Big\} ds$$
$$+ \langle G(X(T) - \mathbb{E}[X(T)]), X(T) - \mathbb{E}[X(T)] \rangle + \langle (G + \bar{G})]\mathbb{E}[X(T)], \mathbb{E}[X(T)] \rangle$$
$$- \langle P(T)(X(T) - \mathbb{E}[X(T)], X(T) - \mathbb{E}[X(T)] \rangle - \langle \Pi(T)\mathbb{E}[X(T)], \mathbb{E}[X(T)] \rangle$$
$$+ \int_0^T \Big\{ \langle \dot{P}(X - \mathbb{E}[X], X - \mathbb{E}[X] \rangle$$
$$+ 2 \langle P\{A(X - \mathbb{E}[X]) + B(u - \mathbb{E}[u])\}, X - \mathbb{E}[X] \rangle$$
$$+ \langle P\{C(X - \mathbb{E}[X]) + (C + \bar{C})\mathbb{E}[X] + D(u - \mathbb{E}[u]) + (D + \bar{D})\mathbb{E}[u]\},$$
$$C(X - \mathbb{E}[X]) + (C + \bar{C})\mathbb{E}[X] + D(u - \mathbb{E}[u]) + (D + \bar{D})\mathbb{E}[u] \rangle$$
$$+ \langle \dot{\Pi}\mathbb{E}[X], \mathbb{E}[X] \rangle + 2 \langle \Pi\{(A + \bar{A})\mathbb{E}[X] + (B + \bar{B})\mathbb{E}[u], \mathbb{E}[X] \rangle \Big\} ds$$
$$= \mathbb{E} \int_0^T \Big\{ \langle (\dot{P} + PA + A^T P + C^T PC + Q)(X - \mathbb{E}[X], X - \mathbb{E}[X] \rangle$$
$$+ 2 \langle u - \mathbb{E}[u], (B^T P + D^T PC)(X - \mathbb{E}[X]) \rangle$$
$$+ \langle (R + D^T PD)(u - \mathbb{E}[u], u - \mathbb{E}[u] \rangle$$
$$+ \langle [\dot{\Pi} + \Pi(A + \bar{A}) + (A + \bar{A})^T \Pi + (C + \bar{C})^T P(C + \bar{C}) + Q + \bar{Q}]\mathbb{E}[X], \mathbb{E}[X] \rangle$$
$$+ 2 \langle \mathbb{E}[u], \{(B + \bar{B})^T \Pi + (D + \bar{D})^T P(C + \bar{C})\}\mathbb{E}[X] \rangle$$
$$+ \langle [R + \bar{R} + (D + \bar{D})^T P(D + \bar{D})]\mathbb{E}[u], \mathbb{E}[u] \rangle \Big\} ds$$
$$= \mathbb{E} \int_0^T \Big\{ \langle [\dot{P} + PA + A^T P + C^T PC + Q$$
$$- (PB + CPD)\Sigma_0^{-1}(B^T P + D^T PC)](X - \mathbb{E}[X]), X - \mathbb{E}[X] \rangle$$
$$+ \Big| \Sigma_0^{\frac{1}{2}} \Big[u - \mathbb{E}[u] + \Sigma_0^{-1}(B^T P + D^T PC)(X - \mathbb{E}[X]) \Big] \Big|^2$$
$$+ \Big\langle \Big[\dot{\Pi} + \Pi(A + \bar{A}) + (A + \bar{A})^T \Pi + (C + \bar{C})^T P(C + \bar{C}) + Q + \bar{Q}$$
$$- \Big(\Pi(B + \bar{B}) + (C + \bar{C})^T P(D + \bar{D}) \Big) \Sigma_1^{-1}$$
$$\cdot \Big((B + \bar{B})^T \Pi + (D + \bar{D})^T P(C + \bar{C}) \Big] \mathbb{E}[X], \mathbb{E}[X] \Big\rangle$$
$$+ \Big| \Sigma_1^{\frac{1}{2}} \Big[\mathbb{E}[u] + \Sigma_1^{-1}\{(B + \bar{B})^T \Pi + (D + \bar{D})^T P(C + \bar{C}\}\mathbb{E}[X] \Big] \Big|^2 \Big\} ds$$
$$= \mathbb{E} \int_0^T \Big\{ \Big| \Sigma_0^{\frac{1}{2}} \Big[u - \mathbb{E}[u] + \Sigma_0^{-1}(B^T P + D^T PC)(X - \mathbb{E}[X]) \Big] \Big|^2$$
$$+ \Big| \Sigma_1^{\frac{1}{2}} \Big[\mathbb{E}[u] + \Sigma_1^{-1}\Big((B + \bar{B})^T \Pi + (D + \bar{D})^T P(C + \bar{C}) \Big) \mathbb{E}[X] \Big] \Big|^2 \Big\} ds \geq 0.$$

Then our claim follows. $\quad\Box$

We see that for Riccati equation (4.9), when

$$\bar{A} = \bar{C} = 0, \qquad \bar{B} = \bar{D} = 0, \qquad \bar{Q} = 0, \qquad \bar{R} = 0, \qquad \bar{G} = 0,$$

one has

$$\Sigma_0 = \Sigma_1 = R + D^T P D,$$

and the Riccati equation for $\Pi(\cdot)$ can be written as

$$\begin{cases} \dot{\Pi} + \Pi A + A^T \Pi + C^T P C + Q \\ \quad - (\Pi B + C^T P D)(R + D^T P D)^{-1}(B^T \Pi + D^T P C) = 0, \\ \Pi(T) = G. \end{cases}$$

Then

$$\begin{aligned} 0 &= (\dot{\Pi} - \dot{P}) + (\Pi - P)A + A^T(\Pi - P) - (\Pi B + C^T P D)\Sigma_0^{-1}(B^T \Pi + D^T P C) \\ &\quad + (PB + C^T P D)\Sigma_0^{-1}(B^T P + D^T P C) \\ &= (\Pi - P)' + (\Pi - P)A + A^T(\Pi - P) - (\Pi - P)B\Sigma_0^{-1}(B^T \Pi + D^T P C) \\ &\quad - (PB + C^T P D)\Sigma_0^{-1}B^T(\Pi - P). \end{aligned}$$

Therefore, by uniqueness, we have

$$\Pi = P.$$

Consequently, the feedback control can be written as

$$\begin{aligned} u^* &= -\Sigma_0^{-1}(B^T P + D^T P C)\Big(X^* - \mathbb{E}[X^*]\Big) \\ &\quad - \Sigma_1^{-1}\Big[(B + \bar{B})^T \Pi + (D + \bar{D})^T P(C + \bar{C})\Big]\mathbb{E}[X^*], \\ &= -\Sigma_0^{-1}(B^T P + D^T P C)\Big(X^* - \mathbb{E}[X^*]\Big) - \Sigma_0^{-1}\Big(B^T P + D^T P C\Big)\mathbb{E}[X^*] \\ &= -(R + D^T P D)^{-1}(B^T P + D^T P C)X^*. \end{aligned}$$

This recovers the result for the classical LQ problem [38].

5. A modification of standard LQ problems. In this section, we are going to look at a special case which was mentioned in the introduction. For convenience, let us rewrite the state equation here:

$$(5.1) \qquad \begin{cases} dX(s) = \Big[A(s)X(s) + B(s)u(s)\Big]ds + \Big[C(s)X(s) + D(s)u(s)\Big]dW(s), \\ X(0) = x. \end{cases}$$

Let us introduce the cost functional:

$$(5.2) \qquad \begin{aligned} J_0(x; u(\cdot)) = \mathbb{E}\bigg[&\int_0^T \Big(\langle\, Q_0(s)X(s), X(s) \,\rangle + \langle\, R_0(s)u(s), u(s) \,\rangle \Big)ds \\ &+ \langle\, G_0 X(T), X(T) \,\rangle \bigg]. \end{aligned}$$

The classical LQ problem can be stated as follows.

Problem (LQ). For any given $x \in \mathbb{R}^n$, find a $u^*(\cdot) \in \mathcal{U}[0, T]$ such that

$$J_0(x; u^*(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}[0, T]} J_0(x; u(\cdot)).$$

The following result is standard (see [38]).

THEOREM 5.1. *Let* (H1) *hold and*

$$Q_0(s) \geq 0, \quad R_0(s) \geq \delta I, \qquad s \in [0,T]; \qquad G_0 \geq 0.$$

Then problem (LQ) *admits a unique optimal pair* $(X_0^*(\cdot), u_0^*(\cdot))$. *Moreover, the following holds:*

$$(5.3) \qquad u_0^*(s) = -\big[R_0(s) + D(s)^T P_0(s) D(s)\big]^{-1} \big[B(s)^T P_0(s) + D(s)^T P_0(s) C(s)\big] X_0^*(s),$$
$$s \in [0,T],$$

where $P_0(\cdot)$ *is the solution to the following Riccati equation:*

$$(5.4) \qquad \begin{cases} \dot{P}_0 + P_0 A + A^T P_0 + C^T P_0 C + Q_0 \\ \quad - (P_0 B + C^T P_0 D)(R_0 + D^T P_0 D)^{-1}(B^T P_0 + D^T P_0 C) = 0, \\ \hspace{6cm} s \in [0,T], \\ P_0(T) = G_0. \end{cases}$$

and $X_0^*(\cdot)$ *is the solution to the following closed-loop system:*

$$\begin{cases} dX_0^*(s) = \Big[A - B(R_0 + D^T P_0 D)^{-1}(B^T P_0 + D^T P_0 C)\Big] X_0^*(s) ds \\ \qquad + \Big[C - D(R_0 + D^T P_0 D)(B^T P_0 + D^T P_0 C)\Big] X_0^*(s) dW(s), \quad s \in [0,T], \\ X_0^*(0) = x. \end{cases}$$

We now introduce the following modified cost functional:

$$\begin{aligned}
\widehat{J}_0(x;u(\cdot)) &= \mathbb{E}\left[\int_0^T \Big(\langle Q_0(s)X(s), X(s)\rangle + \langle R_0(s)u(s), u(s)\rangle\Big) ds + \langle G_0 X(T), X(T)\rangle\right] \\
&\quad + \mathbb{E}\left[\int_0^T \Big(q(s)\mathrm{var}\,[X(s)] + \rho(s)\mathrm{var}\,[u(s)]\Big) ds + g\mathrm{var}\,[X(T)]\right] \\
&= \mathbb{E}\left\{\int_0^T \left[\Big\langle \big(Q_0(s) + q(s)I\big)X(s), X(s)\Big\rangle - q(s)\Big|\mathbb{E}[X(s)]\Big|^2\right.\right. \\
&\qquad + \left\langle \big(R_0(s) + \rho(s)I\big)u(s), u(s)\right\rangle - \rho(s)\Big|\mathbb{E}[u(s)]\Big|^2\bigg] ds \\
&\qquad + \left\langle \big(G_0 + gI\big)X(T), X(T)\right\rangle - g\Big|\mathbb{E}[X(T)]\Big|^2\bigg\}
\end{aligned}$$

with $q(\cdot), \rho(\cdot) \in L^\infty(0,T)$, $g \in [0,\infty)$ such that

$$q(s), \rho(s) \geq 0, \qquad s \in [0,T].$$

Also, of course, we assume that

$$\int_0^T [q(s) + \rho(s)] ds + g > 0.$$

We want to compare problem (LQ) with the following modified LQ problem:

Problem (LQ)′. For any given $x \in \mathbb{R}^n$, find a $u^*(\cdot) \in \mathcal{U}[0,T]$ such that

$$\widehat{J}_0(x; u^*(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}[0,T]} \widehat{J}_0(x; u(\cdot)).$$

We refer to Problem (LQ)′ as a *modified* LQ problem. This is a special case of problem (MF) with

$$\begin{cases} \bar{A} = \bar{C} = 0, \quad \bar{B} = \bar{D} = 0, \\ Q = Q_0 + qI, \quad \bar{Q} = -qI, \quad R = R_0 + \rho I, \quad \bar{R} = -\rho I, \quad G = G_0 + gI, \quad \bar{G} = -gI. \end{cases}$$

Then the Riccati equations are

(5.5)
$$\begin{cases} \dot{P} + PA + A^T P + C^T PC + Q_0 + qI \\ \quad - (PB + C^T PD)(R_0 + \rho I + D^T PD)^{-1}(B^T P + D^T PC) = 0, \\ \hspace{8em} s \in [0,T], \\ P(T) = G_0 + gI, \end{cases}$$

and

(5.6)
$$\begin{cases} \dot{\Pi} + \Pi A + A^T \Pi + C^T PC + Q_0 \\ \quad - (\Pi B + C^T PD)(R_0 + D^T PD)^{-1}(B^T \Pi + D^T PC) = 0, \quad s \in [0,T], \\ \Pi(T) = G_0. \end{cases}$$

The optimal control is given by

$$\begin{aligned} u^* = &-(R_0 + \rho I + D^T PD)^{-1}(B^T P + D^T PC)\Big(X^* - \mathbb{E}[X^*]\Big) \\ &-(R_0 + D^T PD)^{-1}(B^T \Pi + D^T PC)\mathbb{E}[X^*], \end{aligned}$$

and the closed-loop system reads

(5.7)
$$\begin{cases} dX^* = \Big\{ \Big[A - B(R_0 + \rho I + B_1^T PB_1)^{-1}(B^T P + B_1^T PC)\Big]\Big(X^* - \mathbb{E}[X^*]\Big) \\ \hspace{3em} + \Big[A - B(R_0 + D^T PD)^{-1}\Big(B^T \Pi + D^T PC\Big)\Big]\mathbb{E}[X^*]\Big\} ds \\ \hspace{2em} + \Big\{ \Big[C - D\Sigma_0^{-1}(B^T P + D^T PC)\Big]\Big(X^* - \mathbb{E}[X^*]\Big) \\ \hspace{3em} + \Big[C - D\Sigma_1^{-1}\Big(B^T \Pi + D^T PC\Big)\Big]\mathbb{E}[X^*]\Big\} dW(s), \\ X^*(0) = x. \end{cases}$$

By the definition of $u_0^*(\cdot)$ and $u^*(\cdot)$, we know that

(5.8)
$$\langle P_0(0)x, x \rangle = J_0(x; u_0^*(\cdot)) \leq J_0(x; u^*(\cdot))$$

and
(5.9)
$$\langle\, \Pi(0)x, x \,\rangle$$

$$= J_0(x; u^*(\cdot)) + \mathbb{E}\left[\int_0^T \Big(q(s)\text{var}\,[X^*(s)] + \rho(s)\text{var}\,[u^*(s)]\Big)ds + g\text{var}\,[X^*(T)]\right]$$

$$= \widehat{J}_0(x; u^*(\cdot)) \leq \widehat{J}_0(x; u_0^*(\cdot))$$

$$= J_0(x; u_0^*(\cdot)) + \mathbb{E}\left[\int_0^T \Big(q(s)\text{var}\,[X_0^*(s)] + \rho(s)\text{var}\,[u_0^*(s)]\Big)ds + g\text{var}\,[X_0^*(T)]\right]$$

$$= \langle\, P_0(0)x, x \,\rangle + \mathbb{E}\left[\int_0^T \Big(q(s)\text{var}\,[X_0^*(s)] + \rho(s)\text{var}\,[u_0^*(s)]\Big)ds + g\text{var}\,[X_0^*(T)]\right]$$

$$\leq J_0(x; u^*(\cdot)) + \mathbb{E}\left[\int_0^T \Big(q(s)\text{var}\,[X_0^*(s)] + \rho(s)\text{var}\,[u_0^*(s)]\Big)ds + g\text{var}\,[X_0^*(T)]\right].$$

This implies

(5.10)
$$\mathbb{E}\left[\int_0^T \Big(q(s)\text{var}\,[X^*(s)] + \rho(s)\text{var}\,[u^*(s)]\Big)ds + g\text{var}\,[X^*(T)]\right]$$

$$\leq \mathbb{E}\left[\int_0^T \Big(q(s)\text{var}\,[X_0^*(s)] + \rho(s)\text{var}\,[\bar{u}_0(s)]\Big)ds + g\text{var}\,[X_0^*(T)]\right].$$

Hence,

$$J_0(x; u^*(\cdot)) - J_0(x; u_0^*(\cdot))$$

is the price for the decrease of

$$\mathbb{E}\left[\int_0^T \Big(q(s)\text{var}\,[X_0^*(s)] + \rho(s)\text{var}\,[u_0^*(s)]\Big)ds + g\text{var}\,[X_0^*(T)]\right]$$

$$- \mathbb{E}\left[\int_0^T \Big(q(s)\text{var}\,[X^*(s)] + \rho(s)\text{var}\,[u^*(s)]\Big)ds + g\text{var}\,[X^*(T)]\right]$$

for the (weighted) variances of the optimal state-control pair $(X_0^*(\cdot), u_0^*(\cdot))$. Moreover, (5.9) further implies that

$$J_0(x; u^*(\cdot)) - J_0(x; u_0^*(\cdot))$$

$$\leq \mathbb{E}\left[\int_0^T \Big(q(s)\text{var}\,[X_0^*(s)] + \rho(s)\text{var}\,[u_0^*(s)]\Big)ds + g\text{var}\,[X_0^*(T)]\right]$$

$$- \mathbb{E}\left[\int_0^T \Big(q(s)\text{var}\,[X^*(s)] + \rho(s)\text{var}\,[u^*(s)]\Big)ds + g\text{var}\,[X^*(T)]\right].$$

Roughly speaking, the above means that the amount increased in the cost is "covered" by the amount decreased in the weighted variance of the optimal state-control pair.

We now look at a simple case to illustrate the above. Let us look at a one-dimensional controlled linear SDE,

(5.11)
$$\begin{cases} dX(s) = bu(s)ds + X(s)dW(s), \\ X(0) = x, \end{cases}$$

with cost functionals

(5.12) $$J_0(x; u(\cdot)) = \mathbb{E}\left[\int_0^T |u(s)|^2 ds + g_0|X(T)|^2\right]$$

and

(5.13)
$$\widehat{J}_0(x; u(\cdot)) = \mathbb{E}\left\{\int_0^T |u(s)|^2 ds + g_0|X(T)|^2 + g\text{var}\,[X(T)]\right\}$$
$$= \mathbb{E}\left\{\int_0^T |u(s)|^2 ds + (g_0 + g)|X(T)|^2 - g\Big(\mathbb{E}[X(T)]\Big)^2\right\},$$

where $g_0 \geq 0$ and $g > 0$. As above, we refer to the optimal control problem associated with (5.11) and (5.12) as the *standard LQ problem* and to that associated with (5.11) and (5.13) as the *modified LQ problem*. The Riccati equation for the standard LQ problem is

$$\begin{cases} \dot{p}_0(s) + p_0(s) - b^2 p_0(s)^2 = 0, & s \in [0, T], \\ p_0(T) = g_0. \end{cases}$$

A straightforward calculation leads to

$$p_0(s) = \frac{e^{T-s}g_0}{(e^{T-s} - 1)b^2 g_0 + 1} > 0, \qquad s \in [0, T].$$

The optimal control is

$$u_0^*(s) = -bp_0(s)X_0^*(s), \qquad s \in [0, T],$$

and the closed-loop system is

$$\begin{cases} dX_0^*(s) = -b^2 p_0(s)X_0^*(s)ds + X_0^*(s)dW(s), & s \in [0, T], \\ X_0^*(0) = x. \end{cases}$$

Thus,

$$X_0^*(s) = xe^{-b^2 \int_0^s p_0(\tau)d\tau - \frac{1}{2}s + W(s)}, \qquad s \in [0, T].$$

Consequently,

$$\mathbb{E}[X_0^*(T)] = xe^{-b^2 \int_0^T p_0(\tau)d\tau - \frac{1}{2}T + \frac{1}{2}T} = xe^{-b^2 \int_0^T p_0(\tau)d\tau}$$

and

$$\mathbb{E}[X_0^*(T)^2] = x^2 e^{-2b^2 \int_0^T p_0(\tau)d\tau - T + 2T} = x^2 e^{-2b^2 \int_0^T p_0(\tau)d\tau + T}.$$

Hence,

$$\text{var}\,[X_0^*(T)] = \mathbb{E}[X_0^*(T)^2] - \Big(\mathbb{E}[X_0^*(T)]\Big)^2 = x^2 e^{-2b^2 \int_0^T p_0(\tau)d\tau}(e^T - 1).$$

Also, the optimal expected cost is

$$J_0(x; u_0^*(\cdot)) = p_0(0)x^2 = \frac{e^T g_0}{(e^T - 1)b^2 g_0 + 1}x^2.$$

Next, for the modified LQ peoblem, the Riccati equations are

$$\begin{cases} \dot{p}(s) + p(s) - b^2 p(s)^2 = 0, & s \in [0, T], \\ p(T) = g_0 + g, \end{cases}$$

and

$$\begin{cases} \dot{\pi}(s) - b^2 \pi^2(s) + p(s) = 0, & s \in [0, T], \\ \pi(T) = g_0. \end{cases}$$

Clearly,

$$p(s) = \frac{e^{T-s}(g_0 + g)}{(e^{T-s} - 1)b^2(g_0 + g) + 1} > \frac{e^{T-s}g_0}{(e^{T-s} - 1)b^2 g_0 + 1} = p_0(s) > 0, \quad s \in [0, T].$$

We now show that

(5.14) $$p_0(s) < \pi(s) < p(s), \qquad s \in [0, T].$$

In fact,

$$\begin{cases} \dfrac{d}{ds}[\pi(s) - p_0(s)] - b^2[\pi(s) + p_0(s)][\pi(s) - p_0(s)] + p(s) - p_0(s) = 0, & s \in [0, T], \\ \pi(T) - p_0(T) = 0. \end{cases}$$

Thus,

$$\pi(s) - p_0(s) = \int_s^T e^{-\int_s^t b^2[\pi(\tau) + p_0(\tau)]d\tau}[p(t) - p_0(t)]dt > 0, \quad s \in [0, T).$$

Next,

$$\begin{cases} \dfrac{d}{ds}[p(s) - \pi(s)] - b^2[p(s) + \pi(s)][p(s) - \pi(s)] = 0, & s \in [0, T], \\ p(T) - \pi(T) = g. \end{cases}$$

Hence,

$$p(s) - \pi(s) = e^{-\int_s^T b^2[p(\tau) + \pi(\tau)]d\tau}g > 0, \qquad s \in [0, T].$$

This proves (5.14). Note that the optimal control of modified LQ problem is given by

$$u^*(s) = -bp(s)\Big(X^*(s) - \mathbb{E}[X^*(s)]\Big) - b\pi(s)\mathbb{E}[X^*(s)], \qquad s \in [0, T],$$

and the closed-loop system is

$$\begin{cases} dX^*(s) = -\Big[b^2 p(s)\Big(X^*(s) - \mathbb{E}[X^*(s)]\Big) + b^2 \pi(s)\mathbb{E}[X^*(s)]\Big]ds \\ \qquad\qquad + X^*(s)dW(s), & s \in [0, T], \\ X^*(0) = x. \end{cases}$$

Thus,

$$\frac{d}{ds}\Big(\mathbb{E}[X^*(s)]\Big) = -b^2\pi(s)\mathbb{E}[X^*(s)],$$

which leads to

$$\mathbb{E}[X^*(s)] = e^{-b^2\int_0^s \pi(\tau)d\tau}x, \qquad s \in [0, T].$$

On the other hand, by Itô's formula,

$$d[(X^*)^2] = \Big\{ - 2\Big[b^2 p X^*\Big(X^* - \mathbb{E}[X^*]\Big) + b^2\pi X^*\mathbb{E}[X^*]\Big] + (X^*)^2\Big\}ds + [\cdots]dW.$$

Then

$$\begin{aligned}
d\Big(\mathbb{E}[(X^*)^2]\Big) &= \Big\{ - 2b^2 p\Big[\mathbb{E}[(X^*)^2] - \Big(\mathbb{E}[X^*]\Big)^2\Big] - 2b^2\pi\Big(\mathbb{E}[X^*]\Big)^2 + \mathbb{E}[(X^*)^2]\Big\}ds \\
&= \Big\{(1 - 2b^2 p)\mathbb{E}[(X^*)^2] + 2b^2(p - \pi)\Big(\mathbb{E}[\bar{X}]\Big)^2\Big\}ds \\
&= \Big\{(1 - 2b^2 p)\mathbb{E}[(X^*)^2] + 2b^2(p - \pi)e^{-2b^2\int_0^s \pi(\tau)d\tau}x^2\Big\}ds.
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathbb{E}[(X^*(s)^2] &= e^{s - 2b^2\int_0^s p(\tau)d\tau}x^2\left[1 + \int_0^s e^{-t + 2\int_0^t p(\tau)d\tau}2b^2[p(t) - \pi(t)]e^{-2b^2\int_0^t \pi(\tau)d\tau}dt\right] \\
&= e^{s - 2b^2\int_0^s p(\tau)d\tau}x^2\left[1 + \int_0^s e^{-t + 2b^2\int_0^t [p(\tau)-\pi(\tau)]d\tau}2b^2[p(t) - \pi(t)]dt\right] \\
&= e^{s - 2b^2\int_0^s p(\tau)d\tau}x^2\left[e^{-s + 2b^2\int_0^s [p(\tau)-\pi(\tau)]d\tau} + \int_0^s e^{-t + 2b^2\int_0^t [p(\tau)-\pi(\tau)]d\tau}dt\right] \\
&= \left[e^{-2b^2\int_0^s \pi(\tau)d\tau} + \int_0^s e^{s - t - 2b^2\int_t^s p(\tau)d\tau - 2b^2\int_0^t \pi(\tau)d\tau}dt\right]x^2.
\end{aligned}$$

Then

$$\begin{aligned}
\text{var}\,[X^*(s)] &= \left[e^{-2b^2\int_0^s \pi(\tau)d\tau} + \int_0^s e^{s - t - 2b^2\int_t^s p(\tau)d\tau - 2b^2\int_0^s \pi(\tau)d\tau}dt\right]x^2 \\
&\quad - \Big(e^{-2b^2\int_0^t \pi(\tau)d\tau}\Big)x^2 \\
&= \left[\int_0^s e^{s - t - 2b^2\int_t^s p(\tau)d\tau - 2b^2\int_0^t \pi(\tau)d\tau}dt\right]x^2.
\end{aligned}$$

Consequently, noting $p_0(s) < p(s)$, we have

$$\begin{aligned}
\text{var}\,[X^*(T)] &= \left[\int_0^T e^{T - t - 2b^2\int_t^T p(\tau)d\tau - 2b^2\int_0^t p_0(\tau)d\tau}dt\right]x^2 \\
&< \left[\int_0^T e^{T - t - 2b^2\int_t^T p_0(\tau)d\tau - 2b^2\int_0^t p_0(\tau)d\tau}dt\right]x^2 \\
&= \left[\int_0^T e^{T - t - 2b^2\int_0^T p_0(\tau)d\tau}dt\right]x^2 \\
&= \left[e^{-2b^2\int_0^T p_0(\tau)d\tau}\int_0^T e^{T - t}dt\right]x^2 = \text{var}\,[X_0^*(T)].
\end{aligned}$$

On the other hand, we claim that

(5.15) $$p_0(s) > \frac{g_0}{g_0 + g} p(s), \qquad s \in [0, T).$$

In fact, by letting $\widetilde{p} = \frac{g_0}{g_0+g} p$, we have

$$\begin{cases} \widetilde{p}'(s) + \widetilde{p}(s) - b^2 \widetilde{p}(s)^2 - \dfrac{g_0 g}{(g_0 + g)^2} p(s)^2 = 0, \\ \widetilde{p}(T) = g_0. \end{cases}$$

Then

$$\begin{cases} [p_0'(s) - \widetilde{p}'(s)] + [p_0(s) - \widetilde{p}(s)] - b^2 [p_0(s) + \widetilde{p}(s)][p_0(s) - \widetilde{p}(s)] + \dfrac{g_0 g}{(g_0+g)^2} p(s)^2 = 0, \\ p_0(T) - \widetilde{p}(T) = 0. \end{cases}$$

This leads to (5.15). Consequently,

$$\begin{aligned} J_0(x; u^*(\cdot)) + \mathrm{var}\,[X^*(T)] = \widehat{J}_0(x; u^*(\cdot)) &= \pi(0) x^2 \\ &\leq p(0) x^2 \leq \tfrac{g_0+g}{g_0} p_0(0) x^2 = \tfrac{g_0+g}{g_0} J(x; u_0^*(\cdot)). \end{aligned}$$

Therefore,

$$0 \leq J_0(x; u^*(\cdot)) - J_0(x; u_0^*(\cdot)) + g\,\mathrm{var}\,[X^*(T)] \leq \frac{g}{g_0} J(x; u_0^*(\cdot)).$$

The above already shows that it is a very good trade-off to consider the modified LQ problem if one wishes to have a smaller $\mathrm{var}\,[X(T)]$. It is possible to more carefully calculate the price difference $J_0(x; u(\cdot)) - J_0(x; u_0^*(\cdot))$. We omit the details here.

Also, it is possible to calculate the situation of including $\mathrm{var}\,[u(s)]$ and/or $\mathrm{var}\,[X(s)]$ in the integrand of the modified cost functional. The details are omitted here as well.

To conclude this paper, let us make some remarks. In this paper, we have presented some results on the LQ problem for MF-SDEs with deterministic coefficients. Optimal control is represented as a state feedback form via the solutions of two Riccati equations. We may expect that a theory can also be established for a problem in an infinite horizon with constant coefficients. For that, some stabilizability problems might need to be investigated first. Differential Riccati equations will become algebraic Riccati equations. We will present some relevant results along this line in a near future. On the other hand, we note that when $A(s)$ is deterministic, $\mathbb{E}[A(s)X(s)] = A(s)\mathbb{E}[X(s)]$. Throughout this paper, such a fact has played an essential role. Such a fact will be no longer true if $A(s)$ is random. Therefore, for the case of random coefficient case, one might have to find a different approach. The problem is still under investigation. Another issue that we would like to mention is the so-called *time-inconsistency*. It seems to be a common opinion that if a conditional expectation of the state/control process appears in the cost functional/state equation, the corresponding optimal control problem should be time-inconsistent, meaning that an optimal control for a given initial pair (t, x) on time interval $[t, T]$ might not be optimal if the control is restricted to a later time interval, say, $[\tau, T]$ (for $\tau > t$). However, from our rough observations on the problem studied in the current paper, it seems to us that if the coefficients are deterministic, a problem similar to our problem (MF-LQ)

(with $\mathbb{E}$ replaced by $\mathbb{E}_t$) seems to still be time-consistent. We suspect that the problem will be time-inconsistent if the coefficients are random. We will investigate this issue carefully in the near future. The reader is referred to [7, 6, 17, 19, 18, 22, 36, 37] for recent works on time-inconsistent optimal control problems.

REFERENCES

[1] N. U. AHMED AND X. DING, *A semilinear McKean-Vlasov stochastic evolution equation in Hilbert space*, Stochastic Process. Appl., 60 (1995), pp. 65–85.

[2] N. U. AHMED AND X. DING, *Controlled McKean-Vlasov equations*, Commun. Appl. Anal., 5 (2001), pp. 183–206.

[3] N. U. AHMED, *Nonlinear diffusion governed by McKean-Vlasov equation on Hilbert space and optimal control*, SIAM J. Control Optim., 46 (2007), pp. 356–378.

[4] D. ANDERSSON AND B. DJEHICHE, *A maximum principle for SDEs of mean-field type*, Appl. Math. Optim., 63 (2011), pp. 341–356.

[5] R. BELLMAN, R. KALABA, AND G. M. WING, *Invariant imbedding and the reduction of two-point boundary value problems to initial value problems*, Proc. Natl. Acad. Sci. USA, 46 (1960), pp. 1646–1649.

[6] T. BJÖRK AND A. MURGOCI, *A General Theory of Markovian Time Inconsistent Stochastic Control Problem*, working paper.

[7] T. BJÖRK, A. MURGOCI, AND X. Y. ZHOU, *Mean varaiance portfolio optimization with state dependent risk aversion*, Math. Finance, in press.

[8] V. S. BORKAR AND K. S. KUMAR, *McKean-Vlasov limit in portfolio optimization*, Stochastic Anal. Appl., 28 (2010), pp. 884–906.

[9] R. BUCKDAHN, B. DJEHICHE, AND J. LI, *A General Maximum Principle for SDEs of Mean-Field Type*, Appl. Math. Optim., 64 (2011), pp. 197–216.

[10] R. BUCKDAHN, B. DJEHICHE, J. LI, AND S. PENG, *Mean-field backward stochastic differential equations: A limit approach*, Ann. Probab., 37 (2009), pp. 1524–1565.

[11] R. BUCKDAHN, J. LI, AND S. PENG, *Mean-field backward stochastic differential equations and related partial differential equations*, Stochastic Process. Appl., 119 (2009), pp. 3133–3154,

[12] T. CHAN, *Dynamics of the McKean-Vlasov equation*, Ann. Probab., 22 (1994), pp. 431–441.

[13] T. CHIANG, *McKean-Vlasov equations with discontinuous coefficients*, Soochow J. Math., 20 (1994), pp. 507–526.

[14] D. CRISAN AND J. XIONG, *Approximate McKean-Vlasov representations for a class of SPDEs*, Stochastics, 82 (2010), pp. 53–68.

[15] D. A. DAWSON, *Critical dynamics and fluctuations for a mean-field model of cooperative behavior*, J. Statist. Phys., 31 (1983), pp. 29–85.

[16] D. A. DAWSON AND J. GÄRTNER, *Large deviations from the McKean-Vlasov limit for weakly interacting diffusions*, Stochastics, 20 (1987), pp. 247–308.

[17] I. EKELAND AND A. LAZRAK, *The golden rule when preferences are time inconsistent*, Math. Finan. Econ., 4 (2010), pp. 29–55.

[18] I. EKELAND AND T. A. PIRVU, *Investment and consumption without commitment*, Math. Finan. Econ., 2 (2008), pp. 57–86.

[19] I. EKELAND, O. MBODJI, AND T. A. PIRVU, *Time-Consistent Portfolio Management*, SIAM J. Financial Math., 3 (2012), pp. 1–32.

[20] J. GÄRTNER, *On the McKean-Vlasov limit for interacting diffusions*, Math. Nachr., 137 (1988), pp. 197–248.

[21] C. GRAHAM, *McKean-Vlasov Ito-Skorohod equations, and nonlinear diffusions with discrete jump sets*, Stochastic Process. Appl., 40 (1992), pp. 69–82.

[22] Y. HU, H. JIN, AND X. Y. ZHOU, *Time-Inconsistent Stochastic Linear-Quadratic Control*, arXiv:1111.0818v1, 2011.

[23] M. HUANG, R. P. MALHAMÉ, AND P. E. CAINES, *Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle*, Commun. Inf. Systems, 6 (2006), pp. 221–252.

[24] M. KAC, *Foundations of kinetic theory*, in Proceedings of the 3rd Berkeley Symposium on Mathematical Statistics and Probability, vol. 3, 1956, pp. 171–197.

[25] P. M. KOTELENEZ AND T. G. KURTZ, *Macroscopic limit for stochastic partial differential equations of McKean-Vlasov type*, Prob. Theory Related Fields, 146 (2010), pp. 189–222.

[26] J. MA, P. PROTTER, AND J. YONG, *Solving forward-backward stochastic differential equations explicitly: A four-step scheme*, Probab. Theory Related Fields, 98 (1994), pp. 339–359.

[27] J. Ma and J. Yong, *Forward-Backward Stochastic Differential Equations and Their Applications*, Lecture Notes in Math. 1702, Springer-Verlag, New York, 1999.

[28] N. I. Mahmudov and M. A. McKibben, *On a class of backward McKean-Vlasov stochastic equations in Hilbert space: Existence and convergence properties*, Dynam. Systems Appl., 16 (2007), pp. 643–664.

[29] H. P. McKean, *A class of Markov processes associated with nonlinear parabolic equations*, Proc. Natl. Acad. Sci. USA, 56 (1966), pp. 1907–1911.

[30] T. Meyer-Brandis, B. Oksendal, and X. Zhou, *A mean-field stochastic maximum principle via Malliavin calculus*, Stochastics, 84 (2012), pp. 643–666.

[31] L. Mou and J. Yong, *Two-person zero-sum linear quadratic stochastic differential games by a Hilbert space method*, J. Industrial Management Optim., 2 (2006), pp. 95–117.

[32] D. Nualart, *The Malliavin Calculus and Related Topics*, 2nd ed., Springer-Verlag, Berlin, 2006.

[33] J. Y. Park, P. Balasubramaniam, and Y. H. Kang, *Controllability of McKean-Vlasov stochastic integrodifferential evolution equation in Hilbert spaces*, Numer. Funct. Anal. Optim., 29 (2008), pp. 1328–1346.

[34] M. Scheutzow, *Uniqueness and non-uniqueness of solutions of Vlasov-McKean equations*, J. Aust. Math. Soc. Ser. A, 43 (1987), pp. 246–256.

[35] A. Yu. Veretennikov, *On ergodic measures for McKean–Vlasov stochastic equations*, in Stochastic Calculus to Mathematical Finance, Springer, Berlin, 2006, pp. 623–633.

[36] J. Yong, *A deterministic linear quadratic time-inconsistent optimal control problem*, Math. Control Related Fields, 1 (2011), pp. 83–118.

[37] J. Yong, *Time-inconsistent optimal control problems and the equilibrium HJB equation*, Math. Control Related Fields, 2 (2012), pp. 271–329.

[38] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York, 1999.