ESAIM: COCV 30 (2024) 76
https://doi.org/10.1051/cocv/2024064

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

TIME-INCONSISTENT LINEAR QUADRATIC OPTIMAL CONTROL PROBLEM FOR FORWARD–BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

Qi Lü[1,*] and Bowen Ma[2]

Abstract. We study the time-inconsistent linear quadratic optimal control problem for forward–backward stochastic differential equations with potentially indefinite cost weighting matrices for both the state and the control variables. Our research makes two contributions. Firstly, we introduce a novel type of Riccati equation system with parameters and constraint conditions, known as the generalized equilibrium Riccati equation. This equation system offers a comprehensive solution for the closed-loop equilibrium strategy of the problem at hand. Secondly, we establish the well-posedness of the generalized equilibrium Riccati equation for the one-dimensional case, provided certain conditions are met.

Mathematics Subject Classification. 93E20, 49N10.

Received February 24, 2024. Accepted August 19, 2024.

1. Introduction

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a complete filtered probability space, on which a standard one-dimensional Brownian motion $W(\cdot)$ is defined, and $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ is the natural filtration generated by $W(\cdot)$.

For a matrix M, we use $M^\dagger$ to represent the Moore-Penrose inverse of M and $M^\top$ to represent the transpose of M. In addition, for a positive integer n, $\mathbb{S}^n$ denotes the space consisting of all $n \times n$ symmetric matrices, and $\mathbb{S}^n_+$ denotes the space consisting of all positive semidefinite matrices in $\mathbb{S}^n$. Furthermore, $|M|_\infty$ represents the spectral norm of a matrix $M \in \mathbb{R}^{n\times m}$, which is equal to the square root of the largest eigenvalue of $M^\top M$.

We define some Banach spaces for $k \in \mathbb{N}$, $t \in [0, T]$, and $p \in [1, \infty)$, $q \in [1, \infty)$.

Firstly, $L^p_{\mathcal{F}_t}(\Omega; \mathbb{R}^k)$ is the Banach space consisting of all $\mathcal{F}_t$-measurable random variables $\xi : \Omega \to \mathbb{R}^k$ such that $\mathbb{E}|\xi|^p_{\mathbb{R}^k} < \infty$, with the canonical norm.

Next, $L^p_{\mathbb{F}}(\Omega; C([t, T]; \mathbb{R}^k))$ is the Banach space of all $\mathbb{R}^k$-valued $\mathbb{F}$-adapted continuous stochastic processes $\phi(\cdot)$. The norm for this space is defined as $|\phi(\cdot)|_{L^p_{\mathbb{F}}(\Omega;C([t,T];\mathbb{R}^k))} = \left(\mathbb{E} \max_{s\in[t,T]} |\phi(s)|^p_{\mathbb{R}^k}\right)^{1/p}$.

Keywords and phrases: Stochastic linear quadratic control problem, forward–backward stochastic differential equation, time-inconsistency.

[1] School of Mathematics, Sichuan University, Chengdu 610064, PR China.
[2] College of Mathematics and Physics, Chengdu University of Technology, Chengdu 610059, PR China.

[*] Corresponding author: lu@scu.edu.cn

Lastly, $L^p_{\mathbb{F}}(\Omega; L^q(t,T;\mathbb{R}^k))$ is a Banach space consisting of $\mathbb{R}^k$-valued $\mathbb{F}$-adapted stochastic processes $\phi(\cdot)$ defined on the product space $\Omega\times[t,T]$ such that $\mathbb{E}\left(\int_t^T|\phi(s)|^q_{\mathbb{R}^k}\mathrm{d}s\right)^{p/q}<\infty$. The norm of this Banach space is given by the canonical norm. We can also write $L^p_{\mathbb{F}}(\Omega;L^p(t,T;\mathbb{R}^k))$ simply as $L^p_{\mathbb{F}}(t,T;\mathbb{R}^k)$.

Let $T>0$ and $(t,x)\in[0,T)\times L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^n)$. Consider the following controlled *linear forward–backward stochastic differential equation* (FBSDE, for short) on the time horizen $[t,T]$:

$$\begin{cases}\mathrm{d}X(s)=(A(s)X(s)+B(s)u(s))\mathrm{d}s+(C(s)X(s)+D(s)u(s))\mathrm{d}W(s), & s\in[t,T],\\ \mathrm{d}Y(s)=-\big(\widehat{A}(s)X(s)+\widehat{B}(s)u(s)+\widehat{C}(s)Y(s)+\widehat{D}(s)Z(s)\big)\mathrm{d}s+Z(s)\mathrm{d}W(s), & s\in[t,T],\\ X(t)=x,\quad Y(T)=HX(T),\end{cases}\tag{1.1}$$

where $A(\cdot),C(\cdot)\in L^\infty(0,T;\mathbb{R}^{n\times n})$, $B(\cdot),D(\cdot)\in L^\infty(0,T;\mathbb{R}^{n\times k})$, $\widehat{A}(\cdot)\in L^\infty(0,T;\mathbb{R}^{m\times n})$, $\widehat{B}(\cdot)\in L^\infty(0,T;\mathbb{R}^{m\times k})$, and $\widehat{C}(\cdot),\widehat{D}(\cdot)\in L^\infty(0,T;\mathbb{R}^{m\times m})$, $H\in\mathbb{R}^{m\times n}$, and the control $u(\cdot)\in\mathcal{U}[t,T]\triangleq L^2_{\mathbb{F}}(t,T;\mathbb{R}^k)$.

The equation (1.1) consists of a (forward) stochastic differential equation (SDE, for short) and a backward stochastic differential equation (BSDE, for short). Since $Y(\cdot)$ and $Z(\cdot)$ do not appear in the SDE in (1.1), one can first solve the SDE to obtain $X(\cdot)$, and then solve the BSDE to obtain $(Y(\cdot),Z(\cdot))$. Therefore, based on the classical well-posedness of SDEs and BSDEs, it is known that for any $(t,x)\in[0,T)\times L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^n)$ and $u(\cdot)\in\mathcal{U}[t,T]$, the equation (1.1) has a unique solution $(X(\cdot),Y(\cdot),Z(\cdot))\in L^2_{\mathbb{F}}(\Omega;C([t,T];\mathbb{R}^n))\times L^2_{\mathbb{F}}(\Omega;C([t,T];\mathbb{R}^m))\times L^2_{\mathbb{F}}(t,T;\mathbb{R}^m)$ (*e.g.*, [1], Sects. 3.1 and 4.1).

We introduce the following cost functional

$$\begin{aligned}\mathcal{J}(t,x;u(\cdot))=&\frac{1}{2}\mathbb{E}_t\Big[\int_t^T\big(\langle Q(s,t)X(s),X(s)\rangle+\langle R(s,t)u(s),u(s)\rangle+\langle M(s,t)Y(s),Y(s)\rangle\\&+\langle N(s,t)Z(s),Z(s)\rangle\big)\mathrm{d}s+\langle G_1(t)X(T),X(T)\rangle+\langle G_2(t)Y(t),Y(t)\rangle\Big],\end{aligned}\tag{1.2}$$

where $\mathbb{E}_t:=\mathbb{E}(\cdot\,|\,\mathcal{F}_t)$ denotes the conditional expectation with respect to $\mathcal{F}_t$, $G_1(\cdot)\in C([0,T];\mathbb{S}^n)$, $G_2(\cdot)\in C([0,T];\mathbb{S}^m)$, $Q(\cdot,\cdot)\in C([0,T]^2;\mathbb{S}^n)$, $M(\cdot,\cdot),N(\cdot,\cdot)\in C([0,T]^2;\mathbb{S}^m)$ and $R(\cdot,\cdot)\in C([0,T]^2;\mathbb{S}^k)$.

For any $u(\cdot)\in\mathcal{U}[t,T]$, the cost functional $\mathcal{J}(t,x;u(\cdot))$ is well defined. Naturally, we can pose a standard optimal control problem as follows:

Problem (TI-FBSLQ). For any $(t,x)\in[0,T)\times L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^n)$, find a control $\bar{u}(\cdot)\in\mathcal{U}[t,T]$ such that

$$\mathcal{J}(t,x;\bar{u}(\cdot))=\inf_{u(\cdot)\in\mathcal{U}[t,T]}\mathcal{J}(t,x;u(\cdot)).\tag{1.3}$$

Remark 1.1. Recently, Wang *et al.* [2] made an interesting discovery regarding optimal control problems for FBSDEs, that is, they found that these problems exhibit a characteristic of being *time-inconsistent*. In light of this, we refer to the aforementioned problem as the TI-FBSLQ problem, with the abbreviation "TI"signifying the *time-inconsistency*.

For Problem (TI-FBSLQ), a control $\bar{u}(\cdot)\in\mathcal{U}[t,T]$ that satisfies (1.3) is referred to as a "pre-committed optimal control". Although the pre-committed optimal control $\bar{u}(\cdot)$ is optimal for the cost functional $\mathcal{J}(t,x;\cdot)$ for any fixed initial pair (t,x), it may not be practical in practice as it may not remain optimal at later times. There are two main reasons in order:

(1) In the cost functional (1.2), the coefficients $Q(\cdot,t),R(\cdot,t),M(\cdot,t),N(\cdot,t),G_1(t),G_2(t)$ depend on the initial time t. This subjective time-preference leads to time-inconsistent behavior (see [3–6] and the rich reference therein). For example, if we assume:

$$\begin{cases} H=0,\ \widehat{A}(\cdot)=\widehat{C}(\cdot)=0, \widehat{B}(\cdot)=\widehat{D}(\cdot)=0,\\ M(\cdot,\cdot)=0, \quad N(\cdot,\cdot)=0, \quad G_2(\cdot)=0, \end{cases}$$

then Problem (TI-FBSLQ) will reduce to the model in [4, 6], where an illustrative example is provided.

(2) Even if there is no subjective time-preference, *i.e.*, the coefficients $Q(\cdot,t), R(\cdot,t), M(\cdot,t), N(\cdot,t), G_1(t)$, $G_2(t)$ do not depend on the initial time t, Problem (TI-FBSLQ) remains time-inconsistent due to the forward–backward structure of the control system (1.1) (*e.g.*, [2]). For instance, consider the following simple example:

Example 1.2. Let $n=m=k=1$,

$$\begin{cases} \dot{X}(s)=u(s), \quad s\in[t,T],\\ \dot{Y}(s)=u(s), \quad s\in[t,T],\\ X(t)=x,\ Y(T)=2X(T), \end{cases}$$

with cost functional

$$\mathcal{J}(t,x;u(\cdot))=|Y(t)|^2.$$

A direct calculation gives $Y(t)=2x+\int_t^T u(s)\mathrm{d}s$. Hence,

$$\mathcal{J}(t,x;u(\cdot))=\left(2x+\int_t^T u(s)\mathrm{d}s\right)^2.$$

This implies that a control $\bar{u}(\cdot)\in L^2(t,T;\mathbb{R})$ is an optimal control for the initial pair (t,x) if and only if

$$\int_t^T \bar{u}(s)\mathrm{d}s=-2x.$$

As a result, for the initial pair $(t,x)=(0,1)$, an optimal control $\bar{u}(\cdot)$ must satisfy

$$\int_0^T \bar{u}(s)\mathrm{d}s=-2.$$

Therefore, the optimal state satisfies $\overline{X}(T;1,\bar{u})=-1$. Then there exists $0<t^*<T$ such that

$$\overline{X}(t^*;1,\bar{u})=1+\int_0^{t^*}\bar{u}(s)\mathrm{d}s=0. \tag{1.4}$$

For initial pair $(t^*,0)$, we have

$$\mathcal{J}(t^*,0;u(\cdot))=|Y(t^*)|^2=\left(\int_{t^*}^T u(s)\mathrm{d}s\right)^2. \tag{1.5}$$

Clearly, the restriction of the control $\bar{u}(\cdot)$ on $[t^*,T]$ is not an optimal control for (1.5), since from (1.4), we have

$$\int_{t^*}^{T}\bar{u}(s)\mathrm{d}s=-1.$$

Thus, the problem is time-inconsistent.

To address the issue of time inconsistency, we propose the following definition.

Definition 1.3. A matrix-valued function $\Theta(\cdot)\in L^2(0,T;\mathbb{R}^{k\times n})$ is called the closed-loop equilibrium strategy of Problem (TI-FBSLQ) if for any sequence $\{\varepsilon_j\}_{j=1}^{\infty}\subset(0,+\infty)$ converging to 0, $(t,x)\in[0,T)\times L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^n)$ and $v\in L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^k)$,

$$\varliminf_{j\to\infty}\frac{\mathcal{J}\left(t,\overline{X}(t);u^{\varepsilon_j}(\cdot)\right)-\mathcal{J}\left(t,\overline{X}(t);\bar{u}(\cdot)\right)}{\varepsilon_j}\geq 0,\qquad \mathbb{P}\text{-a.s.}\tag{1.6}$$

Here

$$\begin{cases}\mathrm{d}\overline{X}(s)=(A(s)+B(s)\Theta(s))\overline{X}(s)\mathrm{d}s+(C(s)+D(s)\Theta(s))\overline{X}(s)\mathrm{d}W(s), & s\in[t,T],\\ \mathrm{d}\overline{Y}(s)=-\big[\big(\widehat{A}(s)+\widehat{B}(s)\Theta(s)\big)\overline{X}(s)+\widehat{C}(s)\overline{Y}(s)+\widehat{D}(s)\overline{Z}(s)\big]\mathrm{d}s+\overline{Z}(s)\mathrm{d}W(s), & s\in[t,T],\\ \overline{X}(t)=x,\quad \overline{Y}(T)=H\overline{X}(T),\end{cases}\tag{1.7}$$

$$\bar{u}(s)=\Theta(s)\overline{X}(s),\quad u^{\varepsilon_j}(s)=\chi_{[t,t+\varepsilon_j]}(s)v+\Theta(s)X^{\varepsilon_j}(s),\tag{1.8}$$

and $X^{\varepsilon_j}(\cdot)$ is the solution of equation (1.1) corresponding to the control $u^{\varepsilon_j}(\cdot)$.

Remark 1.4. The variation of control $\bar{u}(\cdot)$ in Definition 1.3 is given by $\chi_{[t,t+\varepsilon_j]}v+\Theta(s)X^{\varepsilon_j}(s)$, which is different from the general form $\chi_{[t,t+\varepsilon_j)}u(\cdot)$ used in [3, 4]. Although Definition 1.3 is weaker, it is still useful, as shown in [7, 8]. We adopt this definition to facilitate the analysis of our problem using the variational approach.

In what follows, for simplicity of notations, for $\Theta(\cdot)\in L^2(0,T;\mathbb{R}^{k\times n})$, we denote

$$\begin{cases}A_\Theta(s):=A(s)+B(s)\Theta(s),\\ C_\Theta(s):=C(s)+D(s)\Theta(s),\\ \widehat{A}_\Theta(s):=\widehat{A}(s)+\widehat{B}(s)\Theta(s).\end{cases}$$

To study the closed-loop equilibrium strategy of Problem (TI-FBSLQ), we need to introduce the following assumption:

Assumption 1.5. There exists a constant $\mathcal{C}>0$, such that for any $0\leq t\leq\tau\leq s\leq T$, it holds that

$$\begin{aligned}&|Q(s,t)-Q(s,\tau)|_\infty+|R(s,t)-R(s,\tau)|_\infty+|M(s,t)-M(s,\tau)|_\infty\\ &\ +|N(s,t)-N(s,\tau)|_\infty+|G_1(t)-G_1(\tau)|_\infty+|G_2(t)-G_2(\tau)|_\infty\leq\mathcal{C}|t-\tau|.\end{aligned}\tag{1.9}$$

Here and in what follows, to simplify the expression, we make two conventions when there is no ambiguity: First, we denote by $\mathcal{C}$ a generic constant which may vary from line to line; Second, when a function in the equation involves only a single time variable, we omit it, whereas with two time variables, we explicitly specify them.

For $\theta_0(\cdot) \in L^2(0,T;\mathbb{R}^{k\times n})$, consider the following equation:

$$\begin{cases} \dfrac{\mathrm{d}P_1(s;t)}{\mathrm{d}s} + P_1(s;t)A_\Theta + A_\Theta^\top P_1(s;t) \\ \quad + C_\Theta^\top P_1(s;t)C_\Theta + Q(s,t) + \Theta^\top R(s,t)\Theta = 0, & 0 \le t \le s \le T, \\ \dfrac{\mathrm{d}P_2}{\mathrm{d}s} + P_2 A_\Theta + \widehat{A}_\Theta + \widehat{C}P_2 + \widehat{D}P_2 C_\Theta = 0, & 0 \le t \le s \le T, \\ \dfrac{\mathrm{d}P_3(s;t)}{\mathrm{d}s} + P_3(s;t)A_\Theta + A_\Theta^\top P_3(s;t) + C_\Theta^\top P_3(s;t)C_\Theta \\ \quad + P_2^\top M(s,t)P_2 + C_\Theta^\top P_2^\top N(s,t)P_2 C_\Theta = 0, & 0 \le t \le s \le T, \\ P_1(T;t) = G_1(t), \quad P_2(T) = H, \quad P_3(T;t) = 0, & 0 \le t \le T. \end{cases} \tag{1.10}$$

where

$$\begin{aligned} \Theta(t) = &-\big[R(t,t) + D(t)^\top\big(P_1(t;t) + P_3(t;t) + P_2(t)^\top N(t,t)P_2(t)\big)D(t)\big]^\dagger \\ &\times \big[B(t)^\top\big(P_1(t;t) + P_3(t;t)\big) + D(t)^\top\big(P_1(t;t) + P_3(t;t) + P_2(t)^\top N(t,t)P_2(t)\big)C(t) \\ &\quad + \big(\widehat{B}(t)^\top + B(t)^\top P_2(t)^\top + D(t)^\top P_2(t)^\top \widehat{D}(t)^\top\big)G_2(t)P_2(t)\big] + \theta_0(t) \\ &- \big[R(t,t) + D(t)^\top\big(P_1(t;t) + P_3(t;t) + P_2(t)^\top N(t,t)P_2(t)\big)D(t)\big]^\dagger \\ &\times \big[R(t,t) + D(t)^\top\big(P_1(t;t) + P_3(t;t) + P_2(t)^\top N(t,t)P_2(t)\big)D(t)\big]\theta_0(t). \end{aligned} \tag{1.11}$$

Our first main result is as follows.

Theorem 1.6. *Let Assumption 1.5 hold. Then Problem (TI-FBSLQ) admits a closed-loop equilibrium strategy if and only if there exists $\theta_0(\cdot) \in L^2(0,T;\mathbb{R}^{k\times n})$ such that the equation (1.10) admits a solution $(P_1(\cdot;\cdot), P_2(\cdot), P_3(\cdot;\cdot))$ satisfying*

$$\begin{aligned} &\big[R(\cdot,\cdot) + D(\cdot)^\top\big(P_1(\cdot;\cdot) + P_3(\cdot;\cdot) + P_2(\cdot)^\top N(\cdot,\cdot)P_2(\cdot)\big)D(\cdot)\big]^\dagger \\ &\times\big\{\big[B(\cdot)^\top\big(P_1(\cdot;\cdot) + P_3(\cdot;\cdot)\big) + D(\cdot)^\top\big(P_1(\cdot;\cdot) + P_3(\cdot;\cdot) + P_2(\cdot)^\top N(\cdot,\cdot)P_2(\cdot)\big)C(\cdot)\big] \\ &+\big(\widehat{B}(\cdot)^\top + B(\cdot)^\top P_2(\cdot)^\top + D^\top(\cdot)P_2(\cdot)^\top \widehat{D}(\cdot)^\top\big)G_2(\cdot)P_2(\cdot)\big\} \in L^2(0,T;\mathbb{R}^{k\times n}), \end{aligned} \tag{1.12}$$

$$\begin{aligned} &\mathcal{R}\Big(R(t,t) + D(t)^\top\big(P_1(t;t) + P_3(t;t) + P_2(t)^\top N(t,t)P_2(t)\big)D(t)\Big) \supseteq \\ &\mathcal{R}\Big(B(t)^\top\big(P_1(t;t) + P_3(t;t)\big) + D(t)^\top\big(P_1(t;t) + P_3(t;t) + P_2(t)^\top N(t,t)P_2(t)\big)C(t) \\ &\quad + \big(\widehat{B}(t)^\top + B(t)^\top P_2(t)^\top + D(t)^\top P_2(t)^\top \widehat{D}(t)^\top\big)G_2(t)P_2(t)\Big), \quad \text{a.e. } t \in [0,T], \end{aligned} \tag{1.13}$$

and

$$R(t,t) + D(t)^\top P_1(t;t)D(t) + D(t)^\top P_3(t;t)D(t) + D(t)^\top P_2(t)^\top N(t,t)P_2(t)D(t) \ge 0, \quad \text{a.e. } t \in [0,T]. \tag{1.14}$$

In this case, $\Theta(\cdot)$ given by (1.11) is a closed-loop equilibrium strategy.

Remark 1.7. If

$$H = 0,\ \widehat{A}(\cdot) = \widehat{C}(\cdot) = 0,\ \widehat{B}(\cdot) = \widehat{D}(\cdot) = 0,\ M(\cdot,\cdot) = N(\cdot,\cdot) = G_2(\cdot) = 0, \tag{1.15}$$

then the Problem (TI-FBSLQ) can be reduced to the Problem (TISLQ) that was studied in the reference [4]. The equations (1.10)–(1.11) mentioned in the problem are consistent with the equations derived in [4], Theorem 2.10. It is worth noting that in [4], the equations were obtained using the approach of multiperson differential games. However, in our case, we use a different approach, namely, the variational approach, to obtain the equations (1.10)–(1.11).

Remark 1.8. We will refer to equations (1.10)–(1.14) as the "generalized equilibrium Riccati equation"for Problem (TI-FBSLQ). This term is used because when condition (1.15) is satisfied and the time-preference disappears (meaning that the coefficients $Q(\cdot,t)$, $R(\cdot,t)$, $G_1(t)$ do not depend on the initial time t), our problem can be naturally reduced to a classical indefinite stochastic LQ problem. In this case, equations (1.10)–(1.14) become equivalent to the classical generalized Riccati equation as discussed in references [9–11].

In Theorem 1.6, the solvability of equations (1.10)–(1.12) is dependent on the choice of θ_0. This may appear to be counterintuitive. However, the following example demonstrates that such dependence is indeed necessary in Theorem 1.6.

Example 1.9. Let (1.15) hold, $n = k = 1$, and

$$\begin{cases} A(\cdot) = D(\cdot) = 0,\ B(\cdot) = C(\cdot) = 1,\ R(s,t) = s - t, \\ Q(s,t) \geq 0, \quad G_1(t) = -\displaystyle\int_t^T \mathrm{e}^{-(T-\tau)} Q(\tau,t)\mathrm{d}\tau. \end{cases}$$

When we choose $\theta_0(\cdot) = 0$, the equations (1.10)–(1.14) are solvable. In this case, we can get $P_6(\cdot) = 0$, $P_8(\cdot;\cdot) = 0$, and

$$\begin{cases} \dfrac{\mathrm{d}P_1(s;t)}{\mathrm{d}s} = -P_1(s;t) - Q(s,t), & t \leq s \leq T, \\ P_1(T;t) = G_1(t), & 0 \leq t \leq T. \end{cases}$$

Therefore $P_1(t;t) = 0$, and the constraint conditions (1.12)–(1.14) hold. By Theorem 1.6, the closed-loop equilibrium strategies exist.

When we choose $\theta_0(\cdot) = -\frac{1}{2}$, the equations (1.10)–(1.11) are solvable, but the constraint conditions (1.13) fail. Indeed, in this case, we can get $P_6(\cdot) = 0, P_8(\cdot;\cdot) = 0$, and

$$\begin{cases} \dfrac{\mathrm{d}P_1(s;t)}{\mathrm{d}s} = -Q(s,t) - \dfrac{s-t}{4}, & t \leq s \leq T, \\ P_1(T;t) = G_1(t), & 0 \leq t \leq T. \end{cases}$$

Hence

$$\begin{aligned} P_1(t;t) &= G_1(t) - \int_t^T \big[-Q(\tau,t) - (\tau - t)/4\big]\mathrm{d}\tau \\ &= -\int_t^T \mathrm{e}^{-(T-\tau)} Q(\tau,t)\mathrm{d}\tau + \int_t^T \big[Q(\tau,t) + (\tau-t)/4\big]\mathrm{d}\tau \\ &= \int_t^T \big(1 - \mathrm{e}^{-(T-\tau)}\big) Q(\tau,t)\mathrm{d}\tau + (T-t)^2/8 > 0. \end{aligned}$$

We can see that

$$\mathcal{R}\left(R(t,t)+D(t)^\top P_1(t;t)D(t)\right) \equiv \mathcal{R}(0) \not\supseteq \mathcal{R}\left(B(t)^\top P_1(t;t)+D(t)^\top P_1(t;t)C(t)\right).$$

Consequently, (1.13) does not hold.

This example demonstrates that, in general, for Problem (TI-FBSLQ), the solvability of the equations (1.10)–(1.14) is dependent on the choice of the parameter $\theta_0(\cdot)$, even if its closed-loop equilibrium strategy exists.

Theorem 1.6 establishes the connection between the existence of closed-loop equilibrium strategies and the solvability of the system (1.10)–(1.14). It is natural to inquire about the conditions under which the system (1.10)–(1.14) is solvable. Unfortunately, as of now, we only have the answer to this question for the one-dimensional case, which is the second main result of this paper. To present it, we first make the following assumption:

Assumption 1.10. Let $m=n=k=1$. There exists a constant $\delta>0$, such that

$$\begin{cases} R(t,t)\geq \delta, & t\in[0,T],\\ N(t,t)\geq\delta, & t\in[0,T],\\ D(t)^\top D(t)\geq\delta, & \text{a.e. } t\in[0,T]. \end{cases}$$

Moreover,

$$\begin{cases} Q(s,t)\geq 0,\\ M(s,t)\geq 0, & 0\leq t\leq s\leq T.\\ G_1(t)\geq 0, \end{cases}$$

Remark 1.11. In Assumption 1.10, $D^\top D\geq\delta$ is a strong requirement which essentially helps us to establish *a priori* estimate for $\Theta(\cdot)$, consequently leading to the well-posedness of (1.10)–(1.11). In some sense, this means that the control on the diffusion term plays a positive role in obtaining the well-posedness of (1.10)–(1.11). On the other hand, under Assumptions 1.10, we have

$$R(t,t)+D(t)^\top\left(P_1(t;t)+P_3(t;t)+P_2(t)^\top N(t,t)P_2(t)\right)D(t)\geq\delta, \quad \text{a.e. } t\in[0,T].$$

Then the parameter $\theta_0(\cdot)$ in (1.11) naturally vanishes and the generalized equilibrium Riccati equation (1.10)–(1.14) reduces to an equilibrium Riccati equation (1.10)–(1.11) without the parameter $\theta_0(\cdot)$.

The second main result of this paper is as follows:

Theorem 1.12. *Let Assumptions 1.5–1.10 hold. Then for any $\theta_0(\cdot)\in L^2(0,T;\mathbb{R})$, the system (1.10)–(1.14) admits a unique solution $(P_1(\cdot;\cdot),P_2(\cdot),P_3(\cdot;\cdot))$. Consequently, Problem (TI-FBSLQ) admits a unique equilibrium strategy given by*

$$\begin{aligned}\Theta(t)=&-\left[R(t,t)+D(t)^\top\left(P_1(t;t)+P_3(t;t)+P_2(t)^\top N(t,t)P_2(t)\right)D(t)\right]^\dagger\\ &\times\left[B(t)^\top\left(P_1(t;t)+P_3(t;t)\right)+D(t)^\top\left(P_1(t;t)+P_3(t;t)+P_2(t)^\top N(t,t)P_2(t)\right)\right.\\ &\left.\times C(t)+\left(\widehat{B}(t)^\top+B(t)^\top P_2(t)^\top+D(t)^\top P_2(t)^\top\widehat{D}(t)^\top\right)G_2(t)P_2(t)\right].\end{aligned} \tag{1.16}$$

Since the seminal work [12] on optimal control problems for forward–backward stochastic differential equations (FBSDEs), this area of research has received significant attention. Numerous studies have been conducted, although it is impossible to provide an exhaustive list. Interested readers can refer to [13–16], as well as the references cited therein.

It is worth noting that in the aforementioned literature, the essential time-inconsistency of optimal control problems for FBSDEs has been ignored, with the focus being on pre-committed optimal control. However, a recent study [2] revealed that time-inconsistency is indeed present in these problems. This implies that the pre-committed optimal control obtained at the initial time may not remain optimal at a later time.

The study of time-inconsistent problems can be traced back to the mid-18th century, with works by Hume [17] and Smith [18] analyzing the time-inconsistent behavior of animals and humans. Strotz [19] was the first to mathematically formulate a time-inconsistent problem within an economic context. Since then, time-inconsistent problems have been widely studied in economics and finance (*e.g.*, [20–24]), focusing mainly on discrete dynamic systems, simple ordinary differential equations (ODEs), or stochastic differential equations (SDEs). Inspired by these studies, researchers in control theory started investigating time-inconsistent optimal control problems. Substantial literature has been published on this subject, including [2–8, 25], as well as the references cited therein, for time-inconsistent linear-quadratic (LQ) problems for SDEs. Among these works, there are two main methods employed. The first one is based on *multiperson differential games*. By this method, Yong [6] obtained a unique closed-loop equilibrium strategy for time-inconsistent stochastic LQ problems under *standard conditions* (namely, the control cost weighting matrix R is uniformly positive definite, and the other weighting coefficients Q, G_1 are positive semidefinite). The second one is based on variational approaches and decoupling techniques. Hu *et al.* [7, 8] applied this method to obtain a unique equilibrium control and provided an equivalent characterization of that control for time-inconsistent stochastic LQ problems under *standard conditions* and a singular case where they still require $R + D^\top MD$ to be uniformly positive definite (Here M satisfies a certain Riccati equation). In these works, the proof of the existence of closed-loop equilibrium strategies or equilibrium controls essentially requires $R + D^\top P_1 D$ or $R + D^\top MD$ to be uniformly positive definite. Recently, through multiperson differential games and the establishment of a sharp estimate, Lü and Ma [4] introduced conditions weaker than prior researches, where $R + D^\top P_1 D$ can be singular, to guarantee the existence of closed-loop equilibrium strategies for time-inconsistent stochastic LQ problems. It should be noted that in the aforementioned works, the control systems are SDEs, and the time-inconsistency arises from *time-preferences* or *risk-preferences*. Very recently, Wang *et al.* [2] discovered that the forward–backward structure of the control system can also lead to time-inconsistency. They introduced a general framework for time-inconsistent optimal control problems for FBSDEs, and proved the well-posedness of the *equilibrium Hamilton-Jacobi-Bellman equation* when the diffusion coefficient does not contain the control variable. They also posed a time-inconsistent FBSDE linear-quadratic (TI-FBSLQ) problem but did not provide a proof for the well-posedness of the equilibrium Riccati equation.

Despite the rich literature on optimal control theory for FBSDEs and time-inconsistent optimal control theory for SDEs, there has been limited research on time-inconsistent optimal control theory for FBSDEs, with the exception of the aforementioned study in [2].

The main contributions of the current paper are as follows:

1. We obtained the generalized equilibrium Riccati equation for Problem (TI-FBSLQ), which is a coupled equilibrium Riccati equation system with parameters and constraint conditions. The solvability of this equation fully characterizes the existence of closed-loop equilibrium strategies for Problem (TI-FBSLQ). We also provide an equivalent characterization of the closed-loop equilibrium strategy. To obtain the results, the main difficulties lie in three aspects: First, the cost weighting matrices for the state and the control are allowed to be indefinite. In [4], Lü and Ma utilized multiperson differential games to relax the standard conditions to some extent, and derived a family of equilibrium Riccati equations depending on parameters to obtain closed-loop equilibrium strategies. However, the method in [4] is not suitable for handling the more general indefinite case. To address the indefinite case, we adopt a different approach, the variational approach. Second, the estimate for the martingale term Z in the BSDE is generally not enough for us to handle the problem. To overcome this difficulty, we notice that the BSDE is actually always coupled with a nice SDE, which may imply that Z have some good properties. Hence we

technically use decoupling techniques multiple times to dig the hidden information of Z. Eventually, we introduce fourteen auxiliary equations $\{\mathcal{P}_i\}_{i=1}^{14}$ for decoupling. Third, the time-preferences, namely, $Q(\cdot,t), R(\cdot,t), M(\cdot,t), N(\cdot,t), G_1(t), G_2(t)$ depending on the initial time t, pose another obstacle to obtaining the result. We overcome it by establishing stability estimates for some auxiliary equations generated during the derivation and subsequently obtaining some convergence results.

2. Our study reveals a new and interesting phenomenon. For Problem (TI-FBSLQ), even if closed-loop equilibrium strategies exist, the solvability of the generalized equilibrium Riccati equation depends on the choice of the parameter. This is different from the time-consistent scenario, where the existence of closed-loop optimal control implies the solvability of the generalized Riccati equation.
3. When the state is one-dimensional and under slightly stronger assumptions than standard conditions, we prove the well-posedness of the generalized equilibrium Riccati equation for any chosen parameter. From this result, we can explicitly obtain a closed-loop equilibrium strategy for Problem (TI-FBSLQ). The main difficulty lies in the complicated nature of the coupled differential equation system with non-local terms and constraint conditions. We introduce assumptions to eliminate the constraint conditions and simplify the problem into a system of differential equations with non-local terms. By appropriately applying the Banach fixed-point theorem piecewise, we establish the well-posedness on the entire interval.

The rest of this paper is organized as follows. Section 2 is devoted to the proof of Theorem 1.6 and Section 3 addresses the proof of Theorem 1.12.

2. Proof of Theorem 1.6

The purpose of this section is to provide a proof for Theorem 1.6. To achieve this, it is crucial to demonstrate the following equivalent characterization of closed-loop equilibrium strategies for Problem (TI-FBSLQ).

Theorem 2.1. *Let Assumption 1.5 hold. Then $\Theta(\cdot) \in L^2(0,T;\mathbb{R}^{k\times n})$ is a closed-loop equilibrium strategy of Problem (TI-FBSLQ) if and only if for* a.e. $t \in [0,T]$,

$$R(t,t) + D(t)^\top \mathbf{P_1}(t;t)D(t) + D(t)^\top \mathbf{P_3}(t;t)D(t) + D(t)^\top \mathbf{P_2}(t)^\top N(t,t)\mathbf{P_2}(t)D(t) \geq 0, \tag{2.1}$$

and

$$\begin{aligned} &B(t)^\top \mathbf{P_1}(t;t) + D(t)^\top \mathbf{P_1}(t;t)C_\Theta(t) + B(t)^\top \mathbf{P_3}(t;t) + D(t)^\top \mathbf{P_3}(t;t)C_\Theta(t) \\ &+ R(t,t)\Theta(t) + \big(\widehat{B}(t)^\top + B(t)^\top \mathbf{P_2}(t)^\top + D(t)^\top \mathbf{P_2}(t)^\top \widehat{D}(t)^\top\big)G_2(t)\mathbf{P_2}(t) \\ &+ D(t)^\top \mathbf{P_2}(t)^\top N(t,t)\mathbf{P_2}(t)C_\Theta(t) = 0, \end{aligned} \tag{2.2}$$

where $(\mathbf{P_1}(\cdot;\cdot), \mathbf{P_2}(\cdot), \mathbf{P_3}(\cdot;\cdot))$ satisfies the following equations:

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{P_1}(s;t)}{\mathrm{d}s} + \mathbf{P_1}(s;t)A_\Theta + A_\Theta^\top \mathbf{P_1}(s;t) \\ \quad + C_\Theta^\top \mathbf{P_1}(s;t)C_\Theta + Q(s,t) + \Theta^\top R(s,t)\Theta = 0, & 0 \leq t \leq s \leq T, \\ \dfrac{\mathrm{d}\mathbf{P_2}}{\mathrm{d}s} + \mathbf{P_2}A_\Theta + \widehat{A}_\Theta + \widehat{C}\mathbf{P_2} + \widehat{D}\mathbf{P_2}C_\Theta = 0, & 0 \leq t \leq s \leq T, \\ \dfrac{\mathrm{d}\mathbf{P_3}(s;t)}{\mathrm{d}s} + \mathbf{P_3}(s;t)A_\Theta + A_\Theta^\top \mathbf{P_3}(s;t) + C_\Theta^\top \mathbf{P_3}(s;t)C_\Theta \\ \quad + \mathbf{P_2}^\top M(s,t)\mathbf{P_2} + C_\Theta^\top \mathbf{P_2}^\top N(s,t)\mathbf{P_2}C_\Theta = 0, & 0 \leq t \leq s \leq T, \\ \mathbf{P_1}(T;t) = G_1(t), \quad \mathbf{P_2}(T) = H, \quad \mathbf{P_3}(T;t) = 0, & 0 \leq t \leq T. \end{cases} \tag{2.3}$$

Proof of Theorem 2.1. Since the proof is long, we divide it into eleven steps.

Step 1. In this step, we compute $\mathcal{J}(t,x;u^\varepsilon(\cdot))-\mathcal{J}(t,x;u(\cdot))$.

For any $(t,x)\in[0,T)\times L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^n)$, $v\in L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^k)$ and $\Theta(\cdot)\in L^2(0,T;\mathbb{R}^{k\times n})$, consider the following equations:

$$
\begin{cases}
\mathrm{d}X=A_\Theta X\mathrm{d}s+C_\Theta X\mathrm{d}W(s), & s\in[t,T],\\
\mathrm{d}Y=-\big(\widehat{A}_\Theta X+\widehat{C}Y+\widehat{D}Z\big)\mathrm{d}s+Z\mathrm{d}W(s), & s\in[t,T],\\
X(t)=x,\quad Y(T)=HX(T),
\end{cases}
\tag{2.4}
$$

and

$$
\begin{cases}
\mathrm{d}X^\varepsilon(s;t)=\big(A_\Theta X^\varepsilon(s;t)+\chi_{[t,t+\varepsilon]}Bv\big)\mathrm{d}s+\big(C_\Theta X^\varepsilon(s;t)+\chi_{[t,t+\varepsilon]}Dv\big)\mathrm{d}W(s), & s\in[t,T],\\
\mathrm{d}Y^\varepsilon(s;t)=-\big(\widehat{A}_\Theta X^\varepsilon(s;t)+\chi_{[t,t+\varepsilon]}\widehat{B}v+\widehat{C}Y^\varepsilon(s;t)+\widehat{D}Z^\varepsilon(s;t)\big)\mathrm{d}s & \\
\qquad\qquad +Z^\varepsilon(s;t)\mathrm{d}W(s), & s\in[t,T],\\
X^\varepsilon(t;t)=x,\quad Y^\varepsilon(T;t)=HX^\varepsilon(T;t).
\end{cases}
\tag{2.5}
$$

Let $X_0^\varepsilon:=X^\varepsilon-X, Y_0^\varepsilon:=Y^\varepsilon-Y, Z_0^\varepsilon:=Z^\varepsilon-Z$. Then we have

$$
\begin{cases}
\mathrm{d}X_0^\varepsilon(s;t)=\big(A_\Theta X_0^\varepsilon(s;t)+\chi_{[t,t+\varepsilon]}Bv\big)\mathrm{d}s+\big(C_\Theta X_0^\varepsilon(s;t)+\chi_{[t,t+\varepsilon]}Dv\big)\mathrm{d}W(s), & s\in[t,T],\\
\mathrm{d}Y_0^\varepsilon(s;t)=-\big(\widehat{A}_\Theta X_0^\varepsilon(s;t)+\chi_{[t,t+\varepsilon]}\widehat{B}v+\widehat{C}Y_0^\varepsilon(s;t)+\widehat{D}Z_0^\varepsilon(s;t)\big)\mathrm{d}s & \\
\qquad\qquad +Z_0^\varepsilon(s;t)\mathrm{d}W(s), & s\in[t,T],\\
X_0^\varepsilon(t;t)=0,\quad Y_0^\varepsilon(T;t)=HX_0^\varepsilon(T;t).
\end{cases}
\tag{2.6}
$$

A direct calculation gives that

$$
\begin{aligned}
&\mathcal{J}(t,x;u^\varepsilon(\cdot))-\mathcal{J}(t,x;u(\cdot))\\
&=\frac{1}{2}\mathbb{E}_t\bigg[\int_t^T\Big(\big\langle Q(s,t)X_0^\varepsilon(s;t),X_0^\varepsilon(s;t)\big\rangle+\big\langle R(s,t)\big(\Theta X_0^\varepsilon(s;t)+v\chi_{[t,t+\varepsilon]}\big),\Theta X_0^\varepsilon(s;t)\\
&\qquad\qquad +v\chi_{[t,t+\varepsilon]}\big\rangle+\big\langle M(s,t)Y_0^\varepsilon(s;t),Y_0^\varepsilon(s;t)\big\rangle+\big\langle N(s,t)Z_0^\varepsilon(s;t),Z_0^\varepsilon(s;t)\big\rangle\Big)\mathrm{d}s\\
&\qquad +2\int_t^T\Big(\big\langle Q(s,t)X,X_0^\varepsilon(s;t)\big\rangle+\big\langle R(s,t)\Theta X,\Theta X_0^\varepsilon(s;t)+v\chi_{[t,t+\varepsilon]}\big\rangle\\
&\qquad\qquad +\big\langle M(s,t)Y,Y_0^\varepsilon(s;t)\big\rangle+\big\langle N(s,t)Z,Z_0^\varepsilon(s;t)\big\rangle\Big)\mathrm{d}s+\big\langle G_1(t)X_0^\varepsilon(T;t),X_0^\varepsilon(T;t)\big\rangle\\
&\qquad +\big\langle G_2(t)Y_0^\varepsilon(t;t),Y_0^\varepsilon(t;t)\big\rangle+2\big\langle G_1(t)X(T),X_0^\varepsilon(T;t)\big\rangle+2\big\langle G_2(t)Y(t),Y_0^\varepsilon(t;t)\big\rangle\bigg].
\end{aligned}
$$

Consequently, it holds that

$$
\begin{aligned}
&\mathcal{J}(t,x;u^\varepsilon(\cdot))-\mathcal{J}(t,x;u(\cdot))\\
&=J_1(t,x)+J_2(t,x)+J_3(t,x)+J_4(t,x)+\mathbb{E}_t\int_t^{t+\varepsilon}\big\langle\Theta^\top R(s,t)v,X_0^\varepsilon(s;t)\big\rangle\mathrm{d}s,
\end{aligned}
\tag{2.7}
$$

where

$$\begin{cases} J_1(t,x) := \mathbb{E}_t \displaystyle\int_t^T \big(\big\langle F_1(s,t), X_0^\varepsilon(s;t)\big\rangle + \big\langle F_2(s,t), v\chi_{[t,t+\varepsilon]}\big\rangle\big)\mathrm{d}s \\ \qquad\qquad +\mathbb{E}_t\big\langle G_1(t)X(T), X_0^\varepsilon(T;t)\big\rangle, \\ J_2(t,x) := \dfrac{1}{2}\mathbb{E}_t \displaystyle\int_t^T \big\langle F_1^\varepsilon(s,t), X_0^\varepsilon(s;t)\big\rangle\mathrm{d}s + \dfrac{1}{2}\mathbb{E}_t\big\langle G_1(t)X_0^\varepsilon(T;t), X_0^\varepsilon(T;t)\big\rangle, \\ J_3(t,x) := \dfrac{1}{2}\mathbb{E}_t \displaystyle\int_t^T \big(\big\langle F_3^\varepsilon(s,t), Y_0^\varepsilon(s;t)\big\rangle + \big\langle F_4^\varepsilon(s,t), Z_0^\varepsilon(s;t)\big\rangle\big)\mathrm{d}s \\ \qquad\qquad +\dfrac{1}{2}\mathbb{E}_t\big\langle G_2(t)Y_0^\varepsilon(t;t), Y_0^\varepsilon(t;t)\big\rangle, \\ J_4(t,x) := \mathbb{E}_t \displaystyle\int_t^T \big(\big\langle F_3(s,t), Y_0^\varepsilon(s;t)\big\rangle + \big\langle F_4(s,t), Z_0^\varepsilon(s;t)\big\rangle\big)\mathrm{d}s \\ \qquad\qquad +\mathbb{E}_t\big\langle G_2(t)Y(t), Y_0^\varepsilon(t;t)\big\rangle, \end{cases} \tag{2.8}$$

with

$$\begin{cases} F_1(s,t) \equiv \big(Q(s,t) + \Theta(s)^\top R(s,t)\Theta(s)\big)X(s), \\ F_2(s,t) \equiv \dfrac{1}{2}R(s,t)v + R(s,t)\Theta(s)X(s), \\ F_1^\varepsilon(s,t) \equiv \big(Q(s,t) + \Theta(s)^\top R(s,t)\Theta(s)\big)X_0^\varepsilon(s;t), \\ F_3(s,t) \equiv M(s,t)Y(s), \quad F_4(s,t) \equiv N(s,t)Z(s), \\ F_3^\varepsilon(s,t) \equiv M(s,t)Y_0^\varepsilon(s;t), \quad F_4^\varepsilon(s,t) \equiv N(s,t)Z_0^\varepsilon(s;t). \end{cases} \tag{2.9}$$

Step 2. In this step, we prove the following estimates:

$$\begin{aligned} &\mathbb{E}_t \sup_{s\in[t,T]} |X_0^\varepsilon(s;t)|_{\mathbb{R}^n}^2 \le \mathcal{C}\varepsilon|v|_{\mathbb{R}^k}^2, \quad \text{a.s.}, \\ &\mathbb{E}_t\Big(\sup_{s\in[t,T]} |Y_0^\varepsilon(s;t)|_{\mathbb{R}^m}^2 + \int_t^T |Z_0^\varepsilon(s;t)|_{\mathbb{R}^m}^2\mathrm{d}s\Big) \le \mathcal{C}\varepsilon|v|_{\mathbb{R}^k}^2, \quad \text{a.s.} \end{aligned} \tag{2.10}$$

The inequality (2.10) should be a known result. However, we failed to find an exact reference for it. Hence, we provide a proof here for the convenience of readers.

First, for any $\mathcal{A} \in \mathcal{F}_t$, we have

$$\begin{cases} \mathrm{d}\chi_{\mathcal{A}}X_0^\varepsilon(s;t) = \big(\chi_{\mathcal{A}}A_\Theta X_0^\varepsilon(s;t) + \chi_A\chi_{[t,t+\varepsilon]}Bv\big)\mathrm{d}s \\ \qquad\qquad +\big(\chi_{\mathcal{A}}C_\Theta X_0^\varepsilon(s;t) + \chi_{\mathcal{A}}\chi_{[t,t+\varepsilon]}Dv\big)\mathrm{d}W(s), \quad s\in[t,T], \\ \chi_{\mathcal{A}}X_0^\varepsilon(t;t) = 0. \end{cases}$$

Then,

$$\begin{aligned} &\mathbb{E}\Big(\chi_{\mathcal{A}} \sup_{s\in[t,T]} |X_0^\varepsilon(s;t)|_{\mathbb{R}^n}^2\Big) \\ &= \mathbb{E}\Big(\sup_{s\in[t,T]} |\chi_{\mathcal{A}}X_0^\varepsilon(s;t)|_{\mathbb{R}^n}^2\Big) \end{aligned}$$

$$\begin{aligned}&\leq \mathcal{C}\mathbb{E}\Big[\Big(\int_t^{t+\varepsilon}|\chi_{\mathcal{A}}Bv|_{\mathbb{R}^n}\mathrm{d}s\Big)^2+\int_t^{t+\varepsilon}|\chi_{\mathcal{A}}Dv|^2_{\mathbb{R}^n}\mathrm{d}s\Big]\\&\leq \mathcal{C}\mathbb{E}\big(\chi_{\mathcal{A}}\varepsilon|v|^2_{\mathbb{R}^k}\big),\end{aligned}\tag{2.11}$$

where the constant $\mathcal{C}$ only depends on A_Θ, C_Θ, B, D, T and is independent of $\mathcal{A}$. Since v is $\mathcal{F}_t$-measurable, we have

$$\mathbb{E}_t\sup_{s\in[t,T]}|X_0^\varepsilon(s;t)|^2_{\mathbb{R}^n}\leq \mathcal{C}\varepsilon|v|^2_{\mathbb{R}^k},\quad \text{a.s.}$$

Secondly, we similarly consider the following equation:

$$\begin{cases}\mathrm{d}\chi_{\mathcal{A}}Y_0^\varepsilon(s;t)=-\big(\chi_{\mathcal{A}}\widehat{A}_\Theta X_0^\varepsilon(s;t)+\chi_{\mathcal{A}}\chi_{[t,t+\varepsilon]}\widehat{B}v+\chi_{\mathcal{A}}\widehat{C}Y_0^\varepsilon(s;t)\\ \qquad\qquad +\chi_{\mathcal{A}}\widehat{D}Z_0^\varepsilon(s;t)\big)\mathrm{d}s+\chi_{\mathcal{A}}Z_0^\varepsilon(s;t)\mathrm{d}W(s),\quad s\in[t,T],\\ \chi_{\mathcal{A}}Y_0^\varepsilon(T;t)=\chi_{\mathcal{A}}HX_0^\varepsilon(T;t).\end{cases}$$

Then from (2.11), we have

$$\begin{aligned}&\mathbb{E}\Big[\chi_{\mathcal{A}}\Big(\sup_{s\in[t,T]}|Y_0^\varepsilon(s;t)|^2_{\mathbb{R}^m}+\int_t^T|Z_0^\varepsilon(s;t)|^2_{\mathbb{R}^m}\mathrm{d}s\Big)\Big]\\&=\mathbb{E}\,\Big(\sup_{s\in[t,T]}|\chi_{\mathcal{A}}Y_0^\varepsilon(s;t)|^2_{\mathbb{R}^m}+\int_t^T|\chi_{\mathcal{A}}Z_0^\varepsilon(s;t)|^2_{\mathbb{R}^m}\mathrm{d}s\Big)\\&\leq \mathcal{C}\mathbb{E}\Big[\chi_{\mathcal{A}}|X_0^\varepsilon(T;t)|^2_{\mathbb{R}^n}+\Big(\int_t^T\big|\chi_{\mathcal{A}}\widehat{A}_\Theta X_0^\varepsilon(s;t)+\chi_{\mathcal{A}}\chi_{[t,t+\varepsilon]}\widehat{B}v\big|_{\mathbb{R}^m}\mathrm{d}s\Big)^2\Big]\\&\leq \mathcal{C}\mathbb{E}(\chi_{\mathcal{A}}\varepsilon|v|^2_{\mathbb{R}^k}).\end{aligned}$$

This implies that

$$\mathbb{E}_t\Big(\sup_{s\in[t,T]}|Y_0^\varepsilon(s;t)|^2_{\mathbb{R}^m}+\int_t^T|Z_0^\varepsilon(s;t)|^2_{\mathbb{R}^m}\mathrm{d}s\Big)\leq \mathcal{C}\varepsilon|v|^2_{\mathbb{R}^k},\quad \text{a.s.}$$

In the following four steps, we deal with $J_i(t,x)$, $i=1,\cdots,4$.

Step 3. In this step, we give an estimate of $J_1(t,x)$.

First, consider the following adjoint equation:

$$\begin{cases}\mathrm{d}Y_1(s;t)=-\big[A_\Theta^\top Y_1(s;t)+C_\Theta^\top Z_1(s;t)+\big(Q(s;t)\\ \qquad\qquad +\Theta^\top R(s;t)\Theta\big)X\big]\mathrm{d}s+Z_1(s;t)\mathrm{d}W(s),\quad s\in[t,T],\\ Y_1(T;t)=G_1(t)X(T).\end{cases}\tag{2.12}$$

Applying Ito's formula for $\langle Y_1,X_0^\varepsilon\rangle$, we have

$$J_1(t,x)=\mathbb{E}_t\int_t^{t+\varepsilon}\big\langle B^\top Y_1(s;t)+D^\top Z_1(s;t)+F_2(s,t),v\big\rangle\mathrm{d}s.\tag{2.13}$$

Next, we get rid of the terms containing Y_1 and Z_1 in (2.13). To this end, we introduce the following equation:

$$\begin{cases} \dfrac{\mathrm{d}\mathcal{P}_1(s;t)}{\mathrm{d}s} + \mathcal{P}_1(s;t)A_\Theta + A_\Theta^\top \mathcal{P}_1(s;t) \\ \qquad + C_\Theta^\top \mathcal{P}_1(s;t)C_\Theta + Q(s,t) + \Theta^\top R(s,t)\Theta = 0, \quad s \in [t,T], \\ \mathcal{P}_1(T;t) = G_1(t). \end{cases} \tag{2.14}$$

By (2.4) and (2.14), we have

$$\begin{aligned} \mathrm{d}\big(\mathcal{P}_1 X\big) &= -\big(\mathcal{P}_1 A_\Theta + A_\Theta^\top \mathcal{P}_1 + C_\Theta^\top \mathcal{P}_1 C_\Theta + Q + \Theta^\top R\Theta\big)X\mathrm{d}s + \mathcal{P}_1 A_\Theta X\mathrm{d}s + \mathcal{P}_1 C_\Theta X\mathrm{d}W(s) \\ &= -\big(A_\Theta^\top \mathcal{P}_1 + C_\Theta^\top \mathcal{P}_1 C_\Theta + Q + \Theta^\top R\Theta\big)X\mathrm{d}s + \mathcal{P}_1 C_\Theta X\mathrm{d}W(s). \end{aligned}$$

This, together with $\mathcal{P}_1(T;t) = G_1(t)$, implies that $(\mathcal{P}_1 X, \mathcal{P}_1 C_\Theta X)$ is a solution to (2.12). By the uniqueness of the solution to (2.12), we obtain

$$\begin{cases} Y_1(\cdot;t) = \mathcal{P}_1(\cdot;t)X(\cdot), \\ Z_1(\cdot;t) = \mathcal{P}_1(\cdot;t)C_\Theta(\cdot)X(\cdot). \end{cases} \tag{2.15}$$

This, together with (2.13), implies that

$$\begin{aligned} J_1(t,x) &= \mathbb{E}_t \int_t^{t+\varepsilon} \big\langle B^\top Y_1(s;t) + D^\top Z_1(s;t) + F_2(s,t), v\big\rangle \mathrm{d}s \\ &= \mathbb{E}_t \int_t^{t+\varepsilon} \Big\langle \big(B^\top \mathcal{P}_1(s;t) + D^\top \mathcal{P}_1(s;t)C_\Theta + R(s,t)\Theta\big)X + \frac{1}{2}R(s,t)v, v\Big\rangle \mathrm{d}s. \end{aligned} \tag{2.16}$$

Step 4. In this step, we give an estimate of $J_2(t,x)$.

First, we introduce the following adjoint equation:

$$\begin{cases} \mathrm{d}Y_2(s;t) = -\big[A_\Theta^\top Y_2(s;t) + C_\Theta^\top Z_2(s;t) + \big(Q(s,t) + \Theta^\top R(s,t)\Theta\big)X_0^\varepsilon(s;t)\big]\mathrm{d}s \\ \qquad\qquad\qquad + Z_2(s;t)\mathrm{d}W(s), \qquad s \in [t,T], \\ Y_2(T;t) = G_1(t)X_0^\varepsilon(T;t). \end{cases} \tag{2.17}$$

Applying Ito's formula for $\langle Y_2, X_0^\varepsilon\rangle$, we have

$$J_2(t,x) = \frac{1}{2}\mathbb{E}_t \int_t^{t+\varepsilon} \big\langle B^\top Y_2(s;t) + D^\top Z_2(s;t), v\big\rangle \mathrm{d}s. \tag{2.18}$$

To get rid of the terms containing Y_2 and Z_2 in (2.18), we introduce the following equations:

$$\begin{cases} \dfrac{\mathrm{d}\mathcal{P}_2(s;t)}{\mathrm{d}s} + \mathcal{P}_2(s;t)A_\Theta + A_\Theta^\top \mathcal{P}_2(s;t) + C_\Theta^\top \mathcal{P}_2(s;t)C_\Theta + Q(s,t) + \Theta^\top R(s,t)\Theta = 0, \quad s \in [t,T], \\ \mathcal{P}_2(T;t) = G_1(t), \end{cases} \tag{2.19}$$

and

$$\begin{cases} \dfrac{\mathrm{d}\mathcal{P}_3(s;t)}{\mathrm{d}s} + A_\Theta^\top \mathcal{P}_3(s;t) + \chi_{[t,t+\varepsilon]}\mathcal{P}_2(s;t)B + \chi_{[t,t+\varepsilon]}C_\Theta^\top \mathcal{P}_2(s;t)D = 0, \quad s \in [t,T], \\ \mathcal{P}_3(T;t) = 0. \end{cases} \tag{2.20}$$

From (2.6), (2.19) and (2.20), we get

$$\begin{aligned} &\mathrm{d}\big(\mathcal{P}_2 X_0^\varepsilon + \mathcal{P}_3 v\big) \\ &= -\big(\mathcal{P}_2 A_\Theta + A_\Theta^\top \mathcal{P}_2 + C_\Theta^\top \mathcal{P}_2 C_\Theta + Q + \Theta^\top R \Theta\big) X_0^\varepsilon \mathrm{d}s \\ &\quad -\big(A_\Theta^\top \mathcal{P}_3 v + \chi_{[t,t+\varepsilon]}\mathcal{P}_2 B v + \chi_{[t,t+\varepsilon]} C_\Theta^\top \mathcal{P}_2 D v\big)\mathrm{d}s \\ &\quad + \mathcal{P}_2\big(A_\Theta X_0^\varepsilon + \chi_{[t,t+\varepsilon]} B v\big)\mathrm{d}s + \mathcal{P}_2\big(C_\Theta X_0^\varepsilon + \chi_{[t,t+\varepsilon]} D v\big)\mathrm{d}W(s) \\ &= -\big[A_\Theta^\top\big(\mathcal{P}_2 X_0^\varepsilon + \mathcal{P}_3 v\big) + C_\Theta^\top \mathcal{P}_2\big(C_\Theta X_0^\varepsilon + \chi_{[t,t+\varepsilon]} D v\big) + \big(Q + \Theta^\top R \Theta\big) X_0^\varepsilon\big]\mathrm{d}s \\ &\qquad + \mathcal{P}_2\big(C_\Theta X_0^\varepsilon + \chi_{[t,t+\varepsilon]} D v\big)\mathrm{d}W(s). \end{aligned}$$

This, together with $\mathcal{P}_2(T;t) = G_1(t)$ and $\mathcal{P}_3(T;t) = 0$, implies that $\big(\mathcal{P}_2 X_0^\varepsilon + \mathcal{P}_3 v, \mathcal{P}_2(C_\Theta X_0^\varepsilon + \chi_{[t,t+\varepsilon]} D v)\big)$ is a solution to (2.17). Then, the uniqueness of the solution to (2.17) implies that

$$\begin{cases} Y_2(\cdot;t) = \mathcal{P}_2(\cdot;t) X_0^\varepsilon(\cdot;t) + \mathcal{P}_3(\cdot;t) v, \\ Z_2(\cdot;t) = \mathcal{P}_2(\cdot;t)\big(C_\Theta(\cdot) X_0^\varepsilon(\cdot;t) + D(\cdot) v \chi_{[t,t+\varepsilon]}\big). \end{cases} \tag{2.21}$$

By Gronwall's inequality, we get from (2.20) that

$$\begin{aligned} \sup_{s\in[t,T]} |\mathcal{P}_3(s;t)|_\infty^2 &\le \mathcal{C}\left(\int_t^{t+\varepsilon} \big|\mathcal{P}_2(\tau;t)B + C_\Theta^\top \mathcal{P}_2(\tau;t) D\big|_\infty \mathrm{d}\tau\right)^2 \\ &\le \mathcal{C}\left[\int_t^{t+\varepsilon} (1 + |\Theta|_\infty)\mathrm{d}\tau\right]^2. \end{aligned} \tag{2.22}$$

Combining (2.21), (2.22) and (2.10), we obtain that

$$\begin{aligned} J_2(t,x) &= \frac{1}{2}\mathbb{E}_t \int_t^{t+\varepsilon} \Big\langle B^\top Y_2(s;t) + D^\top Z_2(s;t), v\Big\rangle \mathrm{d}s \\ &= \frac{1}{2}\mathbb{E}_t \int_t^{t+\varepsilon} \Big\langle \big(B^\top \mathcal{P}_2(s;t) + D^\top \mathcal{P}_2(s;t) C_\Theta\big) X_0^\varepsilon(s;t) \\ &\qquad\qquad + \big(B^\top \mathcal{P}_3(s;t) + D^\top \mathcal{P}_2(s;t)\chi_{[t,t+\varepsilon]} D\big) v, v\Big\rangle \mathrm{d}s \\ &= \frac{1}{2}\mathbb{E}_t \int_t^{t+\varepsilon} \langle D^\top \mathcal{P}_2(s;t) D v, v\rangle \mathrm{d}s + \mathcal{C}|v|_{\mathbb{R}^k}^2 o(\varepsilon). \end{aligned} \tag{2.23}$$

Step 5. In this step, we give an estimate of $J_3(t,x)$.

We first introduce the adjoint equation:

$$\begin{cases} \mathrm{d}Y_3(s;t) = \big(\widehat{C}^\top Y_3(s;t) + F_3^\varepsilon(s,t)\big)\mathrm{d}s + \big(\widehat{D}^\top Y_3(s;t) + F_4^\varepsilon(s,t)\big)\mathrm{d}W(s), \quad s \in [t,T], \\ Y_3(t;t) = G_2(t) Y_0^\varepsilon(t;t). \end{cases} \tag{2.24}$$

Applying Ito's formula for $\langle Y_3, Y_0^\varepsilon\rangle$, we have that

$$J_3(t,x) = \frac{1}{2}\mathbb{E}_t\big\langle Y_3(T;t), HX_0^\varepsilon(T;t)\big\rangle + \frac{1}{2}\mathbb{E}_t\int_t^{t+\varepsilon}\big\langle \widehat{B}^\top Y_3(s;t), v\big\rangle \mathrm{d}s$$
$$+\frac{1}{2}\mathbb{E}_t\int_t^T\big\langle \widehat{A}_\Theta^\top Y_3(s;t), X_0^\varepsilon(s;t)\big\rangle \mathrm{d}s.$$

Similar to the proof of the inequality (2.10), we can prove that

$$\mathbb{E}_t \sup_{s\in[t,T]}|Y_3(s;t)|^2 \le \mathcal{C}\varepsilon|v|^2_{\mathbb{R}^k}, \quad \text{a.s.} \tag{2.25}$$

Next, consider the following equation:

$$\begin{cases}\mathrm{d}Y_4(s;t) = -\big(A_\Theta^\top Y_4(s;t) + C_\Theta^\top Z_4(s;t) + \widehat{A}_\Theta^\top Y_3(s;t)\big)\mathrm{d}s + Z_4(s;t)\mathrm{d}W(s), \quad s\in[t,T],\\ Y_4(T;t) = H^\top Y_3(T;t).\end{cases} \tag{2.26}$$

Similar to the proof of the inequality (2.10), we can show that

$$\mathbb{E}_t\left(\sup_{s\in[t,T]}|Y_4(s;t)|^2 + \int_t^T|Z_4(s;t)|^2\mathrm{d}s\right) \le \mathcal{C}\varepsilon|v|^2_{\mathbb{R}^k}, \quad \text{a.s.} \tag{2.27}$$

Applying Ito's formula to $\langle Y_4, X_0^\varepsilon\rangle$, we obtain

$$J_3(t,x) = \frac{1}{2}\mathbb{E}_t\int_t^{t+\varepsilon}\big\langle \widehat{B}^\top Y_3(s;t) + B^\top Y_4(s;t) + D^\top Z_4(s;t), v\big\rangle\mathrm{d}s.$$

From the inequalities (2.25) and (2.27), we find that

$$J_3(t,x) = \frac{1}{2}\mathbb{E}_t\int_t^{t+\varepsilon}\big\langle D^\top Z_4(s;t), v\big\rangle\mathrm{d}s + \mathcal{C}|v|^2_{\mathbb{R}^k}o(\varepsilon). \tag{2.28}$$

In the rest of this step, we handle the term $\frac{1}{2}\mathbb{E}_t\int_t^{t+\varepsilon}\big\langle D^\top Z_4(s;t), v\big\rangle\mathrm{d}s$.

We first introduce the following equations:

$$\begin{cases}\dfrac{\mathrm{d}\mathcal{P}_4}{\mathrm{d}s} + \mathcal{P}_4\widehat{C}^\top + A_\Theta^\top\mathcal{P}_4 + C_\Theta^\top\mathcal{P}_4\widehat{D}^\top + \widehat{A}_\Theta^\top = 0, \quad s\in[t,T],\\ \mathcal{P}_4(T) = H^\top,\end{cases} \tag{2.29}$$

and

$$\begin{cases}\mathrm{d}\mathcal{P}_5(s;t) = -\big(A_\Theta^\top\mathcal{P}_5(s;t) + C_\Theta^\top\mathcal{L}_5(s;t) + \mathcal{P}_4F_3^\varepsilon(s,t) + C_\Theta^\top\mathcal{P}_4F_4^\varepsilon(s,t)\big)\mathrm{d}s \\ \qquad\qquad\quad +\mathcal{L}_5(s;t)\mathrm{d}W(s), \qquad\qquad\qquad\qquad\qquad\qquad s\in[t,T],\\ \mathcal{P}_5(T;t) = 0.\end{cases} \tag{2.30}$$

Combining (2.24), (2.29) and (2.30), we get that

$$
\begin{aligned}
&\mathrm{d}\big(\mathcal{P}_4 Y_3+\mathcal{P}_5\big)\\
&=-\Big(\mathcal{P}_4\widehat{C}^\top+A_\Theta^\top\mathcal{P}_4+C_\Theta^\top\mathcal{P}_4\widehat{D}^\top+\widehat{A}_\Theta^\top\Big)Y_3\mathrm{d}s+\mathcal{P}_4\Big(\widehat{C}^\top Y_3+F_3^\varepsilon\Big)\mathrm{d}s+\mathcal{P}_4\Big(\widehat{D}^\top Y_3+F_4^\varepsilon\Big)\mathrm{d}W(s)\\
&\quad-\big(A_\Theta^\top\mathcal{P}_5+C_\Theta^\top\mathcal{L}_5+\mathcal{P}_4F_3^\varepsilon+C_\Theta^\top\mathcal{P}_4F_4^\varepsilon\big)\mathrm{d}s+\mathcal{L}_5\mathrm{d}W(s)\\
&=-\Big\{A_\Theta^\top\big(\mathcal{P}_4Y_3+\mathcal{P}_5\big)+C_\Theta^\top\Big[\mathcal{P}_4\Big(\widehat{D}^\top Y_3+F_4^\varepsilon\Big)+\mathcal{L}_5\Big]+\widehat{A}_\Theta^\top Y_3\Big\}\mathrm{d}s+\Big[\mathcal{P}_4\Big(\widehat{D}^\top Y_3+F_4^\varepsilon\Big)+\mathcal{L}_5\Big]\mathrm{d}W(s).
\end{aligned}
$$

This, together with $\mathcal{P}_4(T)=H^\top$ and $\mathcal{P}_5(T;t)=0$, implies that $\big(\mathcal{P}_4Y_3+\mathcal{P}_5,\mathcal{P}_4(\widehat{D}^\top Y_3+F_4^\varepsilon)+\mathcal{L}_5\big)$ is a solution to (2.26). From the uniqueness of the solution to (2.26), we find that

$$
\begin{cases}
Y_4(\cdot;t)=\mathcal{P}_4(\cdot)Y_3(\cdot;t)+\mathcal{P}_5(\cdot;t),\\
Z_4(\cdot;t)=\mathcal{P}_4(\cdot)\widehat{D}(\cdot)^\top Y_3(\cdot;t)+\mathcal{P}_4(\cdot)N(\cdot,t)Z_0^\varepsilon(\cdot;t)+\mathcal{L}_5(\cdot;t).
\end{cases}
\tag{2.31}
$$

Combining (2.31) and (2.25), we get that

$$
\begin{aligned}
&\frac{1}{2}\mathbb{E}_t\int_t^{t+\varepsilon}\big\langle D^\top Z_4(s;t),v\big\rangle\mathrm{d}s\\
&=\frac{1}{2}\mathbb{E}_t\int_t^{t+\varepsilon}\big\langle D^\top\big(\mathcal{P}_4N(s,t)Z_0^\varepsilon(s;t)+\mathcal{L}_5(s;t)\big),v\big\rangle\mathrm{d}s+\mathcal{C}|v|^2_{\mathbb{R}^k}o(\varepsilon).
\end{aligned}
\tag{2.32}
$$

Next, we introduce the following equations:

$$
\begin{cases}
\dfrac{\mathrm{d}\mathcal{P}_6}{\mathrm{d}s}+\mathcal{P}_6A_\Theta+\widehat{A}_\Theta+\widehat{C}\mathcal{P}_6+\widehat{D}\mathcal{P}_6C_\Theta=0, & s\in[t,T],\\
\dfrac{\mathrm{d}\mathcal{P}_7(s;t)}{\mathrm{d}s}+\widehat{C}\mathcal{P}_7(s;t)+\chi_{[t,t+\varepsilon]}\mathcal{P}_6B+\chi_{[t,t+\varepsilon]}\widehat{B}+\widehat{D}\mathcal{P}_6D\chi_{[t,t+\varepsilon]}=0, & s\in[t,T],\\
\mathcal{P}_6(T)=H,\quad \mathcal{P}_7(T;t)=0.
\end{cases}
\tag{2.33}
$$

Similar to the proof of (2.31), we can show that

$$
\begin{cases}
Y_0^\varepsilon(\cdot;t)=\mathcal{P}_6(\cdot)X_0^\varepsilon(\cdot;t)+\chi_{[t,t+\varepsilon]}\mathcal{P}_7(\cdot;t)v,\\
Z_0^\varepsilon(\cdot;t)=\mathcal{P}_6(\cdot)\big(C_\Theta(\cdot)X_0^\varepsilon(\cdot;t)+\chi_{[t,t+\varepsilon]}D(\cdot)v\big).
\end{cases}
\tag{2.34}
$$

This, together with the inequality (2.10), yields

$$
\begin{aligned}
&\frac{1}{2}\mathbb{E}_t\int_t^{t+\varepsilon}\Big\langle D^\top\mathcal{P}_4N(s,t)Z_0^\varepsilon(s;t),v\Big\rangle\mathrm{d}s\\
&=\frac{1}{2}\mathbb{E}_t\int_t^{t+\varepsilon}\Big\langle D^\top\mathcal{P}_4N(s,t)\mathcal{P}_6\big(C_\Theta X_0^\varepsilon(s;t)+Dv\chi_{[t,t+\varepsilon]}\big),v\Big\rangle\mathrm{d}s\\
&=\frac{1}{2}\mathbb{E}_t\int_t^{t+\varepsilon}\Big\langle D^\top\mathcal{P}_4N(s,t)\mathcal{P}_6Dv,v\Big\rangle\mathrm{d}s+\mathcal{C}|v|^2_{\mathbb{R}^k}o(\varepsilon).
\end{aligned}
\tag{2.35}
$$

Lastly, by (2.34), we can rewrite the equation (2.30) as

$$\begin{cases} \mathrm{d}\mathcal{P}_5(s;t) = -\big[A_\Theta^\top \mathcal{P}_5(s;t) + C_\Theta^\top \mathcal{L}_5(s;t) + \mathcal{P}_4 M(s,t)\big(\mathcal{P}_6 X_0^\varepsilon(s;t) + \chi_{[t,t+\varepsilon]}\mathcal{P}_7(s;t)v\big) \\ \qquad + C_\Theta^\top \mathcal{P}_4 N(s,t)\mathcal{P}_6\big(C_\Theta X_0^\varepsilon(s;t) + \chi_{[t,t+\varepsilon]}Dv\big)\big]\mathrm{d}s + \mathcal{L}_5(s;t)\mathrm{d}W(s), \qquad s \in [t,T], \\ \mathcal{P}_5(T;t) = 0. \end{cases} \tag{2.36}$$

Let us further introduce the following equations:

$$\begin{cases} \dfrac{\mathrm{d}\mathcal{P}_8(s;t)}{\mathrm{d}s} + \mathcal{P}_8(s;t)A_\Theta + A_\Theta^\top \mathcal{P}_8(s;t) + C_\Theta^\top \mathcal{P}_8(s;t)C_\Theta \\ \quad + \mathcal{P}_4 M(s,t)\mathcal{P}_6 + C_\Theta^\top \mathcal{P}_4 N(s,t)\mathcal{P}_6 C_\Theta = 0, \qquad s \in [t,T], \\ \mathcal{P}_8(T;t) = 0, \end{cases} \tag{2.37}$$

and

$$\begin{cases} \dfrac{\mathrm{d}\mathcal{P}_9(s;t)}{\mathrm{d}s} + A_\Theta^\top \mathcal{P}_9(s;t) + \chi_{[t,t+\varepsilon]}C_\Theta^\top \mathcal{P}_8(s;t)D + \chi_{[t,t+\varepsilon]}\mathcal{P}_8(s;t)B \\ \quad + \chi_{[t,t+\varepsilon]}\mathcal{P}_4 M(s,t)\mathcal{P}_7(s;t) + \chi_{[t,t+\varepsilon]}C_\Theta^\top \mathcal{P}_4 N(s,t)\mathcal{P}_6 D = 0, \qquad s \in [t,T], \\ \mathcal{P}_9(T;t) = 0. \end{cases} \tag{2.38}$$

Using Ito's formula, we have

$$\begin{aligned} &\mathrm{d}\big(\mathcal{P}_8 X_0^\varepsilon + \mathcal{P}_9 v\big) \\ &= -\big(\mathcal{P}_8 A_\Theta + A_\Theta^\top \mathcal{P}_8 + C_\Theta^\top \mathcal{P}_8 C_\Theta + \mathcal{P}_4 M \mathcal{P}_6 + C_\Theta^\top \mathcal{P}_4 N \mathcal{P}_6 C_\Theta\big)X_0^\varepsilon \mathrm{d}s \\ &\quad + \mathcal{P}_8\big(A_\Theta X_0^\varepsilon + \chi_{[t,t+\varepsilon]}Bv\big)\mathrm{d}s + \mathcal{P}_8\big(C_\Theta X_0^\varepsilon + \chi_{[t,t+\varepsilon]}Dv\big)\mathrm{d}W(s) \\ &\quad - \big(A_\Theta^\top \mathcal{P}_9 + \chi_{[t,t+\varepsilon]}C_\Theta^\top \mathcal{P}_8 D + \chi_{[t,t+\varepsilon]}\mathcal{P}_8 B + \chi_{[t,t+\varepsilon]}\mathcal{P}_4 M \mathcal{P}_7 \\ &\qquad + \chi_{[t,t+\varepsilon]}C_\Theta^\top \mathcal{P}_4 N \mathcal{P}_6 D\big)v\mathrm{d}s \\ &= -\big[A_\Theta^\top\big(\mathcal{P}_8 X_0^\varepsilon + \mathcal{P}_9 v\big) + C_\Theta^\top \mathcal{P}_8\big(C_\Theta X_0^\varepsilon + \chi_{[t,t+\varepsilon]}Dv\big) \\ &\qquad + \mathcal{P}_4 M\big(\mathcal{P}_6 X_0^\varepsilon + \chi_{[t,t+\varepsilon]}\mathcal{P}_7 v\big) + C_\Theta^\top \mathcal{P}_4 N \mathcal{P}_6\big(C_\Theta X_0^\varepsilon + \chi_{[t,t+\varepsilon]}Dv\big)\big]\mathrm{d}s \\ &\quad + \mathcal{P}_8\big(C_\Theta X_0^\varepsilon + \chi_{[t,t+\varepsilon]}Dv\big)\mathrm{d}W(s). \end{aligned}$$

This, together with $\mathcal{P}_8(T;t) = 0$ and $\mathcal{P}_9(T;t) = 0$, implies that $\big(\mathcal{P}_8 X_0^\varepsilon + \mathcal{P}_9 v, \mathcal{P}_8(C_\Theta X_0^\varepsilon + \chi_{[t,t+\varepsilon]}Dv)\big)$ is a solution to (2.36). Then, the uniqueness of the solution to (2.36) implies that

$$\begin{cases} \mathcal{P}_5(\cdot,t) = \mathcal{P}_8(\cdot;t)X_0^\varepsilon(\cdot;t) + \mathcal{P}_9(\cdot;t)v, \\ \mathcal{L}_5(\cdot;t) = \mathcal{P}_8(\cdot;t)\big(C_\Theta(s)X_0^\varepsilon(\cdot;t) + \chi_{[t,t+\varepsilon]}D(\cdot)v\big). \end{cases}$$

This, together with the inequalities (2.10), implies that

$$\begin{aligned}&\frac{1}{2}\mathbb{E}_t\int_t^{t+\varepsilon}\big\langle D^\top\mathcal{L}_5(s;t),v\big\rangle\mathrm{d}s\\&=\frac{1}{2}\mathbb{E}_t\int_t^{t+\varepsilon}\Big\langle D^\top\mathcal{P}_8(s;t)\big(C_\Theta X_0^\varepsilon+Dv\chi_{[t,t+\varepsilon]}\big),v\Big\rangle\mathrm{d}s\\&=\frac{1}{2}\mathbb{E}_t\int_t^{t+\varepsilon}\Big\langle D^\top\mathcal{P}_8(s;t)Dv,v\Big\rangle\mathrm{d}s+\mathcal{C}|v|^2_{\mathbb{R}^k}o(\varepsilon).\end{aligned}\tag{2.39}$$

Consequently, from equations (2.28), (2.32), (2.35) and (2.39), we obtain that

$$\begin{aligned}J_3(t,x)&=\frac{1}{2}\mathbb{E}_t\int_t^{t+\varepsilon}\big\langle D^\top Z_4(s;t),v\big\rangle\mathrm{d}s+\mathcal{C}|v|^2_{\mathbb{R}^k}o(\varepsilon)\\&=\frac{1}{2}\mathbb{E}_t\int_t^{t+\varepsilon}\Big\langle\big(D^\top\mathcal{P}_4N(s,t)\mathcal{P}_6D+D^\top\mathcal{P}_8(s;t)D\big)v,v\Big\rangle\mathrm{d}s+\mathcal{C}|v|^2_{\mathbb{R}^k}o(\varepsilon).\end{aligned}\tag{2.40}$$

Step 6. In this step, we give an estimate of $J_4(t,x)$.

We first introduce the following adjoint equation:

$$\begin{cases}\mathrm{d}Y_5(s;t)=\big(\widehat{C}^\top Y_5(s;t)+F_3(s,t)\big)\mathrm{d}s+\big(\widehat{D}^\top Y_5(s;t)+F_4(s,t)\big)\mathrm{d}W(s),\quad s\in[t,T],\\Y_5(t;t)=G_2(t)Y(t).\end{cases}\tag{2.41}$$

Applying Ito's formula for $\langle Y_5,Y_0^\varepsilon\rangle$, we have

$$\begin{aligned}J_4(t,x)=\mathbb{E}_t\big\langle Y_5(T;t),HX_0^\varepsilon(T;t)\big\rangle+\mathbb{E}_t\int_t^{t+\varepsilon}\big\langle\widehat{B}^\top Y_5(s;t),v\big\rangle\mathrm{d}s\\+\mathbb{E}_t\int_t^T\big\langle\widehat{A}_\Theta^\top Y_5(s;t),X_0^\varepsilon(s;t)\big\rangle\mathrm{d}s.\end{aligned}\tag{2.42}$$

Similar to the treatment of $J_3(t,x)$ in Step 5, we introduce another adjoint equation:

$$\begin{cases}\mathrm{d}Y_6(s;t)=-\big(A_\Theta^\top Y_6(s;t)+C_\Theta^\top Z_6(s;t)+\widehat{A}_\Theta^\top Y_5(s;t)\big)\mathrm{d}s+Z_6(s;t)\mathrm{d}W(s),\quad s\in[t,T],\\Y_6(T;t)=H^\top Y_5(T;t).\end{cases}\tag{2.43}$$

Applying Ito's formula for $\langle Y_6,X_0^\varepsilon\rangle$, we get

$$\begin{aligned}J_4(t,x)&=\mathbb{E}_t\int_t^{t+\varepsilon}\big\langle\widehat{B}^\top Y_5(s;t),v\big\rangle\mathrm{d}s+\mathbb{E}_t\int_t^{t+\varepsilon}\big\langle B^\top Y_6(s;t)+D^\top Z_6(s;t),v\big\rangle\mathrm{d}s\\&=\mathbb{E}_t\int_t^{t+\varepsilon}\big\langle\widehat{B}^\top Y_5(s;t)+B^\top Y_6(s;t)+D^\top Z_6(s;t),v\big\rangle\mathrm{d}s.\end{aligned}\tag{2.44}$$

Now, our goal is to represent $Y_5(\cdot;t)$ and $\big(Y_6(\cdot;t), Z_6(\cdot;t)\big)$ by $X(\cdot)$. To this end, consider the following equations:

$$\begin{cases}\dfrac{\mathrm{d}\mathcal{P}_{10}}{\mathrm{d}s}+\mathcal{P}_{10}\widehat{C}^\top+A_\Theta^\top\mathcal{P}_{10}+C_\Theta^\top\mathcal{P}_{10}\widehat{D}^\top+\widehat{A}_\Theta^\top=0, & s\in[t,T],\\ \mathrm{d}\mathcal{P}_{11}(s;t)=-\big(A_\Theta^\top\mathcal{P}_{11}(s;t)+C_\Theta^\top\mathcal{L}_{11}(s;t)+\mathcal{P}_{10}F_3(s,t) & \\ \qquad\qquad +C_\Theta^\top\mathcal{P}_{10}F_4(s,t)\big)\mathrm{d}s+\mathcal{L}_{11}(s;t)\mathrm{d}W(s), & s\in[t,T],\\ \mathcal{P}_{10}(T)=H^\top,\quad \mathcal{P}_{11}(T;t)=0. \end{cases} \tag{2.45}$$

Similar to the proof of (2.31), we can show that

$$\begin{cases}Y_6(\cdot;t)=\mathcal{P}_{10}(\cdot)Y_5(\cdot;t)+\mathcal{P}_{11}(\cdot;t),\\ Z_6(\cdot;t)=\mathcal{P}_{10}(\cdot)\big(\widehat{D}(\cdot)^\top Y_5(\cdot;t)+F_4(\cdot,t)\big)+\mathcal{L}_{11}(\cdot;t).\end{cases} \tag{2.46}$$

From (2.44) and (2.46), we obtain

$$\begin{aligned}J_4(t,x)&=\mathbb{E}_t\int_t^{t+\varepsilon}\Big\langle\widehat{B}^\top Y_5(s;t)+B^\top Y_6(s;t)+D^\top Z_6(s;t),v\Big\rangle\mathrm{d}s\\ &=\mathbb{E}_t\int_t^{t+\varepsilon}\Big\langle\big(\widehat{B}^\top+B^\top\mathcal{P}_{10}+D^\top\mathcal{P}_{10}\widehat{D}^\top\big)Y_5(s;t)+B^\top\mathcal{P}_{11}(s;t)\\ &\qquad\qquad +D^\top\mathcal{P}_{10}F_4(s,t)+D^\top\mathcal{L}_{11}(s;t),v\Big\rangle\mathrm{d}s.\end{aligned} \tag{2.47}$$

Next, we introduce the following equation:

$$\begin{cases}\dfrac{\mathrm{d}\mathcal{P}_{12}}{\mathrm{d}s}+\mathcal{P}_{12}A_\Theta+\widehat{A}_\Theta+\widehat{C}\mathcal{P}_{12}+\widehat{D}\mathcal{P}_{12}C_\Theta=0, \quad s\in[0,T],\\ \mathcal{P}_{12}(T)=H.\end{cases} \tag{2.48}$$

Similar to the proof of (2.31), we can obtain that

$$\begin{cases}Y(\cdot)=\mathcal{P}_{12}(\cdot)X(\cdot),\\ Z(\cdot)=\mathcal{P}_{12}(\cdot)C_\Theta(\cdot)X(\cdot).\end{cases} \tag{2.49}$$

By (2.49) and (2.41), we get that

$$\begin{cases}\mathrm{d}Y_5(s;t)=\big(\widehat{C}^\top Y_5(s;t)+M(s,t)\mathcal{P}_{12}X\big)\mathrm{d}s\\ \qquad\qquad +\big(\widehat{D}^\top Y_5(s;t)+N(s,t)\mathcal{P}_{12}C_\Theta X\big)\mathrm{d}W(s), \quad s\in[t,T],\\ Y_5(t;t)=G_2(t)\mathcal{P}_{12}(t)X(t).\end{cases} \tag{2.50}$$

Now we study the relationship between Y_5 and X. To this end, we introduce the following equation:

$$
\begin{cases}
\mathrm{d}\mathcal{P}_{13}(s;t) = \big[\big(\mathcal{P}_{13}(s;t)C_\Theta - \widehat{D}^\top\mathcal{P}_{13}(s;t) - N(s,t)\mathcal{P}_{12}C_\Theta\big)C_\Theta \\
\qquad\qquad\qquad -\mathcal{P}_{13}(s;t)A_\Theta + \widehat{C}^\top\mathcal{P}_{13}(s;t) + M(s,t)\mathcal{P}_{12}\big]\mathrm{d}s \\
\qquad\qquad\qquad +\big(-\mathcal{P}_{13}(s;t)C_\Theta + \widehat{D}^\top\mathcal{P}_{13}(s;t) + N(s,t)\mathcal{P}_{12}C_\Theta\big)\mathrm{d}W(s), \qquad s\in[t,T], \\
\mathcal{P}_{13}(t;t) = G_2(t)\mathcal{P}_{12}(t).
\end{cases}
\tag{2.51}
$$

Similar to the proof of (2.31), we get that

$$
Y_5(\cdot;t) = \mathcal{P}_{13}(\cdot;t)X(\cdot). \tag{2.52}
$$

Combining (2.47), (2.49) and (2.52), we obtain that

$$
\begin{aligned}
J_4(t,x) &= \mathbb{E}_t\int_t^{t+\varepsilon}\Big\langle\big(\widehat{B}^\top + B^\top\mathcal{P}_{10} + D^\top\mathcal{P}_{10}\widehat{D}^\top\big)Y_5(s;t) \\
&\qquad + B^\top\mathcal{P}_{11}(s;t) + D^\top\mathcal{P}_{10}F_4(s,t) + D^\top\mathcal{L}_{11}(s;t), v\Big\rangle\mathrm{d}s \\
&= \mathbb{E}_t\int_t^{t+\varepsilon}\Big\langle\big[\big(\widehat{B}^\top + B^\top\mathcal{P}_{10} + D^\top\mathcal{P}_{10}\widehat{D}^\top\big)\mathcal{P}_{13}(s;t) + D^\top\mathcal{P}_{10}N(s,t)\mathcal{P}_{12}C_\Theta\big]X \\
&\qquad + B^\top\mathcal{P}_{11}(s;t) + D^\top\mathcal{L}_{11}(s;t), v\Big\rangle\mathrm{d}s.
\end{aligned}
\tag{2.53}
$$

Next, we derive the relationship between $X(\cdot)$ and $\big(\mathcal{P}_{11}(\cdot;t),\mathcal{L}_{11}(\cdot;t)\big)$. First, by (2.49), we can rewrite the second equation of (2.45) as

$$
\begin{cases}
\mathrm{d}\mathcal{P}_{11}(s;t) = -\big(A_\Theta^\top\mathcal{P}_{11}(s;t) + C_\Theta^\top\mathcal{L}_{11}(s;t) + \mathcal{P}_{10}M(s,t)\mathcal{P}_{12}X \\
\qquad\qquad\qquad + C_\Theta^\top\mathcal{P}_{10}N(s,t)\mathcal{P}_{12}C_\Theta X\big)\mathrm{d}s + \mathcal{L}_{11}(s;t)\mathrm{d}W(s), \\
\mathcal{P}_{11}(T;t) = 0.
\end{cases}
\tag{2.54}
$$

Secondly, we introduce the following equation:

$$
\begin{cases}
\dfrac{\mathrm{d}\mathcal{P}_{14}(s;t)}{\mathrm{d}s} + \mathcal{P}_{14}(s;t)A_\Theta + A_\Theta^\top\mathcal{P}_{14}(s;t) + C_\Theta^\top\mathcal{P}_{14}(s;t)C_\Theta \\
\quad + \mathcal{P}_{10}M(s,t)\mathcal{P}_{12} + C_\Theta^\top\mathcal{P}_{10}N(s,t)\mathcal{P}_{12}C_\Theta = 0, \qquad s\in[t,T], \\
\mathcal{P}_{14}(T;t) = 0,
\end{cases}
\tag{2.55}
$$

Similar to the proof of (2.31), we can prove that

$$
\begin{cases}
\mathcal{P}_{11}(\cdot;t) = \mathcal{P}_{14}(\cdot;t)X(\cdot), \\
\mathcal{L}_{11}(\cdot;t) = \mathcal{P}_{14}(\cdot;t)C_\Theta(\cdot)X(\cdot).
\end{cases}
\tag{2.56}
$$

From (2.53) and (2.56), we obtain that

$$\begin{aligned} & J_4(t,x) \\ & = \mathbb{E}_t\int_t^{t+\varepsilon}\Big\langle\big[\big(\widehat{B}^\top + B^\top\mathcal{P}_{10}+ D^\top\mathcal{P}_{10}\widehat{D}^\top\big)\mathcal{P}_{13}(s;t) \\ & \qquad + D^\top\mathcal{P}_{10}N(s,t)\mathcal{P}_{12}C_\Theta\big]X + B^\top\mathcal{P}_{11}(s;t)+ D^\top\mathcal{L}_{11}(s;t), v\Big\rangle\mathrm{d}s \\ & = \mathbb{E}_t\int_t^{t+\varepsilon}\Big\langle\big[\big(\widehat{B}^\top + B^\top\mathcal{P}_{10}+ D^\top\mathcal{P}_{10}\widehat{D}^\top\big)\mathcal{P}_{13}(s;t) \\ & \qquad + D^\top\mathcal{P}_{10}N(s,t)\mathcal{P}_{12}C_\Theta + B^\top\mathcal{P}_{14}(s;t) + D^\top\mathcal{P}_{14}(s;t)C_\Theta\big]X, v\Big\rangle\mathrm{d}s \end{aligned} \tag{2.57}$$

Step 7. We combine the results in Step 3–Step 6 to study $\mathcal{J}(t,x;u^\varepsilon(\cdot)) - \mathcal{J}(t,x;u(\cdot))$.

From (2.7)–(2.9), (2.16), (2.23), (2.40) and (2.57), for any $(t,x)\in[0,T)\times L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^n)$, $v\in L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^k)$ and $\Theta(\cdot)\in L^2(0,T;\mathbb{R}^{k\times n})$, we have that

$$\begin{aligned} & \mathcal{J}(t,x;u^\varepsilon(\cdot)) - \mathcal{J}(t,x;u(\cdot)) \\ & = \Big\langle\Big(\int_t^{t+\varepsilon}\frac{1}{2}R(s,t)\mathrm{d}s + \int_t^{t+\varepsilon}\frac{1}{2}D^\top\mathcal{P}_2(s;t)D\mathrm{d}s + \int_t^{t+\varepsilon}\frac{1}{2}D^\top\mathcal{P}_8(s;t)D\mathrm{d}s \\ & \quad + \int_t^{t+\varepsilon}\frac{1}{2}D^\top\mathcal{P}_4N(s,t)\mathcal{P}_6D\mathrm{d}s\Big)v, v\Big\rangle + \Big\langle\mathbb{E}_t\int_t^{t+\varepsilon}\big[B^\top\mathcal{P}_1(s;t) \\ & \quad + D^\top\mathcal{P}_1(s;t)C_\Theta + R(s,t)\Theta + \big(\widehat{B}^\top + B^\top\mathcal{P}_{10} + D^\top\mathcal{P}_{10}\widehat{D}^\top\big)\mathcal{P}_{13}(s;t) \\ & \quad + D^\top\mathcal{P}_{10}N(s,t)\mathcal{P}_{12}C_\Theta + B^\top\mathcal{P}_{14}(s;t) + D^\top\mathcal{P}_{14}(s;t)C_\Theta\big]X\mathrm{d}s, v\Big\rangle + \mathcal{C}|v|^2_{\mathbb{R}^k}o(\varepsilon). \end{aligned} \tag{2.58}$$

Moreover, by comparing the equation (2.3) with the equations (2.14), (2.19), (2.29), (2.33), (2.37), (2.45), (2.48), (2.55), we see that for any $\Theta\in L^2(0,T;\mathbb{R}^{k\times n})$ and $0\le t\le s\le T$, it holds that

$$\begin{cases} \mathbf{P_1}(s;t) = \mathcal{P}_1(s;t) = \mathcal{P}_2(s;t), \\ \mathbf{P_2}(s)^\top = \mathcal{P}_4(s) = \mathcal{P}_{10}(s) = \mathcal{P}_6(s)^\top = \mathcal{P}_{12}(s)^\top, \qquad s\in[t,T]. \\ \mathbf{P_3}(s;t) = \mathcal{P}_8(s;t) = \mathcal{P}_{14}(s;t), \end{cases} \tag{2.59}$$

Step 8. In this step, we establish some regularity estimates.

To this end, we first make some preparation. For $X(\cdot)$ and $Y_5(\cdot;t) = \mathcal{P}_{13}(\cdot;t)X(\cdot)$ (see Eq. (2.52)), by the equations (2.4), (2.50), we can see that $\mathbb{E}_tX(\cdot;t)$ and $\mathbb{E}_tY_5(\cdot;t)$ satisfy the following equations:

$$\begin{cases} \mathrm{d}\mathbb{E}_tX(s;t) = A_\Theta\mathbb{E}_tX(s;t)\mathrm{d}s, \qquad s\in[t,T], \\ \mathbb{E}_tX(t;t) = x, \end{cases}$$

and

$$\begin{cases} \mathrm{d}\mathbb{E}_tY_5(s;t) = \big(\widehat{C}^\top\mathbb{E}_tY_5(s;t) + M(s,t)\mathcal{P}_{12}\mathbb{E}_tX(s;t)\big)\mathrm{d}s, \qquad s\in[t,T], \\ \mathbb{E}_tY_5(t;t) = G_2(t)\mathcal{P}_{12}(t)x. \end{cases}$$

Consider the following equations:

$$\begin{cases} \dfrac{\mathrm{d}\Psi(s;t)}{\mathrm{d}s} = A_\Theta \Psi(s;t), \qquad s\in[t,T],\\ \Psi(t;t)=I, \end{cases}$$

and

$$\begin{cases} \dfrac{\mathrm{d}\Gamma(s;t)}{\mathrm{d}s} = \widehat{C}^\top \Gamma(s;t) + M(s,t)\mathcal{P}_{12}\Psi(s;t), \qquad s\in[t,T],\\ \Gamma(t;t)=G_2(t)\mathcal{P}_{12}(t). \end{cases}$$

Then we have

$$\begin{cases} \mathbb{E}_t X(s;t) = \Psi(s;t)x,\\ \mathbb{E}_t Y_5(s;t) = \Gamma(s;t)x. \end{cases} \tag{2.60}$$

Rewrite the equation for $\Gamma(\cdot;t)$ as

$$\begin{cases} \dfrac{\mathrm{d}\Gamma(s;t)}{\mathrm{d}s} = \widehat{C}^\top \Gamma(s;t) + M(s,t)\mathcal{P}_{12}\Psi(s;t), \quad s\in[\tau,T],\\ \Gamma(\tau;t)=\Gamma(\tau;t), \end{cases} \tag{2.61}$$

where $0\le t\le \tau\le s\le T$. Then

$$\begin{aligned} &\big|\Gamma(\tau;t)-\Gamma(\tau;\tau)\big|_\infty\\ &\le \big|G_2(t)\mathcal{P}_{12}(t)-G_2(\tau)\mathcal{P}_{12}(\tau)\big|_\infty + \int_t^\tau \big|\widehat{C}^\top\Gamma(s;t)+M(s,t)\mathcal{P}_{12}\Psi(s;t)\big|_\infty \mathrm{d}s\\ &\le \mathcal{C}\big(|t-\tau| + |\mathcal{P}_{12}(t)-\mathcal{P}_{12}(\tau)|_\infty\big), \end{aligned}$$

which, together with Assumptions 1.5, yields

$$\begin{aligned} &|\Gamma(s;t)-\Gamma(s;\tau)|_\infty\\ &\le |\Gamma(\tau;t)-\Gamma(\tau;\tau)|_\infty + \int_\tau^s \big|\widehat{C}^\top\big(\Gamma(\eta;t)-\Gamma(\eta;\tau)\big) + M(\eta,t)\mathcal{P}_{12}\Psi(\eta;t)\\ &\qquad\qquad -M(\eta,\tau)\mathcal{P}_{12}\Psi(\eta;\tau)\big|_\infty \mathrm{d}\eta\\ &\le \mathcal{C}\big(|t-\tau| + |\mathcal{P}_{12}(t)-\mathcal{P}_{12}(\tau)|_\infty\big) + \int_\tau^s \big(\mathcal{C}|\Gamma(\eta;t)-\Gamma(\eta;\tau)|_\infty + \mathcal{C}|t-\tau|\big)\mathrm{d}\eta. \end{aligned}$$

This, together with Gronwall's inequality, implies

$$|\Gamma(s;t)-\Gamma(s;\tau)|_\infty \le \mathcal{C}\big(|t-\tau| + |\mathcal{P}_{12}(t)-\mathcal{P}_{12}(\tau)|_\infty\big).$$

By a similar (but simpler) argument, we can show that for any $0 \le t \le \tau \le s \le T$, it holds that

$$\begin{cases}|\Psi(s;t)-\Psi(s;\tau)|_\infty \le \mathcal{C}|t-\tau|,\\ |\Gamma(s;t)-\Gamma(s;\tau)|_\infty \le \mathcal{C}\big(|t-\tau|+|\mathcal{P}_{12}(t)-\mathcal{P}_{12}(\tau)|_\infty\big),\\ |\mathcal{P}_1(s;t)-\mathcal{P}_1(s;\tau)|_\infty \le \mathcal{C}|t-\tau|,\\ |\mathcal{P}_8(s;t)-\mathcal{P}_8(s;\tau)|_\infty \le \mathcal{C}|t-\tau|.\end{cases} \tag{2.62}$$

Step 9. In this step, we prove the following result:

$$\begin{cases}\displaystyle\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\frac{1}{2}R(s,t)\mathrm{d}s=\frac{1}{2}R(t,t),\quad t\in[0,T],\\ \displaystyle\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}D^\top\mathcal{P}_2(s;t)D\mathrm{d}s=D(t)^\top\mathcal{P}_2(t;t)D(t),\quad \text{a.e. } t\in[0,T],\\ \displaystyle\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}D^\top\mathcal{P}_8(s;t)D\mathrm{d}s=D(t)^\top\mathcal{P}_8(t;t)D(t),\quad \text{a.e. } t\in[0,T],\\ \displaystyle\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}D^\top\mathcal{P}_4N(s,t)\mathcal{P}_6D\mathrm{d}s=D(t)^\top\mathcal{P}_4(t)N(t,t)\mathcal{P}_6(t)D(t),\ \text{a.e. } t\in[0,T],\end{cases} \tag{2.63}$$

and

$$\begin{cases}\displaystyle\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}B^\top\mathcal{P}_1(s;t)\Psi(s;t)\mathrm{d}s=B(t)^\top\mathcal{P}_1(t;t),\\ \displaystyle\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}D^\top\mathcal{P}_1(s;t)C_\Theta\Psi(s;t)\mathrm{d}s=D(t)^\top\mathcal{P}_1(t;t)C_\Theta(t),\\ \displaystyle\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\big[R(s,t)\Theta+B^\top\mathcal{P}_{14}(s;t)\big]\Psi(s;t)\mathrm{d}s=R(t,t)\Theta(t)+B(t)^\top\mathcal{P}_{14}(t;t),\\ \displaystyle\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}D^\top\mathcal{P}_{10}N(s,t)\mathcal{P}_{12}C_\Theta\Psi(s;t)\mathrm{d}s=D(t)^\top\mathcal{P}_{10}(t)N(t,t)\mathcal{P}_{12}(t)C_\Theta(t),\\ \displaystyle\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}D^\top\mathcal{P}_{14}(s;t)C_\Theta\Psi(s;t)\mathrm{d}s=D(t)^\top\mathcal{P}_{14}(t;t)C_\Theta(t),\\ \displaystyle\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\big(\widehat{B}^\top+B^\top\mathcal{P}_{10}+D^\top\mathcal{P}_{10}\widehat{D}^\top\big)\Gamma(s;t)\mathrm{d}s\\ =\big(\widehat{B}(t)^\top+B(t)^\top\mathcal{P}_{10}(t)+D(t)^\top\mathcal{P}_{10}(t)\widehat{D}(t)^\top\big)G_2(t)\mathcal{P}_{12}(t),\end{cases}\quad \text{a.e. } t\in[0,T] \tag{2.64}$$

Since

$$\begin{aligned}&\Big|\frac{1}{\varepsilon}\int_t^{t+\varepsilon}D^\top\mathcal{P}_2(s;t)D\mathrm{d}s-D(t)^\top\mathcal{P}_2(t;t)D(t)\Big|_\infty\\ &\le\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\big|D^\top\big(\mathcal{P}_2(s;t)-\mathcal{P}_2(t;t)\big)D\big|_\infty\mathrm{d}s\\ &\quad+\Big|\frac{1}{\varepsilon}\int_t^{t+\varepsilon}D^\top\mathcal{P}_2(t;t)D\mathrm{d}s-D(t)^\top\mathcal{P}_2(t;t)D(t)\Big|_\infty\end{aligned}$$

$$
\begin{aligned}
&\leq \mathcal{C}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\big|\mathcal{P}_2(s;t)-\mathcal{P}_2(t;t)\big|_\infty \mathrm{d}s\\
&\quad+\Big|\frac{1}{\varepsilon}\int_t^{t+\varepsilon}D^\top\mathcal{P}_2(t;t)D\mathrm{d}s-D(t)^\top\mathcal{P}_2(t;t)D(t)\Big|_\infty\\
&\leq \mathcal{C}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\big|\mathcal{P}_2(s;t)-\mathcal{P}_2(t;t)\big|_\infty \mathrm{d}s+\mathcal{C}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}|D^\top-D(t)^\top|_\infty\mathrm{d}s,
\end{aligned}
$$

by Lebesgue's differentiation theorem, we have

$$
\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}D^\top\mathcal{P}_2(s;t)D\mathrm{d}s=D(t)^\top\mathcal{P}_2(t;t)D(t),\quad \text{a.e. } t\in[0,T].
$$

Other formulas in (2.63) can be proved similarly.

Except for the second equality in (2.64), the other ones can be proved similarly. Since

$$
\begin{aligned}
&\Big|\frac{1}{\varepsilon}\int_t^{t+\varepsilon}D^\top\mathcal{P}_1(s;t)C_\Theta\Psi(s;t)\mathrm{d}s-D(t)^\top\mathcal{P}_1(t;t)C_\Theta(t)\Big|_\infty\\
&\leq \frac{1}{\varepsilon}\int_t^{t+\varepsilon}\big|D^\top\big(\mathcal{P}_1(s;t)-\mathcal{P}_1(t;t)\big)C_\Theta\Psi(s;t)\big|_\infty\mathrm{d}s\\
&\quad+\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\big|D^\top\mathcal{P}_1(t;t)C_\Theta\big(\Psi(s;t)-\Psi(t;t)\big)\big|_\infty\mathrm{d}s\\
&\quad+\Big|\frac{1}{\varepsilon}\int_t^{t+\varepsilon}D^\top\mathcal{P}_1(t;t)C_\Theta\Psi(t;t)\mathrm{d}s-D(t)^\top\mathcal{P}_1(t;t)C_\Theta(t)\Big|_\infty\\
&\leq \mathcal{C}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\big|\mathcal{P}_1(s;t)-\mathcal{P}_1(t;t)\big|_\infty(1+|\Theta|_\infty)\mathrm{d}s\\
&\quad+\mathcal{C}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\big|\Psi(s;t)-\Psi(t;t)\big|_\infty(1+|\Theta|_\infty)\mathrm{d}s\\
&\quad+\Big|\frac{1}{\varepsilon}\int_t^{t+\varepsilon}D^\top\mathcal{P}_1(t;t)C_\Theta\Psi(t;t)\mathrm{d}s-D(t)^\top\mathcal{P}_1(t;t)C_\Theta(t)\Big|_\infty\\
&\leq \mathcal{C}\left[\frac{1}{\varepsilon}\int_t^{t+\varepsilon}(1+|\Theta|_\infty)^2\mathrm{d}s\right]^{1/2}\left[\left(\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\big|\mathcal{P}_1(s;t)-\mathcal{P}_1(t;t)\big|_\infty^2\mathrm{d}s\right)^{1/2}\right.\\
&\qquad\left.+\left(\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\big|\Psi(s;t)-\Psi(t;t)\big|_\infty^2\mathrm{d}s\right)^{1/2}\right]\\
&\quad+\Big|\frac{1}{\varepsilon}\int_t^{t+\varepsilon}D^\top\mathcal{P}_1(t;t)C_\Theta\Psi(t;t)\mathrm{d}s-D(t)^\top\mathcal{P}_1(t;t)C_\Theta(t)\Big|_\infty\\
&\leq \mathcal{C}\left[\frac{1}{\varepsilon}\int_t^{t+\varepsilon}(1+|\Theta|_\infty)^2\mathrm{d}s\right]^{1/2}\left[\left(\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\big|\mathcal{P}_1(s;t)-\mathcal{P}_1(t;t)\big|_\infty^2\mathrm{d}s\right)^{1/2}\right.\\
&\qquad\left.+\left(\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\big|\Psi(s;t)-\Psi(t;t)\big|_\infty^2\mathrm{d}s\right)^{1/2}+\left(\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\big|D^\top-D(t)^\top\big|_\infty^2\mathrm{d}s\right)^{1/2}\right]\\
&\quad+\mathcal{C}\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\big|C_\Theta-C_\Theta(t)\big|_\infty\mathrm{d}s,
\end{aligned}
$$

we get the second equality in (2.64) immediately.

Step 10. In this step, we prove the "only if" part of Theorem 2.1.

Choose any sequence $\{\varepsilon_j\}_{j=1}^{\infty} \subset (0,+\infty)$ satisfying $\lim_{j\to\infty} \varepsilon_j = 0$. By (2.58), for any $(t,x,v) \in [0,T) \times L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^n) \times L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^k)$, we have that

$$
\begin{aligned}
&\varliminf_{j\to\infty} \frac{\mathcal{J}(t,x;u^{\varepsilon_j}(\cdot)) - \mathcal{J}(t,x;u(\cdot))}{\varepsilon_j}\\
&= \varliminf_{j\to\infty} \Big\{ \Big\langle \frac{1}{\varepsilon_j} \Big(\int_t^{t+\varepsilon_j} \frac{1}{2} R(s,t) \mathrm{d}s + \int_t^{t+\varepsilon_j} \frac{1}{2} D^\top \mathcal{P}_2(s;t) D \mathrm{d}s \\
&\qquad + \int_t^{t+\varepsilon_j} \frac{1}{2} D^\top \mathcal{P}_8(s;t) D \mathrm{d}s + \int_t^{t+\varepsilon_j} \frac{1}{2} D^\top \mathcal{P}_4 N(s,t) \mathcal{P}_6 D \mathrm{d}s \Big) v, v \Big\rangle \\
&\quad + \Big\langle \frac{1}{\varepsilon_j} \Big\{ \mathbb{E}_t \int_t^{t+\varepsilon_j} \Big[B^\top \mathcal{P}_1(s;t) + D^\top \mathcal{P}_1(s;t) C_\Theta + R(s,t)\Theta \\
&\qquad + \big(\widehat{B}^\top + B^\top \mathcal{P}_{10} + D^\top \mathcal{P}_{10} \widehat{D}^\top\big) \mathcal{P}_{13}(s;t) + D^\top \mathcal{P}_{10} N(s,t) \mathcal{P}_{12} C_\Theta \\
&\qquad + B^\top \mathcal{P}_{14}(s;t) + D^\top \mathcal{P}_{14}(s;t) C_\Theta \Big] X \mathrm{d}s \Big\}, v \Big\rangle \Big\} \geq 0, \quad \mathbb{P}\text{-a.s.}
\end{aligned}
$$

This, together with (2.52), (2.59), (2.60), (2.63) and (2.64), yields (2.1) and (2.2).

Step 11. In this step, we prove the "if" part of Theorem 2.1.

Choose any sequence $\{\varepsilon_j\}_{j=1}^{\infty} \subset (0,+\infty)$ satisfying $\lim_{j\to\infty} \varepsilon_j = 0$. By (2.1), (2.2) and (2.59), for any $t \in [0,T]$, we have

$$
\frac{1}{\varepsilon_j} \int_t^{t+\varepsilon_j} \big(R(s,s) + D^\top \mathcal{P}_2(s;s) D + D^\top \mathcal{P}_8(s;s) D + D^\top \mathcal{P}_4 N(s,s) \mathcal{P}_6 D \big) \mathrm{d}s \geq 0, \tag{2.65}
$$

and

$$
\begin{aligned}
\frac{1}{\varepsilon_j} \int_t^{t+\varepsilon_j} & \Big[B^\top \mathcal{P}_1(s;s) + D^\top \mathcal{P}_1(s;s) C_\Theta + B^\top \mathcal{P}_{14}(s;s) \\
& + D^\top \mathcal{P}_{14}(s;s) C_\Theta + R(s,s)\Theta + \big(\widehat{B}^\top + B^\top \mathcal{P}_{10} \\
& + D^\top \mathcal{P}_{10} \widehat{D}^\top\big) G_2 \mathcal{P}_{12} + D^\top \mathcal{P}_{10} N(s,s) \mathcal{P}_{12} C_\Theta \Big] \mathrm{d}s = 0.
\end{aligned} \tag{2.66}
$$

On the other hand, by Assumption 1.5 and (2.62), it holds that

$$
\begin{aligned}
&\Big| \frac{1}{\varepsilon_j} \int_t^{t+\varepsilon_j} \big(R(s,s) + D^\top \mathcal{P}_2(s;s) D + D^\top \mathcal{P}_8(s;s) D \\
&+ D^\top \mathcal{P}_4 N(s,s) \mathcal{P}_6 D \big) \mathrm{d}s - \frac{1}{\varepsilon_j} \int_t^{t+\varepsilon_j} \big(R(s,t) + D^\top \mathcal{P}_2(s;t) D \\
&+ D^\top \mathcal{P}_8(s;t) D + D^\top \mathcal{P}_4 N(s,t) \mathcal{P}_6 D \big) \mathrm{d}s \Big|_\infty \leq \mathcal{C} \varepsilon_j,
\end{aligned}
$$

and

$$
\begin{aligned}
&\bigg|\frac{1}{\varepsilon_j}\int_t^{t+\varepsilon_j}\Big[B^\top\mathcal{P}_1(s;s)+D^\top\mathcal{P}_1(s;s)C_\Theta+B^\top\mathcal{P}_{14}(s;s)\\
&\qquad\qquad +D^\top\mathcal{P}_{14}(s;s)C_\Theta+R(s,s)\Theta+\big(\widehat{B}^\top+B^\top\mathcal{P}_{10}\\
&\qquad\qquad +D^\top\mathcal{P}_{10}\widehat{D}^\top\big)G_2\mathcal{P}_{12}+D^\top\mathcal{P}_{10}N(s,s)\mathcal{P}_{12}C_\Theta\Big]\mathrm{d}s\\
&-\frac{1}{\varepsilon_j}\int_t^{t+\varepsilon_j}\Big[B^\top\mathcal{P}_1(s;t)\Psi(s;t)+D^\top\mathcal{P}_1(s;t)C_\Theta\Psi(s;t)\\
&\qquad\qquad +B^\top\mathcal{P}_{14}(s;t)\Psi(s;t)+D^\top\mathcal{P}_{14}(s;t)C_\Theta\Psi(s;t)\\
&\qquad\qquad +R(s,t)\Theta\Psi(s;t)+\big(\widehat{B}^\top+B^\top\mathcal{P}_{10}+D^\top\mathcal{P}_{10}\widehat{D}^\top\big)\Gamma(s;t)\\
&\qquad\qquad +D^\top\mathcal{P}_{10}N(s,t)\mathcal{P}_{12}C_\Theta\Psi(s;t)\Big]\mathrm{d}s\bigg|_\infty\\
&\leq \mathcal{C}\varepsilon_j+\mathcal{C}\int_t^{t+\varepsilon_j}\big(1+|\Theta|_\infty\big)\mathrm{d}s+\mathcal{C}\frac{1}{\varepsilon_j}\int_t^{t+\varepsilon_j}\big|\Gamma(s;s)-\Gamma(s;t)\big|_\infty\mathrm{d}s\\
&\leq \mathcal{C}\varepsilon_j+\mathcal{C}\int_t^{t+\varepsilon_j}\big(1+|\Theta|_\infty\big)\mathrm{d}s+\frac{\mathcal{C}}{\varepsilon_j}\int_t^{t+\varepsilon_j}\big|\mathcal{P}_{12}(t)-\mathcal{P}_{12}\big|_\infty\mathrm{d}s\\
&\leq \mathcal{C}\varepsilon_j+\mathcal{C}\int_t^{t+\varepsilon_j}\big(1+|\Theta|_\infty\big)\mathrm{d}s+\frac{\mathcal{C}}{\varepsilon_j}\int_t^{t+\varepsilon_j}\int_t^s\big(1+|\Theta(\tau)|_\infty\big)\mathrm{d}\tau\mathrm{d}s.
\end{aligned}
$$

Combining (2.65), (2.66) and (2.58), we get

$$
\begin{aligned}
&\varliminf_{j\to\infty}\frac{\mathcal{J}(t,x;u^{\varepsilon_j}(\cdot))-\mathcal{J}(t,x;u(\cdot))}{\varepsilon_j}\\
&=\varliminf_{j\to\infty}\bigg\{\Big\langle\frac{1}{2\varepsilon_j}\Big[\int_t^{t+\varepsilon_j}\big(R(s,s)+D^\top\mathcal{P}_2(s;s)D+D^\top\mathcal{P}_8(s;s)D\\
&\quad +D^\top\mathcal{P}_4N(s,s)\mathcal{P}_6D\big)\mathrm{d}s\Big]v,v\Big\rangle+\Big\langle\frac{1}{\varepsilon_j}\int_t^{t+\varepsilon_j}\Big[B^\top\mathcal{P}_1(s;s)\\
&\quad +D^\top\mathcal{P}_1(s;s)C_\Theta+B^\top\mathcal{P}_{14}(s;s)+D^\top\mathcal{P}_{14}(s;s)C_\Theta+R(s,s)\Theta\\
&\quad +\big(\widehat{B}^\top+B^\top\mathcal{P}_{10}+D^\top\mathcal{P}_{10}\widehat{D}^\top\big)G_2\mathcal{P}_{12}\\
&\quad +D^\top\mathcal{P}_{10}N(s,s)\mathcal{P}_{12}C_\Theta\Big]\mathrm{d}s\ x,v\Big\rangle\bigg\}\\
&\geq\Big\langle\varliminf_{j\to\infty}\frac{1}{2\varepsilon_j}\Big[\int_t^{t+\varepsilon_j}\big(R(s,s)+D^\top\mathcal{P}_2(s;s)D+D^\top\mathcal{P}_8(s;s)D\\
&\quad +D^\top\mathcal{P}_4N(s,s)\mathcal{P}_6D\big)\mathrm{d}s\Big]v,v\Big\rangle+\Big\langle\varliminf_{j\to\infty}\frac{1}{\varepsilon_j}\int_t^{t+\varepsilon_j}\Big[B^\top\mathcal{P}_1(s;s)\\
&\quad +D^\top\mathcal{P}_1(s;s)C_\Theta+B^\top\mathcal{P}_{14}(s;s)+D^\top\mathcal{P}_{14}(s;s)C_\Theta+R(s,s)\Theta\\
&\quad +\big(\widehat{B}^\top+B^\top\mathcal{P}_{10}+D^\top\mathcal{P}_{10}\widehat{D}^\top\big)G_2\mathcal{P}_{12}\\
&\quad +D^\top\mathcal{P}_{10}N(s,s)\mathcal{P}_{12}C_\Theta\Big]\mathrm{d}s\ x,v\Big\rangle\\
&\geq 0,\quad \mathbb{P}\text{-a.s.}
\end{aligned}
$$

This completes the proof. □

Now we are in a position to prove Theorem 1.6.

Proof of Theorem 1.6. The "only if" part. Suppose there exists an equilibrium strategy $\Theta(\cdot)\in L^2(0,T;\mathbb{R}^{k\times n})$. By Theorem 2.1, we get $(\mathbf{P_1}(\cdot;\cdot),\mathbf{P_2}(\cdot),\mathbf{P_3}(\cdot;\cdot))$ as the solution of the equation (2.3). Let $(P_1(\cdot;\cdot),P_2(\cdot),P_3(\cdot;\cdot))=(\mathbf{P_1}(\cdot;\cdot),\mathbf{P_2}(\cdot),\mathbf{P_3}(\cdot;\cdot))$. From the equation (2.2), we get (1.13) and know that there exists a $\theta_0(\cdot)\in L^2(0,T;\mathbb{R}^{k\times n})$ such that (1.11) holds. From (1.11), we find that

$$
\begin{aligned}
&\big[R(t,t)+D(t)^\top\big(P_1(t;t)+P_3(t;t)+P_2(t)^\top N(t,t)P_2(t)\big)D(t)\big]^\dagger\\
&\big[R(t,t)+D(t)^\top\big(P_1(t;t)+P_3(t;t)+P_2(t)^\top N(t,t)P_2(t)\big)D(t)\big]\Theta(t)\\
=&\ \big[R(t,t)+D(t)^\top\big(P_1(t;t)+P_3(t;t)+P_2(t)^\top N(t,t)P_2(t)\big)D(t)\big]^\dagger\\
&\times\big[B(t)^\top\big(P_1(t;t)+P_3(t;t)\big)+D(t)^\top\big(P_1(t;t)+P_3(t;t)+P_2(t)^\top N(t,t)P_2(t)\big)C(t)\\
&\quad+\big(\widehat B(t)^\top+B(t)^\top P_2(t)^\top+D(t)^\top P_2(t)^\top\widehat D^\top(t)\big)G_2(t)P_2(t)\big],
\end{aligned}
$$

which implies (1.12). Lastly, (1.14) follows from (2.1).

The "if" part. Note that the systems (1.10) and (2.3) are consistent. If such a $\theta_0(\cdot)\in L^2(0,T;\mathbb{R}^{k\times n})$ exists, then by (1.12), we know that the $\Theta(\cdot)$ defined by (1.11) is in $L^2(0,T;\mathbb{R}^{k\times n})$. Moreover, by (1.11), (1.13) and (1.14), we see that (2.1) and (2.2) hold. Thus, from Theorem 2.1, such $\Theta(\cdot)$ is a closed-loop equilibrium strategy. □

3. Proof of Theorem 1.12

In this section, we prove Theorem 1.12. To reach this goal, we first reduce the systems (1.10)–(1.12) into an integral equation system, which is more convenient to be handled.

Proposition 3.1. *Let Assumption 1.5 hold. Then Problem (TI-FBSLQ) admits a closed-loop equilibrium strategy if and only if there exists $\theta_0(\cdot)\in L^2(0,T;\mathbb{R}^{k\times n})$ such that the following system*

$$
\begin{cases}
\Phi(s,t)=I+\int_t^s A_\Theta\Phi(r,t)dr+\int_t^s C_\Theta\Phi(r,t)\mathrm{d}W(r), & 0\le t\le s\le T,\\
\widetilde P_2(t)=H+\int_t^T\big(\widetilde P_2A_\Theta+\widehat A_\Theta+\widehat C\widetilde P_2+\widehat D\widetilde P_2C_\Theta\big)\mathrm{d}s, & 0\le t\le T,\\
\widetilde P_1(t)=\mathbb{E}_t\Big[\Phi(T,t)^\top G_1(t)\Phi(T,t)+\int_t^T\Phi(s,t)^\top\big(Q(s,t)+\Theta^\top R(s,t)\Theta\\
\qquad+\widetilde P_2^\top M(s,t)\widetilde P_2+C_\Theta^\top\widetilde P_2^\top N(s,t)\widetilde P_2C_\Theta\big)\Phi(s,t)\mathrm{d}s\Big], & 0\le t\le T,
\end{cases}
\tag{3.1}
$$

where

$$
\begin{aligned}
\Theta(s)=&-\big[R(s,s)+D(s)^\top\big(\widetilde P_1(s)+\widetilde P_2(s)^\top N(s,s)\widetilde P_2(s)\big)D(s)\big]^\dagger\\
&\times\big[B(s)^\top\widetilde P_1(s)+D(s)^\top\big(\widetilde P_1(s)+\widetilde P_2(s)^\top N(s,s)\widetilde P_2(s)\big)C(s)\\
&+\big(\widehat B(s)^\top+B(s)^\top\widetilde P_2(s)^\top+D(s)^\top\widetilde P_2(s)^\top\widehat D(s)^\top\big)G_2(s)\widetilde P_2(s)\big]\\
&+\theta_0(s)-\big[R(s,s)+D(s)^\top\big(\widetilde P_1(s)+\widetilde P_2(s)^\top N(s,s)\widetilde P_2(s)\big)D(s)\big]^\dagger\\
&\times\big[R(s,s)+D(s)^\top\big(\widetilde P_1(s)+\widetilde P_2(s)^\top N(s,s)\widetilde P_2(s)\big)D(s)\big]\theta_0(s),
\end{aligned}
\tag{3.2}
$$

admits a solution $(\widetilde P_1(\cdot),\widetilde P_2(\cdot))$ satisfying

$$
\begin{cases}
\big[R(\cdot,\cdot)+D(\cdot)^\top\big(\widetilde P_1(\cdot)+\widetilde P_2(\cdot)^\top N(\cdot,\cdot)\widetilde P_2(\cdot)\big)D(\cdot)\big]^\dagger \\
\times\big[B(\cdot)^\top\widetilde P_1(\cdot)+D(\cdot)^\top\big(\widetilde P_1(\cdot)+\widetilde P_2(\cdot)^\top N(\cdot,\cdot)\widetilde P_2(\cdot)\big)C(\cdot) \\
\quad+\big(\widehat B(\cdot)^\top+B(\cdot)^\top\widetilde P_2(\cdot)^\top+D(\cdot)^\top\widetilde P_2(\cdot)^\top\widehat D(\cdot)^\top\big)G_2(\cdot)\widetilde P_2(\cdot)\big]\in L^2(0,T;\mathbb{R}^{k\times n}),\\
\mathcal{R}\Big(R(t,t)+D(t)^\top\big(\widetilde P_1(t)+\widetilde P_2(t)^\top N(t,t)\widetilde P_2(t)\big)D(t)\Big)\supseteq\\
\mathcal{R}\Big(B(t)^\top\widetilde P_1(t)+D^\top(t)\big(\widetilde P_1(t)+\widetilde P_2(t)^\top N(t,t)\widetilde P_2(t)\big)C(t)\\
\quad+\big(\widehat B(t)^\top+B(t)^\top\widetilde P_2(t)^\top+D(t)^\top\widetilde P_2(t)^\top\widehat D(t)^\top\big)G_2(t)\widetilde P_2(t)\Big), \qquad \text{a.e. } t\in[0,T],\\
R(t,t)+D(t)^\top\widetilde P_1(t)D(t)+D(t)^\top\widetilde P_2(t)^\top N(t,t)\widetilde P_2(t)D(t)\ge 0, \qquad \text{a.e. } t\in[0,T].
\end{cases}
\tag{3.3}
$$

Proof. The "only if" part. By Theorem 1.6, we know that there exists $\theta_0(\cdot)\in L^2(0,T;\mathbb{R}^{k\times n})$ such that the equations (1.10)–(1.14) admit a solution $(P_1(\cdot;\cdot),P_2(\cdot),P_3(\cdot;\cdot))$. Let

$$
\begin{cases}
\widetilde P(s;t):=P_1(s;t)+P_3(s;t),\\
\widetilde P_1(s):=\widetilde P(s;s),\quad \widetilde P_2(s):=P_2(s).
\end{cases}
\tag{3.4}
$$

Apply Ito's formula to $s\to\Phi(s,t)^\top\widetilde P(s;t)\Phi(s,t)$, we have

$$
\begin{aligned}
&\widetilde P_1(t)-\Phi(T,t)^\top G_1(t)\Phi(T,t)\\
&=\int_t^T\Phi(s,t)^\top\big(Q(s,t)+\Theta^\top R(s,t)\Theta+P_2^\top M(s,t)P_2\\
&\quad+C_\Theta^\top P_2^\top N(s,t)P_2C_\Theta\big)\Phi(s,t)\mathrm{d}s-\int_t^T\big(\Phi(s,t)^\top\widetilde P(s,t)C_\Theta\Phi(s,t)\\
&\quad+\Phi(s,t)^\top C_\Theta^\top\widetilde P(s,t)\Phi(s,t)\big)\mathrm{d}W(s),
\end{aligned}
$$

which implies the last equation in (3.1). Combining (1.10)–(1.14), we see (3.3) holds.

The "if" part. Since there exists $\theta_0(\cdot)\in L^2(0,T;\mathbb{R}^{k\times n})$ such that the equations (3.1)–(3.3) admit a solution $(\widetilde P_1(\cdot),\widetilde P_2(\cdot))$. With $(\widetilde P_1(\cdot),\widetilde P_2(\cdot))$ and $\theta_0(\cdot)$, we choose $\Theta(\cdot)$ as in (3.2), which is in $L^2(0,T;\mathbb{R}^{k\times n})$ by (3.3). Consider the following system:

$$
\begin{cases}
\dfrac{\mathrm{d}P_1^*(s;t)}{\mathrm{d}s}+P_1^*(s;t)A_\Theta+A_\Theta^\top P_1^*(s;t)\\
\quad+C_\Theta^\top P_1^*(s;t)C_\Theta+Q(s,t)+\Theta^\top R(s,t)\Theta=0, & 0\le t\le s\le T,\\
\dfrac{\mathrm{d}P_2^*}{\mathrm{d}s}+P_2^*A_\Theta+\widehat A_\Theta+\widehat CP_2^*+\widehat DP_2^*C_\Theta=0, & 0\le t\le s\le T,\\
\dfrac{\mathrm{d}P_3^*(s;t)}{\mathrm{d}s}+P_3^*(s;t)A_\Theta+A_\Theta^\top P_3^*(s;t)+C_\Theta^\top P_3^*(s;t)C_\Theta\\
\quad+P_2^{*\top}M(s,t)P_2^*+C_\Theta^\top P_2^{*\top}N(s,t)P_2^*C_\Theta=0, & 0\le t\le s\le T,\\
P_1^*(T;t)=G_1(t),P_2^*(T)=H,P_3^*(T;t)=0, & 0\le t\le T.
\end{cases}
\tag{3.5}
$$

Since $P_2^*(\cdot)$ and $\widetilde P_2(\cdot)$ solve the same differential equation, we get that $P_2^*(\cdot)=\widetilde P_2(\cdot)$ by the uniqueness of the solution to the second equation of (3.5).

Set $\widetilde{P}_1^*(s;t) := P_1^*(s;t) + P_3^*(s;t)$. Similar to the proof of the "only if" part, we can obtain a representation of $\widetilde{P}_1^*(t;t)$, and consequently deduce that $\widetilde{P}_1^*(t;t) = \widetilde{P}_1(t)$. Therefore, $\Theta(\cdot)$ can be rewritten as

$$\begin{aligned}\Theta(s) =& -\big[R(s,s) + D(s)^\top\big(P_1^*(s;s) + P_3^*(s;s) + P_2^*(s)^\top N(s,s)P_2^*(s)\big)D(s)\big]^\dagger\\ &\times\big[B(s)^\top\big(P_1^*(s;s) + P_3^*(s;s)\big) + D(s)^\top\big(P_1^*(s;s) + P_3^*(s;s) + P_2^*(s)^\top N(s,s)P_2^*(s)\big)\\ &\quad\times C(s) + \big(\widehat{B}(s)^\top + B(s)^\top P_2^*(s)^\top + D(s)^\top P_2^*(s)^\top \widehat{D}(s)^\top\big)G_2(s)P_2^*(s)\big]\\ &+\theta_0(s) - \big[R(s,s) + D(s)^\top\big(P_1^*(s;s) + P_3^*(s;s) + P_2^*(s)^\top N(s,s)P_2^*(s)\big)D(s)\big]^\dagger\\ &\times\big[R(s,s) + D(s)^\top\big(P_1^*(s;s) + P_3^*(s;s) + P_2^*(s)^\top N(s,s)P_2^*(s)\big)D(s)\big]\theta_0(s).\end{aligned}$$

Combining (3.3) and Theorem 1.6, the sufficiency is proved. □

Now, we prove Theorem 1.12 *via* Proposition 3.1.

Proof of Theorem 1.12. We divide the proof into four steps.

Step 1. In this step, we prove the boundedness of $\Theta(\cdot)$, $\widetilde{P}_1(\cdot)$ and $\widetilde{P}_2(\cdot)$ in (3.1)–(3.2).
By Assumption 1.10 and the last equation in (3.1), we get that $\widetilde{P}_1(\cdot) \geq 0$.
Since $m = n = k = 1$, $\Theta(\cdot)$ can be rewritten as

$$\begin{aligned}\Theta(s) =& \big(R(s,s) + D(s)^2\widetilde{P}_1(s) + D(s)^2N(s,s)\widetilde{P}_2(s)^2\big)^{-1}\\ &\times\big\{\big(B(s) + D(s)C(s)\big)\widetilde{P}_1(s) + \widehat{B}(s)G_2(s)\widetilde{P}_2(s)\\ &\quad+\big[D(s)C(s)N(s,s) + \big(B(s) + D(s)\widehat{D}(s)\big)G_2(s)\big]\widetilde{P}_2(s)^2\big\}\\ =:& \, I_1(s) + I_2(s),\end{aligned}$$

where

$$\begin{aligned}|I_1(s)| &= \left|\frac{\big(B(s) + D(s)C(s)\big)\widetilde{P}_1(s)}{R(s,s) + D(s)^2\widetilde{P}_1(s) + D(s)^2N(s,s)\widetilde{P}_2(s)^2}\right|\\ &\leq \max\left\{\frac{|B(s) + D(s)C(s)|}{\delta}, \frac{|B(s) + D(s)C(s)|}{D(s)^2}\right\} \leq \frac{\mathcal{C}}{\delta}, \qquad \text{a.e. } s \in [0,T],\end{aligned}$$

and

$$\begin{aligned}|I_2(s)| &= \left|\frac{\widehat{B}(s)G_2(s)\widetilde{P}_2(s) + \big[D(s)C(s)N(s,s) + \big(B(s) + D(s)\widehat{D}(s)\big)G_2(s)\big]\widetilde{P}_2(s)^2}{R(s,s) + D(s)^2\widetilde{P}_1(s) + D(s)^2N(s,s)\widetilde{P}_2(s)^2}\right|\\ &\leq \max\left\{\frac{\mathcal{C}}{\delta}, \frac{\mathcal{C}}{D(s)^2\delta}\right\} \leq \max\left\{\frac{\mathcal{C}}{\delta}, \frac{\mathcal{C}}{\delta^2}\right\}, \quad \text{a.e. } s \in [0,T].\end{aligned}$$

Hence, we have

$$|\Theta(t)| < \mathcal{C}, \qquad \text{a.e. } t \in [0,T].$$

From the second equation in (3.1), we have

$$|\widetilde{P}_2(t)| \leq |H| + \int_t^T \Big(\big|A_\Theta + \widehat{C} + \widehat{D}C_\Theta\big| \times \big|\widetilde{P}_2\big| + \big|\widehat{A}_\Theta\big|\Big)\mathrm{d}s,$$

which, together with Gronwall's inequality, implies that

$$\widetilde{P}_2(s) < \mathcal{C}, \qquad s \in [0,T].$$

Similarly, we can get

$$\sup_{t\in[0,T]} \sup_{s\in[t,T]} \mathbb{E}|\Phi(s,t)|^2 < \mathcal{C}.$$

Then,

$$\begin{aligned}\widetilde{P}_1(t) =& \mathbb{E}\Big\{\Phi(T,t)^2 G_1(t) + \int_t^T \Phi(s,t)^2 \big[Q(s,t) + \Theta^2 R(s,t) + (\widetilde{P}_2)^2 M(s,t) + C_\Theta^2 (\widetilde{P}_2)^2 N(s,t)\big]\mathrm{d}s\Big\}\\ &= G_1(t)\mathbb{E}\Phi(T,t)^2 + \int_t^T \big[Q(s,t) + \Theta^2 R(s,t) + (\widetilde{P}_2)^2 M(s,t) + C_\Theta^2 (\widetilde{P}_2)^2 N(s,t)\big]\mathbb{E}\Phi(s,t)^2 \mathrm{d}s\\ &< \mathcal{C}.\end{aligned}$$

Step 2. In this step, we establish estimates for $\widetilde{P}_1(\cdot)$ and $\widetilde{P}_2(\cdot)$.
Given Θ_i $(i=1,2)$, and

$$\begin{cases}\dfrac{\mathrm{d}\widetilde{P}_2^i}{\mathrm{d}s} = -\big(\widetilde{P}_2^i A_{\Theta_i} + \widehat{A}_{\Theta_i} + \widehat{C}\widetilde{P}_2^i + \widehat{D}\widetilde{P}_2^i C_{\Theta_i}\big), \quad s \in [0,T],\\ \widetilde{P}_2^i(T) = H,\end{cases}$$

we have

$$\begin{aligned}&\widetilde{P}_2^1(t) - \widetilde{P}_2^2(t)\\ &= \int_t^T \big[\big(\widetilde{P}_2^1 - \widetilde{P}_2^2\big)A_{\Theta_1} + \widetilde{P}_2^2 B\big(\Theta_1 - \Theta_2\big) + \widehat{B}\big(\Theta_1 - \Theta_2\big)\\ &\qquad + \widehat{C}\big(\widetilde{P}_2^1 - \widetilde{P}_2^2\big) + \widehat{D}\big(\widetilde{P}_2^1 - \widetilde{P}_2^2\big)C_{\Theta_1} + \widehat{D}\widetilde{P}_2^2 D\big(\Theta_1 - \Theta_2\big)\big]\mathrm{d}s.\end{aligned}$$

By Gronwall's inequality, we get that

$$\begin{aligned}\big|\widetilde{P}_2^1(t) - \widetilde{P}_2^2(t)\big| &\leq \int_t^T \big(\mathcal{C}\big|\widetilde{P}_2^1 - \widetilde{P}_2^2\big| + \mathcal{C}\big|\Theta_1 - \Theta_2\big|\big)\mathrm{d}s\\ &\leq \mathcal{C}\int_t^T \big|\Theta_1 - \Theta_2\big|\mathrm{d}s\\ &\leq \mathcal{C}(T-t)\big|\Theta_1(\cdot) - \Theta_2(\cdot)\big|_{L^\infty(t,T;\mathbb{R})},\end{aligned} \tag{3.6}$$

which yields

$$\sup_{s\in[t,T]} \big|\widetilde{P}_2^1(s) - \widetilde{P}_2^2(s)\big| \leq \mathcal{C}(T-t)\big|\Theta_1(\cdot) - \Theta_2(\cdot)\big|_{L^\infty(t,T;\mathbb{R})}. \tag{3.7}$$

Similarly, for $0 \le t \le \tau \le T$, consider the following equation:

$$\begin{cases} \mathrm{d}\Phi^i(s,t) = A_{\Theta_i}\Phi^i(s,t)\mathrm{d}s + C_{\Theta_i}\Phi^i(s,t)\mathrm{d}W(s), \quad s\in[t,\tau], \quad i=1,2, \\ \Phi^i(t,t) = I. \end{cases}$$

By standard estimate for SDEs(*e.g.*, [1], Thm. 3.2), we have

$$\begin{aligned} \mathbb{E}\sup_{s\in[t,\tau]}\big|\Phi^1(s,t)-\Phi^2(s,t)\big|^2 &\le \mathcal{C}\int_t^\tau \big|\Theta_1-\Theta_2\big|^2\mathrm{d}s \\ &\le \mathcal{C}(\tau-t)\big|\Theta_1(\cdot)-\Theta_2(\cdot)\big|^2_{L^\infty(t,\tau;\mathbb{R})}, \end{aligned} \tag{3.8}$$

where the constant $\mathcal{C}$ is independent of the choice of t and τ.

For $i=1,2$, let

$$\begin{aligned} \widetilde{P}_1^i(t) =& \mathbb{E}\Big\{\Phi^i(T,t)^\top G_1(t)\Phi^i(T,t) + \int_t^T \Phi^i(s,t)^\top\big[Q(s,t)+\Theta_i^\top R(s,t)\Theta_i \\ &+ (\widetilde{P}_2^i)^\top M(s,t)\widetilde{P}_2^i + C_{\Theta_i}^\top(\widetilde{P}_2^i)^\top N(s,t)\,\widetilde{P}_2^i\,C_{\Theta_i}\big]\Phi^i(s,t)\mathrm{d}s\Big\}. \end{aligned}$$

From (3.6) and (3.8), we obtain that

$$\begin{aligned} &\big|\widetilde{P}_1^1(t)-\widetilde{P}_1^2(t)\big| \\ &\le \mathbb{E}\Big\{\big|G_1(t)\big(\Phi^1(T,t)^2-\Phi^2(T,t)^2\big)\big| + \int_t^T\Big[\Big|Q(s,t)\big(\Phi^1(s,t)^2-\Phi^2(s,t)^2\big)\Big| \\ &\quad+\Big|R(s,t)\big(\Phi^1(s,t)^2\Theta_1^2-\Phi^2(s,t)^2\Theta_2^2\big)\Big|+\Big|M(s,t)\big(\Phi^1(s,t)^2(\widetilde{P}_2^1)^2-\Phi^2(s,t)^2(\widetilde{P}_2^2)^2\big)\Big| \\ &\quad+\Big|N(s,t)\big(\Phi^1(s,t)^2(\widetilde{P}_2^1)^2C_{\Theta_1}^2-\Phi^2(s,t)^2(\widetilde{P}_2^2)^2C_{\Theta_2}^2\big)\Big|\Big]\mathrm{d}s\Big\} \\ &\le \mathcal{C}\Big\{\Big(\mathbb{E}\big|\Phi^1(T,t)-\Phi^2(T,t)\big|^2\Big)^{1/2} + \int_t^T\Big[\Big(\mathbb{E}\big|\Phi^1(s,t)-\Phi^2(s,t)\big|^2\Big)^{1/2} \\ &\quad+\Big(\mathbb{E}\big|\Phi^1(s,t)\Theta_1-\Phi^2(s,t)\Theta_2\big|^2\Big)^{1/2}+\Big(\mathbb{E}\big|\Phi^1(s,t)\widetilde{P}_2^1-\Phi^2(s,t)\widetilde{P}_2^2\big|^2\Big)^{1/2} \\ &\quad+\Big(\mathbb{E}\big|\Phi^1(s,t)\widetilde{P}_2^1C_{\Theta_1}-\Phi^2(s,t)\widetilde{P}_2^2C_{\Theta_2}\big|^2\Big)^{1/2}\Big]\mathrm{d}s\Big\} \\ &\le \mathcal{C}\Big[\Big(\int_t^T\big|\Theta_1(s)-\Theta_2(s)\big|^2\mathrm{d}s\Big)^{1/2} \\ &\quad+\int_t^T\Big(\int_t^s\big|\Theta_1(\tau)-\Theta_2(\tau)\big|^2\mathrm{d}\tau+\big|\Theta_1(s)-\Theta_2(s)\big|^2+\big|\widetilde{P}_2^1(s)-\widetilde{P}_2^2(s)\big|^2\Big)^{1/2}\mathrm{d}s\Big] \\ &\le \mathcal{C}\Big(\int_t^T\big|\Theta_1(s)-\Theta_2(s)\big|^2\mathrm{d}s\Big)^{1/2} \\ &\le \mathcal{C}(T-t)^{1/2}\big|\Theta_1(\cdot)-\Theta_2(\cdot)\big|_{L^\infty(t,T;\mathbb{R})}. \end{aligned} \tag{3.9}$$

Step 3. Define the mapping $F_0 : L^\infty(t_0, T; \mathbb{R}) \to L^\infty(t_0, T; \mathbb{R})$ as follows (t_0 is to be determined later): for $\Theta(\cdot) \in L^\infty(t_0, T; \mathbb{R})$,

$$\begin{aligned}F_0(\Theta(\cdot))(s) =& - \big[R(s,s) + D(s)^\top \big(\widetilde P_1(s) + \widetilde P_2(s)^\top N(s,s)\widetilde P_2(s)\big)D(s)\big]^\dagger \\ &\times \big[B(s)^\top \widetilde P_1(s) + D(s)^\top \big(\widetilde P_1(s) + \widetilde P_2(s)^\top N(s,s)\widetilde P_2(s)\big)C(s) \\ &\quad + \big(\widehat B(s)^\top + B(s)^\top \widetilde P_2(s)^\top + D(s)^\top \widetilde P_2(s)^\top \widehat D(s)^\top\big)G_2(s)\widetilde P_2(s)\big] \\ &+ \theta_0(s) - \big[R(s,s) + D(s)^\top \big(\widetilde P_1(s) + \widetilde P_2(s)^\top N(s,s)\widetilde P_2(s)\big)D(s)\big]^\dagger \\ &\times \big[R(s,s) + D(s)^\top \big(\widetilde P_1(s) + \widetilde P_2(s)^\top N(s,s)\widetilde P_2(s)\big)D(s)\big]\theta_0(s), \qquad \forall\, s \in [t_0, T],\end{aligned}$$

where $(\widetilde P_1(\cdot), \widetilde P_2(\cdot))$ are the solution to the equation (3.1) by choosing $\Theta(\cdot)$ to be the control strategy on time interval $[t_1, T]$ in the equation (3.2).

Given $\Theta_i(\cdot) \in L^\infty(t_0, T; \mathbb{R})$, $i = 1, 2$,

$$\begin{aligned}&F_0(\Theta_1(\cdot))(s) - F_0(\Theta_2(\cdot))(s) \\ &= \Big\{\Big[\big(R(s,s) + D(s)^2\widetilde P_1^1(s) + D(s)^2\widetilde P_2^1(s)^2N(s,s)\big)^{-1} \\ &\qquad - \big(R(s,s) + D(s)^2\widetilde P_1^2(s) + D(s)^2\widetilde P_2^2(s)^2N(s,s)\big)^{-1}\Big] \\ &\quad \times\Big[\big(B(s) + D(s)C(s)\big)\widetilde P_1^1(s) + \widehat B(s)G_2(s)\widetilde P_2^1(s) \\ &\qquad + \big[D(s)C(s)N(s,s) + \big(B(s) + D(s)\widehat D(s)\big)G_2(s)\big]\widetilde P_2^1(s)^2\Big]\Big\} \\ &\quad + \Big\{\big(R(s,s) + D(s)^2\widetilde P_1^2(s) + D(s)^2\widetilde P_2^2(s)^2N(s,s)\big)^{-1} \\ &\qquad \times\Big[\big(B(s) + D(s)C(s)\big)\big(\widetilde P_1^1(s) - \widetilde P_1^2(s)\big) + \widehat B(s)G_2(s)\big(\widetilde P_2^1(s) - \widetilde P_2^2(s)\big) \\ &\qquad\quad + \big[D(s)C(s)N(s,s) + \big(B(s) + D(s)\widehat D(s)\big)G_2(s)\big]\big(\widetilde P_2^1(s)^2 - \widetilde P_2^2(s)^2\big)\Big]\Big\} \\ &\leq \frac{\mathcal C}{\delta^2}\big(\big|\widetilde P_1^1(s) - \widetilde P_1^2(s)\big| + \big|P_2^2(s)^2 - P_2^1(s)^2\big|\big) \\ &\quad + \frac{\mathcal C}{\delta}\big(\big|\widetilde P_1^1(s) - \widetilde P_1^2(s)\big| + \big|\widetilde P_2^2(s) - \widetilde P_2^1(s)\big| + \big|\widetilde P_2^2(s)^2 - \widetilde P_2^1(s)^2\big|\big) \\ &\leq \mathcal C\big(\big|\widetilde P_1^1(s) - \widetilde P_1^2(s)\big| + \big|\widetilde P_2^1(s) - \widetilde P_2^2(s)\big| + \big|\widetilde P_2^1(s)^2 - \widetilde P_2^2(s)^2\big|\big).\end{aligned} \tag{3.10}$$

By (3.7) and (3.9), we get that

$$\big|F_0(\Theta_1(\cdot))(\cdot) - F_0(\Theta_2(\cdot))(\cdot)\big|_{L^\infty(t_0,T;\mathbb{R})} \leq \mathcal C(T - t_0)^{1/2}\big|\Theta_1(\cdot) - \Theta_2(\cdot)\big|_{L^\infty(t_0,T;\mathbb{R})}. \tag{3.11}$$

Choose t_0 such that $\mathcal C(T - t_0)^{1/2} < 1/2$, *via* Banach fixed-point theorem, we know there is a unique fixed point $\Theta^*(\cdot) \in L^\infty(t_0, T; \mathbb{R})$ for the mapping F_0.

Step 4. Define the mapping $F_1 : L^\infty(t_1, t_0; \mathbb{R}) \to L^\infty(t_1, t_0; \mathbb{R})$ as follows (t_1 is to be determined later): for $\Theta(\cdot) \in L^\infty(t_1, t_0; \mathbb{R})$,

$$\begin{aligned}F_1(\Theta(\cdot))(s) =& \Big\{ - \big[R(s,s) + D(s)^\top \big(\widetilde P_1(s) + \widetilde P_2(s)^\top N(s,s)\widetilde P_2(s)\big)D(s)\big]^\dagger \\ &\times \big[B(s)^\top \widetilde P_1(s) + D(s)^\top \big(\widetilde P_1(s) + \widetilde P_2(s)^\top N(s,s)\widetilde P_2(s)\big)C(s) \\ &\quad + \big(\widehat B(s)^\top + B(s)^\top \widetilde P_2(s)^\top + D(s)^\top \widetilde P_2(s)^\top \widehat D(s)^\top\big)G_2(s)\widetilde P_2(s)\big]\end{aligned}$$

$$+\theta_0(s)-\big[R(s,s)+D(s)^\top\big(\widetilde P_1(s)+\widetilde P_2(s)^\top N(s,s)\widetilde P_2(s)\big)D(s)\big]^\dagger$$
$$\times\big[R(s,s)+D(s)^\top\big(\widetilde P_1(s)+\widetilde P_2(s)^\top N(s,s)\widetilde P_2(s)\big)D(s)\big]\theta_0(s)\Big\}\times\chi_{[t_1,t_0]}(s),\ s\in[t_1,t_0],$$

where $(\widetilde P_1(\cdot),\widetilde P_2(\cdot))$ is the solution of the equation (3.1) by choosing $\Theta(\cdot)\chi_{[t_1,t_0]}(\cdot)+\Theta^*(\cdot)\chi_{[t_0,T]}(\cdot)$ to be the control strategy on time interval $[t_1,T]$ in the equation (3.2).

Given $\Theta_i(\cdot)\in L^\infty(t_1,t_0;\mathbb{R})$, $i=1,2$, similar to the proof of (3.11), we can show that

$$\big|F_1(\Theta_1(\cdot))(\cdot)-F_1(\Theta_2(\cdot))(\cdot)\big|_{L^\infty(t_1,t_0;\mathbb{R})}\le \mathcal{C}(t_0-t_1)^{1/2}\big|\Theta_1(\cdot)-\Theta_2(\cdot)\big|_{L^\infty(t_1,t_0;\mathbb{R})}.$$

By choosing $t_0-t_1=T-t_0$, it follows from Banach fixed-point theorem that there is a unique fixed point $\Theta^*(\cdot)\in L^\infty(t_1,t_0;\mathbb{R})$ for the mapping F_1. Inductively, we obtain the existence and uniqueness of the solution to the equations (3.1)–(3.3). By Proposition 3.1 and Theorem 1.6, we complete the proof. □

Funding

This work is supported by the NSF of China under grants 12025105 and 11931011, by the New Cornerstone Science Foundation, and by the Science Development Project of Sichuan University under grant 2020SCUNL201.

References

[1] Q. Lü and X. Zhang, Mathematical Control Theory for Stochastic Partial Differential Equations. Springer, Cham (2021).

[2] H. Wang, J. Yong and C. Zhou, Optimal controls for forward–backward stochastic differential equations: time-inconsistency and time-consistent solutions. *J. Math. Pures Appl.* **190** (2024) 103603.

[3] F. Dou and Q. Lü, Time-inconsistent linear quadratic optimal control problems for stochastic evolution equations. *SIAM J. Control Optim.* **58** (2020) 485–509.

[4] Q. Lü and B. Ma, Time-inconsistent stochastic linear quadratic control problem with indefinite control weight costs. *Sci. China Math.* **67** (2024), 211–236.

[5] J. Yong, Time-inconsistent optimal control problems, in *Proceedings of 2014 ICM, Section 16. Control Theory and Optimization* (2014) 947–969.

[6] J. Yong, Linear-quadratic optimal control problems for mean-field stochastic differential equations–time-consistent solutions. *Trans. Amer. Math. Soc.* **369** (2017) 5467–5523.

[7] Y. Hu, H. Jin and X. Zhou, Time-inconsistent stochastic linear–quadratic control. *SIAM J. Control Optim.* **50** (2012) 1548–1572.

[8] Y. Hu, H. Jin and X. Zhou, Time-inconsistent stochastic linear-quadratic control: Characterization and uniqueness of equilibrium. *SIAM J. Control Optim.* **55** (2017) 1261–1279.

[9] S. Chen, X. Li and X.Y. Zhou, Stochastic linear quadratic regulators with indefinite control weight costs. *SIAM J. Control Optim.* **36** (1998) 1685–1702.

[10] M.A. Rami, J.B. Moore and X.Y. Zhou, Indefinite stochastic linear quadratic control and generalized differential Riccati equation. *SIAM J. Control Optim.* **40** (2001) 1296–1311.

[11] J. Sun and J. Yong, Stochastic linear-quadratic optimal control theory: open-loop and closed-loop solutions. Springer Briefs in Mathematics (2020).

[12] S. Peng, Backward stochastic differential equations and applications to optimal control. *Appl. Math. Optim.* **27** (1993) 125–144.

[13] M. Hu, S. Ji and X. Xue, A global stochastic maximum principle for fully coupled forward–backward stochastic systems. *SIAM J. Control Optim.* **56** (2018) 4309–4335.

[14] J. Ma and J. Yong, Forward–Backward Stochastic Differential Equations and Their Applications. Springer-Verlag, Berlin (1999).

[15] G. Wang, Z. Wu and J. Xiong, Maximum principles for forward–backward stochastic control systems with correlated state and observation noises. *SIAM J. Control Optim.* **51** (2013) 491–524.

[16] J. Yong, Optimality variational principle for controlled forward–backward stochastic differential equations with mixed initial-terminal conditions. *SIAM J. Control Optim.* **48** (2010) 4119–4156.

[17] D. Hume, A Treatise of Human Nature, 1st edn., 1739, reprint. Oxford University Press, Oxford (1978).

[18] A. Smith, The Theory of Moral Sentiments, 1st edn., 1759, reprint. Oxford University Press, Oxford (1976).

[19] R. Strotz, Myopia and inconsistency in dynamic utility maximization. *Rev. Econ. Stud.* **23** (1955) 165–180.

[20] S. Basak and G. Chabakauri, Dynamic mean-variance asset allocation. *Rev. Finan. Stu.* **23** (2010) 2970–3016.

[21] I. Ekeland and T. Pirvu, Investment and consumption without commitment. *Math. Financ. Econ.* **2** (2008) 57–86.

[22] S.M. Goldman, Consistent plans. *Rev. Econ. Stud.* **47** (1980) 533–537.

[23] B. Peleg and M.E. Yaari, On the existence of a consistent course of action when tastes are changing. *Rev. Econ. Stud.* **40** (1973) 391–401.

[24] R.A. Pollak, Consistent planning. *Rev. Econ. Stud.* **35** (1968) 185–199.

[25] B. Ma, A singular linear quadratic time-inconsistent optimal control problem. *J. Syst. Sci. Complex.* **36** (2023) 1024–1052.