Linear Quadratic Optimal Control Problems of Delayed Backward Stochastic Differential Equations

Weijun Meng · Jingtao Shi

Accepted: 12 April 2021 / Published online: 26 April 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
This paper is concerned with a linear quadratic optimal control problem of delayed backward stochastic differential equations. An explicit representation is derived for the optimal control, which is a linear feedback of the entire past history and the expected value of the future state trajectory in a short period of time. To obtain the optimal feedback, a new class of delayed Riccati equations and delayed-advanced forward-backward stochastic differential equations are introduced. Furthermore, the unique solvability of their solutions are discussed in detail.

Keywords Linear quadratic optimal control · Stochastic differential delayed equation · Delayed backward stochastic differential equation · Time-advanced stochastic differential delayed equation · Delayed Riccati equation · Delayed-advanced forward–backward stochastic differential equation

Mathematics Subject Classification 93E20 · 60H10 · 34K50

1 Introduction

The stochastic control problems with delay have attracted more and more scholars’ attention in recent years, due to their wide applications in various fields such as economics, engineering, information science, and networked communication. Stochastic differential delayed equations (SDDEs, for short) are nice tools to described the
dynamics of some natural and social phenomena (see Mohammed [17,18]). Since then, massive research on related topics has become a desirable and serious endeavor among researchers in stochastic optimal control, differential games, mathematical finance and so on (see [2,3,6,9,10,12,19,20,23,31,33,35,37]). The forward SDDEs characterize the dynamic changes of state processes with given initial state trajectories. However, in the financial investment problems we usually prefer to study the dynamic changes of state processes with specified terminal states. Pardoux and Peng [21] established the general theory of backward stochastic differential equations (BSDEs, for short), whose solution is a pair of adapted processes when the terminal state is given. Since the BSDE itself is a nice dynamic structure, there have been abundant research results about the optimal control problems and differential games of BSDEs, for example [7,11,14,15,22,28–30,36].

Recently, Delong and Imkeller [5] introduced BSDEs with time delayed generators, which are generalization of BSDEs by adding the influence of time delay. Further, Delong [4] studied their applications in finance and insurance. Due to these, it is necessary and urgent to study the optimal control problems of BSDEs with time delayed generators. This is an interesting but challenging topic owing to the influence of the time delay and the backward structure on the controlled systems. A few research can be found in this kind of optimal control problems and their applications. Shi [24] considered the optimal control problem described by a kind of BSDE with time delayed generator and proved a sufficient maximum principle. Shi [25] generalized the above problem to the case driven by Brownian motion and Poisson random measure, by introducing a new class of time-advanced stochastic differential equations (ASDEs, for short) with jumps as the adjoint equation, and gave the sufficient maximum principle. Chen and Huang [1] investigated a stochastic recursive delayed control problem and derived the necessary and sufficient conditions of the maximum principle, by introducing a kind of more general ASDEs as the adjoint equation. Wu and Wang [32] focused on the optimal control problem of delayed BSDEs under partial information and the necessary and sufficient conditions of optimality are obtained. Shi and Wang [27] studied a nonzero sum differential game of BSDEs with time-delayed generator and gave an Arrow’s sufficient condition for the open-loop equilibrium point.

As one of the important special cases of optimal control problems, the linear-quadratic (LQ, for short) optimal control problems have been a hot topic for a long period. However, to our knowledge, the literatures about LQ optimal control problems of delayed BSDEs are very scarce. Although the above literatures have discussed the LQ cases, either the state feedback expressions of the optimal controls are not given, or the controlled systems are very special. Hence this paper aims to study the general LQ optimal control problem of delayed BSDE, which we called delayed backward stochastic linear-quadratic (D-BSLQ, for short) optimal control problem. The main contributions of this paper can be summarized in three aspects.

- Firstly, a general D-BSLQ optimal control problem is proposed and the necessary conditions for the existence of the optimal control are derived. The optimal control can be expressed as a linear feedback of the entire past history and the expected value of the future state trajectory in a short period of time. Furthermore, the optimal cost can be expressed by a delayed Riccati equation and a delayed-advanced
forward-backward stochastic differential equation (DAFBSDE, for short). See Theorems 3 and 4 in Sect. 3. It is worth mentioning that the completion-of-squares technique is invalid in this LQ problem because of the term $E^F_{t-\delta}[\bar{X}(t)] - \bar{X}(t)$ in (3.4), hence only the necessary conditions can be obtained.

- Secondly, it is interesting that a new class of time-advanced SDDEs (ASDDEs, for short) is introduced to seek the state feedback expression of the optimal control, which has not been studied yet although it has considerable study value. See (3.5) in Sect. 3.

- Thirdly, the delayed Riccati equations and the DAFBSDEs mentioned above play very important roles in our problem. Moreover, to the best of our knowledge, this class of delayed Riccati equations and DAFBSDEs have not appeared in the previous literature. Thus, the existence and uniqueness of their solutions are discussed in detail in Sect. 4.

The rest of this paper is organized as follows. In Section 2, the D-BSLQ optimal control problem is formulated and some preliminary results on ASDDEs are given. In Section 3, the main results about the state feedback expression of the optimal control, are presented. Section 4 is devoted to the existence and uniqueness of solutions to the delayed Riccati equations and DAFBSDEs. The proofs of the main results are carried out in Sect. 5. Finally, some concluding remarks are given in Sect. 6.

2 Problems Formulation and Preliminaries

Throughout the paper, $\mathbb{R}^{n\times m}$ is the Euclidean space of all $n \times m$ real matrices, $\mathbb{S}^n$ is the space of all $n \times n$ symmetric matrices, $\mathbb{S}_+^n$ is the subset of $\mathbb{S}^n$ consisting of positive semi-definite matrices, $\mathbb{S}_+^{n}$ is the subset of $\mathbb{S}^n$ consisting of positive definite matrices. We write $\mathbb{R}^{n\times m}$ as $\mathbb{R}^n$ when $m = 1$. The norm in $\mathbb{R}^n$ is denoted by $|\cdot|$ and the inner product is denoted by $\langle \cdot, \cdot \rangle$. The transpose of vectors or matrices is denoted by the superscript $^\top$. $\mathbb{R}^+ \equiv [0, \infty)$ and $\mathbb{N}^+$ is the set of positive integers. $I$ is the identity matrix with appropriate dimension.

Suppose that $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ is a complete filtered probability space, the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ is generated by the one-dimensional standard Brownian motion $\{W(t)\}_{t \geq 0}$, and $\mathbb{E}$ denotes the mathematical expectation with respect to the probability $\mathbb{P}$. Let $T > 0$ be the finite time duration and $\delta > 0$ be a sufficiently small time delay parameter. Denote $\mathcal{F}_{t-\delta} := \mathcal{F}_0$ when $t - \delta < 0$.

First we define the following spaces which will be used in this paper:

\begin{align*}
L^p([0, T]; \mathbb{R}^{n\times n}) &:= \left\{ \text{\mathbb{R}^{n\times n}-valued functicon } \phi(t); \int_0^T |\phi(t)|^p dt < \infty \right\}, \\
L^\infty([0, T]; \mathbb{R}^{n\times n}) &:= \left\{ \text{\mathbb{R}^{n\times n}-valued functicon } \phi(t); \sup_{0 \leq t \leq T} |\phi(t)| dt < \infty \right\}, \\
L^2_{\mathcal{F}_t}(\mathbb{R}^n) &:= \left\{ \text{\mathbb{R}^n-valued } \mathcal{F}_t\text{-measurable random variable } \xi; \mathbb{E}|\xi|^2 < \infty \right\}, \\
C([0, T]; \mathbb{R}^n) &:= \left\{ \text{\mathbb{R}^n-valued continuous functicon } \phi(t); \sup_{0 \leq t \leq T} |\phi(t)| < \infty \right\}.
\end{align*}
\begin{equation}
L^2_{\mathcal{F}}([0, T]; \mathbb{R}^n) := \left\{ \text{measurable } \mathcal{F}_t \text{-adapted process } \phi(t); \; \mathbb{E} \int_0^T |\phi(t)|^2 dt < \infty \right\}.
\end{equation}

\begin{equation}
L^2_{\mathcal{F}}(\Omega; C([0, T]; \mathbb{R}^n)) := \left\{ \text{measurable } \mathcal{F}_t \text{-adapted process } \phi(t); \; \mathbb{E} \left[ \sup_{0 \leq t \leq T} |\phi(t)|^2 \right] < \infty \right\}.
\end{equation}

Next we formulate the optimal control problem which will be studied in this paper. For given \( s \in [0, T] \), let us consider the following controlled linear delayed BSDE:

\begin{equation}
\begin{aligned}
- dY^u(t) &= \left[ A(t)Y^u(t) + \tilde{A}(t)Y^u(t - \delta) + B(t)Z^u(t) + \tilde{B}(t)Z^u(t - \delta) + C(t)u(t) + \tilde{C}(t)u(t - \delta) \right] dt - Z^u(t)dW(t), \quad t \in [s, T], \\
Y^u(T) &= \xi, \quad Y^u(t) = \varphi(t), \quad Z^u(t) = \psi(t), \quad u(t) = \eta(t), \quad t \in [s - \delta, s),
\end{aligned}
\end{equation}

along with the cost functional

\begin{equation}
J(u(\cdot)) = \mathbb{E}\left\{ \tilde{G}Y^u(s - \delta), Y^u(s - \delta) \right\} + \{GY^u(s), Y^u(s)\}
\end{equation}

\begin{equation}
+ \int_s^T \left\{ \{Q(t)Y^u(t), Y^u(t)\} + \{\tilde{Q}(t)Y^u(t - \delta), Y^u(t - \delta)\} \right. \\
\left. + \{R(t)Z^u(t), Z^u(t)\} + \{\tilde{R}(t)Z^u(t - \delta), Z^u(t - \delta)\} \right. \\
\left. + \{N(t)u(t), u(t)\} + \{\tilde{N}(t)u(t - \delta), u(t - \delta)\} \right\} dt,
\end{equation}

where \( A(\cdot), \tilde{A}(\cdot), B(\cdot), \tilde{B}(\cdot), C(\cdot), \tilde{C}(\cdot) \) are deterministic matrix-valued functions, \( \xi \) is an \( \mathcal{F}_T \)-measurable random vector, \( \varphi(\cdot), \psi(\cdot) \) and \( \eta(\cdot) \) are the initial trajectories of the state and the control, respectively. \( G, \tilde{G} \) are symmetric matrices, \( Q(\cdot), \tilde{Q}(\cdot), R(\cdot), \tilde{R}(\cdot), N(\cdot), \tilde{N}(\cdot) \) are deterministic matrix-valued functions, with appropriate dimensions.

The admissible control set is defined as follows:

\[ U[s, T] := \left\{ u : [s - \delta, T] \times \Omega \to \mathbb{R}^d \right\} \text{ for } t \in [s - \delta, s), \; u(t) = \eta(t); \text{ for } t \in [s, T], \]

\[ u(t) \text{ is a } \mathcal{F}_t \text{-predictable process, } \mathbb{E}\int_s^T |u(t)|^2 dt < \infty \right\}.
\]

The LQ optimal control problem of delayed BSDEs, which we call delayed backward stochastic LQ optimal control problem, can be stated as follows:

**Problem (D-BSLQ).** For any \( s \in [0, T], \xi \in L^2_{\mathcal{F}_T}(\mathbb{R}^n), \) to find a \( u^*(\cdot) \in U[s, T] \) such that

\begin{equation}
J(u^*(\cdot)) = \inf_{u(\cdot) \in U[s, T]} J(u(\cdot)).
\end{equation}

Any \( u^*(\cdot) \in U[s, T] \) that achieves the above infimum is called an optimal control and the corresponding solution \( (Y^*(\cdot), Z^*(\cdot)) \equiv (Y^u(\cdot), Z^u(\cdot)) \) is called the optimal state trajectory.
Now we introduce the following assumptions that will be in force throughout the paper.

(A1) The coefficients of (2.1) satisfy the following assumptions: 
\[ A(\cdot), \tilde{A}(\cdot), B(\cdot), \tilde{B}(\cdot) \in L^\infty([0, T]; \mathbb{R}^{n \times n}), \quad C(\cdot), \tilde{C}(\cdot) \in L^\infty([0, T]; \mathbb{R}^{n \times d}). \]

(A2) The initial trajectory of (2.1) satisfies \( \varphi(\cdot), \psi(\cdot), \eta(\cdot) \in L^2_T([s - \delta, s]; \mathbb{R}^n). \)

(A3) The weight coefficients of (2.2) satisfy the following assumptions: 
\[ \begin{align*}
G, \tilde{G} &\in \mathbb{S}^n, \quad Q(\cdot), \tilde{Q}(\cdot) \in L^\infty([0, T]; \mathbb{S}^n), \\
R(\cdot), \tilde{R}(\cdot) &\in L^\infty([0, T]; \mathbb{S}^n), \quad N(\cdot), \tilde{N}(\cdot) \in L^\infty([0, T]; \mathbb{S}^d),
\end{align*} \]

and there exists a constant \( \alpha > 0 \) such that 
\[ G \geq 0, \quad Q(t) + \tilde{Q}(t + \delta) \geq 0, \quad R(t) + \tilde{R}(t + \delta) \geq 0, \quad N(t) + \tilde{N}(t + \delta) \geq \alpha I, \]
a.e. \( t \in [s - \delta, T]. \)

Now we give a result to guarantee the well-posedness of (2.1).

**Theorem 1** Let (A1), (A2) hold and \( \delta \) be sufficiently small. Then for any \( (\xi, u(\cdot)) \in L^2_T(\mathbb{R}^n) \times \mathcal{U}([s, T]), \) the state equation (2.1) admits a unique adapted solution \( (Y^u(\cdot), Z^u(\cdot)) \in L^2_T(\Omega; C([s, T]; \mathbb{R}^n)) \times L^2_T([s, T]; \mathbb{R}^n). \) Moreover, there exists a constant \( K > 0, \) independent of \( \xi \) and \( u(\cdot), \) such that 
\[
\mathbb{E} \left[ \sup_{s \leq t \leq T} |Y^u(t)|^2 + \int_s^T |Z^u(t)|^2 dt \right] 
\leq K \mathbb{E} \left[ |\xi|^2 + \int_s^t |u(t)|^2 dt + \int_{s-\delta}^s \left( |\varphi(t)|^2 + |\psi(t)|^2 + |\eta(t)|^2 \right) dt \right]. 
\]

**Proof** We try to use the contraction mapping theorem to prove the result. For any \( \beta \in \mathbb{R}, \) let 
\[ \mathcal{M}_\beta[s, T] := L^2_T(\Omega; C([s, T]; \mathbb{R}^n)) \times L^2_T([s, T]; \mathbb{R}^n), \]
equipped with the norm 
\[ \left\| (Y^u, Z^u) \right\|_{\mathcal{M}_\beta[s, T]} := \left( \mathbb{E} \left[ \sup_{s \leq t \leq T} |Y^u(t)|^2 e^{\beta h(t)} \right] + \mathbb{E} \int_s^T |Z^u(t)|^2 e^{\beta h(t)} dt \right)^{\frac{1}{2}}, \]
where \( h(t) := \int_s^t \left[ |A(r)| + |\tilde{A}(r)| + |B(r)|^2 + |\tilde{B}(r)|^2 \right] dr, \) \( t \in [s, T]. \) We define the mapping 
\[ T : \mathcal{M}_\beta[s, T] \to \mathcal{M}_\beta[s, T], \]
\[ (y(\cdot), z(\cdot)) \mapsto (Y^u(\cdot), Z^u(\cdot)), \]
apparently it is well-defined. In fact, for any \((y(\cdot), z(\cdot)) \in \mathcal{M}_\beta[s, T]\) and \(y(t) = \varphi(t), z(t) = \psi(t), s - \delta \leq t < s\), consider the following BSDE:

\[
\begin{align*}
-dY^u(t) & = \left[ A(t)y(t) + \tilde{A}(t)y(t - \delta) + B(t)z(t) + \tilde{B}(t)z(t - \delta) + C(t)u(t) \\
+ \tilde{C}(t)u(t - \delta) \right] dt - Z^u(t)dW(t), \quad t \in [s, T], \\
Y^u(T) & = \xi, \quad Y^u(t) = \varphi(t), \quad Z^u(t) = \psi(t), \quad u(t) = \eta(t), \quad t \in [s - \delta, s),
\end{align*}
\]

then by Pardoux and Peng [21], \((Y^u(\cdot), Z^u(\cdot)) = T(y(\cdot), z(\cdot)) \in \mathcal{M}_\beta[s, T]\).

For any \((y_1(\cdot), z_1(\cdot), y_2(\cdot), z_2(\cdot)) \in \mathcal{M}_\beta[s, T]\), and \(y_1(t) = y_2(t) = \varphi(t), z_1(t) = z_2(t) = \psi(t), t \in [s - \delta, s)\), denote

\[
\begin{align*}
(Y^u_1(\cdot), Z^u_1(\cdot)) & = T(y_1(\cdot), z_1(\cdot)), \quad (Y^u_2(\cdot), Z^u_2(\cdot)) = T(y_2(\cdot), z_2(\cdot)), \\
\tilde{Y}(\cdot) & = Y^u_1(\cdot) - Y^u_2(\cdot), \quad \tilde{Z}(\cdot) = Z^u_1(\cdot) - Z^u_2(\cdot), \\
\hat{y}(\cdot) & = y_1(\cdot) - y_2(\cdot), \quad \hat{z}(\cdot) = z_1(\cdot) - z_2(\cdot).
\end{align*}
\]

Then we have

\[
\begin{align*}
-d\hat{Y}(t) & = \left[ A(t)\hat{y}(t) + \tilde{A}(t)\hat{y}(t - \delta) + B(t)\hat{z}(t) + \tilde{B}(t)\hat{z}(t - \delta) \right] dt \\
- \hat{Z}(t)dW(t), \quad t \in [s, T], \\
\hat{Y}(T) & = 0, \quad \hat{Y}(t) = 0, \quad \hat{Z}(t) = 0, \quad t \in [s - \delta, s).
\end{align*}
\]

Applying Itô’s formula to \(e^{\beta h(t)}|\hat{Y}(t)|^2\), we obtain

\[
\begin{align*}
-e^{\beta h(t)}|\hat{Y}(t)|^2 & = \int_t^T e^{\beta h(r)} \left[ \beta|\hat{Y}(r)|^2 (|A(r)| + |\tilde{A}(r)| + |B(r)|^2 + |\tilde{B}(r)|^2) \\
- 2|\hat{Y}(r), A(r)\hat{y}(r)| - 2|\hat{Y}(r), \tilde{A}(r)\hat{y}(r - \delta)| - 2|\hat{Y}(r), B(r)\hat{z}(r)| \\
- 2|\hat{Y}(r), \tilde{B}(r)\hat{z}(r - \delta)| + |\hat{Z}(r)|^2 \right] dr + 2 \int_t^T e^{\beta h(r)}|\hat{Y}(r), \hat{Z}(r)|dW(r).
\end{align*}
\]

Then we deduce

\[
\begin{align*}
e^{\beta h(t)}|\hat{Y}(t)|^2 & + \int_t^T \beta e^{\beta h(r)}|\hat{Y}(r)|^2 (|A(r)| + |\tilde{A}(r)| + |B(r)|^2 + |\tilde{B}(r)|^2) dr \\
+ \int_t^T e^{\beta h(r)}|\hat{Z}(r)|^2 dr + 2 \int_t^T e^{\beta h(r)}|\hat{Y}(r), \hat{Z}(r)|dW(r) \\
\leq & \int_t^T \beta e^{\beta h(r)}(|A(r)| + |\tilde{A}(r)| + |B(r)|^2 + |\tilde{B}(r)|^2)|\hat{Y}(r)|^2 dr \\
+ & \beta^{-1} \int_t^T e^{\beta h(r)} \left[ |A(r)||\hat{y}(r)|^2 + |\tilde{A}(r)||\hat{y}(r - \delta)|^2 + |\hat{z}(r)|^2 + |\hat{z}(r - \delta)|^2 \right] dr \\
\leq & \int_t^T \beta e^{\beta h(r)}(|A(r)| + |\tilde{A}(r)| + |B(r)|^2 + |\tilde{B}(r)|^2)|\hat{Y}(r)|^2 dr \\
+ & \beta^{-1} \int_t^T e^{\beta h(r)} \left[ |A(r)| + e^{\beta h(r + \delta - h(r))}|\tilde{A}(r + \delta)| \right]|\hat{y}(r)|^2 dr
\end{align*}
\]
\[
+ \beta^{-1} \int_{t-\delta}^{t} e^{\beta h(r+\delta)} |\hat{z}(r)|^2 dr + \beta^{-1} \int_{t-\delta}^{t} e^{\beta h(r+\delta)} |\hat{A}(r + \delta)||\hat{y}(r)|^2 dr \\
+ \beta^{-1} \int_{t}^{T} e^{\beta h(r)} (1 + e^{\beta h(r+\delta-h(r))}) |\hat{z}(r)|^2 dr.
\]

(2.5)

And thus we derive

\[
\mathbb{E} \int_{s}^{T} e^{\beta h(r)} |\hat{Z}(r)|^2 dr \\
\leq \beta^{-1} \mathbb{E} \int_{s}^{T} e^{\beta h(r)} (|A(r)| + e^{\beta [h(r+\delta)-h(r)]} |\hat{A}(r + \delta)||\hat{y}(r)|^2 dr \\
+ \beta^{-1} \mathbb{E} \int_{s}^{T} e^{\beta h(r)} (1 + e^{\beta [h(r+\delta)-h(r)]}) |\hat{z}(r)|^2 dr \\
\leq \beta^{-1} \left[ \sup_{s \leq r \leq T} e^{\beta h(r)} |\hat{y}(r)|^2 \int_{s}^{T} (|A(r)| + e^{\beta [h(r+\delta)-h(r)]} |\hat{A}(r + \delta)||\hat{y}(r)|^2 dr \right] \\
+ \beta^{-1} \left[ (1 + \sup_{s \leq r \leq T} e^{\beta [h(r+\delta)-h(r)]}) \mathbb{E} \int_{s}^{T} e^{\beta h(r)} |\hat{z}(r)|^2 dr \leq \tilde{\beta} \| (\hat{y}, \hat{z}) \|^2_{\mathcal{M}_\beta[s,T]}, \right.
\]

(2.6)

where \( \tilde{\beta} := \beta^{-1} \left[ 1 + \sup_{s \leq r \leq T} e^{\beta [h(r+\delta)-h(r)]} \right] \left[ (1 + \int_{s}^{T} (|A(r)| + |\hat{A}(r + \delta)||\hat{y}(r)|^2) dr \right]. \)

On the other hand, by Burkholder-Davis-Gundy’s inequality, we have

\[
\mathbb{E} \left[ \sup_{s \leq t \leq T} \left| \int_{t}^{T} e^{\beta h(r)} (|\hat{Y}(r), \hat{Z}(r)| dW(r)) \right| \right] \\
\leq C \mathbb{E} \left( \int_{s}^{T} e^{2\beta h(r)} |\hat{Y}(r)|^2 |\hat{Z}(r)|^2 dr \right)^{\frac{1}{2}} \\
\leq C \mathbb{E} \left[ \sup_{s \leq t \leq T} e^{\frac{1}{2} \beta h(r)} |\hat{Y}(r)| \left( \int_{s}^{T} e^{\beta h(r)} |\hat{Z}(r)|^2 dr \right)^{\frac{1}{2}} \right] \\
\leq \frac{1}{4} \mathbb{E} \left[ \sup_{s \leq t \leq T} e^{\beta h(r)} |\hat{Y}(r)|^2 \right] + C^2 \mathbb{E} \int_{s}^{T} e^{\beta h(r)} |\hat{Z}(r)|^2 dr,
\]

(2.7)

where \( C \) is a constant. Combining (2.5), (2.6) and (2.7), we obtain

\[
\mathbb{E} \left[ \sup_{s \leq t \leq T} e^{\beta h(t)} |\hat{Y}(t)|^2 \right] \\
\leq 2\tilde{\beta} \| (\hat{y}, \hat{z}) \|^2_{\mathcal{M}_\beta[s,T]} + 4C^2 \mathbb{E} \int_{s}^{T} e^{\beta h(r)} |\hat{Z}(r)|^2 dr + 2\beta^{-1} \mathbb{E} \left[ \sup_{s \leq t \leq T} \int_{t-\delta}^{t} e^{\beta h(r+\delta)} |\hat{A}(r + \delta)||\hat{y}(r)|^2 dr \right]
\]
\[ + 2\beta^{-1}E\left[ \sup_{s \leq t \leq T} \int_{t-\delta}^{t} e^{\delta h(r+\delta)}|\hat{Z}(r)|^2dr \right] \]
\[ \leq 6\beta||\langle \hat{Y}, \hat{Z}\rangle||_{\mathcal{M}_{\beta}[s,T]}^2 + 4C^2E\int_{s}^{T} e^{\delta h(r)}|\hat{Z}(r)|^2dr. \] 
(2.8)

Finally, by (2.6) and (2.8), we deduce
\[ ||\langle \hat{Y}, \hat{Z}\rangle||_{\mathcal{M}_{\beta}[s,T]} \leq \beta(7 + 4C^2)||\langle \hat{Y}, \hat{Z}\rangle||_{\mathcal{M}_{\beta}[s,T]}. \] 
(2.9)

Thus as long as \( \beta(7 + 4C^2) < 1 \), then \( T \) is a contraction mapping. Since \((A1)\) holds and \( \delta \) is sufficiently small, we can indeed choose some sufficiently large \( \beta \) such that (2.9) holds.

As for the estimate (2.4), applying Itô’s formula to \( |Y^u(\cdot)|^2 \), we have
\[ |Y^u(t)|^2 + \int_{t}^{T} |Z^u(r)|^2dr \]
\[ = |\xi|^2 + 2\int_{t}^{T} \langle Y^u(r), A(r)Y^u(r) + \tilde{A}(r)Y^u(r-\delta) + B(r)Z^u(r) + \tilde{B}(r)Z^u(r-\delta) + C(r)u(r) + \tilde{C}(r)u(r-\delta) \rangle dr - 2\int_{t}^{T} \langle Y^u(r), Z^u(r) \rangle dW(r). \] 
(2.10)

Hence it follows that
\[ \mathbb{E}\int_{s}^{T} |Z^u(r)|^2dr \leq K\mathbb{E}\left\{ |\xi|^2 + \int_{s-\delta}^{s} [||\varphi(r)||^2 + ||\psi(r)||^2 + ||\eta(r)||^2]dr + \int_{s}^{T} |u(r)|^2dr + \int_{s}^{T} |Y^u(r)|^2dr \right\}, \] 
(2.11)

where \( K > 0 \) is a generic constant. Combining (2.7), (2.10) and (2.11), we obtain
\[ \mathbb{E}\left[ \sup_{s \leq t \leq T} |Y^u(t)|^2 \right] \leq K\mathbb{E}\left\{ |\xi|^2 + \int_{s-\delta}^{s} [||\varphi(r)||^2 + ||\psi(r)||^2 + ||\eta(r)||^2]dr + \int_{s}^{T} |u(r)|^2dr + \int_{s}^{T} |Y^u(r)|^2dr + \int_{s}^{T} |Z^u(r)|^2dr \right\}. \] 
(2.12)

Finally using the Gronwall’s inequality twice and we complete the proof. \( \square \)

Remark 1 From Theorem 1, we can easily know that under \((A1)-(A2)\), for any \((s, \xi) \in [0, T] \times L^2_{T,T} (\mathbb{R}^n) \) and \( u(\cdot) \in U[s, T] \), the cost functional (2.2) is well-posed and hence Problem (D-BSLQ) makes sense.

Problem 2 From the proof of Theorem 1, the conditions imposed on the coefficients of (2.1) can be relaxed. For example, when \( A(\cdot), \tilde{A}(\cdot) \in L^1([0, T]; \mathbb{R}^{n \times n}), B(\cdot), \tilde{B}(\cdot) \in L^2([0, T]; \mathbb{R}^{n \times n}), C(\cdot), \tilde{C}(\cdot) \in L^2([0, T]; \mathbb{R}^{n \times d}) \), (2.1) is still well-posed.
In the last part of this section we introduce a new class of ASDDEs, and prove its well-posedness. For any given \( s \in [0, T] \), consider the following ASDDE:

\[
\begin{aligned}
    dX(t) &= b(t, X(t), X(t - \delta_1(t)), X(t + \delta_2(t)))dt \\
         &+ \sigma(t, X(t), X(t - \delta_1(t)), X(t + \delta_2(t)))dW(t), \quad t \in [s, T],
\end{aligned}
\]

(2.13)

where \( \delta_1(\cdot), \delta_2(\cdot) \) are \( \mathbb{R}^+ \)-valued continuous functions defined on \([s, T]\) and \( s - K_1 \leq t - \delta_1(t) \leq T, s \leq t + \delta_2(t) \leq K_2 + T \) for all \( t \in [s, T] \). Moreover, we suppose the following assumptions hold.

**H1** There exists a constant \( L > 0 \) such that

\[
\begin{aligned}
    \int_s^T g(t - \delta_1(t))dt &\leq L \int_{s-K_1}^T g(t)dt, \\
    \int_s^T g(t + \delta_2(t))dt &\leq L \int_s^{T+K_2} g(t)dt,
\end{aligned}
\]

for any nonnegative integrable function \( g(\cdot) \).

**H2** Let \( b(t, \omega, x, \theta, \vartheta) : [0, T] \times \Omega \times \mathbb{R}^n \times L^2_{\mathcal{F}_1} (\mathbb{R}^n) \times L^2_{\mathcal{F}_2} (\mathbb{R}^n) \to L^2_{\mathcal{F}_1} (\mathbb{R}^n) \)

and \( \sigma(t, \omega, x, \theta, \vartheta) : [0, T] \times \Omega \times \mathbb{R}^n \times L^2_{\mathcal{F}_1} (\mathbb{R}^n) \times L^2_{\mathcal{F}_2} (\mathbb{R}^n) \to L^2_{\mathcal{F}_1} (\mathbb{R}^n) \), where \( r_1 \in [t - K_1, T], r_2 \in [t, T + K_2] \), satisfy:

\[
\begin{aligned}
    |b(t, x, \theta_1, \vartheta_1) - b(t, \tilde{x}, \tilde{\theta}_1, \tilde{\vartheta}_1)| + |\sigma(t, x, \theta_1, \vartheta_1) - \sigma(t, \tilde{x}, \tilde{\theta}_1, \tilde{\vartheta}_1)| \\
    \leq L(|x - \tilde{x}| + |\theta_1 - \tilde{\theta}_1| + |\vartheta_1 - \tilde{\vartheta}_1|),
\end{aligned}
\]

for any \( t \in [s, T], x, \tilde{x} \in \mathbb{R}^n, \theta, \tilde{\theta} \in L^2_{\mathcal{F}_1}([t - K_1, T]; \mathbb{R}^n), \vartheta, \tilde{\vartheta} \in L^2_{\mathcal{F}_1}([t, T + K_2]; \mathbb{R}^n) \), and for the above constant \( L > 0 \);

**H3** \( \mathbb{E} \int_s^T (|b(t, 0, 0, 0)|^2 + |\sigma(t, 0, 0, 0)|^2)dt < +\infty \).

**H4** The initial and terminal conditions \( \alpha_1(\cdot) \in L^2_{\mathcal{F}_1}([s - K_1, s]; \mathbb{R}^n), \alpha_2(\cdot) \in L^2_{\mathcal{F}_1}([T, T + K_2]; \mathbb{R}^n) \).

**Remark 3** The fact that \( b(\cdot, \ldots), \sigma(\cdot, \ldots) \) are \( \mathcal{F}_t \)-measurable guarantees the adaptability of the solution to the ASDDE (2.13).

Then we have the following result.

**Theorem 2** Let (H1)–(H4) hold, suppose \( K_2 > 0 \) is sufficiently small, then for any \( \zeta \in L^2_{\mathcal{F}_1} (\mathbb{R}^n) \), the ASDDE (2.13) has a unique solution \( X(\cdot) \in L^2_{\mathcal{F}_1}(\Omega; C([s, T]; \mathbb{R}^n)) \).

**Proof** For any \( \beta > 0 \), define

\[
\mathcal{H}_\beta[s, T] := L^2_{\mathcal{F}_1}([s, T]; \mathbb{R}^n),
\]
equipped with the norm
\[ \|X\|_{H_\beta[s,T]} := \left( \mathbb{E} \int_s^T |X(t)|^2 e^{-\beta t} \, dt \right)^{\frac{1}{2}}. \]

Let the mapping \( T : H_\beta[s,T] \to H_\beta[s,T] \), \( x(\cdot) \mapsto X(\cdot) \), apparently it is well-defined. In fact, for any \( x(\cdot) \in H_\beta[s,T] \) and \( x(t) = \alpha_1(t), \ s - K_1 \leq t < s, \ x(t) = \alpha_2(t), \ T < t \leq T + K_2 \), consider the following SDE:

\[
\begin{aligned}
&dX(t) = b(t, X(t), x(t - \delta_1(t)), x(t + \delta_2(t)))dt \\
&+ \sigma(t, X(t), x(t - \delta_1(t)), x(t + \delta_2(t)))dW(t), \quad t \in [s, T], \\
&X(s) = \zeta, \ X(t) = \alpha_1(t), \quad t \in [s - K_1, s), \quad X(t) = \alpha_2(t), \quad t \in (T, T + K_2].
\end{aligned}
\]

Then \( X(\cdot) = T(x(\cdot)) \in H_\beta[s,T] \) by the existence and uniqueness result of the solution to SDE in Yong and Zhou [34]. Next for any \( x_1(\cdot), x_2(\cdot) \in H_\beta[s,T] \), and \( x_1(t) = x_2(t) = \alpha_1(t), \ t \in [s - K_1, s), \ x_1(t) = x_2(t) = \alpha_2(t), \ t \in (T, T + K_2], \) denote

\[
(X_1(\cdot), X_2(\cdot)) = (T(x_1(\cdot)), T(x_2(\cdot))), \quad \hat{X}(\cdot) = X_1(\cdot) - X_2(\cdot), \quad \hat{x}(\cdot) = x_1(\cdot) - x_2(\cdot).
\]

Then we obtain

\[
\begin{aligned}
d\hat{X}(t) &= \left[ b(t, X_1(t), x_1(t - \delta_1(t)), x_1(t + \delta_2(t))) \\
&- b(t, X_2(t), x_2(t - \delta_1(t)), x_2(t + \delta_2(t))) \right]dt \\
&+ \left[ \sigma(t, X_1(t), x_1(t - \delta_1(t)), x_1(t + \delta_2(t))) \\
&- \sigma(t, X_2(t), x_2(t - \delta_1(t)), x_2(t + \delta_2(t))) \right]dW(t), \quad t \in [s, T], \\
\hat{X}(s) &= 0, \quad \hat{X}(t) = 0, \quad t \in [s - K_1, s) \cup (T, T + K_2].
\end{aligned}
\]

Applying Itô’s formula to \( e^{-\beta r} |\hat{X}(r)|^2 \), we get

\[
e^{-\beta r} |\hat{X}(r)|^2 \\
= \int_s^r e^{-\beta t} \left[ -\beta |\hat{X}(t)|^2 + 2\langle \hat{X}(t), b(t, X_1(t), x_1(t - \delta_1(t)), x_1(t + \delta_2(t))) \rangle \\
- b(t, X_2(t), x_2(t - \delta_1(t)), x_2(t + \delta_2(t))) + \sigma(t, X_1(t), x_1(t - \delta_1(t)), x_1(t + \delta_2(t))) \rangle \\
- \sigma(t, X_2(t), x_2(t - \delta_1(t)), x_2(t + \delta_2(t))) \rangle \\
+ \sigma(t, X_2(t), x_2(t - \delta_1(t)), x_2(t + \delta_2(t))) \rangle \\
- \sigma(t, X_2(t), x_2(t - \delta_1(t)), x_2(t + \delta_2(t))) \rangle \\
+ 2 \int_s^r e^{-\beta t} \left[ \langle \hat{X}(t), \sigma(t, X_1(t), x_1(t - \delta_1(t)), x_1(t + \delta_2(t))) \rangle \\
- \sigma(t, X_2(t), x_2(t - \delta_1(t)), x_2(t + \delta_2(t))) \rangle \\
+ \sigma(t, X_2(t), x_2(t - \delta_1(t)), x_2(t + \delta_2(t))) \rangle \\
- \sigma(t, X_2(t), x_2(t - \delta_1(t)), x_2(t + \delta_2(t))) \rangle \\
+ \sigma(t, X_2(t), x_2(t - \delta_1(t)), x_2(t + \delta_2(t))) \rangle \\
- \sigma(t, X_2(t), x_2(t - \delta_1(t)), x_2(t + \delta_2(t))) \rangle \\
+ \sigma(t, X_2(t), x_2(t - \delta_1(t)), x_2(t + \delta_2(t))) \rangle \\
- \sigma(t, X_2(t), x_2(t - \delta_1(t)), x_2(t + \delta_2(t))) \rangle \right]dW(t).
\]
Hence by (H2), we obtain

\[
\mathbb{E} \int_s^T \beta e^{-\beta t} |\dot{X}(t)|^2 dt \\
\leq \mathbb{E} \int_s^T e^{-\beta t} \left[ 2L|\dot{X}(t)|^2 + 2L|\dot{X}(t)||\dot{x}(t - \delta_1(t))| \\
+ 2L|\dot{X}(t)||\dot{x}(t + \delta_2(t))| + 3L^2|\dot{X}(t)|^2 \\
+ 3L^2|\dot{x}(t - \delta_1(t))|^2 + 3L^2|\dot{x}(t + \delta_2(t))|^2 \right] dt \\
\leq \mathbb{E} \int_s^T e^{-\beta t} \left[ (2L + 5L^2)|\dot{X}(t)|^2 + (3L^2 + 1)|\dot{x}(t - \delta_1(t))|^2 \\
+ (3L^2 + 1)|\dot{x}(t + \delta_2(t))|^2 \right] dt.
\]

Noting (H1), we deduce

\[
\mathbb{E} \int_s^T e^{-\beta t} |\dot{X}(t)|^2 dt \\
\leq (3L^2 + 1)(\beta - 2L - 5L^2)^{-1} \mathbb{E} \int_s^T e^{-\beta t} \left[ |\dot{x}(t - \delta_1(t))|^2 + |\dot{x}(t + \delta_2(t))|^2 \right] dt \\
\leq L(3L^2 + 1)(1 + e^{\beta K_2})(\beta - 2L - 5L^2)^{-1} \mathbb{E} \int_s^T e^{-\beta t} |\dot{x}(t)|^2 dt.
\]

Since \(K_2\) is sufficiently small, we can choose some sufficiently large \(\beta\) such that
\(L(3L^2 + 1)(1 + e^{\beta K_2})(\beta - 2L - 5L^2)^{-1} < 1\), then \(T\) is a contraction mapping. Thus ASDDE (2.13) has a unique solution \(X(\cdot) \in L^2_F([s, T]; \mathbb{R}^n)\).

Next we prove that \(X(\cdot) \in L^2_F(\Omega; C([s, T]; \mathbb{R}^n))\). In fact, we have

\[
\mathbb{E} \left[ \sup_{s \leq t \leq T} |X(t)|^2 \right] \\
\leq 3\mathbb{E} |\xi|^2 + 3T \mathbb{E} \int_s^T |b(t, X(t), X(t - \delta_1(t)), X(t + \delta_2(t)))|^2 dt \\
+ 3C \mathbb{E} \int_s^T |\sigma(t, X(t), X(t - \delta_1(t)), X(t + \delta_2(t)))|^2 dt \\
\leq 3\mathbb{E} |\xi|^2 + 12T \mathbb{E} \int_s^T |b(t, 0, 0, 0)|^2 dt + 12C \mathbb{E} \int_s^T |\sigma(t, 0, 0, 0)|^2 dt \\
+ 12L^2(T + C) \mathbb{E} \int_s^T |X(t)|^2 dt + 12L^2(T + C) \mathbb{E} \int_s^T |X(t - \delta_1(t))|^2 dt \\
+ 12L^2(T + C) \mathbb{E} \int_s^T |X(t + \delta_2(t))|^2 dt \\
\leq 3\mathbb{E} |\xi|^2 + 12T \mathbb{E} \int_s^T |b(t, 0, 0, 0)|^2 dt + 12C \mathbb{E} \int_s^T |\sigma(t, 0, 0, 0)|^2 dt
\]
+ 12L^2(T + C)(1 + 2L)\mathbb{E} \int_s^T |X(t)|^2 dt + 12L^2(T + C)\mathbb{E} \int_s^{s-K_1} |\alpha_1(t)|^2 dt
+ 12L^2(T + C)\mathbb{E} \int_T^{T+K_2} |\alpha_2(t)|^2 dt < \infty, \quad (2.14)

where \( C > 0 \) is a constant. Hence we complete the proof. \( \square \)

**Remark 4** From the above proof, it shows that the assumption, \( K_2 > 0 \) is sufficiently small, is to guarantee \( L(3L^2 + 1)(1 + e^{\beta K_2})(\beta - 2L - 5L^2)^{-1} < 1 \), thus \( T \) is a contraction mapping. In fact, let \( \beta = 2L + 5L^2 + 4L(3L^2 + 1) + 1 \), and let \( K_2 > 0 \) satisfy that (i) \( e^{\beta K_2} \approx \beta K_2 + 1 \), (ii) \( \frac{1}{4} \beta K_2 < \frac{1}{3} \), then \( L(3L^2 + 1)(1 + e^{\beta K_2})(\beta - 2L - 5L^2)^{-1} < 1 \).

**Remark 5** In the above proof, we have used the assumption \((H1)\) many times for \( g(t) = e^{-\beta t}|\hat{x}(t)|^2 \) and \( g(t) = |X(t)|^2 \).

**Remark 6** In fact, ASDDEs are a generalization of ASDEs and SDDEs. Let \( \delta_1(\cdot) \equiv 0 \), then ASDDE \((2.13)\) becomes the ASDE. While let \( \delta_2(\cdot) \equiv 0 \), the ASDDE \((2.13)\) becomes the SDDE.

## 3 Necessary Conditions of Optimal Controls

In this section, we present the main results of this paper, which give the necessary condition for optimal controls of the above **Problem (D-BSLQ)**. The detailed proofs of the main theorems are deferred to Sect. 5.

First we introduce the following delayed Riccati equation

\[
\begin{aligned}
\dot{\Sigma}(t) &= -\Sigma(t)A(t)^\top - A(t)\Sigma(t) + 2\Sigma(t)(Q(t) + \tilde{Q}(t + \delta))\Sigma(t) \\
&\quad - B(t)\Sigma(t)\mathcal{M}^{-1}(t)B(t)^\top - C(t)\mathcal{N}^{-1}(t)C(t)^\top \\
&\quad - \tilde{B}(t)\Sigma(t - \delta)\mathcal{M}^{-1}(t - \delta)\tilde{B}(t)^\top - \tilde{C}(t)\mathcal{N}^{-1}(t - \delta)\tilde{C}(t)^\top, \quad t \in [s, T],
\end{aligned}
\]

\[
\Sigma(T) = 0, \quad \Sigma(t) = I, \quad t \in [s - \delta, s), \quad (3.1)
\]

where

\[
\mathcal{M}(t) := 2R(t)\Sigma(t) + 2\tilde{R}(t + \delta)\Sigma(t) + I,
\]

\[
\mathcal{N}(t) := 2N(t) + 2\tilde{N}(t + \delta). \quad (3.2)
\]

The existence and uniqueness of the solution to \((3.1)\) will be discussed in the next section. Let \( \Sigma(\cdot) \) be the solution to \((3.1)\), consider the following equations:

\[
\begin{aligned}
\dot{L}(t) &= L(t)A(t) + A(t)^\top L(t) + 2\left[ Q(t) + \tilde{Q}(t + \delta) \right] \\
&\quad - L(t)B(t)\Sigma(t)\mathcal{M}^{-1}(t)B(t)^\top L(t) - L(t)\tilde{C}(t)\mathcal{N}^{-1}(t)C(t)^\top L(t) \\
&\quad - L(t)\tilde{B}(t)\Sigma(t - \delta)\mathcal{M}^{-1}(t - \delta)\tilde{B}(t)^\top L(t) \\
&\quad - L(t)\tilde{C}(t)\mathcal{N}^{-1}(t - \delta)\tilde{C}(t)^\top L(t), \quad t \in [s, T],
\end{aligned}
\]

\[
L(s) = 2G, \quad (3.3)
\]
Apparently (3.3) is a Riccati equation without delay, (3.4) is a linear DAFBSDE, and their solvability will be addressed in the next section.

Based on (3.1), (3.3) and (3.4), we finally introduce the following ASDDE:

$$\begin{aligned}
    &d\tilde{X}(t) = \left[ A(t)^\top - 2(Q(t) + \tilde{Q}(t + \delta))\Sigma(t) \right] \tilde{X}(t) \\
    &+ 2Q(t) + \tilde{Q}(t + \delta) + \Lambda(t) + E_{\mathcal{F}_t}\left[ (\tilde{A}^\top \tilde{X})|_{t+\delta} \right]dt \\
    &+ \left[ B(t)^\top - 2(R(t) + \tilde{R}(t + \delta))\Sigma(t)M^{-1}(t)B(t)^\top \right] \tilde{X}(t) \\
    &+ E_{\mathcal{F}_t}\left[ (\tilde{B}(t + \delta)^\top - 2(R(t) + \tilde{R}(t + \delta))\Sigma(t)M^{-1}(t) \times \tilde{B}(t + \delta)^\top ) \tilde{X}(t + \delta) \right] + 2[R(t) + \tilde{R}(t + \delta)] \\
    &\times \left[ 2\Sigma(t)(R(t) + \tilde{R}(t + \delta)) + I \right]^{-1} \Gamma(t)\right]dW(t),
\end{aligned}$$

(3.4)

where

$$\begin{aligned}
    &d\Lambda(t) = \left\{ 2\Sigma(t)[Q(t) + \tilde{Q}(t + \delta)] - A(t) \right\} \Lambda(t) \\
    &- B(t)\left[ 2\Sigma(t)[R(t) + \tilde{R}(t + \delta)] + I \right]^{-1} \Gamma(t) - \tilde{A}(t)\Lambda(t - \delta) \\
    &- \tilde{B}(t)\left[ 2\Sigma(t - \delta)[R(t - \delta) + \tilde{R}(t)] + I \right]^{-1} \Gamma(t - \delta) \\
    &+ \left[ \tilde{B}(t)\Sigma(t - \delta)M^{-1}(t - \delta)\tilde{B}(t)^\top + \tilde{C}(t)N^{-1}(t - \delta)\tilde{C}(t)^\top \right] \\
    &\times \left[ E_{\mathcal{F}_{t-\delta}}[\tilde{X}(t) - \tilde{X}(t)] \right] dt + \Gamma(t)dW(t), \quad t \in [s, T],
\end{aligned}$$

$$\tilde{X}(s) = 2G(I + 2\Sigma(s)G)^{-1} \Lambda(s), \quad \tilde{X}(t) = 0, \quad t \in (T, T + \delta),$$

$$\Lambda(T) = -\xi, \quad \Lambda(t) = 0, \quad \Gamma(t) = 0, \quad t \in [s - \delta, s).$$

$$\begin{aligned}
    &dS(t) = S_1(t)dt + S_2(t)dW(t), \quad t \in [s, T], \\
    &S(s) = 0, \quad S(t) = 0, \quad t \in [s - \delta, s) \cup (T, T + \delta),
\end{aligned}$$

(3.5)
Theorem 4 Assume that \( t \in [s, T] \), suppose \( u^* (\cdot) \) is an optimal control and \( (Y^* (\cdot), Z^* (\cdot)) \) is the corresponding optimal state trajectory. Then the following ASDE admits a unique solution \( X^* (\cdot) \in L^2_F (\Omega; C([s, T]; \mathbb{R}^n)) \):

\[
\begin{align*}
&dX^* (t) = \left( A(t)^T X^* (t) - 2[Q(t) + \tilde{Q}(t + \delta)]Y^* (t) + \mathbb{E}^F_t \left[ (\tilde{A}^T X^*)_{|t+\delta} \right] \right) dt \\
&\quad + \left( B(t)^T X^* (t) - 2[R(t) + \tilde{R}(t + \delta)]Z^* (t) + \mathbb{E}^F_t \left[ (\tilde{B}^T X^*)_{|t+\delta} \right] \right) dW(t), \\
&X^* (s) = -2GY^* (s), \quad X^* (t) = 0, \quad t \in (T, T + \delta],
\end{align*}
\]

and the optimal control can be expressed as

\[
u^* (t) = \mathcal{N}^{-1} (t) \left\{ C(t)^T X^* (t) + \mathbb{E}^F_t \left[ (\tilde{C}^T X^* )_{|t+\delta} \right] \right\}, \quad a.e. t \in [s, T], \ P\text{-a.s.} \ (3.7)
\]

Theorem 3 Let \((A1)-(A3)\) hold and \( \delta \) be sufficiently small. Assume \( \tilde{Q}(t) = \tilde{R}(t) = \tilde{N}(t) = 0 \) for \( t \in [T, T + \delta] \), suppose \( u^* (\cdot) \) is an optimal control and \( (Y^* (\cdot), Z^* (\cdot)) \) is the corresponding optimal state trajectory. Then the following ASDE admits a unique solution \( X^* (\cdot) \in L^2_F (\Omega; C([s, T]; \mathbb{R}^n)) \):

\[
\begin{align*}
&dX^* (t) = \left( A(t)^T X^* (t) - 2[Q(t) + \tilde{Q}(t + \delta)]Y^* (t) + \mathbb{E}^F_t \left[ (\tilde{A}^T X^*)_{|t+\delta} \right] \right) dt \\
&\quad + \left( B(t)^T X^* (t) - 2[R(t) + \tilde{R}(t + \delta)]Z^* (t) + \mathbb{E}^F_t \left[ (\tilde{B}^T X^*)_{|t+\delta} \right] \right) dW(t), \\
&X^* (s) = -2GY^* (s), \quad X^* (t) = 0, \quad t \in (T, T + \delta],
\end{align*}
\]

and the optimal control can be expressed as

\[
u^* (t) = \mathcal{N}^{-1} (t) \left\{ C(t)^T X^* (t) + \mathbb{E}^F_t \left[ (\tilde{C}^T X^* )_{|t+\delta} \right] \right\}, \quad a.e. t \in [s, T], \ P\text{-a.s.} \ (3.7)
\]
Then the optimal control can be expressed as

\[ u^*(t) = \mathcal{N}^{-1}(t) \left\{ -C(t)^T L(t) Y^*(t) + C(t)^T S(t) + \mathbb{E}^{\mathcal{F}_t} \left[ (-\bar{C}^T L Y^* + \bar{C}^T S)_{t+\delta} \right] \right\} \]

\( a.e. \ t \in [s, T], \ \mathbb{P}\text{-a.s.} \)

Moreover, the optimal cost is

\[ J(u^*(\cdot)) = \mathbb{E}\left\{ \bar{G} \varphi(s - \delta), \varphi(s - \delta) \right\} + \left\{ \Lambda(s), (I + 2\Sigma(s)G)^{-1}G\Lambda(s) \right\} \]

\[ + \int_{s-\delta}^{s} \left[ [\bar{Q}(t + \delta) \varphi(t), \varphi(t)] + [\bar{R}(t + \delta) \psi(t), \psi(t)] + [\bar{N}(t + \delta) \eta(t), \eta(t)] \right] dt \]

\[ + \int_{s}^{T} \left[ (\bar{R}(t) + \tilde{R}(t + \delta))[2\Sigma(t)R(t) + \bar{R}(t + \delta)] + I \right]^{-1} \Gamma(t), \Gamma(t) \]

\[ + \left[ (\bar{Q}(t) + \tilde{Q}(t + \delta))\Lambda(t), \Lambda(t) \right] + \left[ \bar{X}(t), \left[ \bar{B}(t)\Sigma(t - \delta)M^{-1}(t - \delta)\bar{B}(t)^T + \bar{C}(t)N^{-1}(t - \delta)\bar{C}(t)^T \right][\mathbb{E}^{\mathcal{F}_{t-\delta}}[\bar{X}(t)] - \bar{X}(t)] \right] dt \].

(3.10)

The proofs of Theorems 3 and 4 are deferred to Sect. 5. For the convenience of reading, we state the following corollary for the existence and uniqueness of the solutions to (3.1)–(3.5).

**Corollary 1** Let \( \delta \) be sufficiently small, suppose one of the following two conditions holds:

(i) \( \bar{B}(\cdot) \equiv 0; \)

(ii) \( \bar{B}(\cdot) = 0 \) does not hold, \( B(\cdot) \equiv I, R(\cdot) + \tilde{R}(\cdot + \delta) \equiv 0, (A4) \) holds in Sect. 4.

Then (3.1) admits a unique solution \( \Sigma(\cdot) \in C([s, T]; \mathcal{S}_{n+}^n) \), and hence (3.3) has the unique solution \( L(\cdot) \in C([s, T]; \mathcal{S}_{n+}^n) \).

Furthermore, suppose one of the following two conditions holds:

(iii) The above assumption (i) holds, \( \sup_{s \leq t \leq T} |\bar{C}(t)N^{-1}(t - \delta)\bar{C}(t)^T| \) is sufficiently small;

(iv) The above assumption (ii) holds, \( \sup_{s \leq t \leq T} |\bar{Q}(t) + \bar{Q}(t + \delta) + |G| \) is sufficiently small.

Then (3.4) admits a unique solution \( (\bar{X}(\cdot), \Lambda(\cdot), \Gamma(\cdot)) \in L_{L^2}^2(\Omega; C([s, T]; R^n)) \times L_{\mathcal{F}}^2(\Omega; C([s, T]; R^n)) \times L_{\mathcal{F}}^2([s, T]; R^n) \) and thus (3.5) has the unique solution \( S(\cdot) \in L_{\mathcal{F}}^2(\Omega; C([s, T]; R^n)) \).

**Proof** Suppose (i) hold, then (3.1) is the one without delay, and it becomes the same type of Riccati equations as (3.4) in Lim and Zhou [15] and hence is uniquely solvable. Suppose (ii) holds, by Corollary 2, (3.1) admits a unique solution \( \Sigma(\cdot) \in C([s, T]; \mathcal{S}_{n+}^n) \). Finally, the proof is completed by Proposition 4, Corollary 3 and Theorem 2. \( \square \)
Remark 7 Noting when $\bar{A}(\cdot), \bar{B}(\cdot), \bar{C}(\cdot), \bar{G}(\cdot), \bar{Q}(\cdot), \bar{R}(\cdot), \bar{N}(\cdot) \equiv 0$, Theorem 3 is reduced to Theorem 3.2 in Lim and Zhou [15], for the problem without delay. In this case, the optimal control is a linear state feedback of the entire past history of the state process $(Y^*(\cdot), Z^*(\cdot))$. 

Remark 8 From (3.9), the optimal control $u^*(t)$ explicitly depends on not only the current state $Y^*(t)$ but also the future state $Y^*(t+\delta)$, hence it is a linear feedback of the entire past history and the expected value of the future state trajectory in a short future period of time $\delta$. Noting if $\bar{C}(\cdot) \equiv 0$, then the optimal control $u^*(t)$ only depends on the current state $Y^*(t)$. That is, if the dynamics (2.1) does not depend on $u(t-\delta)$, then the optimal control does not depend on the future trajectory of the controlled stated process, which have already been identified in the literature, see [9], [10], where controlled SDDEs with delayed control are investigated and optimal controls depending on the future state process are derived.

Remark 9 The equality (3.8) plays an important role in deriving the existence and uniqueness of the solution to the stochastic Hamiltonian system (5.3) (see Proposition 5 in Sect. 5). Although it seems a little complex, it is not harsh. For example, considering the one-dimensional case, let $R(\cdot) + \bar{R}(\cdot) = 0$, $C(\cdot) = 0$ or $\bar{C}(\cdot) = 0$, and $\bar{A}(\cdot + \delta) + \bar{B}(\cdot + \delta)B(\cdot) = 0$, then (3.8) holds.

Remark 10 Let $\bar{C} \equiv Q \equiv \bar{Q} = R \equiv \bar{R} \equiv \bar{N} = 0$, then Problem (D-BSLQ) deduces to Application III in [1], however where the optimal feedback regulator is not discussed. Similarly let $B \equiv \bar{B} \equiv \bar{C} \equiv \bar{Q} \equiv \bar{R} \equiv \bar{N} \equiv 0$, then Problem (D-BSLQ) deduces to the LQ problem in Section 4 in [32], however while studying the optimal feedback regulator, there is the restrictive condition $\bar{A} \equiv 0$ imposed, in other words, time delay is not allowed to appear in the control system. Thus, our results in this paper enrich the theory of LQ optimal control problems with delay.

4 Unique Solvability of Delayed Riccati Equations and DAFBSDEs

In this section, we will study the existence and uniqueness of solutions to (3.1)-(3.4). It is clear that when $\bar{B}(\cdot) \equiv 0$, (3.1) becomes the same type of Riccati equation as (3.4) in Lim and Zhou [15] and is uniquely solvable. Hence in this section, we only discuss the case where $\bar{B}(\cdot) \neq 0$.

First we consider the following two equations:

\[
\begin{align*}
\dot{\Sigma}(t) &= -\Sigma(t) A(t)^T - A(t) \Sigma(t) + 2\Sigma(t) \left[ Q(t) + \bar{Q}(t + \delta) \right] \Sigma(t) \\
&\quad - B(t) \Sigma(t) M^{-1}(t) B(t)^T - C(t) N^{-1}(t) C(t)^T \\
&\quad - \bar{B}(t) \Sigma(t - \delta) M^{-1}(t - \delta) \bar{B}(t)^T - \bar{C}(t) N^{-1}(t - \delta) \bar{C}(t)^T, \quad t \in [s, T], \\
\Sigma(T) &= M, \quad \Sigma(t) = I, \quad t \in [s - \delta, s),
\end{align*}
\]

\[(4.1)\]

\[\Sigma(t) \quad t \in [s - \delta, s),\]

\[\Sigma(T) = M, \quad \Sigma(t) = I, \quad t \in [s - \delta, s),\]

\[(4.1)\]
\[
\dot{P}(t) = P(t)A(t) + A(t)\top P(t) - 2\left[Q(t) + \tilde{Q}(t + \delta)\right]
\]
\[
+ P(t)B(t)\left[2R(t) + 2\tilde{R}(t + \delta) + P(t)\right]^{-1}B(t)\top P(t)
\]
\[
+ P(t)C(t)N^{-1}(t)C(t)\top P(t)
\]
\[
+ P(t)\tilde{B}(t)\left[2R(t - \delta) + 2\tilde{R}(t) + P(t - \delta)\right]^{-1}\tilde{B}(t)\top P(t)
\]
\[
P(T) = M^{-1}, \quad P(t) = I, \quad t \in [s - \delta, s),
\]

where \(M(\cdot), N(\cdot)\) are defined in (3.2), and \(M\) is a given \(n \times n\) symmetric matrix. Apparently, if \(M = 0\), then (4.1) reduces to (3.1).

In the following context, we will suppress some time variables \(t\) for simplicity of writing without ambiguity.

**Proposition 1** If the solution \(\Sigma(\cdot)\) to the delayed Riccati equation (4.1) satisfies \(\Sigma(\cdot) \in C([s, T]; S^n_+)\), then the solution is unique.

**Proof** Suppose \(\Sigma_1(\cdot), \Sigma_2(\cdot) \in C([s, T]; S^n_+)\) are two solutions to (4.1). Denote \(\Delta(\cdot) := \Sigma_1(\cdot) - \Sigma_2(\cdot)\), then we have

\[
\dot{\Delta} = \begin{bmatrix}
- A + 2\Sigma_1(Q + \tilde{Q}_{r + \delta}) & \Delta & - A + 2\Sigma_1(Q + \tilde{Q}_{r + \delta}) \\
- 2\Delta(Q + \tilde{Q}_{r + \delta}) & - B\Delta[2(R + \tilde{R}_{r + \delta})\Sigma_1 + I]^{-1}B\top & + B\Sigma_2[2(R + \tilde{R}_{r + \delta})\Sigma_2 + I]^{-1}(2R + 2\tilde{R}_{r + \delta})\Delta[2(R + \tilde{R}_{r + \delta})\Sigma_1 + I]^{-1}B\top \\
- B\Delta[2(R + \tilde{R}_{r + \delta})\Sigma_1 + I]^{-1}\tilde{B}\top + \check{B}\Sigma_2[2(R + \tilde{R}_{r + \delta})\Sigma_2 + I]^{-1}(2R + 2\tilde{R}_{r + \delta})\Delta[2(R + \tilde{R}_{r + \delta})\Sigma_1 + I]^{-1}\tilde{B}\top & \times (2R + 2\tilde{R})\Delta[2(R + \tilde{R})\Sigma_1 + I]^{-1}\tilde{B}\top & \Delta(T) = 0, \quad \Delta(t) = 0, \quad t \in [s - \delta, s).
\end{bmatrix}
\]

For any \(\beta \in \mathbb{R}\), differentiating \(e^{\beta t}|\Delta(t)|^2\), we have

\[
0 = e^{\beta t}|\Delta(t)|^2 + \int_t^T e^{\beta r} \left\{ \beta|\Delta|^2 + 2\Delta, \left[ - A + 2\Sigma_1(Q + \tilde{Q}_{r + \delta})\right] \Delta \right. \\
+ \Delta, \left[ - A + 2\Sigma_1(Q + \tilde{Q}_{r + \delta})\right] \top - 2\Delta(Q + \tilde{Q}_{r + \delta}) \Delta \\
- B\Delta[2(R + \tilde{R}_{r + \delta})\Sigma_1 + I]^{-1}B\top \\
+ B\Sigma_2[2(R + \tilde{R}_{r + \delta})\Sigma_2 + I]^{-1}(2R + 2\tilde{R}_{r + \delta})\Delta[2(R + \tilde{R}_{r + \delta})\Sigma_1 + I]^{-1}B\top \\
- \check{B}\Delta[2(R + \tilde{R}_{r + \delta})\Sigma_1 + I]^{-1}\tilde{B}\top + \check{B}\Sigma_2[2(R + \tilde{R}_{r + \delta})\Sigma_2 + I]^{-1}(2R + 2\tilde{R}_{r + \delta})\Delta[2(R + \tilde{R}_{r + \delta})\Sigma_1 + I]^{-1}\tilde{B}\top \\
\times (2R + 2\tilde{R})\Delta[2(R + \tilde{R})\Sigma_1 + I]^{-1}\tilde{B}\top \right\} dr.
\]
Hence we obtain
\[
\beta \int_s^T e^{\beta r} |\Delta(r)|^2 dr \leq 2 \int_s^T e^{\beta r} |\Delta(r)||\dot{\Delta}(r)| dr
\]
\[
\leq \sup_{s \leq r \leq T} \left\{ 4|A| + 8|\Sigma_1|(Q + \tilde{Q}_{r+\delta}) + 4|\Delta|(Q + \tilde{Q}_{r+\delta}) + 2|B|^2 \right\}
\]
\[
+ 4|B|^2 |\Sigma_2|(R + \tilde{R}_{r+\delta}) \right\}
\]
\[
\leq \int_s^T e^{\beta r} |\Delta(r)|^2 dr + \sup_{s \leq r \leq T} \left\{ |\tilde{B}|^2 + 2|\tilde{B}|^2 |\Sigma_2| (R + \tilde{R}_{r+\delta}) \right\} \left\{ \int_s^T e^{\beta r} |\Delta(r)|^2 dr + \int_s^T e^{\beta r} |\Delta(r - \delta)|^2 dr \right\}.
\]

Noting
\[
\int_s^T e^{\beta r} |\Delta(r - \delta)|^2 dr = e^{\beta \delta} \int_{s-\delta}^{T-\delta} e^{\beta r} |\Delta(r)|^2 dr \leq e^{\beta \delta} \int_s^T e^{\beta r} |\Delta(r)|^2 dr.
\]

Thus we derive
\[
\beta \int_s^T e^{\beta r} |\Delta(r)|^2 dr
\]
\[
\leq \left[ \sup_{s \leq r \leq T} \left\{ 4|A| + 8|\Sigma_1|(Q + \tilde{Q}_{r+\delta}) + 4|\Delta|(Q + \tilde{Q}_{r+\delta}) + 2|B|^2 \right\} \right]
\]
\[
+ \sup_{s \leq r \leq T} \left\{ |\tilde{B}|^2 + 2|\tilde{B}|^2 |\Sigma_2| (R + \tilde{R}_{r+\delta}) \right\} (1 + e^{\beta \delta}) \right\}
\]
\[
\leq \int_s^T e^{\beta r} |\Delta(r)|^2 dr.
\]

Since $\Sigma_1(\cdot)$, $\Sigma_2(\cdot)$ are continuous on $[s, T]$, $\Delta(\cdot)$ are uniformly bounded. Recall (A1), (A3) and $\delta$ is sufficiently small, we can choose $\beta$ sufficiently large such that
\[
\beta > \sup_{s \leq r \leq T} \left\{ 4|A| + 8|\Sigma_1|(Q + \tilde{Q}_{r+\delta}) + 4|\Delta|(Q + \tilde{Q}_{r+\delta}) + 2|B|^2 \right\}
\]
\[
+ \sup_{s \leq r \leq T} \left\{ |\tilde{B}|^2 + 2|\tilde{B}|^2 |\Sigma_2| (R + \tilde{R}_{r+\delta}) \right\} (1 + e^{\beta \delta}),
\]

then we deduce
\[
\int_s^T e^{\beta r} |\Delta(r)|^2 dr = 0.
\]

And it follows that
\[
\Delta(t) = 0, \quad a.e. \ t \in [s, T]. \quad (4.4)
\]

Finally, by (4.3) and (4.4), we get
\[
e^{\beta t} |\Delta(t)|^2 = 0,
\]
thus \( \Delta(t) = 0 \) for all \( s \leq t \leq T \). The proof is complete. \( \square \)

If we can prove the existence of solutions to the delayed Riccati equation (4.1), then we can obtain the unique solvability of it. However, this is an arduous work and we can only deal with the special case so far in this paper.

Let \( R(\cdot) + \tilde{R}(\cdot + \delta) \equiv 0 \), now (4.1) and (4.2) become:

\[
\begin{aligned}
\dot{\Sigma} &= -\Sigma A^T - A \Sigma + 2 \Sigma (Q + \tilde{Q}|_{t+\delta}) \Sigma - B \Sigma B^T - C \Sigma^{-1} C^T \\
&- \bar{B} \Sigma|_{t-\delta} \bar{B}^T - \tilde{C} \Sigma^{-1}|_{t-\delta} \tilde{C}^T, \quad t \in [s, T], \\
\Sigma(T) &= M, \quad \Sigma(t) = I, \quad t \in [s - \delta, s), \\
(4.5)
\end{aligned}
\]

\[
\begin{aligned}
\dot{\Phi}(t) &= \left[ \hat{A}(t)^T \Phi(t) + \mathbb{E}^{\mathcal{F}_t}\left[ (\hat{B}^T \Phi)|_{t+\delta} \right] \right] dt \\
&+ \left[ \Phi(t) - \mathbb{E}^{\mathcal{F}_t}\left[ (\hat{B}^T \Phi)|_{t+\delta} \right] \right] dW(t), \quad t \in [s, T], \\
\Phi(s) &= I, \quad \Phi(t) = 0, \quad t \in (T, T + \delta). \\
(4.8)
\end{aligned}
\]

By Theorem 2, the ASDE (4.8) has a unique solution \( \Phi(\cdot) \in L^2_{\mathcal{F}}(\Omega; C([s, T]; \mathbb{R}^{n \times n})) \).

In the following, we aim to obtain an explicit solution to (4.8). To this end, let \( \Pi(\cdot) \) be an exponential martingale, which satisfies

\[
\begin{aligned}
d\Pi(t) &= \gamma(t) \Pi(t) dW(t), \quad t \in [s, T + \delta], \\
\Pi(s) &= I,
\end{aligned}
\]

and \( \Phi(t) = \Upsilon(t) \Pi(t) \), where \( \gamma(\cdot) \) and \( \Upsilon(\cdot) \) are deterministic functions to be determined. Applying Itô’s formula to \( \Phi(\cdot) \), we obtain

\[
\begin{aligned}
\dot{\Upsilon}(t) &= \hat{A}(t)^T \Upsilon(t) + \hat{B}(t + \delta)^T \Upsilon(t + \delta), \\
\Upsilon(t) \gamma(t) &= \Upsilon(t) - \hat{B}(t + \delta)^T \Upsilon(t + \delta), \quad t \in [s, T]. \\
(4.9)
\end{aligned}
\]
Hence we derive the following *time-advanced ordinary differential equation* (AODE, for short):

\[
\begin{align*}
\dot{Y}(t) &= \hat{A}(t)^\top Y(t) + \hat{B}(t + \delta)^\top Y(t + \delta), \quad t \in [s, T], \\
Y(s) &= I, \quad Y(t) = 0, \quad t \in (T, T + \delta].
\end{align*}
\]

(4.10)

Since the coefficients of (4.10) are bounded, it admits a unique solution $Y(\cdot) \in C([s, T]; \mathbb{R}^{n \times n})$. Then by (4.9), we can get $\gamma(\cdot)$. Furthermore, if $Y(\cdot) > 0$, then $\Phi(\cdot) > 0$.

**Lemma 1** The linear delayed matrix-valued differential equation (4.7) admits a unique solution $\hat{\Sigma}(\cdot) \in C([s, T]; S^n)$. Furthermore, if $\hat{H}(\cdot) \geq 0$, $\hat{F}(\cdot) \geq 0$, $\hat{M} \geq 0$ and $\hat{A}(\cdot), \hat{B}(\cdot)$ such that the solution to the AODE (4.10) $Y(\cdot) > 0$, then $\hat{\Sigma}(\cdot) \in C([s, T]; S^n_+)$. 

**Proof** Since the coefficients of (4.7) are bounded, the unique solvability of (4.7) can be proved in the same method as Theorem 1. Next we prove the second result. Applying Itô’s formula to $\Phi(\cdot)^\top \hat{\Sigma}(\cdot) \Phi(\cdot)$, we have

\[
\Phi(t)^\top \hat{\Sigma}(t) \Phi(t) = \Phi(T)^\top \hat{M} \Phi(T) + \int_t^T \Phi(r)^\top \left( [\hat{B}(r) \hat{\Sigma}(r - \delta) \hat{B}(r)^\top + \hat{H}(r)] \Phi(r) \\
- \mathbb{E}^{\mathcal{F}_r}[(\Phi^\top \hat{B})|_{r+\delta}] \hat{\Sigma}(r) \mathbb{E}^{\mathcal{F}_r}[(\hat{B}^\top \Phi)|_{r+\delta}] \right) dr - \int_t^T \{ \cdots \} dW(r).
\]

Noting $\Phi(t) = 0$ for $t \in (T, T + \delta]$, we get

\[
\mathbb{E} \int_t^T (\Phi^{-1}(t))^\top \Phi(r)^\top \hat{B}(r) \hat{\Sigma}(r - \delta) \hat{B}(r) \Phi(r) \Phi^{-1}(t) dr
\]

\[
= \mathbb{E} \int_t^{t-\delta} (\Phi^{-1}(t))^\top (\Phi^\top \hat{B})|_{r+\delta} \hat{\Sigma}(r) (\hat{B}^\top \Phi)|_{r+\delta} \Phi^{-1}(t) dr
\]

\[+
\mathbb{E} \int_t^T (\Phi^{-1}(t))^\top (\Phi^\top \hat{B})|_{r+\delta} \hat{\Sigma}(r) (\hat{B}^\top \Phi)|_{r+\delta} \Phi^{-1}(t) dr.
\]

Since $Y(\cdot) > 0$, $\Phi(\cdot) > 0$, we derive

\[
\hat{\Sigma}(t) = \mathbb{E} \left[(\Phi^{-1}(t))^\top \Phi(T)^\top \hat{M} \Phi(T) \Phi^{-1}(t) + \int_t^T (\Phi^{-1}(t))^\top \Phi(r)^\top \hat{H}(r) \Phi(r) \Phi^{-1}(t) dr
\]

\[+
\int_{t-\delta}^t (\Phi^{-1}(t))^\top (\Phi^\top \hat{B})|_{r+\delta} \hat{\Sigma}(r) (\hat{B}^\top \Phi)|_{r+\delta} \Phi^{-1}(t) dr \right].
\]

Consequently $\hat{\Sigma}(s) \geq 0$ if $\hat{H}(\cdot) \geq 0$, $\hat{M} \geq 0$ and $\hat{F}(\cdot) \geq 0$. Similarly, for fixed $t \in (s, T]$, $\hat{\Sigma}(t) \geq 0$ since $\hat{\Sigma}(r) \geq 0$ for $r \in [t - \delta, t)$. The proof is complete.

**Remark 11** In fact, from the proof of Lemma 1, it follows that if $\hat{M} > 0$, then $\hat{\Sigma}(\cdot) > 0$. 

Springer
Remark 12 If we consider the solvability of the following equation:

$$
\begin{align*}
\hat{\Sigma}(t) &= - \hat{\Sigma}(t) \hat{A}(t) \hat{\Sigma}(t) - \hat{A}(t) \hat{\Sigma}(t) - B(t) \hat{\Sigma}(t) B(t)^T \quad & t \in [s, T], \\
\hat{\Sigma}(T) &= \hat{M}, \quad \hat{\Sigma}(t) = \hat{F}(t), \quad t \in [s - \delta, s), \\
\end{align*}
$$

then we can’t find \( \Phi(\cdot) \) satisfying some equation such that \( \hat{\Sigma}(\cdot) \geq 0 \). Hence we only give the solvability in the special case: \( B(\cdot) = I \) (as Lemma 1).

Next we continue to study the existence of the solution to (4.5). Let \( B(\cdot) = I \) and \( M > 0 \). Denote

$$
\begin{align*}
\hat{A}(t) &:= A(t) - 2 \Sigma(t) \left[ Q(t) + \bar{Q}(t + \delta) \right], \\
\hat{H}(t) &:= C(t) N^{-1}(t) C(t)^T + \bar{C}(t) N^{-1}(t - \delta) \bar{C}(t)^T + 2 \Sigma(t) \left[ Q(t) + \bar{Q}(t + \delta) \right] \Sigma(t). \quad (4.11)
\end{align*}
$$

Apparently (4.5) is equivalent to the following equation:

$$
\begin{align*}
\dot{\Sigma} &= - \Sigma \hat{A}^T - \hat{A} \Sigma - \Sigma - \bar{B} \Sigma |_{t = \delta} \bar{B}^T - \hat{H}, \quad t \in [s, T], \\
\Sigma(T) &= M, \quad \Sigma(t) = I, \quad t \in [s - \delta, s). \quad (4.12)
\end{align*}
$$

Next we construct the iterative scheme as follows. For \( i = 0, 1, 2, \cdots \), set

$$
\begin{align*}
\Sigma^0(t) &= M \text{ for } t \in [s, T], \quad \Sigma^0(t) = I \text{ for } t \in [s - \delta, s), \\
\hat{A}_i(t) &= A(t) - 2 \Sigma_i(t) \left[ Q(t) + \bar{Q}(t + \delta) \right], \\
\Psi_i(t) &= 2 \left[ Q(t) + \bar{Q}(t + \delta) \right] \Sigma^i(t), \\
\hat{H}_i(t) &= C(t) N^{-1}(t) C(t)^T + \bar{C}(t) N^{-1}(t - \delta) \bar{C}(t)^T + 2 \Sigma^i(t) \left[ Q(t) + \bar{Q}(t + \delta) \right] \Sigma^i(t), \quad (4.13)
\end{align*}
$$

and \( \Sigma^{i+1}(\cdot) \) be the solution to

$$
\begin{align*}
\hat{\Sigma}^{i+1} &= - \Sigma^{i+1} \hat{A}_i^T - \hat{A}_i \Sigma^{i+1} - \Sigma^{i+1} - \bar{B} \Sigma^{i+1} |_{t = \delta} \bar{B}^T - \hat{H}_i, \quad t \in [s, T], \\
\Sigma^{i+1}(T) &= M, \quad \Sigma^{i+1}(t) = I, \quad t \in [s - \delta, s). \quad (4.14)
\end{align*}
$$

Let \( \Phi_i(\cdot) \) and \( \Upsilon_i(\cdot) \) be the solutions to ASDE (4.8) and AODE (4.10) associated with \( \hat{A}_i(\cdot) \) and \( \bar{B}(\cdot) \), respectively. Suppose \( \Upsilon_i(\cdot) > 0 \), then by Lemma 1, we get \( \Sigma^i(\cdot) \in C([s, T]; S^n_{+}) \). Denote \( \Delta^i(\cdot) := \Sigma^i(\cdot) - \Sigma^{i+1}(\cdot) \), then we have

$$
\begin{align*}
-\hat{A}_i &= \Delta^i \hat{A}_i + \hat{A}_i \Delta^i + \Delta^i + \bar{B} \Delta^i |_{t = \delta} \bar{B}^T + \Sigma^i (\hat{A}_{i-1} - \hat{A}_i)^T \\
+ (\hat{A}_{i-1} - \hat{A}_i) \Sigma^i + \hat{H}_{i-1} - \hat{H}_i, \quad t \in [s, T], \\
\Delta^i(T) &= 0, \quad \Delta^i(t) = 0, \quad t \in [s - \delta, s).
\end{align*}
$$
Denote $\Theta_i(t) := 2\left[ Q(t) + \hat{Q}(t + \delta) \right]^{\frac{1}{2}} (\Sigma^i(t) - \Sigma^{i-1}(t))$, noting
\[
\begin{cases}
\hat{H}_{i-1}(t) - \hat{H}_i(t) = \frac{1}{2} \Theta_i(t)^\top \Theta_i(t) - \frac{1}{2} \Psi_i(t)^\top \Theta_i(t) - \frac{1}{2} \Theta_i(t)^\top \Psi_i(t), \\
\hat{A}_{i-1}(t) - \hat{A}_i(t) = \Theta_i(t)^\top \left[ Q(t) + \hat{Q}(t + \delta) \right]^{\frac{1}{2}}.
\end{cases}
\]
Hence we obtain
\[
-\left[ \hat{A}_i + \Delta^i \hat{A}_i^\top + \hat{A}_i \Delta^i + \Delta^i + \bar{B} \Delta^i |_{t=0} \bar{B}^\top \right] = \frac{1}{2} \Theta_i^\top \Theta_i \geq 0.
\]
Since $\Delta^i(T) = 0$ and $\Delta^i(t) = 0$ for $t \in [s - \delta, s)$, it follows that $\Delta^i(\cdot) \geq 0$. Thus $\{\Sigma^i(\cdot)\}$ is a decreasing sequence in $C([s, T]; \bar{S}^+_n)$. Therefore it has a limit, denoting by $\Sigma(\cdot)$. Clearly $\Sigma(\cdot) \in C([s, T]; S^+_n)$ is the solution to (4.12), hence (4.5). Furthermore let $\Phi(\cdot)$ and $\Upsilon(\cdot)$ be the solutions to ASDE (4.8) and AODE (4.10) associated with $\hat{A}(\cdot)$ and $\bar{B}(\cdot)$, respectively. Suppose $\Upsilon(\cdot) > 0$, then by Remark 11, $\Sigma(\cdot) > 0$ owing to $M > 0$.

If $M \geq 0$, we repeat the above step using $M \geq 0$ instead of $M > 0$. Then we get $\Sigma(\cdot) \geq 0$. Finally we summarize the above results as follows:

(A4): For $i = 0, 1, 2, \cdots$, the solution to AODE (4.10) $\Upsilon_i(\cdot) > 0$, where $\hat{A}_i(\cdot)$ is given by (4.13) and (4.14).

(A5): The solution to AODE (4.10) $\Upsilon(\cdot) > 0$, where $\hat{A}(\cdot)$ is given by (4.11) and (4.12).

**Proposition 2** Let $M \geq 0$ be $n \times n$ symmetric matrix, suppose (A4) holds and $B(\cdot) = I$, then (4.5) has the unique solution $\Sigma(\cdot) \in C([s, T]; S^+_n)$. Furthermore, suppose (A5) holds and $M > 0$, then $\Sigma(\cdot) \in C([s, T]; \bar{S}^+_n)$.

**Remark 13** The condition (A4), imposed to the coefficients of (4.5), is mainly to guarantee $\Sigma(\cdot) \geq 0$.

**Remark 14** Although (A4) and (A5) seem a little complex, (4.10) is a linear ODE, and we have rich results to study its solution. Therefore it is not difficult to study the specific conditions imposed to the coefficients of (4.5) to satisfy (A4) and (A5). However it is not our major job in this paper, so we leave this question to the interested readers, please refer to [8, 16, 26] and references therein.

**Proposition 3** Let $M > 0$ be an $n \times n$ symmetric matrix, suppose (A4) and (A5) hold, $B(\cdot) = I$, then (4.6) is uniquely solvable with the solution $P(\cdot) \in C([s, T]; \bar{S}^+_n)$.

**Proof** By Proposition 2, (4.5) is uniquely solvable with the solution $\Sigma(\cdot) \in C([s, T]; \bar{S}^+_n)$, hence $P(\cdot) = \Sigma^{-1}(\cdot)$ is well-defined. By calculating, it is indeed a solution to (4.6). As for the uniqueness, suppose $P_1(\cdot), P_2(\cdot)$ are two solutions to (4.6), then $P_1^{-1}(\cdot), P_2^{-1}(\cdot)$ are two solutions to (4.5). Thus by the uniqueness of solutions to (4.5), we have $P_1(\cdot) = P_2(\cdot)$. The proof is complete.

**Corollary 2** Let $B(\cdot) = I$, $R(\cdot) + \bar{R}(\cdot + \delta) = 0$, and suppose (A4) holds, then (3.1) admits the unique solution $\Sigma(\cdot) \in C([s, T]; S^+_n)$. 

\[ \square \] Springer
Proof Proposition 2 implies this corollary.

Next we focus on the Riccati equation (3.3).

Proposition 4 Let \( \Sigma(\cdot) \) be the solution to (3.1), then the Riccati equation (3.3) is uniquely solvable, and

(i) if \( G > 0 \), then \( L(\cdot) \in C([s, T]; S^m_+) \);

(ii) if \( G \geq 0 \), then \( L(\cdot) \in C([s, T]; S^m_+) \).

Proof By making the time reversing transformation \( \tau = T - t + s \), we have

\[
\begin{aligned}
\dot{L} &= - A^T L - LA - 2(Q + \tilde{Q})_{|t+\delta} + L \left\{ B \Sigma \mathcal{M}^{-1} B^T + C N^{-1} C^T \right\}
+ \tilde{C} N^{-1}_{|t-\delta} \tilde{C}^T + \tilde{B} (\Sigma \mathcal{M}^{-1})_{|t-\delta} \tilde{B}^T \right\} L, \quad t \in [s, T],
\end{aligned}
\]

(4.15)

When \( M_i = \frac{1}{i} I, i \in \mathbb{N}^+ \), let \( \sigma_i(\cdot) \) and \( p_i(\cdot) \) denote the solutions to (4.1) and (4.2), respectively. Noting

\[
\Sigma(t)\mathcal{M}^{-1}(t) = \Sigma(t)[2(R(t) + \tilde{R}(t + \delta))\Sigma(t) + I]^{-1}
= \lim_{i \to \infty} \sigma_i(t)[2(R(t) + \tilde{R}(t + \delta))\sigma_i(t) + I]^{-1} = \lim_{i \to \infty} [2(R(t) + \tilde{R}(t + \delta)) + p_i(t)]^{-1},
\]

thus \( B(t)\Sigma(t)\mathcal{M}^{-1}(t)B(t)^T + C(t)N^{-1}(t)C(t)^T + \tilde{C}(t)N^{-1}(t - \delta)\tilde{C}(t)^T + \tilde{B}(t)\Sigma(t - \delta)\mathcal{M}^{-1}(t - \delta)\tilde{B}(t)^T \) is symmetric, for all \( t \in [s, T] \). Therefore (4.15) is a standard Riccati equation and the proof is completed. □

Before considering the existence and uniqueness of the solution to the DAFBSDE (3.4), we would like to study the following general equation:

\[
\begin{aligned}
dX(t) &= h(t, X(t), \Lambda(t), \Gamma(t), X(t + \delta))dt + g(t, X(t), \Lambda(t), \Gamma(t), X(t + \delta))dW(t),
d\Lambda(t) &= f(t, X(t), \Lambda(t), \Gamma(t), \mathbb{E}_T^{\mathcal{F}_t}[X(t)], \Lambda(t), \Gamma(t), \Lambda(t - \delta), \Gamma(t - \delta))dt
- \Gamma(t)dW(t), \quad t \in [s, T],
X(s) &= \Phi(\Lambda(s)), \quad X(t) = 0, \quad t \in (T, T + \delta],
\Lambda(T) &= \zeta, \quad \Lambda(t) = \Gamma(t) = 0, \quad t \in [s - \delta, s),
\end{aligned}
\]

(4.16)

where \( h(t, \omega, x, \lambda, \gamma, x_{\delta+}) : \Omega \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to L^2_{\mathcal{F}_t}(\mathbb{R}^n) \), \( g(t, \omega, x, \lambda, \gamma, x_{\delta+}) : \Omega \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to L^2_{\mathcal{F}_t}(\mathbb{R}^n) \), \( f(t, \omega, x, \lambda, \gamma, \theta, \lambda_{\delta+}, \gamma_{\delta+}) : \Omega \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to L^2_{\mathcal{F}_t}(\mathbb{R}^n) \), \( \Phi(\lambda) : \mathbb{R}^n \to \mathbb{R}^n \) satisfy the following conditions:

(H5) For any \( t \in [s, T] \), any \( x, x', \lambda, \lambda', \gamma, \gamma', x_{\delta+}, x'_{\delta+} \in \mathbb{R}^n \), there exist constants \( L_1, L_2, L_3 \) such that

\[
\begin{aligned}
|h(t, \omega, x, \lambda, \gamma, x_{\delta+}) - h(t, \omega, x', \lambda', \gamma', x'_{\delta+})| + |g(t, \omega, x, \lambda, \gamma, x_{\delta+})
- g(t, \omega, x', \lambda', \gamma', x'_{\delta+})| &
\leq L_1|x - x'| + L_2|\lambda - \lambda'| + L_2|\gamma - \gamma'| + L_3 \mathbb{E}_T^{\mathcal{F}_t}[|x_{\delta+} - x'_{\delta+}|].
\end{aligned}
\]
(H6) For any \( t \in [s, T] \), any \( x, x', \lambda, \lambda', \gamma, \gamma', \lambda_\delta, \lambda'_\delta, \gamma_\delta, \gamma'_\delta \in \mathbb{R}^n \), \( \theta, \theta' \in \mathbb{L}^2_{\mathcal{F}_{t-}}(\mathbb{R}^n) \), there exist constants \( L_4, L_5, L_6, L_7 \) such that

\[
|f(t, \omega, x, \lambda, \gamma, \theta, \lambda_\delta, \gamma_\delta) - f(t, \omega, x', \lambda', \gamma', \theta', \lambda'_\delta, \gamma'_\delta)| \leq L_4|x - x'| + L_5|\lambda - \lambda'| + L_5|\gamma - \gamma'| + L_6|\theta - \theta'| + L_7|\lambda_\delta - \lambda'_\delta| + L_7|\gamma_\delta - \gamma'_\delta|.
\]

(H7) For any \( \lambda, \lambda' \in \mathbb{R}^n \), there exists a constant \( L_8 \) such that

\[
|\Phi(\lambda) - \Phi(\lambda')| \leq L_8|\lambda - \lambda'|.
\]

(H8) \( \mathbb{E} \int_s^T (|h(t, 0, 0, 0, 0)|^2 + |\sigma(t, 0, 0, 0, 0)|^2 + |f(t, 0, 0, 0, 0, 0)|^2)dt < +\infty. \)

Then we can give the following result.

**Theorem 5** Let (H5)-(H8) hold and \( \delta \) be sufficiently small, and suppose either \( L_4, L_6 \) are sufficiently small, or \( L_2, L_8 \) are sufficiently small, then for any \( \zeta \in \mathbb{L}^2_{\mathcal{F}_{T}}(\mathbb{R}^n) \), the general DAFBSDE (4.16) has a unique solution \((X(\cdot), \Lambda(\cdot), \Gamma(\cdot)) \in \mathbb{L}^2_{\mathcal{F}}(\Omega; C([s, T]; \mathbb{R}^n)) \times \mathbb{L}^2_{\mathcal{F}}(\Omega; C([s, T]; \mathbb{R}^n)) \times \mathbb{L}^2_{\mathcal{F}}([s, T]; \mathbb{R}^n)\). \)

**Proof** For any \( x(\cdot) \in \mathbb{L}^2_{\mathcal{F}}(\Omega; C([s, T]; \mathbb{R}^n)) \), consider the following equation:

\[
\begin{align*}
\{ & dX(t) = h(t, X(t), \Lambda(t), \Gamma(t), X(t + \delta))dt + g(t, X(t), \Lambda(t), \Gamma(t), X(t + \delta))dW(t), \\
& -d\Lambda(t) = f(t, x(t), \Lambda(t), \Gamma(t), \mathbb{E}\mathcal{F}_{t-}[x(t)], \Lambda(t - \delta), \Gamma(t - \delta))dt \\
& \quad - \Gamma(t)dW(t), \quad t \in [s, T], \\
& X(s) = \Phi(\Lambda(s)), \quad X(t) = 0, \quad t \in (T, T + \delta], \\
& \Lambda(T) = \zeta, \quad \Lambda(t) = \Gamma(t) = 0, \quad t \in [s - \delta, s).
\end{align*}
\]

(4.17)

For given \( x(\cdot) \in \mathbb{L}^2_{\mathcal{F}}(\Omega; C([s, T]; \mathbb{R}^n)) \), since \( \delta \) is sufficiently small, the second equation of (4.17) admits the unique solution \((\Lambda(\cdot), \Gamma(\cdot)) \in \mathbb{L}^2_{\mathcal{F}}(\Omega; C([s, T]; \mathbb{R}^n)) \times \mathbb{L}^2_{\mathcal{F}}([s, T]; \mathbb{R}^n)\) by Theorem 2.1 in [1]. Then by Theorem 2 in Sect. 2, the first equation of (4.17) has the unique solution \( X(\cdot) \in \mathbb{L}^2_{\mathcal{F}}(\Omega; C([s, T]; \mathbb{R}^n)) \). Hence the mapping

\[
T : \mathbb{L}^2_{\mathcal{F}}(\Omega; C([s, T]; \mathbb{R}^n)) \to \mathbb{L}^2_{\mathcal{F}}(\Omega; C([s, T]; \mathbb{R}^n)),
\]

\[
x(\cdot) \mapsto X(\cdot),
\]

is well-defined. For any \( x(\cdot), x'(\cdot) \in \mathbb{L}^2_{\mathcal{F}}(\Omega; C([s, T]; \mathbb{R}^n)) \), denote \((X(\cdot), \Lambda(\cdot), \Gamma(\cdot)) \) and \((X'(\cdot), \Lambda'(\cdot), \Gamma'(\cdot)) \) are the solutions to (4.17) along with \( x(\cdot), x'(\cdot) \), respectively. Denote \( \hat{x}(\cdot) = x(\cdot) - x'(\cdot), \; \hat{X}(\cdot) = X(\cdot) - X'(\cdot), \; \hat{\Lambda}(\cdot) = \Lambda(\cdot) - \Lambda'(\cdot), \; \hat{\Gamma}(\cdot) = \)
\[ \Gamma(\cdot) - \Gamma'(\cdot). \] Applying Ito’s formula to \(|\hat{\Lambda}(\cdot)|^2\), we have

\[
|\dot{\hat{\Lambda}}(t)|^2 + \int_t^T |\dot{\hat{\Gamma}}(r)|^2 dr = -2 \int_t^T \{\hat{\Lambda}(r), \hat{\Gamma}(r)dW(r)\} \\
+ 2 \int_t^T \{\hat{\Lambda}(r), f(r, x(r), \Lambda(r), \Gamma(r), E^{\mathcal{F}_{r-\delta}}[x(r)], \Lambda(r-\delta), \Gamma(r-\delta)) \\
- f(r, x'(r), \Lambda'(r), \Gamma'(r), E^{\mathcal{F}_{r-\delta}}[x'(r)], \Lambda'(r-\delta), \Gamma'(r-\delta))\}dr.
\]

Taking the expectation of both sides of (4.18) and by \(G\)undy’s inequality, we have

\[
E \sup_{s \leq t \leq T} |\hat{\Lambda}(r)|^2 \\
\leq \left( \frac{1}{\epsilon} + 6\epsilon L_5^2 + 6\epsilon L_7^2 \right) E \int_s^T |\dot{\hat{\Lambda}}(r)|^2 dr + 12\epsilon (L_4^2 + L_6^2) E \int_s^T |\dot{\hat{x}}(r)|^2 dr,
\]

where \(\epsilon > 0\) is an arbitrary small constant. On the other hand, by Burkholder-Davis-
Gundy’s inequality, we have

\[
E \sup_{s \leq t \leq T} |\hat{\Lambda}(r)|^2 \\
\leq E \int_s^T \left[ \left( \frac{2}{\epsilon} + 12\epsilon L_5^2 + 12\epsilon L_7^2 \right) |\dot{\hat{\Lambda}}(r)|^2 + 12\epsilon (L_4^2 + L_6^2) |\dot{\hat{x}}(r)|^2 \right] dr + 4C^2 E \int_s^T |\dot{\hat{\Gamma}}(r)|^2 dr,
\]

where \(C\) is a generic constant. Applying Grownwall’s inequality to (4.20), we deduce

\[
E \sup_{s \leq t \leq T} |\hat{\Lambda}(r)|^2 \\
\leq K_1 E \int_s^T \left[ (4C^2 + 12\epsilon L_5^2 + 12\epsilon L_7^2) |\dot{\hat{\Gamma}}(r)|^2 + 12\epsilon (L_4^2 + L_6^2) |\dot{\hat{x}}(r)|^2 \right] dr \\
\leq K_1 (4C^2 + 12\epsilon L_5^2 + 12\epsilon L_7^2) \left( 1 - 6\epsilon L_5^2 - 6\epsilon L_7^2 \right)^{-1} \left( \frac{1}{\epsilon} + 6\epsilon L_5^2 + 6\epsilon L_7^2 \right) E \times \int_s^T |\dot{\hat{\Lambda}}(t)|^2 dt \\
+ K_1 \left[ 6\epsilon (4C^2 + 12\epsilon L_5^2 + 12\epsilon L_7^2) (1 - 6\epsilon L_5^2 - 6\epsilon L_7^2)^{-1} (L_4^2 + L_6^2) \\
+ 12\epsilon (L_4^2 + L_6^2) \right] E \int_s^T |\dot{\hat{x}}(t)|^2 dt,
\]

where \(K_1 = \exp \left\{ T \left( \frac{2}{\epsilon} + 12\epsilon L_5^2 + 12\epsilon L_7^2 \right) \right\} \).
Again applying Grownwall’s inequality to (4.21), it follows that

\[ \mathbb{E} \sup_{t \leq T} |\hat{\Lambda}(t)|^2 \leq K_1 K_2 \left[ 6\epsilon\left(4C^2 + 12\epsilon L_5^2 + 12\epsilon L_7^2\right)(1 - 6\epsilon L_5^2 - 6\epsilon L_7^2)^{-1}(L_4^2 + L_6^2) \right. \\
+ 12\epsilon(L_4^2 + L_6^2) \mathbb{E} \int_s^T |\hat{\chi}(t)|^2 \, dt, \tag{4.22} \]

where

\[ K_2 = \exp \left\{ K_1 T\left(4C^2 + 12\epsilon L_5^2 + 12\epsilon L_7^2\right)(1 - 6\epsilon L_5^2 - 6\epsilon L_7^2)^{-1}\left(\frac{1}{\epsilon} + 6\epsilon L_5^2 + 6\epsilon L_7^2\right) \right\}. \]

Substituting (4.22) into (4.19), we derive

\[ \mathbb{E} \int_s^T |\hat{\Gamma}(r)|^2 \, dr \leq \left\{ (1 - 6\epsilon L_5^2 - 6\epsilon L_7^2)^{-1}\left(\frac{1}{\epsilon} + 6\epsilon L_5^2 + 6\epsilon L_7^2\right)TK_1K_2 \left[ 6\epsilon\left(4C^2 + 12\epsilon L_5^2 + 12\epsilon L_7^2\right) \right. \\
\times (1 - 6\epsilon L_5^2 - 6\epsilon L_7^2)^{-1}(L_4^2 + L_6^2) + 12\epsilon(L_4^2 + L_6^2) \right] \\
+ 6\epsilon(1 - 6\epsilon L_5^2 - 6\epsilon L_7^2)^{-1}(L_4^2 + L_6^2) \mathbb{E} \int_s^T |\hat{\chi}(t)|^2 \, dt. \tag{4.23} \]

Hence by (4.22) and (4.23), it follows that

\[ \mathbb{E} \sup_{t \leq T} |\hat{\Lambda}(t)|^2 + \mathbb{E} \int_s^T |\hat{\Gamma}(t)|^2 \, dt \leq K_3(L_4^2 + L_6^2)\mathbb{E} \int_s^T |\hat{\chi}(t)|^2 \, dt, \tag{4.24} \]

where

\[ K_3 = (1 - 6\epsilon L_5^2 - 6\epsilon L_7^2)^{-1}\left(\frac{1}{\epsilon} + 6\epsilon L_5^2 + 6\epsilon L_7^2\right)TK_1K_2 \left[ 6\epsilon\left(4C^2 + 12\epsilon L_5^2 + 12\epsilon L_7^2\right) \right. \\
\times (1 - 6\epsilon L_5^2 - 6\epsilon L_7^2)^{-1} + 12\epsilon \right] + 6\epsilon(1 - 6\epsilon L_5^2 - 6\epsilon L_7^2)^{-1}. \]

Recall the first equation of (4.16), by (H5), (H7) we have

\[ \mathbb{E} \sup_{t \leq T} |\hat{X}(t)|^2 \leq 3\mathbb{E} \left( \int_s^T |h(t, X(t), \Lambda(t), \Gamma(t), X(t + \delta)) - h(t, X'(t), \Lambda'(t), \Gamma'(t), X'(t + \delta))| \, dt \right)^2 \\
+ 3\mathbb{E} \int_s^T |g(t, X(t), \Lambda(t), \Gamma(t), X(t + \delta)) - g(t, X'(t), \Lambda'(t), \Gamma'(t), X'(t + \delta))|^2 \, dt \\
+ 3\mathbb{E} |\Phi(\Lambda(s)) - \Phi(\Lambda'(s))|^2. \]
Finally combining (4.24) and (4.26), we deduce

\[ \mathbb{E} \sup_{s \leq t \leq T} |\hat{X}(t)|^2 \leq K_4 \left\{ 3L_8^2 \mathbb{E} |\hat{\Lambda}(s)|^2 + 12L_2^2(T + 1) \mathbb{E} \int_s^T |\hat{\Lambda}(t)|^2 dt + 12L_2^2(T + 1) \mathbb{E} \int_s^T |\hat{\Gamma}(t)|^2 dt \right\}, \]

(4.26)

where

\[ K_4 = \exp \left\{ 12T^2L_1^2 + 12T^2L_3^2 + 12TL_4^2 + 12TL_6^2 \right\}. \]

Then applying Gronwall’s inequality to (4.25), it follows that

\[ \mathbb{E} \sup_{s \leq t \leq T} |\hat{X}(t)|^2 \leq K_4 \left\{ 3L_8^2 \mathbb{E} |\hat{\Lambda}(s)|^2 + 12L_2^2(T + 1) \mathbb{E} \int_s^T |\hat{\Lambda}(t)|^2 dt + 12L_2^2(T + 1) \mathbb{E} \int_s^T |\hat{\Gamma}(t)|^2 dt \right\}, \]

(4.26)

where

\[ K_4 = \exp \left\{ 12T^2L_1^2 + 12T^2L_3^2 + 12TL_4^2 + 12TL_6^2 \right\}. \]

Finally combining (4.24) and (4.26), we deduce

\[ \mathbb{E} \sup_{s \leq t \leq T} |\hat{X}(t)|^2 \leq K_4 K_3 T \left( L_4^2 + L_6^2 \right) \left\{ 3L_8^2 + 12L_2^2T(T + 1) + 12L_2^2(T + 1) \right\} \mathbb{E} \sup_{s \leq t \leq T} |\hat{x}(t)|^2. \]

(4.27)

Since either \( L_4, L_6 \) are sufficiently small, or \( L_2, L_8 \) are sufficiently small, \( T \) is a contraction mapping. Thus the proof is completed. \( \Box \)

Remark 15 In the above proof, \( K_1 \approx e^{2T \epsilon} \), \( K_2 \approx \exp \left\{ \frac{4TK_1C^2}{(1-6\epsilon L_5^2-6\epsilon L_7^2)^2} \right\} \), \( K_3 \approx \frac{24TK_1K_2C^2}{(1-6\epsilon L_5^2-6\epsilon L_7^2)^2} \), where \( \epsilon > 0 \) is sufficiently small such that \( 1 - 6\epsilon L_5^2 - 6\epsilon L_7^2 < 1 \). As long as \( K_4K_3T \left( L_4^2 + L_6^2 \right) \left\{ 3L_8^2 + 12L_2^2T(T + 1) + 12L_2^2(T + 1) \right\} < 1 \), then \( T \) is a contraction mapping, and thus Theorem 5 holds.

By Theorem 5, we get the following result immediately.

Corollary 3 Let \( \delta \) be sufficiently small, suppose one of the following two conditions holds:

(i) \( \sup_{s \leq t \leq T} |\tilde{B}(t)\Sigma(t - \delta)\mathcal{M}^{-1}(t - \delta)\tilde{B}(t)^\top + \tilde{C}(t)\mathcal{N}^{-1}(t - \delta)\tilde{C}(t)^\top| \) is sufficiently small;

(ii) \( \sup_{s \leq t \leq T} |Q(t) + \tilde{Q}(t + \delta)| + \sup_{s \leq t \leq T} |R(t) + \tilde{R}(t + \delta)| + |G| \) is sufficiently small.

Then the DAFBSDE (3.4) has a unique solution \((\tilde{X}(\cdot), \Lambda(\cdot), \Gamma(\cdot)) \in L_2^2(\Omega; C([s, T]; \mathbb{R}^n)) \times L_2^2(\Omega; C([s, T]; \mathbb{R}^n)) \times L_2^2([s, T]; \mathbb{R}^n)\).
5 Proofs of Theorems 3 and 4

In this section, we proceed to complete the proofs of Theorem 3 and Theorem 4. The basic idea is first to give the optimal control expression (3.7) based on (3.6) and then to find the optimal linear state feedback (3.9).

First we give the proof of Theorem 3.

Proof For any \( u(\cdot) \in \mathcal{U}[s, T] \) with \( u(t) = \eta(t) \) for \( t \in [s - \delta, s) \), define \( u^\varepsilon(\cdot) = u^*(\cdot) + \varepsilon (u(\cdot) - u^*(\cdot)), \varepsilon \in [0, 1] \). Suppose \( (Y^u(\cdot), Z^u(\cdot)) \) and \( (Y^\varepsilon(\cdot), Z^\varepsilon(\cdot)) \) are state trajectories corresponding to \( u(\cdot) \) and \( u^\varepsilon(\cdot) \), respectively. Then we have

\[
0 \leq J(u^\varepsilon(\cdot)) - J(u^*(\cdot)) = 2\varepsilon \mathbb{E} \left\{ \left[ GY^*(s), Y^u(s) - Y^*(s) \right] + \int_s^T \left[ \left( Q(t)Y^*(t), Y^u(t) - Y^*(t) \right) \right. \right.
\]
\[
+ \left. \left( \tilde{Q}(t)Y^*(t - \delta), Y^u(t - \delta) - Y^*(t - \delta) \right) + \left( R(t)Z^*(t), Z^u(t) - Z^*(t) \right) \right] \left( \tilde{R}(t)Z^*(t - \delta), Z^u(t - \delta) - Z^*(t - \delta) \right) \right. \right.
\]
\[
+ \left. \left( \tilde{N}(t)u^*(t - \delta), u(t - \delta) - u^*(t - \delta) \right) \right] dt + \varepsilon^2 \left\{ \cdots \right\}. \tag{5.1}
\]

Applying Itô’s formula to \( \left( X^*(\cdot), Y^u(\cdot) - Y^*(\cdot) \right) \), from (5.1) we deduce

\[
0 \leq J(u^\varepsilon(\cdot)) - J(u^*(\cdot)) = 2\varepsilon \mathbb{E} \left\{ \left\{ - C(t)^\top X^*(t) + 2N(t)u^*(t) - \tilde{C}(t + \delta)^\top X^*(t + \delta) \right. \right.
\]
\[
+ \left. 2\tilde{N}(t + \delta)u^*(t), u(t) - u^*(t) \right] dt + \varepsilon^2 \left\{ \cdots \right\}. \tag{5.2}
\]

Dividing both sides of (5.2) by \( \varepsilon \), we derive

\[
0 \leq \frac{1}{\varepsilon} \left\{ J(u^\varepsilon(\cdot)) - J(u^*(\cdot)) \right\} = 2\mathbb{E} \left\{ \left( - C(t)^\top X^*(t) + 2N(t)u^*(t) - \tilde{C}(t + \delta)^\top X^*(t + \delta) \right. \right.
\]
\[
+ \left. 2\tilde{N}(t + \delta)u^*(t), u(t) - u^*(t) \right] dt + \varepsilon \left\{ \cdots \right\}.
\]

Letting \( \varepsilon \to 0 \), due to the arbitrariness of \( u(\cdot) \), we complete the proof. \( \square \)
Now we continue to prove Theorem 4. First we introduce the stochastic Hamiltonian system as follows:

\[
\begin{aligned}
    dX^*(t) &= \left\{ A(t)^\top X^*(t) - 2(Q(t) + \tilde{Q}(t + \delta))Y^*(t) + \mathbb{E}^{F_t}[(\tilde{A}^\top X^*)_{t+\delta}] \right\} dt \\
    &\quad + \left\{ B(t)^\top X^*(t) - 2[R(t) + \tilde{R}(t + \delta)]Z^*(t) + \mathbb{E}^{F_t}[(\tilde{B}^\top X^*)_{t+\delta}] \right\} dW(t), \\
    -dY^*(t) &= \left\{ A(t)Y^*(t) + \tilde{A}(t)Y^*(t - \delta) + B(t)Z^*(t) + \tilde{B}(t)Z^*(t - \delta) \\
    &\quad + C(t)N^{-1}(t) \left\{ C(t)^\top X^*(t) + \mathbb{E}^{F_t}[(\tilde{C}^\top X^*)_{t+\delta}] \right\} \\
    &\quad + \tilde{C}(t)N^{-1}(t - \delta) \left\{ C(t - \delta)^\top X^*(t - \delta) + \mathbb{E}^{F_{t-\delta}}[(\tilde{C}(t)^\top X^*(t))_{t}] \right\} \right\} dt \\
    &\quad - \left[ 2\Sigma(t)(R(t) + \tilde{R}(t + \delta)) + I \right]^{-1} \Gamma(t), \\
    X^*(s) &= -2GY^*(s), \\
    Y^*(t) &= \xi, \\
    Z^*(t) &= \phi(t), \\
    t \in [s, T].
\end{aligned}
\]

(5.3)

We have the following three propositions.

**Proposition 5** Suppose \( \tilde{A}(t) = \tilde{B}(t) = \tilde{C}(t) = 0 \) for \( t \in [s, s + \delta] \) and (3.8) holds, and suppose \( \Sigma(\cdot) \in C([s, T]; S_+^n), L(\cdot) \in C([s, T]; S_+^n), (\bar{X}(\cdot), \Lambda(\cdot), \Gamma(\cdot)) \in L^{2}_{F_T}(\Omega; C([s, T]; R^n)) \times L^{2}_{F_T}(\Omega; C([s, T]; R^n)) \times L^{2}_{F_T}([s, T]; R^n) \) are the solutions to (3.1), (3.3), (3.4), respectively. Then the stochastic Hamilton system (5.3) is uniquely solvable. Moreover, for \( t \in [s, T] \), the following relationships hold:

\[
\begin{aligned}
    Y^*(t) &= \Sigma(t)X^*(t) - \Lambda(t), \\
    Z^*(t) &= \Sigma(t)M^{-1}(t) \left\{ B(t)^\top X^*(t) + \mathbb{E}^{F_t}[(\tilde{B}^\top X^*)_{t+\delta}] \right\} \\
    &\quad - \left[ 2\Sigma(t)(R(t) + \tilde{R}(t + \delta)) + I \right]^{-1} \Gamma(t), \\
    Y^*(s) &= -(I + 2\Sigma(s)G)^{-1}\Lambda(s).
\end{aligned}
\]

(5.4)

**Proof** For the existence of solutions to (5.3), define \( \tilde{Y}(t) := \Sigma(t)\tilde{X}(t) - \Lambda(t) \) for \( t \in [s, T] \), noting (3.8), we get

\[
\begin{aligned}
    -\Sigma(t)\mathbb{E}^{F_t}[(\tilde{A}^\top \tilde{X})_{t+\delta}] &= B(t)\Sigma(t)M^{-1}(t)\mathbb{E}^{F_t}[(\tilde{B}^\top \tilde{X})_{t+\delta}] \\
    &\quad + C(t)N^{-1}(t)\mathbb{E}^{F_t}[(\tilde{C}^\top \tilde{X})_{t+\delta}],
\end{aligned}
\]

and similarly

\[
\begin{aligned}
    -\tilde{A}(t)\Sigma(t - \delta)\tilde{X}(t - \delta) &= \tilde{B}(t)\Sigma(t - \delta)M^{-1}(t - \delta)B(t - \delta)^\top \tilde{X}(t - \delta) \\
    &\quad + \tilde{C}(t)N^{-1}(t - \delta)C(t - \delta)^\top \tilde{X}(t - \delta).
\end{aligned}
\]
Applying Itô’s formula, we obtain

\[
-d\tilde{Y}(t) = \left\{ A(t)\tilde{Y}(t) + \tilde{A}(t)(\tilde{Y}(t - \delta) + B(t)\tilde{Z}(t) + \tilde{B}(t)\tilde{Z}(t - \delta)
+ C(t)\mathcal{N}^{-1}(t)\left\{ C(t)^\top \tilde{X}(t) + \mathbb{E}^\mathcal{F}_t[(\tilde{C}^\top \tilde{X})_{\mid t+\delta}] \right\} + \tilde{C}(t)\mathcal{N}^{-1}(t - \delta)
\times \left\{ C(t - \delta)^\top \tilde{X}(t - \delta) + \mathbb{E}^\mathcal{F}_{t-\delta}[(\tilde{C}(t)^\top \tilde{X}(t))_{\mid t+\delta}] \right\}\right\} dt - \tilde{Z}(t)dW(t),
\]

where

\[
\tilde{Z}(t) = \Sigma(t)\mathcal{M}^{-1}(t)\left\{ B(t)^\top \tilde{X}(t) + \mathbb{E}^\mathcal{F}_t[(\tilde{B}^\top \tilde{X})_{\mid t+\delta}] \right\} - \left[ 2\Sigma(t)(R(t) + \tilde{R}(t + \delta)) + \Gamma(t) \right]^{-1} \Gamma(t).
\]

Substituting \(\tilde{Y}(t) = \Sigma(t)\tilde{X}(t) - \Lambda(t)\) for \(t \in [s, \ T]\) and \(5.5)\) into \(3.4)\), and letting \(\tilde{Y}(s) = -(I + 2\Sigma(s)G)^{-1}\Lambda(s)\), it follows that

\[
\left\{ \begin{align*}
\frac{d\tilde{X}(t)}{dt} &= \left\{ A(t)^\top \tilde{X}(t) - 2(Q(t) + \tilde{Q}(t + \delta))\tilde{Y}(t) + \mathbb{E}^\mathcal{F}_t[(\tilde{A}^\top \tilde{X})_{\mid t+\delta}] \right\} dt \\
&\quad + \left\{ B(t)^\top \tilde{X}(t) - 2(R(t) + \tilde{R}(t + \delta))\tilde{Z}(t) + \mathbb{E}^\mathcal{F}_t[(\tilde{B}^\top \tilde{X})_{\mid t+\delta}] \right\} dW(t), \quad t \in [s, \ T], \\
\tilde{X}(s) &= -2G\tilde{Y}(s), \quad \tilde{X}(0) = 0, \quad t \in (T, \ T + \delta).
\end{align*} \right.
\]

Let \(\tilde{Y}(t) = \varphi(t), \tilde{Z}(t) = \psi(t)\) for \(t \in [s - \delta, \ s]\), noting \(\tilde{Y}(T) = \Sigma(T)\tilde{X}(T) - \Lambda(T) = \xi, (\tilde{X}(\cdot), \tilde{Y}(\cdot), \tilde{Z}(\cdot))\) is a solution to the stochastic Hamiltonian system \(5.3)\) and the relationships in \(5.4)\) hold.

As for the uniqueness of solutions to \(5.3)\). Suppose \((X_1(\cdot), Y_1(\cdot), Z_1(\cdot)), (X_2(\cdot), Y_2(\cdot), Z_2(\cdot))\) are two solutions to \(5.3)\). Denote \(\hat{X}(\cdot) = X_1(\cdot) - X_2(\cdot), \hat{Y}(\cdot) = Y_1(\cdot) - Y_2(\cdot), \hat{Z}(\cdot) = Z_1(\cdot) - Z_2(\cdot)\), then \((\hat{X}(\cdot), \hat{Y}(\cdot), \hat{Z}(\cdot))\) is a solution to the following:

\[
\left\{ \begin{align*}
\frac{d\hat{X}(t)}{dt} &= \left\{ A(t)^\top \hat{X}(t) - 2(Q(t) + \tilde{Q}(t + \delta))\hat{Y}(t) + \mathbb{E}^\mathcal{F}_t[(\tilde{A}^\top \hat{X})_{\mid t+\delta}] \right\} dt \\
&\quad + \left\{ B(t)^\top \hat{X}(t) - 2(R(t) + \tilde{R}(t + \delta))\hat{Z}(t) + \mathbb{E}^\mathcal{F}_t[(\tilde{B}^\top \hat{X})_{\mid t+\delta}] \right\} dW(t), \\
-d\hat{Y}(t) &= \left\{ A(t)\hat{Y}(t) + \tilde{A}(t)(\hat{Y}(t - \delta) + B(t)\hat{Z}(t) + \tilde{B}(t)\hat{Z}(t - \delta)
+ C(t)\mathcal{N}^{-1}(t)\left\{ C(t)^\top \hat{X}(t) + \mathbb{E}^\mathcal{F}_t[(\tilde{C}^\top \hat{X})_{\mid t+\delta}] \right\} \right\} dt \\
&\quad + \left\{ C(t - \delta)^\top \hat{X}(t - \delta) + \mathbb{E}^\mathcal{F}_{t-\delta}[(\tilde{C}(t)^\top \hat{X}(t))_{\mid t+\delta}] \right\} dt \\
&\quad - \hat{Z}(t)dW(t), \quad t \in [s, \ T], \\
\hat{X}(s) &= -2G\hat{Y}(s), \quad \hat{X}(0) = 0, \quad t \in (T, \ T + \delta), \\
\hat{Y}(T) &= 0, \quad \hat{Y}(t) = 0, \quad \hat{Z}(t) = 0, \quad t \in [s - \delta, \ s).
\end{align*} \right.
\]
Since \( \tilde{C}(t) = 0 \) for \( t \in [s, s + \delta] \), we have
\[
\mathbb{E} \int_s^T \langle \hat{X}(t), C(t) \mathcal{N}^{-1}(t) \mathbb{E}^{\mathcal{F}_t}[(\hat{C}^\top \hat{X})]_{t+\delta} \rangle dt
\]
\[
= \mathbb{E} \int_s^T \langle \hat{X}(t - \delta), C(t - \delta) \mathcal{N}^{-1}(t - \delta) \tilde{C}(t)^\top \hat{X}(t) \rangle dt,
\]
\[
\mathbb{E} \int_s^T \langle \hat{X}(t), \bar{A}(t) \bar{Y}(t - \delta) \rangle dt = \mathbb{E} \int_s^T \langle \bar{A}(t + \delta)^\top \hat{X}(t + \delta), \bar{Y}(t) \rangle dt,
\]
\[
\mathbb{E} \int_s^T \langle \hat{X}(t), \bar{B}(t) \hat{Z}(t - \delta) \rangle dt = \mathbb{E} \int_s^T \langle \bar{B}(t + \delta)^\top \hat{X}(t + \delta), \hat{Z}(t) \rangle dt,
\]
\[
\mathbb{E} \int_s^T \langle \hat{X}(t), \tilde{C}(t) \mathcal{N}^{-1}(t - \delta) \tilde{C}(t)^\top \hat{X}(t) \rangle dt
\]
\[
= \mathbb{E} \int_s^T \langle \tilde{C}(t + \delta)^\top \hat{X}(t + \delta), \mathcal{N}^{-1}(t) \tilde{C}(t + \delta)^\top \hat{X}(t + \delta) \rangle dt.
\]
Applying Itô’s formula to \( \langle \hat{X}(\cdot), \bar{Y}(\cdot) \rangle \), we derive
\[
2\langle G\tilde{Y}(s), \bar{Y}(s) \rangle = \mathbb{E} \int_s^T \left[ -2\langle (Q(t) + \tilde{Q}(t + \delta)) \bar{Y}(t), \bar{Y}(t) \rangle
\right.
\]
\[
- \langle \hat{X}(t), C(t) \mathcal{N}^{-1}(t) C(t)^\top \hat{X}(t) + \tilde{C}(t) \mathcal{N}^{-1}(t - \delta) \tilde{C}(t)^\top \hat{X}(t) \rangle
\]
\[
+ 2C(t) \mathcal{N}^{-1}(t) \tilde{C}(t + \delta)^\top \hat{X}(t + \delta) \rangle - 2\langle \hat{Z}(t), (R(t) + \bar{R}(t + \delta)) \bar{Z}(t) \rangle \left. \right] dt
\]
\[
= \mathbb{E} \int_s^T \left[ -2\langle (Q(t) + \tilde{Q}(t + \delta)) \bar{Y}(t), \bar{Y}(t) \rangle - 2\langle \hat{Z}(t), (R(t) + \bar{R}(t + \delta)) \bar{Z}(t) \rangle
\right.
\]
\[
- \langle \mathcal{N}^{-1}(t) [C(t)^\top \hat{X}(t) + \tilde{C}(t + \delta)^\top \hat{X}(t + \delta)], C(t)^\top \hat{X}(t) + \tilde{C}(t + \delta)^\top \hat{X}(t + \delta) \rangle \right] dt.
\]
Due to \( R(\cdot) + \bar{R}(\cdot + \delta) \geq 0, Q(\cdot) + \tilde{Q}(\cdot + \delta) \geq 0, G \geq 0 \) and \( \mathcal{N}(\cdot) > 0 \), it follows that
\[
C(t)^\top \hat{X}(t) + \tilde{C}(t + \delta)^\top \hat{X}(t + \delta) = 0, \quad a.e. \ t \in [s, T], \quad \mathbb{P}\text{-a.s.} \quad (5.8)
\]
Substituting (5.8) into (5.7), we obtain
\[
\begin{cases}
-d\hat{Y}(t) = [A(t)\hat{Y}(t) + \bar{A}(t)\bar{Y}(t - \delta) + B(t)\hat{Z}(t) + \bar{B}(t)\bar{Z}(t - \delta)] dt \\
- \hat{Z}(t) dW(t), \quad t \in [s, T],
\end{cases}
\]
\[
\hat{Y}(T) = 0, \quad \bar{Y}(t) = 0, \quad \hat{Z}(t) = 0, \quad t \in [s - \delta, s).
\]
By the uniqueness of solutions to the linear delayed BSDE (5.9), we have \( (\hat{Y}(\cdot), \hat{Z}(\cdot)) \equiv (0, 0) \). Substituting it into (5.7), we deduce \( \hat{X}(\cdot) \equiv 0 \). Hence the proof is completed. \( \Box \)

**Proposition 6** Let \((X^*(\cdot), Y^*(\cdot), Z^*(\cdot))\) be the solution to the stochastic Hamiltonian system (5.3), suppose \( \bar{A}(t) = \bar{B}(t) = \tilde{C}(t) = 0 \) for \( t \in [s, s + \delta] \) and \( \tilde{Q}(t) = \bar{R}(t) = 0 \).
\[ \dot{N}(t) = 0 \text{ for } t \in [T, T + \delta]. \]

Let
\[ u^*(t) := \mathcal{N}^{-1}(t) \left\{ C(t)^\top X^*(t) + \mathbb{E}\mathcal{F}_t^\delta [(\tilde{C}^\top X^*)_{t+\delta}] \right\}. \tag{5.10} \]

Then \((Y^*(\cdot), Z^*(\cdot))\) is the solution to the state equation (2.1) associated with the control \(u^*(\cdot)\). Furthermore, the corresponding cost can be expressed as
\[ J(u^*(\cdot)) = \mathbb{E} \left\{ \left[ \tilde{G} \varphi(s - \delta), \varphi(s - \delta) \right] + \left[ \Lambda(s), (I + 2\Sigma(s)G)^{-1}G\Lambda(s) \right] \right. \]
\[ + \int_{s - \delta}^s \left[ (\tilde{Q}(t + \delta)\varphi(t), \varphi(t)) + (\tilde{R}(t + \delta)\psi(t), \psi(t)) \right] dt \]
\[ + \int_s^T \left[ ((R(t) + \tilde{R}(t + \delta))[2\Sigma(t)(R(t) + \tilde{R}(t + \delta)) + I]^{-1}\Gamma(t), \Gamma(t)) \right] \]
\[ + \left[ (Q(t) + \tilde{Q}(t + \delta))\Lambda(t), \Lambda(t) \right] \]
\[ + \left[ \tilde{X}(t), \left[ \tilde{B}(t)\Sigma(t - \delta)\mathcal{M}^{-1}(t - \delta)\tilde{B}(t)^\top + \tilde{C}(t)\mathcal{N}^{-1}(t - \delta)\tilde{C}(t)^\top \right][\mathbb{E}\mathcal{F}_t^\delta [\tilde{X}(t)] - \tilde{X}(t)] \right] \right\} dt. \tag{5.11} \]

**Proof** The first conclusion is easy to verify and we mainly prove the second conclusion. Applying Itô’s formula to \((X^*(\cdot), Y^*(\cdot))\), we have
\[ \frac{1}{2} \mathbb{E} \langle X^*(T), \xi \rangle + \mathbb{E} \langle GY^*(s), Y^*(s) \rangle \]
\[ = \frac{1}{2} \mathbb{E} \int_s^T \left[ -2\langle Y^*(t), (Q(t) + \tilde{Q}(t + \delta))Y^*(t) \rangle - \langle X^*(t), C(t)\mathcal{N}^{-1}(t)C(t)^\top X^*(t) \right] \]
\[ + \langle C(t)\mathcal{N}^{-1}(t)\tilde{C}(t + \delta)^\top X^*(t + \delta) + \tilde{C}(t)\mathcal{N}^{-1}(t - \delta)C(t - \delta)^\top X^*(t - \delta) \]) \]
\[ + \langle \tilde{C}(t)\mathcal{N}^{-1}(t - \delta)\tilde{C}(t)^\top X^*(t) \rangle - 2\langle Z^*(t), (R(t) + \tilde{R}(t + \delta))Z^*(t) \rangle \right] dt. \]

Hence the corresponding cost functional associated with the control \(u^*(\cdot)\) becomes
\[ J(u^*(\cdot)) = \mathbb{E} \left\{ \left[ \tilde{G}Y^*(s - \delta), Y^*(s - \delta) \right] - \frac{1}{2} \langle X^*(T), \xi \rangle \right. \]
\[ + \int_s^T \left[ -\langle Y^*(t), (Q(t) + \tilde{Q}(t + \delta))Y^*(t) \rangle - \frac{1}{2} \langle X^*(t), C(t)\mathcal{N}^{-1}(t)C(t)^\top X^*(t) \right] \]
\[ + \frac{1}{2} \langle C(t)\mathcal{N}^{-1}(t)\tilde{C}(t + \delta)^\top X^*(t + \delta) + \frac{1}{2} \tilde{C}(t)\mathcal{N}^{-1}(t - \delta)C(t - \delta)^\top X^*(t - \delta) \]
\[ + \frac{1}{2} \langle \tilde{C}(t)\mathcal{N}^{-1}(t - \delta)\tilde{C}(t)^\top X^*(t) \rangle - 2\langle Z^*(t), (R(t) + \tilde{R}(t + \delta))Z^*(t) \rangle \]
\[ + \langle Q(t)Y^*(t), Y^*(t) \rangle + \langle \tilde{Q}(t)Y^*(t - \delta), Y^*(t - \delta) \rangle \]
\[
+ \{R(t)Z^*(t), Z^*(t)\} + \{\bar{R}(t)Z^*(t - \delta), Z^*(t - \delta)\} \\
+ \{N(t)u^*(t), u^*(t)\} + [\bar{N}(t)u^*(t - \delta), u^*(t - \delta)]dt \}.
\]

Since \(\bar{Q}(t) = \bar{R}(t) = \bar{N}(t) = 0\) for \(t \in [T, T + \delta]\), we get
\[
\begin{align*}
\mathbb{E} \int_s^T \langle \bar{Q}(t)Y^*(t - \delta), Y^*(t - \delta) \rangle dt \\
= \mathbb{E} \int_s^{s + \delta} \langle \bar{Q}(t + \delta)\varphi(t), \varphi(t) \rangle dt + \mathbb{E} \int_s^T \langle \bar{Q}(t + \delta)Y^*(t), Y^*(t) \rangle dt, \\
\mathbb{E} \int_s^T \langle \bar{R}(t)Z^*(t - \delta), Z^*(t - \delta) \rangle dt \\
= \mathbb{E} \int_s^{s + \delta} \langle \bar{R}(t + \delta)\psi(t), \psi(t) \rangle dt + \mathbb{E} \int_s^T \langle \bar{R}(t + \delta)Z^*(t), Z^*(t) \rangle dt, \\
\mathbb{E} \int_s^T \langle \bar{N}(t)u^*(t - \delta), u^*(t - \delta) \rangle dt \\
= \mathbb{E} \int_s^{s + \delta} \langle \bar{N}(t + \delta)\eta(t), \eta(t) \rangle dt + \mathbb{E} \int_s^T \langle \bar{N}(t + \delta)u^*(t), u^*(t) \rangle dt.
\end{align*}
\]

Noting \(\bar{C}(t) = 0\) for \(t \in [s, s + \delta]\), we have
\[
\begin{align*}
\mathbb{E} \int_s^T \langle X^*(t), \bar{C}(t)N^{-1}(t - \delta)C(t - \delta)^\top X^*(t - \delta) \rangle dt \\
= \mathbb{E} \int_s^T \langle \bar{C}(t + \delta)^\top X^*(t + \delta), N^{-1}(t)C(t)^\top X^*(t) \rangle dt, \\
\mathbb{E} \int_s^T \langle X^*(t), \bar{C}(t)N^{-1}(t - \delta)\bar{C}(t)^\top X^*(t) \rangle dt \\
= \mathbb{E} \int_s^T \langle \bar{C}(t + \delta)^\top X^*(t + \delta), N^{-1}(t)\bar{C}(t + \delta)^\top X^*(t + \delta) \rangle dt.
\end{align*}
\]

Thus it follows that
\[
J(u^*(\cdot)) = \mathbb{E}\left\{ [G\bar{Y}^*(s - \delta), Y^*(s - \delta)] - \frac{1}{2} [X^*(T), \xi] + \int_{s - \delta}^s \left[ \langle \bar{Q}(t + \delta)\varphi(t), \varphi(t) \rangle \\
+ \langle \bar{R}(t + \delta)\psi(t), \psi(t) \rangle + [\bar{N}(t + \delta)\eta(t), \eta(t)] \right] dt \right\}.
\]

Next applying Itô’s formula to \(X^*(\cdot), \Lambda(\cdot)\), we obtain
\[
- \mathbb{E} \langle X^*(T), \xi \rangle = \mathbb{E} \langle X^*(s), \Lambda(s) \rangle + \mathbb{E} \int_s^T \left\{ [\Lambda(t), 2(\bar{Q}(t) + \bar{Q}(t + \delta))(-Y^*(t) \\
+ \Sigma(t)X^*(t))] + [X^*(t), -B(t)[2\Sigma(t)(R(t) + \bar{R}(t + \delta)) + I]^{-1}\Gamma(t) \right\} dt.
\]
- $\tilde{B}(t)[2\Sigma(t - \delta)(R(t - \delta) + \tilde{R}(t)) + I]^{-1}\Gamma(t - \delta) \\
+ [\tilde{B}(t)\Sigma(t - \delta)\mathcal{M}^{-1}(t - \delta)\tilde{B}(t) + \tilde{C}(t)\mathcal{N}^{-1}(t - \delta)\tilde{C}(t)^\top] \\
\times \left[\mathbb{E}^{F_t-i}[\tilde{X}(t) - \tilde{X}(t)] + \left\{\Gamma(t), B(t)^\top X^*(t) + \tilde{B}(t + \delta)^\top X^*(t + \delta) - 2(R(t) + \tilde{R}(t + \delta))Z^*(t)\right\}\right]dt.

(5.13)

Substituting (5.4) and (5.13) into (5.12), we derive (5.11). The proof is complete. □

**Proposition 7** Let $(X^*(\cdot), Y^*(\cdot), Z^*(\cdot))$ be the solution to the stochastic Hamiltonian system (5.3) and $L(\cdot)$, $S(\cdot)$ be the solutions to (3.3) and (3.5), respectively. Suppose $\tilde{A}(t) = \tilde{B}(t) = \tilde{C}(t) = 0$ for $t \in [s, s + \delta] \cup [T, T + \delta]$, then the following relationship holds:

$$X^*(t) = -L(t)Y^*(t) + S(t), \quad t \in [s, T], \ \mathbb{P}\text{-a.s.}$$

(5.14)

**Proof** Let us assume for the time being that $S(\cdot)$ be the solution to the following equation:

$$dS(t) = \begin{cases} 
\left\{A(t)^\top - L(t)B(t)\Sigma(t)\mathcal{M}^{-1}(t)B(t)^\top - L(t)C(t)\mathcal{N}^{-1}(t)C(t)^\top\right\}S(t) \\
+ \mathbb{E}^{F_t}[\tilde{A}(t)S(t)]_{t+\delta} - L(t)B(t)\Sigma(t)\mathcal{M}^{-1}(t)E^{F_t}[\tilde{B}(t)S(t)]_{t+\delta} \\
+ L(t)B(t)\left[2\Sigma(t)[R(t) + \tilde{R}(t + \delta)] + I\right]^{-1}\Gamma(t) \\
- L(t)(B(t)\Sigma(t - \delta)\mathcal{M}^{-1}(t - \delta)(B(t)S(t))_{t+\delta} - L(t)\tilde{B}(t)\Sigma(t - \delta)\mathcal{M}^{-1}(t - \delta) \\
x \tilde{B}(t)^\top S(t) + L(t)\tilde{B}(t)\left[2\Sigma(t - \delta)[R(t - \delta) + \tilde{R}(t)] + I\right]^{-1}\Gamma(t - \delta) \\
- L(t)C(t)\mathcal{N}^{-1}(t)L(t)^\top [(\tilde{C}(t)S(t)]_{t+\delta} - L(t)\tilde{C}(t)(\mathcal{N}^{-1}C(t)^\top S(t))_{t+\delta} \\
- L(t)\tilde{C}(t)\mathcal{N}^{-1}(t - \delta)\tilde{C}(t)^\top S(t) + \left\{L(t)B(t)\Sigma(t)\mathcal{M}^{-1}(t)\tilde{B}(t + \delta)^\top + L(t)C(t)\mathcal{N}^{-1}(t)\tilde{C}(t + \delta)^\top - \tilde{A}(t + \delta)^\top\right\}L(t + \delta)E^{F_t}[Y^*(t + \delta)] \\
+ L(t)\left\{\tilde{B}(t)\Sigma(t - \delta)\mathcal{M}^{-1}(t - \delta)(\mathcal{M}^{-1}L(t))_{t+\delta} - \tilde{A}(t) \\
+ \tilde{C}(t)(\mathcal{N}^{-1}C^\top L(t)]_{t+\delta} \right\}Y^*(t - \delta) - L(t)\tilde{B}(t)\Sigma(t - \delta)\mathcal{M}^{-1}(t - \delta)\tilde{B}(t)^\top \\
+ \tilde{C}(t)\mathcal{N}^{-1}(t - \delta)\tilde{C}(t)^\top [E^{F_t-i}(\tilde{X}(t)) - \tilde{X}(t)]\right\}dt \\
+ \left\{I + L(t)\Sigma(t)\mathcal{M}^{-1}(t)\left\{B(t)^\top S(t) + E^{F_t}[\tilde{B}(t)S(t)]_{t+\delta}\right\}\right\} \\
- \left\{L(t) - 2R(t) - 2\tilde{R}(t + \delta)\right\}[2\Sigma(t)[R(t) + \tilde{R}(t + \delta)] + I]^{-1}\Gamma(t) \\
- I + L(t)\Sigma(t)\mathcal{M}^{-1}(t)\tilde{B}(t)^\top L(t)E^{F_t}[Y^*(t + \delta)] \\
- I + L(t)\Sigma(t)\mathcal{M}^{-1}(t)B(t)^\top L(t)Y^*(t)]dW(t), \quad t \in [s, T],
\end{cases}

(5.15)

Note that (3.5) and (5.15) are very similar. In fact, it will be shown later that (3.5) and (5.15) have the same solutions. However, (5.15) is easy to deal.
Recall we have shown $X^*(\cdot) = \tilde{X}(\cdot)$, which satisfies (3.4). Applying Itô’s formula to $L(\cdot)Y^*(\cdot)$ and using (5.4), (3.4), (5.10), we derive

$$
d[−X^*(t) − L(t)Y^*(t) + S(t)] = \left\{ \begin{array}{l} A(t)^\top − L(t)B(t)S(t)M^{-1}(t)B(t)^\top − L(t)C(t)N^{-1}(t)C(t)^\top \\
−L(t)\tilde{B}(t)\Sigma(t − δ)M^{-1}(t − δ)\tilde{B}(t)^\top − L(t)\tilde{C}(t)N^{-1}(t − δ)\tilde{C}(t)^\top \\
x[−X^*(t) − L(t)Y^*(t) + S(t)] \\
+ [−L(t)B(t)S(t)M^{-1}(t)\tilde{B}(t + δ)^\top − L(t)C(t)N^{-1}(t)\tilde{C}(t + δ)^\top \\
+ \tilde{A}(t + δ)^\top]E_{\tilde{T}}[(-X^* − LY^* + S)_{|t + δ}] \\
+ [−L(t)\tilde{B}(t)\Sigma(t − δ)M^{-1}(t − δ)B(t − δ)^\top − L(t)\tilde{C}(t)N^{-1}(t − δ)C(t − δ)^\top \\
x[(-X^* − LY^* + S)_{|t − δ}]] \right\} dt
\right.
$$

and $−X^*(s) − L(s)Y^*(s) + S(s) = 2GY^*(s) − 2GY^*(s) = 0$. Let $X^*(t) = L(t) = 0$ for $t \in [s − δ, s]$, then we have $−X^*(t) − L(t)Y^*(t) + S(t) = 0$ for $t \in [s − δ, s]$. Let $L(t) = 0$ for $t \in (T, T + δ]$, then it yields $−X^*(t) − L(t)Y^*(t) + S(t) = 0$ for $t \in (T, T + δ]$. Thus by the unique solvability of the ASDDE (5.16), we obtain

$$
−X^*(t) − L(t)Y^*(t) + S(t) = 0, \quad t \in [s, T],
$$

where $S(\cdot)$ is the solution to (5.15). Substituting (5.17) into (5.4), it follows that

$$
Y^*(t) = (I + \Sigma(t)L(t))^{-1}(\Sigma(t)S(t) + \Lambda(t)), \quad t \in [s, T].
$$

Finally, substituting (5.18) into (5.15), it follows that $S(\cdot)$ is the unique solution to (3.5). The proof is complete.

Finally we give the proof of Theorem 4.

**Proof** It apparently follows from Theorem 3, Propositions 6 and 7.

### 6 Concluding Remarks

In this paper, we have discussed the D-BSLQ optimal control problem. It is, in fact, an optimal control problem where the state equation is a controlled linear delayed BSDE and the cost is a quadratic functional. This kind of optimal control problem has three interesting characteristics worthy of being emphasized. Firstly, the necessary conditions for the existence of the optimal control are derived, where the optimal control is a linear feedback of the entire past history and the expected value of the future state trajectory in a short period of time, and the optimal cost is expressed by a
delayed Riccati equation and a DAFBSDE, which is different from the BSLQ optimal control problem without delay (Lim and Zhou [15]) and has not been considered in the literatures with delay (for example, Chen and Huang [1]). Secondly, a new class of ASDDEs is introduced to seek the state feedback expression of the optimal control, which has not been studied yet. Thirdly, the existence and uniqueness of the delayed Riccati equations and the DAFBSDEs mentioned above are discussed in detail, which have not appeared in the previous literature.

Differential games of delayed BSDEs and forward-backward SDEs (FBSDEs, for short) are well worth studying, considering both Nash and Stackelberg equilibria ([27–30,33,36]). The solvability of related delayed Riccati equations are more complicated, and some numerical methods are desirable. We will consider these topics in the future.

Acknowledgements The authors would like to thank the editor and two anonymous referees for their constructive and insightful comments for improving the quality of this work. Many thanks for discussion and suggestions with Professor Li Chen at China University of Mining and Technology (Beijing).

References

1. Chen, L., Huang, J.H.: Stochastic maximum principle for controlled backward delayed system via advanced stochastic differential equation. J. Optim. Theory Appl. 167, 1112–1135 (2015)
2. Chen, L., Wu, Z.: Maximum principle for the stochastic optimal control problem with delay and application. Automatica 46, 1074–1080 (2010)
3. Chen, L., Wu, Z., Yu, Z.Y.: Delayed stochastic linear-quadratic control problem and related applications. J. Appl. Math. (2012)
4. Delong, L.: Applications of time-delayed backward stochastic differential equations to pricing, hedging and management of financial and insurance risks. Appl. Math. 39, 463–488 (2012)
5. Delong, L., Imkeller, P.: Backward stochastic differential equation with time delayed generators-results and counterexamples. Ann. Appl. Probab. 20, 1512–1536 (2010)
6. Du, H., Huang, J.H., Qin, Y.L.: A stochastic maximum principle for delayed mean-field stochastic differential equations and its applications. IEEE Trans. Autom. Control 58, 3212–3217 (2013)
7. Dokuchaev, N., Zhou, X.Y.: Stochastic controls with terminal contingent conditions. J. Math. Anal. Appl. 238, 143–165 (1999)
8. Gopalsamy, K.: Nonoscillation in systems of linear differential equations with delayed and advanced arguments. J. Math. Anal. Appl. 140, 374–380 (1989)
9. Huang, J.H., Li, X., Shi, J.T.: Forward-backward linear quadratic stochastic optimal control problem with delay. Syst. Control Lett. 61, 623–630 (2012)
10. Huang, J.H., Shi, J.T.: Maximum principle for optimal control of fully coupled forward-backward stochastic differential delayed equations. ESAIM. COCV 18, 1073–1096 (2012)
11. Huang, J.H., Wang, G.C., Xiong, J.: A maximum principle for partial information backward stochastic control problem with applications. SIAM J. Control Optim. 48, 2106–2117 (2009)
12. Li, N., Wang, Y., Wu, Z.: An indefinite stochastic linear quadratic optimal control problem with delay and related forward-backward stochastic differential equations. J. Optim. Theory Appl. 179, 722–744 (2018)
13. Li, N., Wang, G.C., Wu, Z.: Linear quadratic optimal control problem for time-delay stochastic system with recursive utility under full and partial information. Automatica 121, 109169 (2020)
14. Li, X., Sun, J.R., Xiong, J.: Linear quadratic optimal control problems for mean-field backward stochastic differential equations. Appl. Math. Optim. 80, 2123–250 (2019)
15. Lim, A.E.B., Zhou, X.Y.: Linear-quadratic control of backward stochastic differential equations. IEEE Trans. Autom. Control 40, 450–474 (2001)
16. Li, X.Y., Zhu, D.M.: Oscillation and nonoscillation of advanced differential equations with variable coefficients. J. Math. Anal. Appl 269, 462–488 (2002)
17. Mohammed, S.E.A.: Stochastic Functional Differential Equations. Pitman, Boston (1984)
18. Mohammed, S.E.A.: Stochastic differential equations with memory: theory, examples and applications. Progress in Probability, Stochastic Analysis and Related Topics 6. Birkhäuser, The Geido Workshop (1996)

19. Øksendal, B., Sulem, A.: A maximum principle for optimal control of stochastic systems with delay, with applications to finance. In: Menaldi, J.M., Rofman, E., Sulem, A. (eds.) Optimal Control and Partial Differential Equations, pp. 64–79. IOS Press, Amsterdam (2000)

20. Øksendal, B., Sulem, A., Zhang, T.S.: Optimal control of stochastic delay equations and time-advanced backward stochastic differential equations. Adv. Appl. Probab. 43, 572–596 (2011)

21. Pardoux, E., Peng, S.G.: Adapted solution of a backward stochastic differential equation. Syst. Control Lett. 14, 55–61 (1990)

22. Peng, S.G.: Backward stochastic differential equations and applications to optimal control. Appl. Math. Optim. 27, 125–144 (1993)

23. Peng, S.G., Yang, Z.: Anticipated backward stochastic differential equations. Ann. Probab. 37, 877–902 (2009)

24. Shi, J.T.: Optimal control of backward stochastic differential equations with time delayed generators. In: Proc. 30th Chinese Control Conf., pp. 1285–1289, Yantai, July 22–24 (2011)

25. Shi, J.T.: Optimal control of BSDEs with time delayed generators driven by Brownian motions and Poisson random measures. In: Proc. 32nd Chinese Control Conf., pp. 1575–1580, Xi’an, July 26–28 (2013)

26. Shan, S.M., Wiener, J.: Advanced differential equations with piecewise constant argument deviations. Inter. J. Math. Math. Sci. 6, 671–703 (2007)

27. Shi, J.T., Wang, G.C.: A nonzero sum differential game of BSDE with time-delayed generator and applications. IEEE Trans. Autom. Control 61, 1959–1964 (2016)

28. Wang, G.C., Xiao, H., Xiong, J.: A kind of LQ non-zero sum differential game of backward stochastic differential equations with asymmetric information. Automatica 97, 346–352 (2018)

29. Wang, G.C., Yu, Z.Y.: A Pontryagins maximum principle for non-zero sum differential games of BSDEs with applications. IEEE Trans. Autom. Control 55, 1742–1747 (2010)

30. Wang, G.C., Yu, Z.Y.: A partial information non-zero sum differential game of backward stochastic differential equations with applications. Automatica 48, 342–352 (2012)

31. Wu, S., Shu, L.: Partially observed linear quadratic control problem with delay via backward separation method. Optim. Control Appl. Meth. 38, 814–828 (2017)

32. Wu, S., Wang, G.C.: Optimal control problem of backward stochastic differential delay equation under partial information. Syst. Control Lett. 82, 71–78 (2015)

33. Xu, J.J., Shi, J.T., Zhang, H.S.: A leader-follower stochastic linear quadratic differential game with time delay. Sci. China Inf. Sci. 61, 112202:1–112202:13 (2018)

34. Yong, J.M., Zhou, X.Y.: Stochastic Controls: Hamiltonian Systems and HJB Equations. Springer-Verlag, New York (1999)

35. Yu, Z.Y.: The stochastic maximum principle for optimal control problems of delay systems involving continuous and impulse controls. Automatica 48, 2420–2432 (2012)

36. Yu, Z.Y., Ji, S.L.: Linear-quadratic non-zero sum differential game of backward stochastic differential equations. In: Proc. 27th Chinese Control Conference, pp. 562–566, Kunming, China, July 16–18 (2008)

37. Zhang, H.S., Xu, J.J.: Control for Itô stochastic systems with input delay. IEEE Trans. Autom. Control 62, 350–365 (2017)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.