Closed-loop solvability of delayed control problems: A stochastic Volterra system approach *

Weijun Meng† Tianxiao Wang‡§ Ji-Feng Zhang¶

September 30, 2025

Abstract: A general and new stochastic linear quadratic optimal control problem is studied, where the coefficients are allowed to be time-varying, and both state delay and control delay can appear simultaneously in the state equation and the cost functional. The closed-loop outcome control of this delayed problem is given by a new Riccati system whose solvability is carefully established. To this end, a novel method is introduced to transform the delayed problem into a control problem driven by a stochastic Volterra integral system without delay. This method offers several advantages: it bypasses the difficulty of decoupling the forward delayed state equation and the backward anticipated adjoint equation, avoids the introduction of infinite-dimensional spaces and unbounded control operators, and ensures that the closed-loop outcome control depends only on past state and control, without relying on future state or complex conditional expectation calculations. Finally, several particular important stochastic systems are discussed. It is found that the model can cover a class of stochastic integro-differential systems, whose closed-loop solvability has not been available before.

Keywords: Stochastic delay optimal control, time-varying coefficients, closed-loop solvability, Riccati equation

Mathematics Subject Classification: 93E20, 60H10, 34K50, 49N10

1 Introduction

1.1 Delayed optimal control problems

In this paper, we consider an optimal control problem where the state equation is described as:

$$\begin{cases} dx(t) = \big[A_1(t)x(t) + A_2(t)y(t) + A_3(t)z(t) + B_1(t)u(t) + B_2(t)\nu(t) + B_3(t)\mu(t) + b(t)\big]dt \\ \qquad\qquad +\big[C_1(t)x(t) + C_2(t)y(t) + C_3(t)z(t) + D_1(t)u(t) + \sigma(t)\big]dW(t), \quad t \in (t_0, T), \\ x(t) = \xi(t-t_0), \quad u(t) = \varsigma(t-t_0), \quad t \in [t_0-\delta, t_0], \end{cases} \tag{1.1}$$

and the cost functional is defined as:

$$J(t_0, \xi(\cdot), \varsigma(\cdot); u(\cdot)) = \mathbb{E}\int_{t_0}^{T}\Big[x(t)^\top Q_1(t)x(t) + y(t)^\top Q_2(t)y(t) + z(t)^\top Q_3(t)z(t) + u(t)^\top R_1(t)u(t) + \nu(t)^\top R_2(t)\nu(t)\Big]dt. \tag{1.2}$$

*This work is supported by the China National Key Research and Development Program (2024YFA1012800), the National Natural Science Foundation of China (62433020, 12371449, 12501618, 12471419).

†*School of Mathematics & Statistics, Nanjing University of Science and Technology, Nanjing 210094, China, E-mail: mengwj@njust.edu.cn*

‡*School of Mathematics, Sichuan University, Chengdu 610065, China, E-mail: wtxiao2014@scu.edu.cn*

§*Corresponding author.*

¶*School of Automation and Electrical Engineering, Zhongyuan University of Technology, Zhengzhou 450007; State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, and school of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100149, China, E-mail: jif@iss.ac.cn*

Here $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a complete filtered probability space, $\mathbb{E}[\cdot]$ denotes the expectation, and the filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t\geqslant 0}$ is generated by a one-dimensional standard Brownian motion $\{W(t)\}_{t\geqslant 0}$. In the above, $x(\cdot) \in \mathbb{R}^n$ is state, $u(\cdot) \in \mathbb{R}^m$ is control, $y(\cdot)$ and $\nu(\cdot)$ are the pointwise delays of the state and the control, respectively, $z(\cdot)$ and $\mu(\cdot)$ are the corresponding (extended) distributed delays. They are defined as follows:

$$y(t) \equiv x(t-\delta), \quad z(t) \equiv \int_{t_0}^{t} F(t,s)x(s)ds, \ \ \nu(t) \equiv u(t-\delta), \ \ \mu(t) \equiv \int_{t_0}^{t} \tilde{F}(t,s)u(s)ds, \quad t \in (t_0, T). \tag{1.3}$$

In addition, $\xi(\cdot)$ and $\varsigma(\cdot)$ are called the initial trajectories of the state and the control, respectively. Under proper conditions (see Section 2), the cost functional (1.2) is well-defined, and the optimal control problem is stated as follows:

Problem (P). To find a control $u^*(\cdot)$ such that (1.1) is satisfied and (1.2) is minimized, i.e.,

$$J(t_0, \xi(\cdot), \varsigma(\cdot); u^*(\cdot)) = \inf_{u(\cdot)\in L^2_{\mathbb{F}}(t_0,T;\mathbb{R}^m)} J(t_0, \xi(\cdot), \varsigma(\cdot); u(\cdot)) = V(t_0, \xi(\cdot), \varsigma(\cdot)),$$

where $L^2_{\mathbb{F}}(t_0, T; \mathbb{R}^m)$ is the Hilbert space consisting of $\mathbb{F}$-adapted processes $\phi(\cdot)$ such that $\mathbb{E}\int_0^T |\phi(t)|^2 dt < \infty$. Any $u^*(\cdot) \in L^2_{\mathbb{F}}(t_0, T; \mathbb{R}^m)$ that achieves the above infimum is called an *open-loop optimal control* for $(t_0, \xi(\cdot), \varsigma(\cdot))$ and the corresponding solution $x^*(\cdot)$ is called the *open-loop optimal trajectory*. $(x^*(\cdot), u^*(\cdot))$ is called an *open-loop optimal pair*. $V(t_0, \xi(\cdot), \varsigma(\cdot))$ is called the value function. In the special case, when $b(\cdot)$ and $\sigma(\cdot)$ vanish, we denote the corresponding delayed linear quadratic (LQ) problem, the cost functional, and the value function by Problem (P_0), $J_0(t_0, \xi(\cdot), \varsigma(\cdot); u(\cdot))$ and $V_0(t_0, \xi(\cdot), \varsigma(\cdot))$, respectively.

As to the above (1.1), it is an extension of the following stochastic differential delay equation (SDDE):

$$\begin{cases} dx(t) = \big[A_1(t)x(t) + A_2(t)y(t) + A_3(t)\widetilde{z}(t) + B_1(t)u(t) + B_2(t)\nu(t) + B_3(t)\widetilde{\mu}(t) + \widetilde{b}(t)\big]dt \\ \qquad\qquad + \big[C_1(t)x(t) + C_2(t)y(t) + C_3(t)\widetilde{z}(t) + D_1(t)u(t) + \widetilde{\sigma}(t)\big]dW(t), \quad t \in (t_0, T), \\ x(t) = \xi(t-t_0), \quad u(t) = \varsigma(t-t_0), \quad t \in [t_0-\delta, t_0], \\ \widetilde{z}(t) \equiv \displaystyle\int_{t-\delta}^{t} G_1(t,s)x(s)ds, \ \ \widetilde{\mu}(t) \equiv \int_{t-\delta}^{t} G_2(t,s)u(s)ds, \quad t \in (t_0, T), \end{cases} \tag{1.4}$$

where $y(\cdot)$ and $\nu(\cdot)$ are defined in (1.3), and G_1, G_2 are bounded. In fact,

$$\widetilde{z}(t) = \widetilde{G}_1(t) + \int_{t_0}^{t} \widehat{G}_1(t,s)x(s)ds, \quad \widetilde{\mu}(t) = \widetilde{G}_2(t) + \int_{t_0}^{t} \widehat{G}_2(t,s)u(s)ds,$$

where

$$\begin{aligned} \widetilde{G}_1(t) &\equiv \int_{t-\delta}^{t_0} G_1(t,s)\xi(s-t_0)ds\mathbf{1}_{[t_0,t_0+\delta)}(t), \\ \widetilde{G}_2(t) &\equiv \int_{t-\delta}^{t_0} G_2(t,s)\varsigma(s-t_0)ds\mathbf{1}_{[t_0,t_0+\delta)}(t), \\ \widehat{G}_1(t,s) &\equiv \big[\mathbf{1}_{[t_0,t_0+\delta)}(t) + \mathbf{1}_{[t_0+\delta,T]}(t)\mathbf{1}_{[t-\delta,t)}(s)\big]G_1(t,s), \\ \widehat{G}_2(t,s) &\equiv \big[\mathbf{1}_{[t_0,t_0+\delta)}(t) + \mathbf{1}_{[t_0+\delta,T]}(t)\mathbf{1}_{[t-\delta,t)}(s)\big]G_2(t,s). \end{aligned}$$

Therefore, (1.4) can be seen as a special case of (1.1) with

$$b = \widetilde{b} + A_3\widetilde{G}_1 + B_3\widetilde{G}_2, \quad \sigma = \widetilde{\sigma} + C_3\widetilde{G}_1, \quad F = \widehat{G}_1, \quad \tilde{F} = \widehat{G}_2.$$

Based on this fact, in the following we name (1.1) *the (extended) controlled SDDE*. More details about SDDE (1.4) can be referred to [45, 46]. Notice that here we allow $t_0 + \delta \leqslant T$.

1.2 Motivations

In the real world, many challenges across disciplines like economics, finance, aerospace, and network communication can be framed as optimal control problems [10, 12, 22]. Moreover, the evolution of certain phenomena hinges not only on present conditions but also on their historical trajectories. Consequently, optimal control problem of stochastic control systems containing state delays and control delays, like the above (1.4), is an important issue in control theory. The relevant optimal control problems have attracted enormous attention of the optimization and engineering communities in the last decades. We refer to the monographes by Bensoussan–Da Prato–Delfour–Mitter [6], Chang [11], Hale [27], Kolmanovskii–Myshkis [32], Kolmanovskii–Shaikhet [33], Meng–Shi–Yong [43], Zhang–Xie [58].

To motivate the current study, let us make some careful comments to the existing papers from the viewpoints of frameworks, methodologies, and conclusions.

(I) The existing frameworks on stochastic delayed systems seem scattered and decentralized, many of which can not fully cover each other.

• Time-invariant versus time-varying coefficients: Such a difference happens even for the deterministic controlled system. For example, [35, 37, 38, 40, 41, 44, 57, 59] and the relevant papers therein are devoted to the time-invariant case, while [3, 4, 5, 14, 19, 20, 39, 60] and the relevant papers therein are concerned with the time-varying one. We observe that the appearance of both cases in the literature is attributed to the complexity of delayed problems and the methodology limitation developed accordingly.

• State delay versus control delay: We found that the inclusion of delayed terms indeed brings significant differences to the corresponding study. When only state variables contain the delayed terms, we refer to [17, 19, 20, 31] for the deterministic case and [24, 25, 38, 39] for the stochastic case. On the other hand, when delayed terms are only added to control variables, we refer to [31, 56] for the deterministic case and [37, 38, 54, 59] for the stochastic case. In fact, the reason for such a difference lies in the fact that the methods used to handle state delay and control delay are fundamentally different.

• Pointwise delay versus distributed delay: We found that delay types also affect the analogue study. For example, [14, 20, 37, 38, 59] are concerned with the pointwise delay in state or control, while [12, 24] are devoted to the case with distributed delay. The poinwise delay is usually more challenging than the distributed delay because the distributed delay can be dealt with using the derivative formula or the Itô formula, but the pointwise delay cannot.

• State (or control) dependence versus constants in diffusion terms: In the literature, we found that whether the diffusion terms depend on the state/control variables, or even their delayed terms, brings essential differences. We refer to [21, 24, 37, 59] for the former case while [4, 10, 12] for the simple constant case. In our opinion, the related reason is that this dependence occurs within the stochastic integral terms, causing the previously used methods for handling Lebesgue integral terms to fail.

• Delayed terms in cost functionals or not: Eventually, let us point out the different frameworks caused by the dependence of delayed terms in cost functionals. Even for deterministic controlled systems, pointwise delay in cost functionals makes weight operators unbounded, while distributed delay makes their processing highly dependent on the given weight coefficients (see the Introduction part of [35] for detailed explanations). When cost functionals contain state/control delay terms, we refer to [35, 42, 43, 59], while most other existing studies do not touch this topic.

Based on the above classifications of the models, we pose the first question:

(Q1): *Is it possible to provide the study in general framework that covers the above models?*

(II) After careful observations of the existing papers, we found that there are several approaches treating LQ problems for stochastic delay systems. We list them out and make some proper comments as follows.

• Maximum principle: It is a very natural method to study optimal control for stochastic delayed systems. We refer to [15, 55, 59] for more details. To obtain the explicit forms of optimal controls, one has to decouple the resulting forward-backward stochastic systems, which is the challenging part. A success example is [59], where the system only contains pointwise control delays, and a class of Riccati equation is established by decoupling. The difficulty in decoupling lies in the fact that the delay effect disrupts the classical decoupling relation and makes Itô formula fail to work [5].

• Dynamic programming: This method is very effective in obtaining feedback forms of optimal controls, see [3, 6, 34, 37, 38] for more details. However, the main difficulty for delayed problems is the lack of Markovianity, and one has to pay the cost of working with intrinsically infinite-dimensional HJB equation.

• State space method: This method is to lift state equations to Banach or Hilbert spaces (depending on the regularity of the data). It is developed in the deterministic case by [19, 31] and then extended in the stochastic case by [23, 24, 25, 44]. This method always comes at a cost of moving problems to infinite dimension. Another limitation in the stochastic case lies in the appearance of unbounded operators in state equations [44].

• Variation of constants formula method: By using this method, one can transform the original delayed system into another integral system of Volterra type without delay which is solved by fundamental solution. We refer to [30, 35, 36, 42] for more details. This works well in deterministic controlled systems. However, when diffusion terms contain state delay terms, it becomes unclear to go through the whole procedure since anticipated stochastic integral may be involved.

After the above analysis, different methods are used to handle different models, but each has its own drawbacks. Therefore, we raise the second question:

(Q2)*: Is it possible to come up with a new approach to bypass the disadvantages in the existing papers?*

(III) Last but not the least important, let us make some comments on the obtained conclusions in terms of feedback controls and Riccati systems in the literature. We separate them into the following cases.

• Riccati systems versus Fredholm systems: To obtain feedback controls, Riccati systems are one of the most popular and important tools. We refer to [3, 4, 10, 13, 15, 16, 18, 37, 43, 44, 53, 59] and the related papers therein. On the other hand, according to the works [35, 36], Fredholm systems also play important roles. In other words, the helpful systems for establishing the feedbacks are by no means unique.

• The challenges of Riccati systems: This point can be explained in the following manners. Firstly, the successful introduction of Riccati systems only hold in specific settings, e.g., the time-invariant case [31, 37, 44, 59], the only state delay case [15, 20, 38], the only control delay case [10, 38, 59]. Secondly, even for the papers containing Riccati systems, verifying the coincidences among them seems quite involved and technical, such as [15] and [44], or [37] and [59]. Thirdly, even though Riccati systems are derived, they are quite challenging to further discuss their solvability. Along this line we mention the works of [15, 44] for some progress in the stochastic setting. Fourthly, as to the papers obtaining the solvability, different works require different assumptions [3, 17]. Hence it seems difficult to give a unified precondition.

• The challenges of feedback controls: In the literature, open-loop optimal controls and closed-loop optimal controls are two main notions for representing the explicit forms. For the former one, there are some relevant papers [13, 44, 55, 60]. Since it is not easy to decouple forward-backward adjoint systems, the closed-loop representations of open-loop optimal controls are not successfully given until some recent works of [15, 44, 59]. As to the later one, there are some works treating the closed-loop solvability. We refer to [43, 44]. However, in our opinion, there are some limitations. Their closed-loop solvability depends on the transformed problems, thus lacks generality, and their methods are not applicable to time-varying coefficients.

To sum up the above, we observe that that different models have different version of Riccati system, among which the difference analysis is lacking and the relevant solvability is not obtained. Therefore, we come up with the third question:

(Q3)*: Is it possible to derive a unified Riccati system and feedback control for the general model* (1.1) *and* (1.2)*, and verify the consistency with the existing papers?*

Now let us return back to Problem (P) and discuss another special case where $A_2, B_2, C_2, Q_2, R_2 = 0$. Here we then arrive at the LQ optimal control problem for (a class of) stochastic integro-differential equations (or stochastic differential equations with memory). This type of optimal control problems naturally emerge in many different application scenarios, and they are particularly common when studying the optimal performances of systems in response to specific inputs. In such cases, the systems' responses do not occur immediately but rather appear after a certain period of time. However, in contrast with the previous discussion on SDDEs, for optimal control problems involving equations with memory, just a few isolated results are available at the moment. We refer to Carlier–Tahraoui [8, 9] for the finite dimensional case and Cannarsa–Frankowska–Marchini [7] for the infinite dimensional study. In the LQ framework, Pandolfi [48] studied a simple integro-differential model in the finite-dimensional space and obtained the optimal synthesis via Riccati-type equations for the first time (commented by [2]). Other enhanced results appeared recently in [26, 49]. For the extension to the infinite dimension, we refer to Acquistapace–Bucci [1, 2]. These results hold in the deterministic controlled system framework. Due to the lack of relevant study on stochastic systems, we raise the fourth question:

(Q4)*: It is possible to provide some systematic study in stochastic integro-differential systems to fill the gaps left by the existing literature?*

1.3 Contributions and novelties

In this paper, our goal is to treat the general framework with the state equation (1.1) and the cost functional (1.2) by employing new methodologies, and give positive answers to the aforementioned four questions. To begin with, we propose a general definition of the closed-loop solvability for Problem (P). To show its sufficiency, we separate the procedures into four parts. We firstly adapt the transformation procedures, introduced in Meng–Shi–Wang–Zhang [42], into our framework and end up with a LQ problem for a stochastic Volterra integral system without delay. Secondly, by borrowing the ideas developed in Hamaguchi–Wang [29] and Gong–Wang [26], we introduce and discuss a class of Riccati systems and backward stochastic adjoint systems. Then, by the previous Riccati systems, we explicitly construct the desired closed-loop strategy, the resulting closed-loop outcome control, and prove its optimality. Finally, we make detailed comparisons/coincidences with the existing study. The contributions and innovations of this paper are summarized as follows.

- A general yet new stochastic LQ optimal control problem is studied. On the one hand, it covers stochastic differential delay system where the coefficients are time-varying, both state delay and control delay can appear in the drift term, the diffusion term, and the cost functional. On the other hand, it covers a class of stochastic integro-differential equations where memory terms enter into the drift, diffusion terms, and the cost functional. This gives a nice response to (Q1), (Q4).
- A new transformation method is employed for the above general framework, and the advantages are shown in three aspects. Firstly, it avoids the complicated decoupling procedures for forward-backward systems. Secondly, it can handle time-varying coefficients case where traditional infinite dimensional lifting methods fail. Thirdly, in our approach there are no unbounded control operators, and there is no need to use infinite-dimensional analysis theories such as operator semigroups. This gives a nice response to (Q2).
- By our main conclusions, we find the following four advantages and new facts. Firstly, we give the solvability of the Riccati system corresponding to Problem (P) and their coincidences with the existing literature in particular cases. Secondly, we explicitly construct the unique closed-loop outcome control, which does not rely on the future state and avoids complex tools of conditional expectations. Then, we only impose integrability conditions on the coefficients of Problem (P), and there is no need of continuity or even differentiability assumptions. Finally, even for the deterministic control system (1.1), we present, for the first time, results regarding the integro-differential part and the cost functional (1.2) with pointwise/distributed delays. This gives a nice response to (Q3), (Q4).

The rest part is organized as follows. In Section 2, we formulate the control problem studied in the paper. In Section 3, we transform the original delayed control system into a Volterra integral control system without delay, and then study the closed-loop solvability of the original delayed optimal control problem. In Section 4, we discuss several important cases to compare our results with the previous ones. In Section 5, we give some concluding remarks. Finally, we provide the proofs of the main results in Appendix.

2 Preliminary

In this section, we give some preliminaries and formulate the control problem studied in this paper.

Let $T>0$ be a given finite time duration, define $\triangle_2(0,T) \equiv \left\{(t,s)\in(0,T)^2 \mid T>t>s>0\right\}$, $\square_3(0,T) \equiv \left\{(s_1,s_2,t)\in(0,T)^3 \mid t<(s_1\wedge s_2)\right\}$. I is the identity matrix with appropriate dimension. $\mathbb{S}^n$ is the set of all $n\times n$ symmetric matrices. Next we define the following spaces which will be used in this paper. Denote by $L^\infty(0,T;\mathbb{R}^n)$ the Banach space consisting of $\mathbb{R}^n$-valued variables $\phi(\cdot)$ such that $\sup\limits_{0\leqslant t\leqslant T}|\phi(t)|<\infty$, by $L^p(0,T;\mathbb{R}^n)$ the Banach space consisting of $\mathbb{R}^n$-valued variables $\phi(\cdot)$ such that $\int_0^T|\phi(t)|^p dt<\infty$, where p is an integer. Denote by $L^2_{\mathcal{F}_t}(\Omega;\mathbb{R}^n)$ the Hilbert space consisting of $\mathbb{R}^n$-valued $\mathcal{F}_t$-measurable random variables ϕ such that $\mathbb{E}|\phi|^2<\infty$, by $L^2_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$ the Banach space consisting of $\mathbb{R}^n$-valued

$\mathbb{F}$-adapted continuous processes $\phi(\cdot)$ such that $\mathbb{E}\big[\sup_{0\leqslant t\leqslant T}|\phi(t)|^2\big] < \infty$, by $L^{2,p}_{\mathbb{F}}(\triangle_2(0,T);\mathbb{R}^{n\times n})$ the Banach space of $\mathbb{R}^{n\times n}$-valued and measurable processes ϕ on $\triangle_2(0,T)$ such that $\phi(t,\cdot)$ is $\mathbb{F}$-progressively measurable on $(0,t)$ for each $t\in(0,T)$, and $\mathbb{E}\big[\int_0^T\Big(\int_0^t|\phi(t,s)|^pds\Big)^{\frac{2}{p}}dt\big]^{\frac12} < \infty$. For $p=2$, we simply denote $L^2_{\mathbb{F}}(\triangle_2(0,T);\mathbb{R}^{n\times n}) \equiv L^{2,2}_{\mathbb{F}}(\triangle_2(0,T);\mathbb{R}^{n\times n})$. Denote by $\mathscr{L}^2(\triangle_2(0,T);\mathbb{R}^{n\times n})$ the set of $\phi\in L^2(\triangle_2(0,T);\mathbb{R}^{n\times n})$ satisfying $\operatorname*{ess\,sup}_{t\in(0,T)}\Big(\int_t^T|\phi(s,t)|^2ds\Big)^{\frac12}<\infty$, and for any $\varepsilon>0$, there exists a finite partition $\{a_i\}_{i=0}^m$ of $(0,T)$ with $0=a_0<a_1<\cdots<a_m=T$ such that $\operatorname*{ess\,sup}_{t\in(a_i,a_{i+1})}\big(\int_t^{a_{i+1}}|\phi(s,t)|^2ds\big)^{\frac12}<\varepsilon$, for each $i\in\{0,1,\ldots,m-1\}$. Denote by $L^{2,2,1}(\square_3(0,T);\mathbb{R}^{n\times n})$ the Banach space of $\mathbb{R}^{n\times n}$-valued deterministic functions ϕ on $\square_3(0,T)$ such that $(\int_0^T\int_0^T(\int_0^{s_1\wedge s_2}|\phi(s_1,s_2,t)|dt)^2ds_1ds_2)^{1/2}<\infty$.

For any $0<t_0<T$, consider the following SDDE:

$$\begin{cases} d\bar{x}(t)=[\bar{A}_1(t)\bar{x}(t)+\bar{A}_2(t)\bar{y}(t)+\bar{A}_3(t)\bar{z}(t)+\bar{b}(t)]dt \\ \qquad\qquad +[\bar{C}_1(t)\bar{x}(t)+\bar{C}_2(t)\bar{y}(t)+\bar{C}_3(t)\bar{z}(t)+\bar{\sigma}(t)]dW(t), \quad t\in(t_0,T), \\ \bar{x}(t)=\bar{\xi}(t-t_0), \quad t\in[t_0-\delta,t_0], \end{cases} \tag{2.1}$$

where $\bar{\xi}(\cdot)\in C([-\delta,0];\mathbb{R}^n)$ is the initial trajectory of the state, $\delta>0$ is the constant delay time,

$$\bar{y}(t)\equiv\bar{x}(t-\delta), \quad \bar{z}(t)\equiv\int_{t_0}^t\bar{F}(t,s)\bar{x}(s)ds, \quad t\in(t_0,T).$$

As to the coefficients of (2.1), we impose the following conditions:

(H1) $\bar{A}_1(\cdot),\bar{A}_2(\cdot),\bar{A}_3(\cdot)\in L^2(0,T;\mathbb{R}^{n\times n})$, $\bar{C}_1(\cdot),\bar{C}_2(\cdot),\bar{C}_3(\cdot)\in L^\infty(0,T;\mathbb{R}^{n\times n})$,
$\bar{F}(\cdot,\cdot)\in L^{2,1}(\Delta_2(0,T);\mathbb{R}^{n\times n})$, $\bar{b}(\cdot)\in L^2_{\mathbb{F}}(\Omega;L^1(0,T;\mathbb{R}^n))$, $\bar{\sigma}(\cdot)\in L^2_{\mathbb{F}}(0,T;\mathbb{R}^n)$.

The following proposition, whose proof is given in the Appendix, guarantees its solvability. In contrast with e.g. [13, 14, 15, 60] where $\bar{A}_i(\cdot)$ is assumed to be bounded, $i=1,2,3$, we slightly relax them into proper integrable conditions.

Proposition 2.1. *Let* (H1) *hold. Then, SDDE* (2.1) *admits a unique solution* $\bar{x}(\cdot)\in L^2_{\mathbb{F}}\big(\Omega;C([t_0,T];\mathbb{R}^n)\big)$.

For later usefulness, for any $0<t_0<T$, we consider the linear SVIE:

$$\check{X}(t)=\check{\varphi}(t)+\int_{t_0}^t\Big[\check{A}(t,s)X(s)+\check{b}(t,s)\Big]ds+\int_{t_0}^t\Big[\check{C}(t,s)X(s)+\check{\sigma}(t,s)\Big]dW(s), \quad t\in(t_0,T), \tag{2.2}$$

with the coefficients satisfying:

(H2) $\check{\varphi}(\cdot)\in L^2_{\mathbb{F}}(0,T;\mathbb{R}^n)$, $\check{A}(\cdot,\cdot)\in L^2(\Delta_2(0,T);\mathbb{R}^{n\times n})$, $\check{C}(\cdot,\cdot)\in\mathscr{L}^2(\triangle_2(0,T);\mathbb{R}^{n\times n})$,
$\check{b}(\cdot,\cdot)\in L^{2,1}_{\mathbb{F}}(\triangle_2(0,T);\mathbb{R}^n)$, $\check{\sigma}(\cdot,\cdot)\in L^2_{\mathbb{F}}(\triangle_2(0,T);\mathbb{R}^n)$.

Then, we present the following result as Lemma A.3 in [28].

Proposition 2.2. *Let* (H2) *hold. Then, SVIE* (2.2) *admits a unique solution* $\check{X}(\cdot)\in L^2_{\mathbb{F}}(t_0,T;\mathbb{R}^n)$.

After the above preparations, let us return back to the state equation (1.1), the coefficients of which satisfy the following assumptions.

$$\textbf{(A1)}\begin{cases} A_1(\cdot),A_2(\cdot),A_3(\cdot)\in L^2(0,T;\mathbb{R}^{n\times n}), \ B_1(\cdot),B_2(\cdot),B_3(\cdot)\in L^\infty(0,T;\mathbb{R}^{n\times m}), \ b(\cdot)\in L^2_{\mathbb{F}}(\Omega;L^1(0,T;\mathbb{R}^n)), \\ C_1(\cdot),C_2(\cdot),C_3(\cdot)\in L^\infty(0,T;\mathbb{R}^{n\times n}), \quad D_1(\cdot)\in L^\infty(0,T;\mathbb{R}^{n\times m}), \quad \sigma(\cdot)\in L^2_{\mathbb{F}}(0,T;\mathbb{R}^n), \\ F(\cdot,\cdot)\in L^\infty(\Delta_2(0,T);\mathbb{R}^{n\times n}), \quad \tilde{F}(\cdot,\cdot)\in L^\infty(\Delta_2(0,T);\mathbb{R}^{n\times m}), \ \xi(\cdot)\in C([-\delta,0];\mathbb{R}^n), \\ Q_1(\cdot),Q_2(\cdot),Q_3(\cdot)\in L^\infty(0,T;\mathbb{S}^n), \quad R_1(\cdot),R_2(\cdot)\in L^\infty(0,T;\mathbb{S}^m), \ \varsigma(\cdot)\in L^2(-\delta,0;\mathbb{R}^m). \end{cases}$$

In terms of Proposition 2.1, SDDE (1.1) admits a unique solution, the cost functional (1.2) is well-defined, and hence it becomes natural to pose Problem (P) in the Introduction.

Due to the particular LQ structure, we are interested in the closed-loop optimal control. To this end, we first look at the closed-loop strategy. For any given $t_0 \in [0, T)$, define

$$\mathbb{L} \equiv L^2(t_0, T; \mathbb{R}^{m\times n}) \times L^2(\Delta_2(t_0, T); \mathbb{R}^{m\times n}) \times L^\infty(t_0, T; \mathbb{R}^{m\times n}) \times L^2(\Delta_2(t_0, T); \mathbb{R}^{m\times m}) \times L^2_{\mathbb{F}}(t_0, T; \mathbb{R}^m).$$

In the following, for any $(K_1(\cdot), K_2(\cdot,\cdot), K_3(\cdot), K_4(\cdot,\cdot), v(\cdot)) \in \mathbb{L}$, we call it a *closed-loop strategy* on $[t_0, T]$. Later we will use this closed-loop strategy, which does not depend on the initial data but only the given coefficients, to construct a closed-loop control on $[t_0, T]$. To introduce the closed-loop state, let us consider the SDDE:

$$\begin{cases} dx(t) = \big[A_1(t)x(t) + A_2(t)y(t) + A_3(t)z(t) + B_1(t)u(t) + B_2(t)\nu(t) + B_3(t)\mu(t) + b(t)\big]dt \\ \qquad\qquad + \big[C_1(t)x(t) + C_2(t)y(t) + C_3(t)z(t) + D_1(t)u(t) + \sigma(t)\big]dW(t), \quad t \in (t_0, T), \\ x(t) = \xi(t - t_0), \quad u(t) = \varsigma(t - t_0), \quad t \in [t_0 - \delta, t_0], \\ u(t) = K_1(t)x(t) + \displaystyle\int_{t_0}^t K_2(t,s)x(s)ds + K_3(t)x(t-\delta) + \int_{t_0}^t K_4(t,s)u(s)ds + v(t), \quad t \in (t_0, T). \end{cases} \tag{2.3}$$

We call (2.3) a *closed-loop system* under $(K_1(\cdot), K_2(\cdot,\cdot), K_3(\cdot), K_4(\cdot,\cdot), v(\cdot))$ corresponding to $(t_0, \xi(\cdot), \varsigma(\cdot))$, and call $x(\cdot)$, $u(\cdot)$ the corresponding *closed-loop state* and *closed-loop outcome control*, respectively. The proof of the following result is given in the Appendix.

Proposition 2.3. *Let* (A1) *hold. Then, for any given* $t_0 \in [0, T)$ *and* $(K_1(\cdot), K_2(\cdot,\cdot), K_3(\cdot), K_4(\cdot,\cdot), v(\cdot)) \in \mathbb{L}$*, the closed-loop system* (2.3) *admits a unique solution* $x(\cdot) \in L^2_{\mathbb{F}}\big(\Omega; C([t_0, T]; \mathbb{R}^n)\big)$ *on* $[t_0, T]$.

At last, let us introduce the following notions.

Definition 2.4. *For any given* $t_0 \in [0, T)$*, the closed-loop strategy* $(K_1^*(\cdot), K_2^*(\cdot,\cdot), K_3^*(\cdot), K_4^*(\cdot,\cdot), v^*(\cdot)) \in \mathbb{L}$ *is called an* optimal closed-loop strategy *of Problem* (P) *if*

$$J\big(t_0, \xi, \varsigma; (K_1^*, K_2^*, K_3^*, K_4^*, v^*)\big) \leqslant J\left(t_0, \xi, \varsigma; u\right), \tag{2.4}$$

for any $(\xi, \varsigma) \in C([-\delta, 0]; \mathbb{R}^n) \times L^2(-\delta, 0; \mathbb{R}^m)$ *and any control* $u(\cdot) \in L^2_{\mathbb{F}}(t_0, T; \mathbb{R}^m)$*. If there (uniquely) exists an optimal closed-loop strategy on* $[t_0, T]$*, Problem* (P) *is said to be* (uniquely) closed-loop solvable on $[t_0, T]$.

Remark 2.5. *Inspired by [3, 19, 31, 50, 51, 59], the closed-loop strategy for deterministic Problem* (P) *was introduced in [43], and then extended to the stochastic setting in [44]. Here our version is more general due to the appearance of* $K_3(\cdot)$*. In the system* (2.3)*,* $K_1(\cdot)$*,* $K_2(\cdot,\cdot)$*,* $K_3(\cdot)$*,* $K_4(\cdot,\cdot)$ *represent respectively the gain coefficients of the current state, the distributed state delay, the pointwise state delay and the distributed control delay.*

3 The closed-loop solvability of Problem (P)

In this section, we firstly transform the delayed Problem (P) into another optimal control problem driven by a (finite dimensional) SVIE without delay. Based on such a transformation, we then introduce and discuss the desired Riccati system for Problem (P). Eventually, we construct an explicit closed-loop strategy and prove its optimality in the sense of Definition 2.4.

3.1 Problem transformation

To begin with, let us take a closer look at the state $x(\cdot)$. Since $\varsigma(\cdot)$ is the initial trajectory of the control, for either $t \in \big[t_0, (t_0+\delta)\wedge T\big]$ or $t \in \big((t_0+\delta)\wedge T, t_0+\delta\big]$ we get

$$\int_{t_0}^t B_2(s)u(s-\delta)ds = \int_{t_0}^t B_2(s)\varsigma(s-\delta-t_0)ds = \int_{t_0}^{t_0+\delta} \mathbf{1}_{[t_0,t)}(s)B_2(s)\varsigma(s-\delta-t_0)ds.$$

For $t\in\big(t_0+\delta,(t_0+\delta)\vee T\big]$, we have

$$\int_{t_0}^{t}B_2(s)u(s-\delta)ds=\int_{t_0}^{t_0+\delta}B_2(s)u(s-\delta)ds+\int_{t_0+\delta}^{t}B_2(s)u(s-\delta)ds$$
$$=\int_{t_0}^{t_0+\delta}\mathbf{1}_{[t_0,t)}(s)B_2(s)\varsigma(s-\delta-t_0)ds+\int_{t_0}^{t}\mathbf{1}_{[t_0,t-\delta)}(s)B_2(s+\delta)u(s)ds.$$

To sum up, for $t\in[t_0,T]$, we have

$$\int_{t_0}^{t}B_2(s)u(s-\delta)ds=\int_{t_0}^{t_0+\delta}\mathbf{1}_{[t_0,t)}(s)B_2(s)\varsigma(s-\delta-t_0)ds+\int_{t_0}^{t}\mathbf{1}_{[t_0,t-\delta)}(s)B_2(s+\delta)u(s)ds.$$

As to the $\mu(\cdot)$ term in $x(\cdot)$, by the Fubini theorem, for $t\in[t_0,T]$, we obtain

$$\int_{t_0}^{t}B_3(s)\mu(s)ds=\int_{t_0}^{t}B_3(s)\int_{t_0}^{s}\tilde F(s,r)u(r)drds=\int_{t_0}^{t}\int_{r}^{t}B_3(s)\tilde F(s,r)dsu(r)dr.$$

Thus we deduce

$$\begin{aligned}x(t)=\xi(0)&+\int_{t_0}^{t_0+\delta}B_2(s)\varsigma(s-\delta-t_0)\mathbf{1}_{[t_0,t)}(s)ds+\int_{t_0}^{t}\Big[A_1(s)x(s)+A_2(s)y(s)\\&+A_3(s)z(s)+\Big(B_1(s)+B_2(s+\delta)\mathbf{1}_{[t_0,t-\delta)}(s)+\int_{s}^{t}B_3(r)\tilde F(r,s)dr\Big)u(s)+b(s)\Big]ds\\&+\int_{t_0}^{t}\big[C_1(s)x(s)+C_2(s)y(s)+C_3(s)z(s)+D_1(s)u(s)+\sigma(s)\big]dW(s),\quad t\in[t_0,T].\end{aligned}\tag{3.1}$$

Next let us turn to the term of $y(\cdot)$. For $t\in\big[t_0+\delta,(t_0+2\delta)\wedge T\big]$ and $t\in\big((t_0+2\delta)\wedge T,t_0+2\delta\big]$, we have

$$\int_{t_0}^{t-\delta}B_2(s)u(s-\delta)ds=\int_{t_0}^{t-\delta}B_2(s)\varsigma(s-\delta-t_0)ds=\int_{t_0}^{t_0+\delta}\mathbf{1}_{[t_0,t-\delta)}(s)B_2(s)\varsigma(s-\delta-t_0)ds.$$

For $t\in\big(t_0+2\delta,(t_0+2\delta)\vee T\big]$, we have

$$\int_{t_0}^{t-\delta}B_2(s)u(s-\delta)ds=\int_{t_0}^{t_0+\delta}B_2(s)u(s-\delta)ds+\int_{t_0+\delta}^{t-\delta}B_2(s)u(s-\delta)ds$$
$$=\int_{t_0}^{t_0+\delta}\mathbf{1}_{[t_0,t-\delta)}(s)B_2(s)\varsigma(s-\delta-t_0)ds+\int_{t_0}^{t}\mathbf{1}_{[t_0,t-2\delta)}(s)B_2(s+\delta)u(s)ds.$$

Therefore, for $t\in[t_0+\delta,T]$,

$$\begin{aligned}\int_{t_0}^{t-\delta}B_2(s)u(s-\delta)ds\mathbf{1}_{(t_0+\delta,\infty)}(t)=\Big[&\int_{t_0}^{t_0+\delta}\mathbf{1}_{[t_0,t-\delta)}(s)B_2(s)\varsigma(s-\delta-t_0)ds\\&+\int_{t_0}^{t}\mathbf{1}_{[t_0,t-2\delta)}(s)B_2(s+\delta)u(s)ds\Big]\mathbf{1}_{(t_0+\delta,\infty)}(t).\end{aligned}$$

As to the $\mu(\cdot)$ term in $y(\cdot)$, for $t\in[t_0+\delta,T]$, it follows from the Fubini theorem that

$$\int_{t_0}^{t-\delta}B_3(s)\mu(s)ds=\int_{t_0}^{t-\delta}B_3(s)\int_{t_0}^{s}\tilde F(s,r)u(r)drds=\int_{t_0}^{t-\delta}B_3(r)\int_{t_0}^{r}\tilde F(r,s)u(s)dsdr$$
$$=\int_{t_0}^{t-\delta}\int_{s}^{t-\delta}B_3(r)\tilde F(r,s)dru(s)ds.$$

Hence for the pointwise state delay, we have

$$
\begin{aligned}
y(t) &= \xi(t-\delta-t_0)\mathbf{1}_{[t_0,t_0+\delta]}(t) + \mathbf{1}_{(t_0+\delta,\infty)}(t)\Big\{\xi(0) + \int_{t_0}^{(t_0+\delta)\wedge(t-\delta)} B_2(s)\varsigma(s-t_0-\delta)ds\Big\} \\
&\quad + \int_{t_0}^{t} \mathbf{1}_{[t_0,t-\delta)}(s)\Big(A_1(s)x(s) + A_2(s)y(s) + A_3(s)z(s) + \Big[B_1(s) + B_2(s+\delta)\mathbf{1}_{[t_0,t-2\delta)}(s) \\
&\quad + \int_{s}^{t-\delta} B_3(r)\tilde{F}(r,s)dr\Big]u(s) + b(s)\Big)ds + \int_{t_0}^{t}\mathbf{1}_{[t_0,t-\delta)}(s)\big[C_1(s)x(s) + C_2(s)y(s) \\
&\quad + C_3(s)z(s) + D_1(s)u(s) + \sigma(s)\big]dW(s), \quad t\in[t_0,T].
\end{aligned} \tag{3.2}
$$

Eventually, let us treat the term of $z(\cdot)$. For $t\in[t_0,T]$,

$$
\begin{aligned}
z(t) = \int_{t_0}^{t} F(t,s)x(s)ds = \int_{t_0}^{t} F(t,s)\Big[\xi(0) + \int_{t_0}^{s}\big(A_1(r)x(r) + A_2(r)y(r) + A_3(r)z(r) \\
+ B_1(r)u(r) + B_2(r)\nu(r) + B_3(r)\mu(r) + b(r)\big)dr \\
+ \int_{t_0}^{s}\big(C_1(r)x(r) + C_2(r)y(r) + C_3(r)z(r) + D_1(r)u(r) + \sigma(r)\big)dW(r)\Big]ds.
\end{aligned} \tag{3.3}
$$

By Fubini theorem,

$$
\int_{t_0}^{t}\int_{t_0}^{s} F(t,s)B_2(r)\nu(r)drds = \int_{t_0}^{t}\Big[\int_{r}^{t} F(t,s)ds\Big]B_2(r)u(r-\delta)dr.
$$

Notice that for $t\in[t_0,t_0+\delta)$,

$$
\int_{t_0}^{t}\int_{t_0}^{s} F(t,s)B_2(r)\nu(r)drds = \int_{t_0}^{t}\Big[\int_{r}^{t} F(t,s)ds\Big]B_2(r)\varsigma(r-\delta-t_0)dr,
$$

while for $t\in\big[t_0+\delta,(t_0+\delta)\vee T\big)$, one has

$$
\begin{aligned}
&\int_{t_0}^{t}\int_{t_0}^{s} F(t,s)B_2(r)\nu(r)drds = \int_{t_0}^{t}\Big[\int_{r}^{t} F(t,s)ds\Big]B_2(r)u(r-\delta)dr \\
&= \int_{t_0}^{t_0+\delta}\Big[\int_{r}^{t} F(t,s)ds\Big]B_2(r)\varsigma(r-\delta-t_0)dr + \int_{t_0}^{t-\delta}\Big[\int_{r+\delta}^{t} F(t,s)ds\Big]B_2(r+\delta)u(r)dr.
\end{aligned}
$$

Therefore, for $t\in[t_0,T]$, we obtain that

$$
\begin{aligned}
&\int_{t_0}^{t}\int_{t_0}^{s} F(t,s)B_2(r)\nu(r)drds \\
&= \int_{t_0}^{(t_0+\delta)\wedge t}\Big[\int_{r}^{t} F(t,s)ds\Big]B_2(r)\varsigma(r-\delta-t_0)dr + \int_{t_0\wedge(t-\delta)}^{t-\delta}\Big[\int_{r+\delta}^{t} F(t,s)ds\Big]B_2(r+\delta)u(r)dr.
\end{aligned} \tag{3.4}
$$

As to the $\mu(\cdot)$ term in $z(\cdot)$, by the Fubini theorem again, for $t\in[t_0,T]$,

$$
\begin{aligned}
&\int_{t_0}^{t}\int_{t_0}^{s} F(t,s)B_3(r)\mu(r)drds = \int_{t_0}^{t}\int_{t_0}^{s} F(t,s)B_3(r)\Big[\int_{t_0}^{r}\tilde{F}(r,\alpha)u(\alpha)d\alpha\Big]drds \\
&= \int_{t_0}^{t}\int_{t_0}^{s}\int_{\alpha}^{s} F(t,s)B_3(r)\tilde{F}(r,\alpha)u(\alpha)drd\alpha ds = \int_{t_0}^{t}\Big[\int_{\alpha}^{t}\int_{\alpha}^{s} F(t,s)B_3(r)\tilde{F}(r,\alpha)drds\Big]u(\alpha)d\alpha.
\end{aligned} \tag{3.5}
$$

Hence, from (3.3)–(3.5) and by adding some indicative functions, we derive

$$
\begin{aligned}
z(t) &= \int_{t_0}^{t} F(t,s)\Big(\xi(0) + \int_{t_0}^{t_0+\delta} B_2(r)\varsigma(r-\delta-t_0)\mathbf{1}_{[0,s)}(r)dr\Big)ds + \int_{t_0}^{t}\mathcal{E}(t,s)\Big[A_1(s)x(s) + A_2(s)y(s) \\
&\quad + A_3(s)z(s)\Big]ds + \int_{t_0}^{t}\Big[\int_{s}^{t} F(t,r)\Big(B_1(s) + B_2(s+\delta)\mathbf{1}_{[t_0,r-\delta)}(s) + \int_{s}^{r} B_3(\theta)\tilde{F}(\theta,s)d\theta\Big)dr\Big]u(s)ds \\
&\quad + \int_{t_0}^{t}\mathcal{E}(t,s)b(s)ds + \int_{t_0}^{t}\mathcal{E}(t,s)\big[C_1(s)x(s) + C_2(s)y(s) + C_3(s)z(s) + D_1(s)u(s) + \sigma(s)\big]dW(s).
\end{aligned} \tag{3.6}
$$

Here and next, for $T \geqslant t > s \geqslant t_0$, we define $\mathcal{E}(\cdot,\cdot)$ and some other notations as follows:

$$
\begin{aligned}
&X(t) \equiv \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix}, \qquad\qquad \mathcal{E}(t,s) \equiv \int_s^t F(t,r)dr, \quad \mathcal{E}(s,t) \equiv 0, \\
&\varphi(t) \equiv \begin{bmatrix} \xi(0) + \int_{t_0}^{t_0+\delta} B_2(s)\varsigma(s-t_0-\delta)\mathbf{1}_{[t_0,t)}(s)ds \\ \xi(t-t_0-\delta)\mathbf{1}_{[t_0,t_0+\delta]}(t) + [\xi(0) + \int_{t_0}^{(t_0+\delta)\wedge(t-\delta)} B_2(s)\varsigma(s-t_0-\delta)ds]\mathbf{1}_{(t_0+\delta,\infty)}(t) \\ \int_{t_0}^t F(t,s)[\xi(0) + \int_{t_0}^{t_0+\delta} B_2(r)\varsigma(r-t_0-\delta)\mathbf{1}_{[t_0,s)}(r)dr]ds \end{bmatrix}, \\
&A(t,s) \equiv \begin{bmatrix} A_1(s) & A_2(s) & A_3(s) \\ \mathbf{1}_{(\delta,\infty)}(t-s)A_1(s) & \mathbf{1}_{(\delta,\infty)}(t-s)A_2(s) & \mathbf{1}_{(\delta,\infty)}(t-s)A_3(s) \\ \mathcal{E}(t,s)A_1(s) & \mathcal{E}(t,s)A_2(s) & \mathcal{E}(t,s)A_3(s) \end{bmatrix}, \\
&B(t,s) \equiv \begin{bmatrix} B_1(s) + B_2(s+\delta)\mathbf{1}_{(\delta,\infty)}(t-s) + \int_s^t B_3(r)\tilde{F}(r,s)dr \\ \mathbf{1}_{(\delta,\infty)}(t-s)[B_1(s) + B_2(s+\delta)\mathbf{1}_{(2\delta,\infty)}(t-s) + \int_s^{t-\delta} B_3(r)\tilde{F}(r,s)dr] \\ \mathcal{E}(t,s)B_1(s) + \mathcal{E}(t,s+\delta)B_2(s+\delta) + \int_s^t \mathcal{E}(t,\theta)B_3(\theta)\tilde{F}(\theta,s)d\theta \end{bmatrix}, \\
&C(t,s) \equiv \begin{bmatrix} C_1(s) & C_2(s) & C_3(s) \\ \mathbf{1}_{(\delta,\infty)}(t-s)C_1(s) & \mathbf{1}_{(\delta,\infty)}(t-s)C_2(s) & \mathbf{1}_{(\delta,\infty)}(t-s)C_3(s) \\ \mathcal{E}(t,s)C_1(s) & \mathcal{E}(t,s)C_2(s) & \mathcal{E}(t,s)C_3(s) \end{bmatrix}, \\
&D(t,s) \equiv \begin{bmatrix} D_1(s) \\ \mathbf{1}_{(\delta,\infty)}(t-s)D_1(s) \\ \mathcal{E}(t,s)D_1(s) \end{bmatrix}, \ \tilde{b}(t,s) \equiv \begin{bmatrix} b(s) \\ \mathbf{1}_{(\delta,\infty)}(t-s)b(s) \\ \mathcal{E}(t,s)b(s) \end{bmatrix}, \ \tilde{\sigma}(t,s) \equiv \begin{bmatrix} \sigma(s) \\ \mathbf{1}_{(\delta,\infty)}(t-s)\sigma(s) \\ \mathcal{E}(t,s)\sigma(s) \end{bmatrix}.
\end{aligned}
\tag{3.7}
$$

Based on (3.1), (3.2) and (3.6), $X(\cdot)$ satisfies the following SVIE:

$$
\begin{aligned}
X(t) = \varphi(t) &+ \int_{t_0}^t \big[A(t,s)X(s) + B(t,s)u(s) + \tilde{b}(t,s)\big]ds \\
&+ \int_{t_0}^t \big[C(t,s)X(s) + D(t,s)u(s) + \tilde{\sigma}(t,s)\big]dW(s), \quad t \in (t_0,T).
\end{aligned}
\tag{3.8}
$$

Denote

$$
Q(t) \equiv \begin{bmatrix} Q_1(t) & 0 & 0 \\ 0 & Q_2(t) & 0 \\ 0 & 0 & Q_3(t) \end{bmatrix}, \quad R(t) \equiv R_1(t) + R_2(t+\delta)\mathbf{1}_{[0,T-\delta)}(t).
$$

Then, the cost functional (1.2) becomes

$$
\begin{aligned}
J(t_0,\xi(\cdot),\varsigma(\cdot);u(\cdot)) = \mathbb{E}&\int_{t_0}^T \big[X(t)^\top Q(t)X(t) + u(t)^\top R(t)u(t)\big]dt \\
+\mathbb{E}&\int_{t_0}^{(t_0+\delta)\wedge T} \varsigma(t-t_0-\delta)^\top R_2(t)\varsigma(t-t_0-\delta)dt.
\end{aligned}
\tag{3.9}
$$

To sum up the above arguments, we have the following result whose proof is given in the Appendix.

Proposition 3.1. *Let Assumption* (A1) *hold. Then,* $u^* \in L^2_{\mathbb{F}}(t_0,T;\mathbb{R}^m)$ *is an optimal control of Problem* (P) *if and only if* $u^*(\cdot)$ *minimizes* (3.9) *subject to SVIE* (3.8)*.*

Remark 3.2. *The idea of transforming the delayed system into another one without delay is popular in the existing literature.*

For example, in [18, 19, 43, 44], the authors transformed the original delayed systems into infinite-dimensional evolution control systems without delay. However, in some cases they had to treat the unbounded control operator which would bring essential difficulties. More importantly, it seems quite complex to transform the obtained operator-valued Riccati equation into the (finite dimensional) matrix-valued case.

In contrast, in the current paper we put $(x(\cdot), y(\cdot), z(\cdot))$ *together to construct a new state* $X(\cdot)$ *and then transform the original system equivalently to a controlled SVIE. Similar procedure also appeared in [42]. On* *the one hand, this helps us to utilize the developed LQ theory for SVIEs in [29]. On the other hand, the advantage lies in the fact that $X(\cdot)$ is still finite dimensional which helps us to bypass the complicated operator language. Moreover, by our approach there is no need to face the challenge of decoupling the forward delayed system and the backward anticipated system.*

We point out that the transformation to Volterra systems also appeared in e.g. [30, 36] via the constant variation formula. Nevertheless, it is not clear whether such an approach still works when the diffusion term in (1.1) contains the pointwise state delays and distributed state delays.

3.2 The solvability of the new Riccati system

To study the closed-loop solvability of Problem (P), we introduce a new Riccati system based on the above transformation and the existing study in [29]. To this end, we need some preparations.

Firstly, we denote by $\Pi(0,T)$ the set of pairs $P=\left(P^{(1)},P^{(2)}\right)$ with $P^{(1)}:(0,T)\to\mathbb{R}^{(3n)\times(3n)}$ and $P^{(2)}:\square_3\left(t_0,T\right)\to\mathbb{R}^{(3n)\times(3n)}$ such that

(i) $P^{(1)}\in L^{\infty}\left(0,T;\mathbb{S}^{3n}\right)$;
(ii) for a.e. $(s_1,s_2)\in(0,T)^2, t\mapsto P^{(2)}\left(s_1,s_2,t\right)$ is absolutely continuous on $(0,s_1\wedge s_2)$;
(iii) $(s_1,s_2)\mapsto P^{(2)}\left(s_1,s_2,s_1\wedge s_2\right)\equiv\lim\limits_{t\uparrow(s_1\wedge s_2)}P^{(2)}\left(s_1,s_2,t\right)$ belongs to $L^2\left((0,T)^2;\mathbb{R}^{(3n)\times(3n)}\right)$;
(iv) $(s_1,s_2,t)\mapsto\dot{P}^{(2)}\left(s_1,s_2,t\right)\equiv\frac{\partial P^{(2)}}{\partial t}\left(s_1,s_2,t\right)$ belongs to $L^{2,2,1}\left(\square_3(0,T);\mathbb{R}^{(3n)\times(3n)}\right)$;
(v) for a.e. $(s_1,s_2,t)\in\square_3(0,T)$, it holds that $P^{(2)}\left(s_1,s_2,t\right)=P^{(2)}\left(s_2,s_1,t\right)^{\top}$.

Secondly, for later notational usefulness, we define several coefficients as follows:

$$
\begin{aligned}
&\Pi(s,t,\theta)\equiv\begin{bmatrix}\frac{1}{s-t}I & \frac{1}{s-t}\mathbf{1}_{(\delta,\infty)}(s-t)I & I\\ \frac{1}{s-t}\mathbf{1}_{(\delta,\infty)}(s-t)I & \frac{1}{s-t}\mathbf{1}_{(2\delta,\infty)}(s-t)I & \mathbf{1}_{(\delta,\infty)}(s-\theta)I\\ \frac{1}{s-t}\mathcal{E}(s,t) & \frac{1}{s-t}\mathcal{E}(s,t+\delta) & \mathcal{E}(s,\theta)\end{bmatrix},\Upsilon(s,t)\equiv\begin{bmatrix}I\\ \mathbf{1}_{(\delta,\infty)}(s-t)I\\ \mathcal{E}(s,t)\end{bmatrix},\\
&\mathcal{A}(t)\equiv\left(A_1(t),A_2(t),A_3(t)\right),\quad\mathcal{B}(\theta,t)\equiv\begin{bmatrix}B_1(t)\\ B_2(t+\delta)\\ B_3(\theta)\widetilde{F}(\theta,t)\end{bmatrix},\quad\mathcal{C}(t)\equiv\left(C_1(t),C_2(t),C_3(t)\right).
\end{aligned}
\tag{3.10}
$$

Given these coefficients, we introduce the following system satisfied by a pair of $(P^{(1)},P^{(2)})$:

$$
\begin{cases}
P^{(1)}(t)=Q(t)+\mathcal{C}(t)^{\top}\mathcal{G}_1(t)\mathcal{C}(t)-\mathcal{C}(t)^{\top}\mathcal{G}_1(t)D_1(t)\mathcal{R}(t)^{-1}D_1(t)^{\top}\mathcal{G}_1(t)\mathcal{C}(t), 0<t<T,\\
P^{(2)}(s,t,r)=P^{(2)}(s,t,t\wedge s)-\displaystyle\int_r^{t\wedge s}\int_\tau^T\int_\tau^T\mathcal{G}_3(s,\tau,\theta_1)\mathcal{B}(\theta_1,\tau)\mathcal{R}(\tau)^{-1}\mathcal{B}(\theta_2,\tau)^{\top}\\
\qquad\qquad\times\mathcal{G}_3(t,\tau,\theta_2)^{\top}d\theta_1d\theta_2d\tau, 0<r<(s\wedge t)<T,\\
P^{(2)}(\bar{s},t,t)=P^{(2)}(t,\bar{s},t)^{\top}=\mathcal{G}_2(\bar{s},t)\mathcal{A}(t)-\displaystyle\int_t^T\mathcal{G}_3(\bar{s},t,\theta)\mathcal{B}(\theta,t)\mathcal{R}(t)^{-1}D_1(t)^{\top}\mathcal{G}_1(t)\mathcal{C}(t)d\theta, 0<t<\bar{s}<T,
\end{cases}
\tag{3.11}
$$

where for $0<t<(s\wedge\bar{s}\wedge\theta)<T$,

$$
\begin{aligned}
&\mathcal{R}(t)\equiv R_1(t)+R_2(t+\delta)\mathbf{1}_{[0,T-\delta)}(t)+D_1(t)^{\top}\mathcal{G}_1(t)D_1(t),\\
&\mathcal{G}_1(t)\equiv\mathcal{G}_1\left(t;P^{(1)},P^{(2)}\right)\equiv\int_t^T\Upsilon(s,t)^{\top}P^{(1)}(s)\Upsilon(s,t)ds+\int_t^T\int_t^T\Upsilon(s_1,t)^{\top}P^{(2)}(s_1,s_2,t)\Upsilon(s_2,t)ds_1ds_2,\\
&\mathcal{G}_2(\bar{s},t)\equiv\mathcal{G}_2\left(\bar{s},t;P^{(1)},P^{(2)}\right)\equiv P^{(1)}(\bar{s})\Upsilon(\bar{s},t)+\int_t^T P^{(2)}(\bar{s},r,t)\Upsilon(r,t)dr,\\
&\mathcal{G}_3(s,t,\theta)\equiv\mathcal{G}_3\left(s,t,\theta;P^{(1)},P^{(2)}\right)\equiv P^{(1)}(s)\mathbf{1}_{(t,s)}(\theta)\Pi(s,t,\theta)+\int_\theta^T P^{(2)}(s,r,t)\Pi(r,t,\theta)dr.
\end{aligned}
\tag{3.12}
$$

In this paper, we name (3.11) the desired Riccati system, explicitly depending on A_i, B_i, C_i, D_1, based on (at least) the following aspects. Firstly, it is consistent with the Riccati system for stochastic Volterra system in [29]. Secondly, in particular cases we will show in Subsection 4.1, 4.2, 4.5 that it can reduce to the Riccati systems in the existing literature. Thirdly, just like the existing Riccati systems, we will use the above (3.11) to construct closed-loop control as well. To guarantee its well-posedness, we need the following standard assumption.

(A2) There exists a constant $\lambda > 0$ such that for all $t \in (0, T)$, $R_1(t) + R_2(t+\delta)\mathbf{1}_{[0,T-\delta)}(t) \geqslant \lambda I$, $Q_i(t) \geqslant 0$, $i = 1, 2, 3$.

The following proposition gives its solvability.

Theorem 3.3. *Let* (A1)–(A2) *hold. Then, the Riccati system* (3.11) *admits a unique solution* $(P^{(1)}, P^{(2)}) \in \Pi(0, T)$ *such that* $\mathcal{R}(\cdot) \geqslant \beta I$ *for some constant* $\beta > 0$.

Remark 3.4. *As to the solvability of the above Riccati system, we make the following comments.*

We first look at the particular case of the SDDE (1.4) *and compare it with the study in [44]. In terms of their framework, both the cost functional and the diffusion term can depend on the distributed control delay which is out of our scope. However, to derive the Riccati system they have to assume that* $C_2, b, \sigma, R_2, Q_2 = 0$. *In addition, to obtain the solvability, they further require that* $D_1 = 0$ *and all the coefficients are time-invariant or continuous. It is worth mentioning that these assumptions are not needed here. Moreover, the methodologies developed in both papers are essentially different.*

Next we turn to another particular case of the state equation (1.1) *with* $A_2, B_2, C_2, Q_2, R_2 = 0$, *and arrive at a LQ problem for a stochastic integro-differential system. Even though there are some positive results of the Riccati system (see [48, 49]) in deterministic scenario, the extension to the stochastic setting is still open. Here we fill this blank in a nice manner.*

At last we point out two interesting facts even for the deterministic system (1.1). *Firstly, in contrast with the relevant literature on differential delay systems (e.g. [3, 19, 31, 43]), the corresponding Riccati systems and their solvability appear for the first time since both the pointwise delays and the distributed delays are allowed to appear simultaneously in the cost functional. Secondly, when all the pointwise delayed terms disappear, the corresponding integro-differential system and cost functional can cover those in [48, 49], and thus the corresponding result in the above Theorem 3.3 is also new.*

Next we introduce the following backward system:

$$\begin{cases} d\eta(t,s) = -\Big\{\mathcal{G}_2(t,s)b(s) + \Gamma^*(t,s)^\top D_1(s)^\top \mathcal{G}_1(s)\sigma(s) + \Gamma^*(t,s)^\top \displaystyle\int_s^T \Big[D_1(s)^\top \Upsilon(r,s)^\top \zeta(r,s) \\ \qquad + \Big(\displaystyle\int_s^r \mathcal{B}(\theta,s)^\top \Pi(r,s,\theta)^\top d\theta\Big)\eta(r,s)\Big]dr\Big\}ds + \zeta(t,s)dW(s), \ 0 < s < t < T, \\ \eta(t,t) = \mathcal{C}(t)^\top \mathcal{G}_1(t)\sigma(t) + \Xi^*(t)^\top D_1(t)^\top \mathcal{G}_1(t)\sigma(t) + \displaystyle\int_t^T \Big\{\Big[\Big(\displaystyle\int_t^r \Pi(r,t,\theta)\mathcal{B}(\theta,t)^\top d\theta\Big)\Xi^*(t) \\ \qquad + \Upsilon(r,t)\mathcal{A}(t)\Big]^\top \eta(r,t) + \Big[\Upsilon(r,t)\mathcal{C}(t) + \Upsilon(r,t)D_1(t)\Xi^*(t)\Big]^\top \zeta(r,t)\Big\}dr, \ 0 < t < T, \end{cases} \tag{3.13}$$

where $\Xi^*(\cdot)$, $\Gamma^*(\cdot,\cdot)$ are defined by

$$\begin{aligned} &\Xi^*(t) = -\mathcal{R}(t)^{-1} D_1(t)^\top \mathcal{G}_1(t)\mathcal{C}(t), \ 0 < t < T, \\ &\Gamma^*(s,t) = -\mathcal{R}(t)^{-1}\int_t^T \mathcal{B}(\theta,t)^\top \mathcal{G}_3(s,t,\theta)^\top d\theta, \quad 0 < t < s < T, \end{aligned} \tag{3.14}$$

and $\Pi(\cdot,\cdot,\cdot)$, $\Upsilon(\cdot,\cdot)$, $\mathcal{R}(\cdot)$, $\mathcal{G}_1(\cdot)$, $\mathcal{G}_3(\cdot,\cdot,\cdot)$ are defined as (3.10), (3.12). In terms of [29], we name it the Type-II extended backward SVIE in our scenario. To study its well-posedness, we introduce the following space. Denote by $L^2_{\mathbb{F},c}(\triangle_2(0,T);\mathbb{R}^{3n})$ the set of $\eta \in L^2_{\mathbb{F}}(\triangle_2(0,T);\mathbb{R}^{3n})$ such that $s \mapsto \eta(t,s)$ is uniformly continuous on $(0,t)$ with the limits defined by $\eta(t,t) \equiv \lim_{s\uparrow t}\eta(t,s)$ and $\eta(t,0) \equiv \lim_{s\downarrow 0}\eta(t,s)$ for a.e. $t \in (0,T)$, a.s., and $\eta(\cdot,\cdot)$ satisfies $\mathbb{E}(\int_0^T \sup\limits_{s\in[0,t]} |\eta(t,s)|^2 \, dt)^{\frac{1}{2}} < \infty$.

Theorem 3.5. *Let* (A1)–(A2) *hold. Then, the Type-II extended backward SVIE* (3.13) *admits a unique solution* $(\eta, \zeta) \in L^2_{\mathbb{F},c}(\triangle_2(0,T);\mathbb{R}^{3n}) \times L^2_{\mathbb{F}}(\triangle_2(0,T);\mathbb{R}^{3n})$.

3.3 The closed-loop solvability of Problem (P)

In this part we give the explicit form of the optimal closed-loop strategy and the sufficient conditions for the closed-loop solvability of Problem (P).

Given $\Pi(\cdot,\cdot,\cdot)$, $\Upsilon(\cdot,\cdot)$, $\mathcal{R}(\cdot)$, $\mathcal{G}_1(\cdot)$, $\mathcal{G}_3(\cdot,\cdot,\cdot)$ in (3.10) and (3.12), $P=(P^{(1)},P^{(2)})$ and (η,ζ) being the solutions to (3.11) and (3.13), respectively, we make the following conventions. For $i=1,3$, $t_0\leqslant t\leqslant T$, we denote

$$K_i^*(t)=-\mathcal{R}(t)^{-1}\Big\{K_i^{(1)}(t)+\int_t^T\int_t^T\mathcal{B}(\theta,t)^\top\mathcal{G}_3(\alpha,t,\theta)^\top K_i^{(2)}(\alpha,t)d\alpha d\theta\Big\},\tag{3.15}$$

while for $i=2,4$ and $t_0\leqslant s<t\leqslant T$,

$$K_i^*(t,s)=-\mathcal{R}(t)^{-1}\Big\{K_i^{(1)}(t,s)+\int_t^T\int_t^T\mathcal{B}(\theta,t)^\top\mathcal{G}_3(\alpha,t,\theta)^\top K_i^{(2)}(t,\alpha,s)d\alpha d\theta\Big\}.\tag{3.16}$$

In addition, for $t_0\leqslant t\leqslant T$, let

$$v^*(t)=-\mathcal{R}(t)^{-1}\Big\{v^{(1)}(t)+\int_t^T\int_t^T\mathcal{B}(\theta,t)^\top\mathcal{G}_3(\alpha,t,\theta)^\top v^{(2)}(t,\alpha)d\alpha d\theta\Big\}.\tag{3.17}$$

In the above, each pair of $(K_i^{(1)},K_i^{(2)})$ and $(v^{(1)},v^{(2)})$ have the following representations:

$$\begin{aligned}
&K_1^{(1)}(t)=D_1(t)^\top\mathcal{G}_1(t)C_1(t),\ K_1^{(2)}(\alpha,t)=\Upsilon(\alpha,t),\ K_3^{(1)}(t)=D_1(t)^\top\mathcal{G}_1(t)C_2(t),\ K_3^{(2)}(\alpha,t)=0,\\
&K_2^{(1)}(t,s)=D_1(t)^\top\mathcal{G}_1(t)C_3(t)F(t,s)+\Big(\int_t^T\mathcal{B}(\theta,t)^\top\mathcal{G}_3(s+\delta,t,\theta)^\top d\theta\Big)(0,I,0)^\top\mathbf{1}_{[t-\delta,T-\delta]}(s),\\
&K_2^{(2)}(t,\alpha,s)=(0,0,I)^\top F(\alpha,s),\quad K_4^{(1)}(t,s)=0,\\
&K_4^{(2)}(t,\alpha,s)=\Big(I,\mathbf{1}_{(s+2\delta,\infty)}(\alpha)\mathbf{1}_{(0,T-\delta)}(t)I,\int_{s+\delta}^{\alpha}F(\alpha,\theta')^\top d\theta'\Big)^\top\mathbf{1}_{(s+\delta,T)}(\alpha)B_2(s+\delta)\mathbf{1}_{[t-\delta,T-\delta]}(s)\\
&\quad+\Big(\tilde{\mathcal{F}}(t,s,\alpha),\tilde{\mathcal{F}}(t,s,\alpha-\delta),\int_t^\alpha\tilde{\mathcal{F}}(t,s,\beta)F(\alpha,\beta)^\top d\beta\Big)^\top,\quad \tilde{\mathcal{F}}(t,s,\tau)\equiv\int_t^\tau\tilde{F}(\theta',s)^\top B_3(\theta')^\top d\theta',
\end{aligned}\tag{3.18}$$

and

$$\begin{aligned}
&v^{(1)}(t)=D_1(t)^\top\mathcal{G}_1(t)\sigma(t)+\int_t^T D_1(t)^\top\Upsilon(\alpha,t)^\top\zeta(\alpha,t)d\alpha+\int_t^T\int_\theta^T\mathcal{B}(\theta,t)^\top\Pi(\alpha,t,\theta)^\top\eta(\alpha,t)d\alpha d\theta,\\
&v^{(2)}(t,\alpha)=(0,I,0)^\top\xi(\alpha-\delta-t_0)\mathbf{1}_{[t_0,t_0+\delta]}(\alpha)+\int_{t-\delta}^{t_0}\mathbf{1}_{(\theta'+\delta,\infty)}(\alpha)\Big(I,\mathbf{1}_{(\theta'+2\delta,\infty)}(\alpha)\mathbf{1}_{(0,T-\delta)}(t)I,\\
&\qquad\int_{\theta'+\delta}^{\alpha}F(\alpha,\beta)^\top d\beta\Big)^\top B_2(\theta'+\delta)\varsigma(\theta'-t_0)d\theta'\mathbf{1}_{[t_0,t_0+\delta]}(t).
\end{aligned}\tag{3.19}$$

At this moment, we present the main result of the current section.

Theorem 3.6. *Let (A1)–(A2) hold and $t_0\in[0,T)$ be given. Then, the five-tuple $(K_1^*(\cdot),K_2^*(\cdot,\cdot),K_3^*(\cdot),K_4^*(\cdot,\cdot),v^*(\cdot))$ given by (3.15)–(3.19) is an optimal closed-loop strategy, and the following $u^*(\cdot)$ is the unique optimal closed-loop outcome control of Problem (P) on $[t_0,T]$:*

$$u^*(t)=K_1^*(t)x^*(t)+\int_{t_0}^t K_2^*(t,s)x^*(s)ds+K_3^*(t)x^*(t-\delta)+\int_{t_0}^t K_4^*(t,s)u^*(s)ds+v^*(t),\quad t\in[t_0,T].\tag{3.20}$$

By the above theorem and Theorem 5.4 in [29], we deduce the following result for Problem (P_0).

Corollary 3.7. *Let (A1)–(A2) hold and $t_0\in[0,T)$ be given. Then, the five-tuple $(K_1^*(\cdot),K_2^*(\cdot,\cdot),K_3^*(\cdot),K_4^*(\cdot,\cdot),v^*(\cdot))$ with $v^{(1)}(\cdot)=0$ is an optimal closed-loop strategy, and the above (3.20) is the unique optimal closed-loop outcome control of Problem (P_0). In addition, the value function is given by*

$$V_0(t_0,\xi,\varsigma)=\int_{t_0}^T\langle P^{(1)}(t)\varphi(t),\varphi(t)\rangle dt+\int_{t_0}^T\int_{t_0}^T\langle P^{(2)}(t_1,t_2,t_0)\varphi(t_2),\varphi(t_1)\rangle dt_1dt_2,$$

for any $(t_0,\xi,\varsigma)\in[0,T)\times C([-\delta,0];\mathbb{R}^n)\times L^2(-\delta,0;\mathbb{R}^m)$.

Remark 3.8. *We see that Theorem 3.6 gives a sufficient condition of the closed-loop solvability in terms of K_i^* (i=1,2,3,4) which are explicitly and clearly constructed. At this moment, we are not sure about its necessity. However, Lemma A.1 in the Appendix actually gives a new necessary condition in terms of the so-called causal feedback strategy developed in [29]. On the other hand, Theorem 3.6 is also true if the standard assumption* (A2) *is relaxed properly (see [44]). For simplicity we prefer not to pursue these generalities.*

Remark 3.9. *There have been lots of works on closed-loop outcome controls of delayed control systems. However, they either contain only state delays [15, 36, 38, 39], or only control delays [13, 31, 38, 47, 54, 59], or work in deterministic systems [31, 36], or have time-invariant coefficients [31, 36, 38, 44, 47, 54, 59], or have no delay in the cost functional [3, 18, 19, 31, 55, 59], or have no solvability of the associated Riccati systems [13, 15, 31, 36, 38, 39, 47, 54, 59]. In this sense, Theorem 3.6 gives a unified treatment of the existing papers with distinctive methods. Moreover, the closed-loop outcome control is explicitly constructed without any continuity or even differentiability assumptions, and does not rely on the future state and avoids complex tools of conditional expectations [55, 59].*

4 Important cases

In this section, we discuss five special yet important stochastic control systems and make relevant comparisons with the existing literature.

4.1 Case I: Stochastic control systems with control delays only

Consider the state equation

$$\begin{cases} dx(t) =\big[A_1(t)x(t) + B_1(t)u(t) + B_2(t)\nu(t) + B_3(t)\mu(t)\big]dt \\ \qquad\qquad + \big[C_1(t)x(t) + D_1(t)u(t)\big]dW(t), \quad t \in [t_0, T], \\ x(t_0) =\xi_0, u(t) = 0, \quad t \in [t_0 - \delta, t_0], \end{cases} \tag{4.1}$$

along with the cost functional

$$J(t_0, \xi_0; u(\cdot)) = \mathbb{E}\int_{t_0}^{T} \big[x(t)^\top Q_1(t)x(t) + u(t)^\top R_1(t)u(t)\big]dt.$$

Also for $\theta, \alpha \in [-\delta, 0]$ and $t, \theta', r \in [t_0, T]$ such that $t \leqslant (\theta' \wedge r)$, we define

$$\mathcal{P}_1(t, \theta', r) \equiv \big(I, 0, 0\big)\Big(\int_{\theta'\vee r}^{T} P^{(1)}(s)ds + \int_{r}^{T}\int_{\theta'}^{T} P^{(2)}(s, \alpha, t)d\alpha ds\Big)\big(I, 0, 0\big)^\top, \tag{4.2}$$

$$\mathcal{S}_0(t) \equiv \mathcal{P}_1(t, t, t), \tag{4.3}$$

$$\mathcal{S}_1(t, \theta) \equiv B_2(t{+}\delta{+}\theta)^\top\mathcal{P}_1(t, t, t{+}\delta{+}\theta) +\int_t^T \tilde{F}(\theta', t{+}\theta)^\top B_3(\theta')^\top\mathcal{P}_1(t, \theta', t)^\top d\theta', \tag{4.4}$$

$$\begin{aligned}\mathcal{S}_2(t, \theta, \alpha) \equiv{}& B_2(t{+}\delta{+}\theta)^\top\Big[\mathcal{P}_1(t, t{+}\delta{+}\theta, t{+}\delta{+}\alpha)^\top B_2(t{+}\delta{+}\alpha){+}\int_t^T\mathcal{P}_1(t,\theta', t{+}\delta{+}\theta)B_3(\theta')\tilde{F}(\theta', t{+}\alpha)d\theta'\Big]\\ &+\int_t^T\tilde{F}(\theta', t{+}\theta)^\top B_3(\theta')^\top\Big[\mathcal{P}_1(t, \theta', t{+}\delta{+}\alpha)^\top B_2(t{+}\delta{+}\alpha){+}\int_t^T\mathcal{P}_1(t,\theta',\beta)^\top B_3(\beta)\tilde{F}(\beta, t{+}\alpha)d\beta\Big]d\theta'.\end{aligned} \tag{4.5}$$

In this part, we will show that $\mathcal{S}_0(\cdot)$ and $\mathcal{S}_1(\cdot, \cdot)$ satisfy

$$\begin{cases} \frac{d}{dt}\mathcal{S}_0(t){+}A_1(t)^\top\mathcal{S}_0(t){+}\mathcal{S}_0(t)A_1(t){+}Q_1(t) + C_1(t)^\top\mathcal{S}_0(t)C_1(t) -\big[B_1(t)^\top\mathcal{S}_0(t) \\ +\mathcal{S}_1(t,0){+}\ D_1(t)^\top\mathcal{S}_0(t)C_1(t)\big]^\top\mathcal{R}(t)^{-1}\big[B_1(t)^\top\mathcal{S}_0(t){+}\mathcal{S}_1(t,0){+}D_1(t)^\top\mathcal{S}_0(t)C_1(t)\big]{=}0, \text{a.e. } t \in [t_0, T], \\ \mathcal{S}_0(T) = 0, \end{cases} \tag{4.6}$$

$$
\begin{cases}
\Big(\frac{\partial}{\partial t}-\frac{\partial}{\partial \theta}\Big)\mathcal{S}_1(t,\theta)+\tilde F(t,t+\theta)^\top B_3(t)^\top \mathcal{S}_0(t)+\mathcal{S}_1(t,\theta)A_1(t)-\big[\mathcal{S}_1(t,\theta)B_1(t)\\
\quad +\mathcal{S}_2(t,\theta,0)\big]\mathcal{R}(t)^{-1}\big[B_1(t)^\top\mathcal{S}_0(t)+\mathcal{S}_1(t,0)+D_1(t)^\top\mathcal{S}_0(t)C_1(t)\big]=0,\quad \text{a.e. } t\in[t_0,T],\theta\in[-\delta,0],\\
\mathcal{S}_1(T,\theta)=0,\ \mathcal{S}_1(t,-\delta)=B_2(t)^\top\mathcal{S}_0(t)+\int_t^T \tilde F(\theta',t-\delta)^\top B_3(\theta')^\top \mathcal{P}_1(t,\theta',t)^\top d\theta'.
\end{cases}
\tag{4.7}
$$

Moreover, $\mathcal{S}_2(\cdot,\cdot,\cdot)$ satisfies

$$
\begin{cases}
\Big(\frac{\partial}{\partial t}-\frac{\partial}{\partial \theta}-\frac{\partial}{\partial \alpha}\Big)\mathcal{S}_2(t,\theta,\alpha)+\tilde F(t,t+\theta)^\top B_3(t)^\top\mathcal{S}_1(t,\alpha)^\top+\mathcal{S}_1(t,\theta)B_3(t)^\top\tilde F(t,t+\alpha)\\
\quad -\big[\mathcal{S}_1(t,\theta)B_1(t)+\mathcal{S}_2(t,\theta,0)\big]\mathcal{R}(t)^{-1}\big[B_1(t)^\top\mathcal{S}_1(t,\alpha)^\top+\mathcal{S}_2(t,0,\alpha)\big]=0, \text{a.e. } t\in[t_0,T],\theta,\alpha\in[-\delta,0],\\
\mathcal{S}_2(t,\theta,-\delta)=\mathcal{S}_1(t,\theta)B_2(t)+B_2(t+\delta+\theta)^\top\int_t^T\mathcal{P}_1(t,\alpha,t+\delta+\theta)B_3(\alpha)\tilde F(\alpha,t-\delta)d\alpha\\
\quad +\int_t^T\int_t^T\tilde F(\alpha,t+\theta)^\top B_3(\alpha)^\top\mathcal{P}_1(t,\alpha,\beta)^\top B_3(\beta)\tilde F(\beta,t-\delta)d\beta d\alpha,\\
\mathcal{S}_2(t,-\delta,\theta)=\mathcal{S}_2(t,\theta,-\delta)^\top,\ \mathcal{S}_2(T,\theta,\alpha)=0.
\end{cases}
\tag{4.8}
$$

In addition, the optimal closed-loop outcome control (3.20) is represented as follows:

$$
\begin{aligned}
u^*(t)=&-\mathcal{R}(t)^{-1}\Big\{\big[B_1(t)^\top\mathcal{S}_0(t)+D_1(t)^\top\mathcal{S}_0(t)C_1(t)+\mathcal{S}_1(t,0)\big]x^*(t)+\int_{t\vee(t_0+\delta)}^{(t+\delta)\wedge T}\big[B_1(t)^\top\mathcal{S}_1(t,r-\delta-t)^\top\\
&+\mathcal{S}_2(t,0,r-\delta-t)\big]u^*(r-\delta)dr\Big\}+\mathcal{R}(t)^{-1}\Big(\int_{t\vee(t_0+\delta)}^{(t+\delta)\wedge T}-\int_{t_0+\delta}^{t+\delta}\Big)\int_t^T\Big\{B_1(t)^\top\mathcal{P}_1(t,t,\theta')\\
&+B_2(t+\delta)^\top\mathcal{P}_1(t,t+\delta,\theta')^\top+\int_t^T\tilde F(\theta,t)^\top B_3(\theta)^\top\mathcal{P}_1(t,\theta,\theta')^\top d\theta\Big\}\\
&\times B_3(\theta')\tilde F(\theta',r-\delta)u^*(r-\delta)d\theta' dr,\ \text{ a.e. } t\in[t_0,T],
\end{aligned}
\tag{4.9}
$$

where $\mathcal{R}(t)=R_1(t)+D_1(t)^\top\mathcal{S}_0(t)D_1(t)$.

We state the main result of this subsection as follows.

Corollary 4.1. *Let* (A1)-(A2) *hold and* $A_2,A_3,C_2,C_3,Q_2,Q_3,R_2,b,\sigma=0$. *Then,* $\mathcal{S}_0(\cdot)$, $\mathcal{S}_1(\cdot,\cdot)$ *and* $\mathcal{S}_2(\cdot,\cdot,\cdot)$, *defined by* (4.3)–(4.5), *satisfy the coupled Riccati equations* (4.6)–(4.8), *and the process in* (4.9) *is the optimal closed-loop outcome control.*

Remark 4.2. *When Problem* (P) *contains only control delays, we obtain the optimal closed-loop outcome control* (4.9) *by the coupled Riccati equations* (4.6)–(4.8). *Furthermore, if the diffusion term disappears in* (4.1), *then* (4.6)–(4.8) *essentially reduce to (2.33)-(2.38) in [31]. Compared with [13, 31, 38, 47, 54, 59], we successfully obtain the solvability of the desired Riccati system.*

4.2 Case II: Stochastic control systems with state delays only

Consider the state equation

$$
\begin{cases}
dx(t)=\big[A_1x(t)+A_2y(t)+B_1u(t)\big]dt+\big[C_1x(t)+C_2y(t)+D_1u(t)\big]dW(t),\quad t\in[t_0,T],\\
x(t)=\xi(t-t_0),\quad t\in[t_0-\delta,t_0],
\end{cases}
$$

along with the cost functional

$$
J(t_0,\xi(\cdot);u(\cdot))=\mathbb{E}\int_{t_0}^T\big[x(t)^\top Q_1x(t)+u(t)^\top R_1u(t)\big]dt.
$$

Here the coefficients are time-invariant. For $t_0 \leqslant t \leqslant \theta \leqslant T$, we define

$$\begin{aligned}\mathcal{P}_2(t) \equiv & \int_t^T \big(I, \mathbf{1}_{(\delta,\infty)}(s-t)I, 0\big)P^{(1)}(s)\big(I, \mathbf{1}_{(\delta,\infty)}(s-t)I, 0\big)^\top ds \\ & + \int_t^T \int_t^T \big(I, \mathbf{1}_{(\delta,\infty)}(s_1-t)I, 0\big)P^{(2)}(s_1, s_2, t)\big(I, \mathbf{1}_{(\delta,\infty)}(s_2-t)I, 0\big)ds_1 ds_2,\end{aligned} \tag{4.10}$$

$$\mathcal{P}_3(t,\theta) \equiv \Big[\big(I, \mathbf{1}_{(\delta,\infty)}(\theta-t)I, 0\big)P^{(1)}(\theta)^\top + \int_t^T \big(I, \mathbf{1}_{(\delta,\infty)}(r-t)I, 0\big)P^{(2)}(\theta, r, t)^\top dr\Big](0, I, 0)^\top. \tag{4.11}$$

In this part, we will show that $\mathcal{P}_2(\cdot)$ and $\mathcal{P}_3(\cdot,\cdot)$ satisfy the following coupled Riccati equations. More precisely, for $t \in (T-\delta, T], \theta \in (t, T]$, we have

$$\begin{cases} -\dot{\mathcal{P}}_2(t) = \mathcal{P}_2(t)A_1 + A_1^\top \mathcal{P}_2(t) + C_1^\top \mathcal{P}_2(t)C_1 + Q_1 \\ \quad -\big(B_1^\top \mathcal{P}_2(t) + D_1^\top \mathcal{P}_2(t)C_1\big)^\top \big(R_1 + D_1^\top \mathcal{P}_2(t)D_1\big)^{-1}\big(B_1^\top \mathcal{P}_2(t) + D_1^\top \mathcal{P}_2(t)C_1\big), \\ -\dfrac{\partial \mathcal{P}_3(t,\theta)}{\partial t} = A_1^\top \mathcal{P}_3(t,\theta) - \big(B_1^\top \mathcal{P}_2(t) + D_1^\top \mathcal{P}_2(t)C_1\big)^\top \big(R_1 + D_1^\top \mathcal{P}_2(t)D_1\big)^{-1} B_1^\top \mathcal{P}_3(t,\theta), \\ \mathcal{P}_3(t,t) = \mathcal{P}_2(t)A_2 + C_1^\top \mathcal{P}_2(t)C_2 - \big(B_1^\top \mathcal{P}_2(t) + D_1^\top \mathcal{P}_2(t)C_1\big)^\top \big(R_1 + D_1^\top \mathcal{P}_2(t)D_1\big)^{-1} D_1^\top \mathcal{P}_2(t)C_2, \\ \mathcal{P}_2(T) = 0, \end{cases} \tag{4.12}$$

while for $t \in [0, T-\delta], \theta \in (t, t+\delta]$,

$$\begin{cases} -\dot{\mathcal{P}}_2(t) = \mathcal{P}_2(t)A_1 + A_1^\top \mathcal{P}_2(t) + C_1^\top \mathcal{P}_2(t)C_1 + C_2^\top \mathcal{P}_2(t+\delta)C_2 + Q_1 \\ \quad +\mathcal{P}_3(t,t+\delta) + \mathcal{P}_3(t,t+\delta)^\top - \big(B_1^\top \mathcal{P}_2(t) + D_1^\top \mathcal{P}_2(t)C_1\big)^\top \big(R_1 + D_1^\top \mathcal{P}_2(t)D_1\big)^{-1}\big(B_1^\top \mathcal{P}_2(t) \\ \quad +D_1^\top \mathcal{P}_2(t)C_1\big) - \big(D_1^\top \mathcal{P}_2(t+\delta)C_2\big)^\top \big(R_1 + D_1^\top \mathcal{P}_2(t)D_1\big)^{-1} D_1^\top \mathcal{P}_2(t+\delta)C_2, \\ -\dfrac{\partial \mathcal{P}_3(t,\theta)}{\partial t} = A_1^\top \mathcal{P}_3(t,\theta) - \big(B_1^\top \mathcal{P}_2(t) + D_1^\top \mathcal{P}_2(t)C_1\big)^\top \big(R_1 + D_1^\top \mathcal{P}_2(t)D_1\big)^{-1} B_1^\top \mathcal{P}_3(t,\theta) \\ \quad +\mathcal{P}_3(\theta, t+\delta)^\top A_2 - \big(B_1^\top \mathcal{P}_3(\theta, t+\delta)\big)^\top \big(R_1 + D_1^\top \mathcal{P}_2(t)D_1\big)^{-1} D_1 \mathcal{P}_2(t)C_2 \\ \quad -\displaystyle\int_t^\theta \big(B_1^\top \mathcal{P}_3(s, t+\delta)\big)^\top \big(R_1 + D_1^\top \mathcal{P}_2(t)D_1\big)^{-1} B_1^\top \mathcal{P}_3(s,\theta)ds, \\ \mathcal{P}_3(t,t) = \mathcal{P}_2(t)A_2 + C_1^\top \mathcal{P}_2(t)C_2 - \big(B_1^\top \mathcal{P}_2(t) + D_1^\top \mathcal{P}_2(t)C_1\big)^\top \big(R_1 + D_1^\top \mathcal{P}_2(t)D_1\big)^{-1} D_1^\top \mathcal{P}_2(t)C_2. \end{cases} \tag{4.13}$$

We state the main result of this subsection as follows.

Corollary 4.3. *Let* (A1)–(A2) *hold with* $A_3, B_2, B_3, C_3, Q_2, Q_3, R_2, b, \sigma = 0$. *Then,* $\mathcal{P}_2(\cdot)$ *and* $\mathcal{P}_3(\cdot,\cdot)$, *defined by* (4.10)–(4.11), *satisfy the coupled Riccati equation* (4.12)–(4.13). *In this case, the following process is the optimal closed-loop outcome control*

$$\begin{aligned} u^*(t) = & -\big(R_1 + D_1^\top \mathcal{P}_2(t)D_1\big)^{-1}\Big\{\big(B_1^\top \mathcal{P}_2(t) + D_1^\top \mathcal{P}_2(t)C_1\big)x^*(t) + \int_{t\vee(t_0+\delta)}^{(t+\delta)\wedge T} B_1^\top \mathcal{P}_3(t,s)x^*(s-\delta)ds \\ & +D_1^\top \mathcal{P}_2(t)C_2 x^*(t-\delta) + \int_t^{T\wedge(t_0+\delta)} B_1^\top \mathcal{P}_3(t,s)\xi(s-t_0-\delta)ds\Big\}, \ t \in [t_0, T]. \end{aligned} \tag{4.14}$$

Remark 4.4. *When Problem* (P) *contains only state delays, the coupled Riccati equations* (4.12)–(4.13) *are the same as (3)–(12) in [38], the optimal closed-loop outcome control* (4.14) *coincides with (13) in [38], and Corollary 4.3 is similar to Theorem 1 in [38]. Compared with [36, 38, 39], one advantage lies in the positive result on the solvability of the Riccati system.*

4.3 Case III: Stochastic control systems with pointwise delays only

Consider the state equation

$$\begin{cases} dx(t) = \big[A_1(t)x(t) + A_2(t)y(t) + B_1(t)u(t) + B_2(t)\nu(t)\big]dt \\ \qquad\qquad +\big[C_1(t)x(t) + C_2(t)y(t) + D_1(t)u(t)\big]dW(t), \quad t \in (t_0, T), \\ x(t) = \xi(t-t_0), \ u(t) = \varsigma(t-t_0), \ t \in [t_0-\delta, t_0], \end{cases}$$

along with the cost functional

$$J(t_0,\xi(\cdot),\varsigma(\cdot);u(\cdot)) = \mathbb{E}\int_{t_0}^{T}\left[x(t)^\top Q_1(t)x(t) + y(t)^\top Q_2(t)y(t) + u(t)^\top R_1(t)u(t) + \nu(t)^\top R_2(t)\nu(t)\right]dt.$$

For $0<t<(s\wedge\bar{s})<T$, we define $\mathcal{G}_1(\cdot)$, $\mathcal{G}_2(\cdot,\cdot)$ as in (3.12), and

$$\begin{aligned}
&Q(t)=\begin{bmatrix}Q_1(t) & 0\\ 0 & Q_2(t)\end{bmatrix},\quad \mathcal{C}(t)=\big(C_1(t),C_2(t)\big),\quad \mathcal{A}(t)=\big(A_1(t),A_2(t)\big),\quad \mathcal{B}(t)=\begin{bmatrix}B_1(t)\\ B_2(t+\delta)\end{bmatrix},\\
&\Upsilon(s,t)=\begin{bmatrix}I\\ \mathbf{1}_{(\delta,\infty)}(s-t)I\end{bmatrix},\qquad \mathcal{R}(t)=R_1(t)+R_2(t+\delta)\mathbf{1}_{[0,T-\delta)}(t)+D_1(t)^\top\mathcal{G}_1(t)D_1(t),\\
&\mathcal{G}_3(s,t)=P^{(1)}(s)\Big(\Upsilon(s,t),\mathbf{1}_{(\delta,\infty)}(s-t)\Upsilon(s,t+\delta)\Big)\\
&\qquad\qquad+\int_t^T P^{(2)}(s,r,t)\Big(\Upsilon(r,t),\mathbf{1}_{(\delta,\infty)}(r-t)\Upsilon(r,t+\delta)\Big)dr.
\end{aligned}\tag{4.15}$$

We consider the Riccati system (3.11), where $P^{(2)}(\cdot,\cdot,\cdot)$ satisfies

$$\begin{cases}P^{(2)}(s,t,r)=P^{(2)}(s,t,t\wedge s)-\displaystyle\int_r^{t\wedge s}\mathcal{G}_3(s,\tau)\mathcal{B}(\tau)\mathcal{R}(\tau)^{-1}\mathcal{B}(\tau)^\top\mathcal{G}_3(t,\tau)^\top d\tau,\ 0<r<(s\wedge t)<T,\\ P^{(2)}(\bar{s},t,t)=P^{(2)}(t,\bar{s},t)^\top=\mathcal{G}_2(\bar{s},t)\mathcal{A}(t)-\mathcal{G}_3(\bar{s},t)\mathcal{B}(t)\mathcal{R}(t)^{-1}D_1(t)^\top\mathcal{G}_1(t)\mathcal{C}(t),\ 0<t<\bar{s}<T.\end{cases}\tag{4.16}$$

With notations in (4.15), we consider the closed-loop strategy:

$$\begin{aligned}
&K_i^*(t)=-\mathcal{R}(t)^{-1}\Big\{K_i^{(1)}(t)+\mathcal{B}(t)^\top\int_t^T\mathcal{G}_3(\alpha,t)^\top K_i^{(2)}(\alpha,t)d\alpha\Big\},\ i=1,3,\ t_0\leqslant t\leqslant T,\\
&K_i^*(t,s)=-\mathcal{R}(t)^{-1}\Big\{K_i^{(1)}(t,s)+\mathcal{B}(t)^\top\int_t^T\mathcal{G}_3(\alpha,t)^\top K_i^{(2)}(t,\alpha,s)d\alpha\Big\},\ i=2,4,\ \ t_0\leqslant s<t\leqslant T,\\
&v^*(t)=-\mathcal{R}(t)^{-1}\Big\{v^{(1)}(t)+\mathcal{B}(t)^\top\int_t^T\mathcal{G}_3(\alpha,t)^\top v^{(2)}(t,\alpha)d\alpha\Big\},\ t_0\leqslant t\leqslant T,
\end{aligned}\tag{4.17}$$

where $K_1^{(1)},K_1^{(2)},K_3^{(1)},K_3^{(2)},K_4^{(1)}$ have the same forms as in (3.18), and

$$\begin{aligned}
&K_2^{(1)}(t,s)=\mathcal{B}(t)^\top\mathcal{G}_3(s+\delta,t)^\top(0,I)^\top\mathbf{1}_{[t-\delta,T-\delta]}(s),\qquad\qquad K_2^{(2)}(t,\alpha,s)=0,\\
&K_4^{(2)}(t,\alpha,s)=\big(I,\mathbf{1}_{(s+2\delta,\infty)}(\alpha)\mathbf{1}_{(0,T-\delta)}(t)I\big)^\top\mathbf{1}_{(s+\delta,T)}(\alpha)B_2(s+\delta)\mathbf{1}_{[t-\delta,T-\delta]}(s),\\
&v^{(1)}(t)=0,\quad v^{(2)}(t,\alpha)=(0,I)^\top\xi(\alpha-\delta-t_0)\mathbf{1}_{[t_0,t_0+\delta]}(\alpha)+\int_{t-\delta}^{t_0}\mathbf{1}_{(\theta'+\delta,\infty)}(\alpha)\\
&\times\big(I,\mathbf{1}_{(\theta'+2\delta,\infty)}(\alpha)\mathbf{1}_{(0,T-\delta)}(t)I\big)^\top B_2(\theta'+\delta)\varsigma(\theta'-t_0)d\theta'\mathbf{1}_{[t_0,t_0+\delta]}(t).
\end{aligned}$$

As a result of Theorem 3.6, we derive the following result.

Corollary 4.5. *Let* (A1)–(A2) *hold with* $A_3,B_3,C_3,Q_3,b,\sigma=0$. *Then, all the strategies in* (4.17) *and the process* $u^*(\cdot)$ *in the same form of* (3.20) *are optimal.*

Remark 4.6. *We make some comparisons with the existing literature. Firstly, in contrast with [38, 44], our advantages lie in the facts that our coefficients are allowed to be time-variant and the cost functional depends on both the pointwise state delays and pointwise control delays. Secondly, even when these features disappear, our framework is still general than that in [38] (except the disappearance of pointwise control delays in the diffusion term), and the solvability of the Riccati system is given. In addition, [44] requires the diffusion term to be independent of the pointwise state delays and control variables, while we drop this assumption here. Thirdly, even though the coefficients in [15] are time-variant, both their state equations and cost functional are still particular cases of ours. Eventually, even for deterministic systems, our coefficients are allowed to be integrable, but not necessarily differentiable as in [16].*

4.4 Case IV: Stochastic control systems with distributed delays only

Consider the state equation

$$\begin{cases} dx(t) = [A_1(t)x(t) + A_3(t)z(t) + B_1(t)u(t) + B_3(t)\mu(t)]dt \\ \qquad\qquad +[C_1(t)x(t) + C_3(t)z(t) + D_1(t)u(t)]dW(t), \quad t \in (t_0, T), \\ x(t) = \xi(t-t_0),\ u(t) = \varsigma(t-t_0),\ t \in [t_0-\delta, t_0], \end{cases}$$

along with the cost functional

$$J(t_0, \xi(\cdot), \varsigma(\cdot); u(\cdot)) = \mathbb{E}\int_{t_0}^T \left[x(t)^\top Q_1(t)x(t) + z(t)^\top Q_3(t)z(t) + u(t)^\top R_1(t)u(t)\right] dt.$$

In this case, we consider the Riccati system (3.11), where for $0 < t < (s \wedge \bar{s} \wedge \theta) < T$,

$$\begin{aligned} &Q(t) = \begin{bmatrix} Q_1(t) & 0 \\ 0 & Q_3(t) \end{bmatrix}, \quad \mathcal{C}(t) = \big(C_1(t), C_3(t)\big), \quad \mathcal{A}(t) = \big(A_1(t), A_3(t)\big), \quad \mathcal{B}(\theta, t) = \begin{bmatrix} B_1(t) \\ B_3(\theta)\tilde{F}(\theta, t) \end{bmatrix}, \\ &\Upsilon(s,t) = \begin{bmatrix} I \\ \mathcal{E}(s,t) \end{bmatrix}, \quad \mathcal{R}(t) = R_1(t) + D_1(t)^\top \mathcal{G}_1(t) D_1(t), \quad \Pi(s,t,\theta) = \begin{bmatrix} \frac{1}{s-t}I & I \\ \frac{1}{s-t}\mathcal{E}(s,t) & \mathcal{E}(s,\theta) \end{bmatrix}, \end{aligned} \tag{4.18}$$

and $\mathcal{G}_1(\cdot)$, $\mathcal{G}_2(\cdot,\cdot)$, $\mathcal{G}_3(\cdot,\cdot,\cdot)$ are defined in (3.12). With the notations in (4.18), we consider the closed-loop strategy (3.15)–(3.19), where $K_1^{(1)}, K_1^{(2)}, K_3^{(2)}, K_4^{(1)}$ have the same forms as in (3.18), $v^{(1)}$, $v^{(2)} = 0$, and

$$\begin{aligned} &K_3^{(1)}(t) = 0, \quad K_2^{(1)}(t,s) = D_1(t)^\top \mathcal{G}_1(t) C_3(t) F(t,s), \quad K_2^{(2)}(t,\alpha,s) = (0, I)^\top F(\alpha, s), \\ &K_4^{(2)}(t,\alpha,s) = \Big(\tilde{\mathcal{F}}(t,s,\alpha), \int_t^\alpha \tilde{\mathcal{F}}(t,s,\beta) F(\alpha,\beta)^\top d\beta\Big)^\top, \quad \tilde{\mathcal{F}}(t,s,\tau) \equiv \int_t^\tau \tilde{F}(\theta', s)^\top B_3(\theta')^\top d\theta'. \end{aligned} \tag{4.19}$$

Then, as a result of Theorem 3.6, we derive the following result.

Corollary 4.7. *Let* (A1)–(A2) *hold and* $A_2, B_2, C_2, Q_2, R_2, b, \sigma = 0$. *Then, all the closed-loop strategies* (3.15)–(3.16) *with* (4.19) *and the following process are optimal:*

$$u^*(t) = K_1^*(t)x^*(t) + \int_{t_0}^t K_2^*(t,s)x^*(s)ds + \int_{t_0}^t K_4^*(t,s)u^*(s)ds, \ t \in [t_0, T]. \tag{4.20}$$

Remark 4.8. *LQ problems for deterministic integro-differential equations were treated in [48], see also [26, Subsection 5.5]. Both of them derived the Riccati systems and the closed-loop optimal controls in the spirit of* (4.20). *However, they require that* $A_1, B_3 \equiv 0$, $A_3 \equiv 1$, $F(\cdot,\cdot)$ *is the convolution kernel,* $Q_3 \equiv 0$, Q_1, R_1 *are time-invariant.*

As to the case of stochastic integro-differential equations (SVIDEs), let us point out the discussion in [26, Subsection 5.3] where the SVIDE is directly regarded as a special SVIE. However, there are no closed-loop controls in the spirit of (4.20) *and the corresponding Riccati systems. To our best knowledge, the above Corollary 4.7 appears for the first time.*

4.5 Case V: Stochastic control systems without delay

In this subsection, let us look at the particular SDEs case. To this end, we consider

$$\begin{cases} \dot{\mathcal{P}}(t) + \mathcal{P}(t)A_1(t) + A_1(t)^\top \mathcal{P}(t) + C_1(t)^\top \mathcal{P}(t)C_1(t) + Q_1(t) - \big(\mathcal{P}(t)B_1(t) + C_1(t)^\top \mathcal{P}(t)D_1(t)\big) \\ \quad \times \mathcal{R}(t)^{-1}\big(B_1(t)^\top \mathcal{P}(t) + D_1(t)^\top \mathcal{P}(t)C_1(t)\big) = 0, \quad 0 < t < T, \\ \mathcal{P}(T) = 0, \end{cases} \tag{4.21}$$

and the following backward stochastic differential equation:

$$\begin{cases} d\widetilde{\eta}(t) = -\big\{\big[A_1^\top(t) - \big(\mathcal{P}(t)B_1(t) + C_1^\top(t)\mathcal{P}(t)D_1(t)\big)\mathcal{R}(t)^{-1}B_1^\top(t)\big]\widetilde{\eta}(t) \\ \qquad + \big[C_1^\top(t) - \big(\mathcal{P}(t)B_1(t) + C_1^\top(t)\mathcal{P}(t)D_1(t)\big)\mathcal{R}(t)^{-1}D_1^\top(t)\big]\tilde{\zeta}(t) \\ \qquad + \big[C_1^\top(t) - \big(\mathcal{P}(t)B_1(t) + C_1^\top(t)\mathcal{P}(t)D_1(t)\big)\mathcal{R}(t)^{-1}D_1^\top(t)\big]\mathcal{P}(t)\sigma(t) \\ \qquad + \mathcal{P}(t)b(t)\big\}dt + \tilde{\zeta}(t)dW(t), \quad 0 < t < T, \\ \widetilde{\eta}(T) = 0, \end{cases} \tag{4.22}$$

where $\mathcal{R}(\cdot) = R_1(\cdot) + D_1(\cdot)^\top \mathcal{P}(\cdot) D_1(\cdot)$.

On the other hand, given $P = (P^{(1)}, P^{(2)})$, (η, ζ) satisfying the Riccati system (3.11) and the backward SVIE (3.13), respectively, we define

$$\begin{aligned} &\mathcal{P}(t) = (I,0,0)\Big(\int_t^T P^{(1)}(s)ds + \int_t^T \int_t^T P^{(2)}(s_1,s_2,t)ds_1 ds_2 \Big)(I,0,0)^\top, \ 0<t<T, \\ &\widetilde{\eta}(t) = \int_t^T (I,0,0)\eta(s,t)ds, \qquad \tilde{\zeta}(t) = \int_t^T (I,0,0)\zeta(s,t)ds, \ 0<t<T. \end{aligned} \tag{4.23}$$

Then, we have the following result.

Corollary 4.9. *Let* (A1)-(A2) *hold such that* $A_2, A_3, B_2, B_3, C_2, C_3, Q_2, Q_3, R_2 = 0$. *Then, the above* $\mathcal{P}(\cdot)$ *and* $(\widetilde{\eta}(\cdot), \widetilde{\zeta}(\cdot))$ *are the unique solutions to* (4.21) *and* (4.22), *respectively. In addition, the five-tuple* $(K_1^*(\cdot), 0, 0, 0, v^*(\cdot))$ *is the optimal closed-loop strategy of Problem* (P), *where*

$$\begin{aligned} &K_1^*(t) = -\mathcal{R}(t)^{-1}\left(B_1(t)^\top \mathcal{P}(t) + D_1(t)^\top \mathcal{P}(t) C_1(t)\right), \ t \in [t_0, T], \\ &v^*(t) = -\mathcal{R}(t)^{-1}\Big(B_1^\top(t)\widetilde{\eta}(t) + D_1^\top(t)\tilde{\zeta}(t) + D_1^\top(t)\mathcal{P}(t)\sigma(t)\Big), \ t \in [t_0, T]. \end{aligned}$$

Remark 4.10. *When delays appear in Problem* (P), *the optimal closed-loop strategy* $(K_1^*(\cdot), v^*(\cdot))$, *the equation* (4.21) *and the equation* (4.22) *reduce to that in [52].*

5 Concluding remarks

In this paper we study a general stochastic LQ optimal control problem, where the coefficients are time-varying, and both state delays and control delays can appear in the state equation and the cost functional. We put the original state process and its delay processes together for dimension expansion, and use the Volterra integral system without delay to describe the new process, then transform the original delayed problem into the control problem without delay. Based on the equivalent problem, we propose the closed-loop solvability of the delayed problem, and assure it by the solvability of the Riccati system and the extended backward SVIEs. Furthermore, we derive the optimal closed-loop outcome control and obtain the solvability of the associated Riccati system. Finally, we study several important stochastic systems and find that our results are consistent with those in the existing literature.

References

[1] P. Acquistapace and F. Bucci. Riccati-based solution to the optimal control of linear evolution equations with finite memory. *Evol. Equ. Control Theory*, 13:2666, 2024.

[2] P. Acquistapace and F. Bucci. Optimal synthesis control for evolution equations subject to nonlocal inputs. *J. Optim. Theory Appl.*, 205(37):https://doi.org/10.1007/s10957–025–02661–0, 2025.

[3] Y. Alekal, P. Brunovsky, D. Chyung, and E. Lee. The quadratic problem for systems with time delays. *IEEE Trans. Automat. Control*, 16(6):673–687, 1971.

[4] A. Bagchi. Control of linear stochastic time delayed systems. *J. Math. Anal. Appl.*, 63:244–252, 1978.

[5] T. Basar, J. Moon, J. Yong, and H. Zhang. Nonstandard linear-quadratic decision making. In *2023 62nd IEEE Conference on Decision and Control (CDC)*, pages 8901–8920, Marina Bay Sands, Singapore, December 13–15, 2023.

[6] A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter. *Representation and Control of Infinite Dimensional Systems*. Birkhäuser, Boston, 2nd edition, 2007.

[7] P. Cannarsa, H. Frankowska, and E. Marchini. Optimal control for evolution equations with memory. *J. Evol. Equ.*, 13:197–227, 2013.

[8] G. Carlier and R. Tahraoui. Hamilton-Jacobi-Bellman equations for the optimal control of a state equation with memory. *ESAIM: Control Optim. Calc. Var.*, 16:744–763, 2010.

[9] G. Carlier and R. Tahraoui. On some optimal control problems governed by a state equation with memory. *ESAIM: Control Optim. Calc. Var.*, 14:725–743, 2018.

[10] R. Carmona, J. P. Fouque, S. M. Mousavi, and L. H. Sun. Systemic risk and stochastic games with delay. *J. Optim. Theory Appl.*, 179(2):366–399, 2018.

[11] M. Chang. *Stochastic Control of Hereditary Systems and Applications.* Springer, 2008.

[12] L. Chen and Z. Wu. Stochastic optimal control problem in advertising model with delay. *J. Syst. Sci. Complex*, 33:968–987, 2020.

[13] L. Chen and Z. Wu. The quadratic problem for stochastic linear control systems with delay. In *Proc. of the 30th Chinese Control Conference*, pages 1344–1349, Yantai, China, July 22-24, 2011.

[14] L. Chen, Z. Wu, and Z. Yu. Delayed stochastic linear-quadratic control problem and related applications. *J. Appl. Math.*, 835319:1–22, 2012.

[15] L. Chen and Y. Zhang. Linear quadratic optimal control problems of state delay systems under full and partial information. *Systems & Control Lett.*, 176:105528, 2023.

[16] M. C. Delfour. The linear quadratic optimal control problem for hereditary differential systems: Theory and numerical solution. *Appl. Math. Optim.*, 3(2):101–162, 1976.

[17] M. C. Delfour. State theory of linear hereditary differential systems. *J. Math. Anal. Appl.*, 60:8–35, 1977.

[18] M. C. Delfour. The linear-quadratic optimal control problem with delays in state and control variables: A state space approach. *SIAM J. Control Optim.*, 24(5):835–883, 1986.

[19] M. C. Delfour and S. K. Mitter. Controllability, observability and optimal feedback control of affine hereditary differential systems. *SIAM J. Control Optim.*, 10:298–328, 1972.

[20] D. H. Eller, J. K. Aggarwal, and H. T. Banks. Optimal control of linear time-delay systems. *IEEE Trans. Automat. Control*, 16(6):678687, 1969.

[21] G. Fabbri and S. Federico. On the infinite-dimensional representation of stochastic controlled systems with delayed control in the diffusion term. *Math. Econ. Lett.*, 2:33–43, 2014.

[22] S. Federico. A stochastic control problem with delay arising in a pension fund model. *Finance Stoch.*, 15:421–459, 2011.

[23] F. De Feo, S. Federico, and A. Swiech. Optimal control of stochastic delay differential equations and applications to path-dependent financial and economic models. *SIAM J. Control Optim.*, 62:1490–1520, 2024.

[24] F. De Feo and A. Swiech. Optimal control of stochastic delay differential equations: Optimal feedback controls. *J. Differential Equation*, 420:450–508, 2025.

[25] F. Flandoli. Solution and control of a bilinear stochastic delay equation. *SIAM J. Control Optim.*, 28:936–949, 1990.

[26] J. Gong and T. Wang. Causal feedback strategies for controlled stochastic volterra systems: A unified treatment. *SIAM J. Control Optim.*, 63(3):1950–1980, 2025.

[27] J. Hale. *Theory of Functional Differential Equations.* Springer-Verlag New York, 1977.

[28] Y. Hamaguchi and T. Wang. Linear-quadratic stochastic Volterra controls I: Causal feedback strategies. *Stochastic Process. Appl.*, 176, 2024.

[29] Y. Hamaguchi and T. Wang. Linear-quadratic stochastic Volterra controls II: Optimal strategies and Riccati-Volterra equations. *ESAIM: Control Optim. Calc. Var.*, 30, 2024.

[30] J. Huang, X. Li, and T. Wang. Mean-field linear-quadratic-gaussian (LQG) games for stochastic integral systems. *IEEE Trans. Automat. Control*, 61:2670–2675, 2016.

[31] A. Ichikawa. Quadratic control of evolution equations with delays in control. *SIAM J. Control Optim.*, 20:645–668, 1982.

[32] V. Kolmanovskii and A. Myshkis. *Applied Theory of Functional Differential Equations.* Springer Dordrecht, 1992.

[33] V. Kolmanovskii and L. Shaikhet. *Control of Systems with Aftereffect.* AMS, Providence, 1996.

[34] B. Larssen. Dynamic programming in stochastic control of systems with delay. *Stoch. Stochastic Rep.*, 74:651–673, 2002.

[35] E. B. Lee and Y. You. Optimal syntheses for infinite-dimensional linear delayed state-output systems: A semicausality approach. *Appl. Math. Optim.*, 19:113–136, 1989.

[36] E. B. Lee and Y. You. Quadratic optimization for infinite-dimensional linear differential difference type systems: Syntheses via the Fredholm equation. *SIAM J. Control Optim.*, 28(2):265–293, 1990.

[37] W. Lefebvre and E. Miller. Linear-quadratic stochastic delayed control and deep learning resolution. *J. Optim. Theory Appl.*, 191:134–168, 2021.

[38] X. Liang, J. Xu, and H. Zhang. Solution to stochastic LQ control problem for Itô systems with state delay or input delay. *Systems & Control Lett.*, 113:86–92, 2018.

[39] A. Lindquist. Optimal control of linear stochastic systems with applications to time lag systems. *Inform. Sciences*, 5:81–126, 1973.

[40] T. Ma, J. Xu, and H. Zhang. Explicit solution to forward and backward stochastic differential equations with state delay and its application to optimal control. *Control Theory Technol.*, 20:303–315, 2022.

[41] X. Ma, Q. Qi, X. Li, and H. Zhang. Optimal control and stabilization for linear continuous-time mean-field systems with delay. *IET Control Theory Appl.*, 16:283–300, 2022.

[42] W. Meng, J. Shi, T. Wang, and J. Zhang. A general maximum principle for optimal control of stochastic differential delay systems. *SIAM J. Control Optim.*, 63(1):175–205, 2025.

[43] W. Meng, J. Shi, and J. Yong. *Time-Delayed Linear Quadratic Optimal Control Problems.* SpringerBriefs on PDEs and Data Science, Singapore, 2025.

[44] W. Meng, J. Shi, J. Zhang, and Y. Zhao. Necessary and sufficient conditions of open-loop and closed-loop solvability for delayed stochastic LQ optimal control problems. *SIAM J. Control Optim.*, 63(3):1736–1767, 2025.

[45] S. E. A. Mohammed. *Stochastic Functional Differential Equations.* Pitman, 1984.

[46] S. E. A. Mohammed. Stochastic differential equations with memory: theory, examples and applications. In *Stochastic Analysis and Related Topics 6, Proceedings of the 6th Oslo-Silivri Workshop Geilo 1996, Progress in Probability 42, L. Decreusefond, Jon Gjerde, B. Oksendal and A. S. Üstünel, eds., Birkhäuser, Boston*, pages 1–77, 1998.

[47] J. Moon. State and control path-dependent stochastic optimal control with jumps. *IEEE Trans. Automat. Control*, 67(9):4555–4567, 2022.

[48] L. Pandolfi. The quadratic regulator problem and the Riccati equation for a process governed by a linear Volterra integro-differential equations. *IEEE Trans. Automat. Control*, 63:15171522, 2018.

[49] L. Pandolfi. The quadratic tracking problem for systems with persistent memory in $\mathbb{R}^d$. *Evol. Equ. Control Theory*, 14:185209, 2025.

[50] J. Sun, X. Li, and J. Yong. Open-loop and closed-loop solvabilities for stochastic linear quadratic optimal control problems. *SIAM J. Control Optim.*, 54:2274–2308, 2016.

[51] J. Sun and J. Yong. Linear quadratic stochastic differential games: open-loop and closed-loop saddle points. *SIAM J. Control Optim.*, 52:4082–4121, 2014.

[52] J. Sun and J. Yong. *Stochastic Linear-Quadratic Optimal Control Theory: Open-Loop and Closed-Loop Solutions*. Springer Briefs in Mathematics, Springer, Cham, 2020.

[53] K. Uchida, E. Shimemura, T. Kubo, and N. Abe. The linear-quadratic optimal control approach to feedback control design for systems with delay. *Automatica*, 24(6):773–780, 1988.

[54] H. Wang and H. Zhang. LQ control for Itô-type stochastic systems with input delays. *Automatica*, 49:3538–3549, 2013.

[55] J. Xu, J. Shi, and H. Zhang. A leader-follower stochastic linear quadratic differential game with time delay. *Sci. China Inf. Sci.*, 61, 2018.

[56] Y. You. Retarded control problems of distributed parameter systems. *Acta Math. Appl. Sin.*, 6(3):347–356, 1983.

[57] Y. You. Quadratic integral games and causal synthesis. *T. Am. Math. Soc.*, 352(6):2737–2764, 1999.

[58] H. Zhang and L. Xie. *Control and Estimation of Systems with Input/Output Delays*. Lecture Notes in Control and Information Sciences, Springer-Verlag Berlin Heidelberg, 2007.

[59] H. Zhang and J. Xu. Control for Itô stochastic systems with input delay. *IEEE Trans. Automat. Control*, 62:350–365, 2017.

[60] S. Zhang, J. Xiong, and J. Shi. A linear-quadratic optimal control problem of stochastic differential equations with delay and partial information. *Systems & Control Lett.*, 157, 2021.

Appendix

The proof of Proposition 2.1:

Proof. Let $\beta > 0$, $T_0 \in (t_0, T]$ be undetermined constants, and

$$\mathcal{M}_\beta[t_0, T_0] \equiv L^2_{\mathbb{F}}(\Omega; C([t_0, T_0]; \mathbb{R}^n)),$$

equipped with the norm

$$||x(\cdot)||_{\mathcal{M}_\beta[t_0,T_0]} \equiv \Big(\mathbb{E}\big[\sup_{t_0 \leqslant t \leqslant T_0} e^{\beta h(t)}|x(t)|^2\big]\Big)^{\frac{1}{2}}.$$

Here

$$h(t) \equiv \int_{t_0}^t \Big(|\bar{A}_1(s)|^2 + |\bar{A}_2(s)|^2 + |\bar{A}_3(s)|^2\Big)ds, \quad t \in [t_0, T_0].$$

For any $\bar{x}(\cdot) \in \mathcal{M}_\beta[t_0, T_0]$ with $\bar{x}(t) = \bar{\xi}(t - t_0)$, $t \in [t_0 - \delta, t_0]$, consider the mapping

$$\begin{aligned} \mathcal{T} : \mathcal{M}_\beta[t_0, T_0] &\to \mathcal{M}_\beta[t_0, T_0], \\ \bar{x}(\cdot) &\mapsto \bar{X}(\cdot), \end{aligned}$$

where $\bar{X}(\cdot)$ is the solution to the following SDDE:

$$\begin{cases} d\bar{X}(t) = [\bar{A}_1(t)\bar{x}(t) + \bar{A}_2(t)\bar{y}(t) + \bar{A}_3(t)\bar{z}(t) + \bar{b}(t)]dt \\ \qquad\qquad +[\bar{C}_1(t)\bar{x}(t) + \bar{C}_2(t)\bar{y}(t) + \bar{C}_3(t)\bar{z}(t) + \bar{\sigma}(t)]dW(t), \quad t \in (t_0, T), \\ \bar{X}(t) = \bar{\xi}(t - t_0), \quad t \in [t_0 - \delta, t_0]. \end{cases} \tag{A.1}$$

For any $\bar{x}_1(\cdot), \bar{x}_2(\cdot) \in \mathcal{M}_\beta[t_0, T_0]$ and $\bar{x}_1(t), \bar{x}_2(t) = \bar{\xi}(t-t_0)$, $t \in [t_0-\delta, t_0]$, denote

$$\bar{X}_1(\cdot) = \mathcal{T}(\bar{x}_1(\cdot)), \bar{X}_2(\cdot) = \mathcal{T}(\bar{x}_2(\cdot)), \hat{X}(\cdot) = \bar{X}_1(\cdot) - \bar{X}_2(\cdot), \hat{x}(\cdot) = \bar{x}_1(\cdot) - \bar{x}_2(\cdot).$$

Then, applying Itô formula to $s \mapsto e^{-\beta h(s)}|\hat{X}(s)|^2$ on $[t_0, t]$, we have

$$\begin{aligned} e^{-\beta h(t)}|\hat{X}(t)|^2 \leqslant & \int_{t_0}^t e^{-\beta h(s)} \Big[(1-\beta)|\hat{X}(s)|^2\big(|\bar{A}_1(s)|^2 + |\bar{A}_2(s)|^2 + |\bar{A}_3(s)|^2\big) \\ & + \big(2|\bar{C}_1(s)|^2+1\big)|\hat{x}(s)|^2 + \big(2|\bar{C}_2(s)|^2+1\big)|\hat{y}(s)|^2 + \big(2|\bar{C}_3(s)|^2+1\big)|\hat{z}(s)|^2\Big]ds \\ & + \int_{t_0}^t 2\big\langle e^{-\beta h(s)}\hat{X}(s), \bar{C}_1(s)\hat{x}(s) + \bar{C}_2(s)\hat{y}(s) + \bar{C}_3(s)\hat{z}(s)\big\rangle dW(s). \end{aligned}$$

Notice that

$$\begin{aligned} & \mathbb{E}\int_{t_0}^{T_0} e^{-\beta h(s)}\big(2|\bar{C}_3(s)|^2+1\big)|\hat{z}(s)|^2 ds \\ & \leqslant \mathbb{E}\Big[\sup_{t_0 \leqslant s \leqslant T_0}\Big\{e^{-\beta h(s)}|\hat{x}(s)|^2\Big(2|\bar{C}_3(s)|^2+1\Big)\Big\}\int_{t_0}^{T_0}\Big(\int_{t_0}^s |\bar{F}(s,r)|dr\Big)^2 ds\Big], \end{aligned}$$

and

$$\begin{aligned} & \mathbb{E}\Big[\sup_{t_0 \leqslant t \leqslant T_0}\Big|2\int_{t_0}^t \langle e^{-\beta h(s)}\hat{X}(s), \bar{C}_1(s)\hat{x}(s) + \bar{C}_2(s)\hat{y}(s) + \bar{C}_3(s)\hat{z}(s)\rangle dW(s)\Big|\Big] \\ & \leqslant \mathbb{E}\Big(\int_{t_0}^{T_0} 4e^{-2\beta h(s)}|\hat{X}(s)|^2\big|\bar{C}_1(s)\hat{x}(s) + \bar{C}_2(s)\hat{y}(s) + \bar{C}_3(s)\hat{z}(s)\big|^2 ds\Big)^{\frac{1}{2}} \\ & \leqslant \frac{1}{2}\mathbb{E}\sup_{t_0 \leqslant t \leqslant T_0}\Big\{e^{-\beta h(s)}|\hat{X}(s)|^2\Big\} + \Big(4(T_0-t_0)\sup_{0 \leqslant s \leqslant T}\Big(|\bar{C}_1(s)|^2 + |\bar{C}_2(s)|^2\Big) \\ & \quad + 4\sup_{0 \leqslant s \leqslant T}|\bar{C}_3(s)|^2\int_{t_0}^{T_0}\Big(\int_{t_0}^s |\bar{F}(s,r)|dr\Big)^2 ds\Big)\mathbb{E}\sup_{t_0 \leqslant s \leqslant T_0}\Big\{e^{-\beta h(s)}|\hat{x}(s)|^2\Big\}. \end{aligned}$$

Let $\beta > 1$. Then, we deduce

$$\begin{aligned} \mathbb{E}\sup_{t_0 \leqslant t \leqslant T_0}\{e^{-\beta h(t)}|\hat{X}(t)|^2\} \leqslant & 2\Big[6(T_0-t_0)\sup_{t_0 \leqslant s \leqslant T_0}\Big\{1 + |\bar{C}_1(s)|^2 + |\bar{C}_2(s)|^2\Big\} \\ & + 3\sup_{t_0 \leqslant s \leqslant T_0}\Big(2|\bar{C}_3(s)|^2+1\Big)\int_{t_0}^{T_0}\Big(\int_{t_0}^s |\bar{F}(s,r)|dr\Big)^2 ds\Big]\mathbb{E}\sup_{t_0 \leqslant s \leqslant T_0}\big[e^{-\beta h(s)}|\hat{x}(s)|^2\big]. \end{aligned}$$

Since $\bar{C}_1(\cdot), \bar{C}_2(\cdot), \bar{C}_3(\cdot) \in L^\infty(0,T;\mathbb{R}^{n\times n})$ and $\bar{F}(\cdot,\cdot) \in L^{2,1}(\Delta_2(0,T);\mathbb{R}^{n\times n})$, we can choose T_0 to be small enough such that

$$12(T_0-t_0)\sup_{t_0 \leqslant s \leqslant T_0}\Big\{1 + |\bar{C}_1(s)|^2 + |\bar{C}_2(s)|^2\Big\} + 6\sup_{t_0 \leqslant s \leqslant T_0}\Big(2|\bar{C}_3(s)|^2+1\Big)\int_{t_0}^{T_0}\Big(\int_{t_0}^s |\bar{F}(s,r)|dr\Big)^2 ds < 1.$$

Then, $\mathcal{T}$ is a contraction mapping on $[t_0, T_0]$ and (A.1) admits a unique solution on $[t_0, T_0]$. Since terminal time T is finite, by repeating the above steps on $[t_0, T]$, we obtain a unique solution of (A.1) on $[t_0, T]$. ∎

The proof of Proposition 2.3:

Proof. Without loss of generality, assume that $T = t_0 + n\delta$ with an integer n. Then, on $[t_0, t_0+\delta]$, the closed-loop system (2.3) becomes

$$\begin{aligned} x(t) = & \xi(0) + \int_{t_0}^t B_2(s)\varsigma(s-\delta-t_0)ds + \int_{t_0}^t \Big[A_1(s)x(s) + A_2(s)\xi(s-t_0-\delta) \\ & + A_3(s)z(s) + \Big(B_1(s) + B_2(s+\delta)\mathbf{1}_{[t_0,t-\delta)}(s) + \int_s^t B_3(r)\tilde{F}(r,s)dr\Big)u(s) + b(s)\Big]ds \\ & + \int_{t_0}^t \big[C_1(s)x(s) + C_2(s)\xi(s-t_0-\delta) + C_3(s)z(s) + D_1(s)u(s) + \sigma(s)\big]dW(s), \end{aligned} \tag{A.2}$$

and

$$u(t)=K_1(t)\xi(0)+K_1(t)\int_{t_0}^{t}B_2(s)\varsigma(s-\delta-t_0)ds+K_1(t)\int_{t_0}^{t}\Big[A_1(s)x(s)+A_2(s)\xi(s-t_0-\delta)$$
$$+A_3(s)z(s)+\Big(B_1(s)+B_2(s+\delta)\mathbf{1}_{[t_0,t-\delta)}(s)+\int_s^t B_3(r)\tilde F(r,s)dr\Big)u(s)+b(s)\Big]ds$$
$$+K_1(t)\int_{t_0}^{t}\big[C_1(s)x(s)+C_2(s)\xi(s-t_0-\delta)+C_3(s)z(s)+D_1(s)u(s)+\sigma(s)\big]dW(s)$$
$$+\int_{t_0}^{t}K_2(t,s)x(s)ds+K_3(t)\xi(t-\delta-t_0)+\int_{t_0}^{t}K_4(t,s)u(s)ds+v(t). \tag{A.3}$$

By Fubini theorem, we have

$$\int_{t_0}^{t}\int_{t_0}^{s}F(t,s)B_2(r)\nu(r)drds$$
$$=\int_{t_0}^{(t_0+\delta)\wedge t}\Big[\int_r^t F(t,s)ds\Big]B_2(r)\varsigma(r-\delta-t_0)dr+\int_{t_0\wedge(t-\delta)}^{t-\delta}\Big[\int_{r+\delta}^{t}F(t,s)ds\Big]B_2(r+\delta)u(r)dr, \tag{A.4}$$

and

$$\int_{t_0}^{t}\int_{t_0}^{s}F(t,s)B_3(r)\mu(r)drds=\int_{t_0}^{t}\int_{t_0}^{s}F(t,s)B_3(r)\Big[\int_{t_0}^{r}\tilde F(r,\alpha)u(\alpha)d\alpha\Big]drds$$
$$=\int_{t_0}^{t}\int_{t_0}^{s}\int_{\alpha}^{s}F(t,s)B_3(r)\tilde F(r,\alpha)u(\alpha)drd\alpha ds=\int_{t_0}^{t}\Big[\int_{\alpha}^{t}\int_{\alpha}^{s}F(t,s)B_3(r)\tilde F(r,\alpha)drds\Big]u(\alpha)d\alpha. \tag{A.5}$$

Hence, from (A.4)–(A.5) and by adding some indicative functions, we derive

$$z(t)=\int_{t_0}^{t}F(t,s)\Big(\xi(0)+\int_{t_0}^{s}B_2(r)\varsigma(r-\delta-t_0)dr\Big)ds+\int_{t_0}^{t}\mathcal{E}(t,s)[A_1(s)x(s)$$
$$+A_2(s)\xi(s-t_0-\delta)+A_3(s)z(s)]ds+\int_{t_0}^{t}\int_s^t F(t,r)\big(B_1(s)+B_2(s+\delta)\mathbf{1}_{[0,r-\delta)}(s)$$
$$+\int_s^r B_3(\theta)\tilde F(\theta,s)d\theta\big)dru(s)ds+\int_{t_0}^{t}\mathcal{E}(t,s)b(s)ds+\int_{t_0}^{t}\mathcal{E}(t,s)[C_1(s)x(s)$$
$$+C_2(s)\xi(s-t_0-\delta)+C_3(s)z(s)+D_1(s)u(s)+\sigma(s)]dW(s), \tag{A.6}$$

where $\mathcal{E}(t,s)\equiv\mathbf{1}_{[0,t)}(s)\int_s^t F(t,r)dr$. Let $\mathbf{X}(\cdot)\equiv\begin{bmatrix}x(\cdot)\\ z(\cdot)\\ u(\cdot)\end{bmatrix}$. Then, by (A.2), (A.3) and (A.6), we have

$$\mathbf{X}(t)=\boldsymbol{\varphi}(t)+\int_{t_0}^{t}\mathbf{A}(t,s)\mathbf{X}(s)ds+\int_{t_0}^{t}\mathbf{C}(t,s)\mathbf{X}(s)dW(s),\quad t\in(t_0,t_0+\delta), \tag{A.7}$$

where

$$\mathbf{A}(t,s)\equiv\begin{bmatrix}A_1(s) & A_3(s) & B_1(s)+B_2(s+\delta)\mathbf{1}_{[t_0,t-\delta)}(s)+\int_s^t B_3(r)\tilde F(r,s)dr\\ \mathcal{E}(t,s)A_1(s) & \mathcal{E}(t,s)A_3(s) & \int_s^t F(t,r)(B_1(s)+B_2(s+\delta)\mathbf{1}_{[0,r-\delta)}(s)+\int_s^r B_3(\theta)\tilde F(\theta,s)d\theta)dr\\ K_1(t)A_1(s)+K_2(t,s) & K_1(t)A_3(s) & K_1(t)(B_1(s)+B_2(s+\delta)\mathbf{1}_{[t_0,t-\delta)}(s)+\int_s^t B_3(r)\tilde F(r,s)dr)+K_4(t,s)\end{bmatrix},$$

$$\mathbf{C}(t,s)\equiv\begin{bmatrix}C_1(s) & C_3(s) & D_1(s)\\ \mathcal{E}(t,s)C_1(s) & \mathcal{E}(t,s)C_3(s) & \mathcal{E}(t,s)D_1(s)\\ K_1(t)C_1(s) & K_1(t)C_3(s) & K_1(t)D_1(s)\end{bmatrix},$$

and

$$\varphi(t) \equiv \begin{bmatrix} \xi(0)+\int_{t_0}^{t}\big(A_2(s)\xi(s-\delta-t_0)+B_2(s)\varsigma(s-\delta-t_0)+b(s)\big)ds \\ +\int_{t_0}^{t}\big(C_2(s)\xi(s-t_0-\delta)+\sigma(s)\big)dW(s) \\ \\ \int_{t_0}^{t}F(t,s)\Big(\xi(0)+\int_{t_0}^{s}B_2(r)\varsigma(r-\delta-t_0)dr\Big)ds \\ +\int_{t_0}^{t}\mathcal{E}(t,s)(A_2(s)\xi(s-\delta-t_0)+b(s))ds+\int_{t_0}^{t}\mathcal{E}(t,s)(C_2(s)\xi(s-\delta-t_0)+\sigma(s))dW(s) \\ \\ K_1(t)\xi(0)+K_1(t)\int_{t_0}^{t}(A_2(s)\xi(s-\delta-t_0)+B_2(s)\varsigma(s-\delta-t_0)+b(s))ds \\ +K_1(t)\int_{t_0}^{t}(C_2(s)\xi(s-t_0-\delta)+\sigma(s))dW(s)+K_3(t)\xi(t-\delta-t_0)+v(t) \end{bmatrix}.$$

Notice that $K_1(\cdot)\in L^2(t_0,T;\mathbb{R}^{m\times n})$ and $\sigma(\cdot)\in L^2_{\mathbb{F}}(0,T;\mathbb{R}^n)$. Then, we obtain

$$\begin{aligned}\mathbb{E}\int_{t_0}^{T}|K_1(t)\int_{t_0}^{t}\sigma(s)dW(s)|^2dt &\leqslant \int_{t_0}^{T}|K_1(t)|^2\mathbb{E}\Big|\int_{t_0}^{t}\sigma(s)dW(s)\Big|^2dt \\ &\leqslant \int_{t_0}^{T}|K_1(t)|^2dt\mathbb{E}\int_{t_0}^{T}|\sigma(s)|^2ds<\infty.\end{aligned}$$

In addition, from (A1), $\xi(\cdot)\in C([-\delta,0];\mathbb{R}^n)$ and $\varsigma(\cdot)\in L^2(-\delta,0;\mathbb{R}^m)$, we have $\varphi(\cdot)\in L^2_{\mathbb{F}}(t_0,T;\mathbb{R}^{2n+m})$. Similarly, by (A1), the Holder inequality and $(K_1(\cdot),K_2(\cdot,\cdot),K_3(\cdot),K_4(\cdot,\cdot),v(\cdot))\in\mathbb{L}$, we obtain $\mathbf{A}(\cdot,\cdot)\in L^2(\Delta_2(t_0,T);\mathbb{R}^{(2n+m)\times(2n+m)})$, $\mathbf{C}(\cdot,\cdot)\in\mathscr{L}^2(\triangle_2(t_0,T);\ \mathbb{R}^{(2n+m)\times(2n+m)})$. Thus by Proposition 2.2, (A.7) admits a unique solution $\mathbf{X}(\cdot)\in L^2_{\mathbb{F}}(t_0,t_0+\delta;\mathbb{R}^{2n+m})$, which implies that $x(\cdot)\in L^2_{\mathbb{F}}(t_0,t_0+\delta;\mathbb{R}^n)$ and $u(\cdot)\in L^2_{\mathbb{F}}(t_0,t_0+\delta;\mathbb{R}^m)$. Furthermore, by the definition of the solution of the first equation in (2.3), $x(\cdot)\in L^2_{\mathbb{F}}(\Omega;C([t_0,t_0+\delta];\mathbb{R}^n))$, and this implies the existence of the solution to (2.3). As for the uniqueness, since (A.7) and (2.3) are equivalent, it can be obtained from the uniqueness of the solution to (A.7). Hence, the closed-loop system (2.3) admits a unique solution on $[t_0,t_0+\delta]$. Then the same steps are repeated on $[t_0+\delta,t_0+2\delta]$, $[t_0+2\delta,t_0+3\delta]$ and so on. The terminal time T is finite, thus (2.3) admits a unique solution on $[t_0,T]$. ∎

The proof of Proposition 3.1:

Proof. To show the equivalence, it is sufficient to prove that under (A1), the $X(\cdot)$ defined in (3.7) is the unique solution of SVIE (3.8). In terms of Proposition 2.1 and Proposition 2.2, we only need to prove that

$$\begin{cases} A(\cdot,\cdot)\in L^2(\Delta_2(0,T);\mathbb{R}^{(3n)\times(3n)}),\ B(\cdot,\cdot)\in L^2(\Delta_2(0,T);\mathbb{R}^{(3n)\times m}),\ C(\cdot,\cdot)\in\mathscr{L}^2\left(\triangle_2(0,T);\mathbb{R}^{(3n)\times(3n)}\right), \\ D(\cdot,\cdot)\in\mathscr{L}^2\left(\triangle_2(0,T);\mathbb{R}^{(3n)\times m}\right),\quad Q(\cdot)\in L^\infty\left(0,T;\mathbb{S}^{3n}\right),\quad R(\cdot)\in L^\infty\left(0,T;\mathbb{S}^m\right), \\ \tilde{b}(\cdot,\cdot)\in L^{2,1}_{\mathbb{F}}\left(\triangle_2(0,T);\mathbb{R}^{3n}\right),\qquad \tilde{\sigma}(\cdot,\cdot)\in L^2_{\mathbb{F}}\left(\triangle_2(0,T);\mathbb{R}^{3n}\right). \end{cases}$$

Notice that

$$\int_0^T\int_0^t|\mathcal{E}(t,s)A_1(s)|^2dsdt=\int_0^T\int_0^t|\int_s^tF(t,r)drA_1(s)|^2dsdt\leqslant\int_0^T|A_1(s)|^2ds\int_0^T\Big(\int_0^t|F(t,r)|dr\Big)^2dt.$$

Then, by $A_1(\cdot)\in L^2(0,T;\mathbb{R}^{n\times n})$ and $F(\cdot,\cdot)\in L^\infty(\Delta_2(0,T);\mathbb{R}^{n\times n})$, we have $\mathcal{E}(\cdot,\cdot)A_1(\cdot)\in L^2(\Delta_2(0,T);\mathbb{R}^{n\times n})$. In terms of Assumption (A1), $A(\cdot,\cdot)\in L^2(\Delta_2(0,T);\mathbb{R}^{(3n)\times(3n)})$. Similarly, one has $C(\cdot,\cdot)\in L^2(\Delta_2(0,T);\mathbb{R}^{(3n)\times(3n)})$. By the fact of $C_1(\cdot),C_2(\cdot),C_3(\cdot)\in L^\infty(0,T;\mathbb{R}^{n\times n})$, we deduce

$$\operatorname*{ess\,sup}_{s\in(0,T)}\left(\int_s^T|\mathcal{E}(t,s)C_1(s)|^2dt\right)^{\frac{1}{2}}\leqslant M\Big(\int_0^T(\int_0^t|F(t,r)|dr)^2dt\Big)^{\frac{1}{2}}<\infty,$$

where M is a generic constant. For any $\varepsilon > 0$, by $F(\cdot,\cdot) \in L^\infty(\Delta_2(0,T);\mathbb{R}^{n\times n})$, there exists a finite partition $\{a_i\}_{i=0}^m$ of $(0,T)$ with $0 = a_0 < a_1 < \cdots < a_m = T$ such that

$$\operatorname*{ess\,sup}_{t\in(a_i,a_{i+1})} \Big(\int_t^{a_{i+1}} |\mathcal{E}(s,t)C_1(t)|^2 ds\Big)^{\frac{1}{2}} \leqslant M\Big(\int_{a_i}^{a_{i+1}} \Big(\int_{a_i}^{s} |F(s,r)|dr\Big)^2 ds\Big)^{\frac{1}{2}} < \varepsilon,$$

which implies that $C(\cdot,\cdot) \in \mathscr{L}^2(\triangle_2(0,T);\mathbb{R}^{(3n)\times(3n)})$. Using the Holder inequality and Assumption (A1), we can prove the integrability of $B(\cdot,\cdot)$, $D(\cdot,\cdot)$, $Q(\cdot)$, $R(\cdot)$, $\tilde{b}(\cdot,\cdot)$ and $\tilde{\sigma}(\cdot,\cdot)$, thus the proof is completed. ∎

The proof of Theorem 3.3:

Proof. Inspired by [29], let us introduce the following Riccati system:

$$\begin{cases} P^{(1)}(t) = Q(t) + (C^\top \ltimes P \rtimes C)(t) \\ \qquad\qquad -(C^\top \ltimes P \rtimes D)(t)(R(t) + (D^\top \ltimes P \rtimes D)(t))^{-1}(D^\top \ltimes P \rtimes C)(t), \quad 0 < t < T, \\ P^{(2)}(s,t,t) = P^{(2)}(t,s,t)^\top \\ = (P\rtimes A)(s,t) - (P\rtimes B)(s,t)(R(t) + (D^\top \ltimes P \rtimes D)(t))^{-1}(D^\top \ltimes P \rtimes C)(t), \quad 0 < t < s < T, \\ \dot{P}^{(2)}(s_1,s_2,t) = (P\rtimes B)(s_1,t)(R(t) + (D^\top \ltimes P \rtimes D)(t))^{-1}(B^\top \ltimes P)(s_2,t), \quad 0 < t < (s_1\wedge s_2) < T, \end{cases} \tag{A.8}$$

where for each $M_1 : \triangle_2(0,T) \to \mathbb{R}^{d_1\times(3n)}$, $M_2 : \triangle_2(0,T) \to \mathbb{R}^{(3n)\times d_2}$ with $d_1, d_2 \in \mathbb{N}$,

$$(M_1 \ltimes P)(s,t) \equiv M_1(s,t)P^{(1)}(s) + \int_t^T M_1(r,t)P^{(2)}(r,s,t)\mathrm{d}r, \quad 0 < t < s < T, \tag{A.9}$$

$$(P \rtimes M_2)(s,t) \equiv P^{(1)}(s)M_2(s,t) + \int_t^T P^{(2)}(s,r,t)M_2(r,t)\mathrm{d}r, \quad 0 < t < s < T, \tag{A.10}$$

$$(M_1\ltimes P\rtimes M_2)(t) \equiv \int_t^T M_1(s,t)P^{(1)}(s)M_2(s,t)ds + \int_t^T\int_t^T M_1(s_1,t)\, P^{(2)}(s_1,s_2,t)M_2(s_2,t)\, ds_1 ds_2, 0<t<T. \tag{A.11}$$

By Corollary 6.7 in [29], the equation (A.8) admits a unique solution $(P^{(1)}, P^{(2)}) \in \Pi(0,T)$ such that $R+(D^\top \ltimes P \rtimes D) > \beta I$ for some constant $\beta > 0$. Next we will equivalently transform (A.8) into (3.11). To this end, we first show that

$$(C^\top \ltimes P \rtimes C)(t) = (C_1(t), C_2(t), C_3(t))^\top \mathcal{G}_1(t)(C_1(t), C_2(t), C_3(t)), \tag{A.12}$$
$$(C^\top \ltimes P \rtimes D)(t) = (C_1(t), C_2(t), C_3(t))^\top \mathcal{G}_1(t)D_1(t), \tag{A.13}$$
$$(D^\top \ltimes P \rtimes D)(t) = D_1(t)^\top \mathcal{G}_1(t)D_1(t), \tag{A.14}$$
$$(D^\top \ltimes P \rtimes C)(t) = D_1(t)^\top \mathcal{G}_1(t)(C_1(t), C_2(t), C_3(t)). \tag{A.15}$$

In fact, notice that

$$C(s,t) = \Upsilon(s,t)(C_1(t), C_2(t), C_3(t)),$$

and

$$(C^\top \ltimes P \rtimes C)(t) = \int_t^T C(s,t)^\top P^{(1)}(s)C(s,t)ds + \int_t^T \int_t^T C(s_1,t)^\top P^{(2)}(s_1,s_2,t)C(s_2,t)ds_1 ds_2.$$

Then, (A.12) holds. Similarly, by $D(s,t) = \Upsilon(s,t)D_1(t)$, we obtain (A.13)–(A.15).

Next it is time to treat the terms $(P\rtimes A)$ and $(P\rtimes B)$. From (A.10) and $A(s,t) = \Upsilon(s,t)(A_1(t), A_2(t), A_3(t))$, we have

$$(P \rtimes A)(s,t) = \Big[P^{(1)}(s)\Upsilon(s,t) + \int_t^T P^{(2)}(s,r,t)\Upsilon(r,t)dr\Big](A_1(t), A_2(t), A_3(t)). \tag{A.16}$$

On the other hand, we observe that

$$B(t,s) = \int_s^t \Pi(t,s,\theta)(B_1(s)^\top, B_2(s+\delta)^\top, (B_3(\theta)\tilde{F}(\theta,s))^\top)^\top d\theta.$$

Then it is easy to see,

$$(P\rtimes B)(s,t)=\int_t^s P^{(1)}(s)\Pi(s,t,\theta)(B_1(t)^\top,B_2(t+\delta)^\top,(B_3(\theta)\tilde F(\theta,t))^\top)^\top d\theta$$
$$+\int_t^T\int_t^r P^{(2)}(s,r,t)\Pi(r,t,\theta)(B_1(t)^\top,B_2(t+\delta)^\top,(B_3(\theta)\tilde F(\theta,t))^\top)^\top d\theta dr. \tag{A.17}$$

Hence after some direct calculations, together with (A.12)–(A.17), we see that (A.8) can be written as (3.11), which completes the proof of Theorem 3.3. ∎

The proof of Theorem 3.5:

Proof. Introduce the following Type-II extended backward SVIE:

$$\begin{cases} d\eta(t,s)=-\Big\{(P\rtimes\tilde b)(t,s)+\Gamma^*(t,s)^\top(D^\top\ltimes P\rtimes\tilde\sigma)(s)+\Gamma^*(t,s)^\top\int_s^T B(r,s)^\top\eta(r,s)dr\\ \qquad+\check\Gamma(t,s)^\top\int_s^T D(r,s)^\top\zeta(r,s)dr\Big\}ds+\zeta(t,s)dW(s),\quad 0<s<t<T,\\ \eta(t,t)=(C^\top\ltimes P\rtimes\tilde\sigma)(t)+\Xi^*(t)^\top(D^\top\ltimes P\rtimes\tilde\sigma)(t)+\int_t^T(A(r,t)+B(r,t)\Xi^*(t))^\top\eta(r,t)dr\\ \qquad+\int_t^T(C(r,t)+D(r,t)\Xi^*(t))^\top\zeta(r,t)dr,\quad 0<t<T,\end{cases} \tag{A.18}$$

where

$$\Xi^*(t)\equiv-(R(t)+(D^\top\ltimes P\rtimes D)(t))^{-1}(D^\top\ltimes P\rtimes C)(t),\ 0<t<T, \tag{A.19}$$
$$\Gamma^*(s,t)\equiv-(R(t)+(D^\top\ltimes P\rtimes D)(t))^{-1}(B^\top\ltimes P)(s,t),\ 0<t<s<T. \tag{A.20}$$

Under Assumptions (A1)–(A2), it follows from Theorem 3.2 in [29] that Equation (A.18) admits a unique solution $(\eta,\zeta)\in L^2_{\mathbb{F},c}(\triangle_2(0,T);\mathbb{R}^{3n})\times L^2_{\mathbb{F}}(\triangle_2(0,T);\mathbb{R}^{3n})$. After a careful observation, we have that

$$(P\rtimes\tilde b)(t,s)=\Big[P^{(1)}(t)\Upsilon(t,s)+\int_s^T P^{(2)}(t,r,s)\Upsilon(r,s)dr\Big]b(s),$$
$$(C^\top\ltimes P\rtimes\tilde\sigma)(t)=(C_1(t),C_2(t),C_3(t))^\top\mathcal{G}_1(t)\sigma(t),$$

and

$$(D^\top\ltimes P\rtimes\tilde\sigma)(s)=D_1(s)^\top\mathcal{G}_1(s)\sigma(s),\quad C(t,s)=\Upsilon(t,s)(C_1(s),C_2(s),C_3(s)),$$
$$D(t,s)=\Upsilon(t,s)D_1(s),\quad B(t,s)=\int_s^t\Pi(t,s,\theta)(B_1(s)^\top,B_2(s+\delta)^\top,(B_3(\theta)\tilde F(\theta,s))^\top)^\top d\theta.$$

Then, we see that (A.18) can be rewritten as (3.13). It is then easy to see the conclusion of Theorem 3.5. ∎

Before proving Theorem 3.6, according to [29], we need some auxiliary results about the new control problem with the state equation (3.8) and the cost functional (3.9).

To begin with, for any given $t_0\in[0,T)$, we consider the following system of the new control problem:

$$\begin{cases} X^{t_0,\xi,\varsigma}(t)=\varphi(t)+\int_{t_0}^t\big[A(t,s)X^{t_0,\xi,\varsigma}(s)+B(t,s)u^{t_0,\xi,\varsigma}(s)+\tilde b(t,s)\big]ds\\ \qquad+\int_{t_0}^t\big[C(t,s)X^{t_0,\xi,\varsigma}(s)+D(t,s)u^{t_0,\xi,\varsigma}(s)+\tilde\sigma(t,s)\big]dW(s),\quad t_0<t<T,\\ \Theta^{t_0,\xi,\varsigma}(s,t)=\varphi(s)+\int_{t_0}^t\big[A(s,r)X^{t_0,\xi,\varsigma}(r)+B(s,r)u^{t_0,\xi,\varsigma}(r)+\tilde b(s,r)\big]dr\\ \qquad+\int_{t_0}^t\big[C(s,r)X^{t_0,\xi,\varsigma}(r)+D(s,r)u^{t_0,\xi,\varsigma}(r)+\tilde\sigma(s,r)\big]dW(r),\quad t_0<t<s<T,\\ u^{t_0,\xi,\varsigma}(t)=\Xi(t)X^{t_0,\xi,\varsigma}(t)+\int_t^T\Gamma(s,t)\Theta^{t_0,\xi,\varsigma}(s,t)ds+\omega(t),\quad t_0<t<T.\end{cases} \tag{A.21}$$

In terms of [29], we call any triplet $(\Xi, \Gamma, \omega) \in \mathcal{S}(t_0, T) \equiv L^\infty(t_0, T; \mathbb{R}^{m\times n}) \times L^2(\triangle_2(t_0, T); \mathbb{R}^{m\times n}) \times L^2_{\mathbb{F}}(t_0, T; \mathbb{R}^m)$ the *causal feedback strategy*. For any $(\Xi, \Gamma, \omega) \in \mathcal{S}(t_0, T)$, $\xi \in C([-\delta, 0]; \mathbb{R}^n)$ and $\varsigma \in L^2(-\delta, 0; \mathbb{R}^m)$, let the triplet $(X^{t_0,\xi,\varsigma}, \Theta^{t_0,\xi,\varsigma}, u^{t_0,\xi,\varsigma})$ be the solution to the system (A.21) and write $u^{t_0,\xi,\varsigma} = (\Xi, \Gamma, w)[t_0, \xi, \varsigma]$. A causal feedback strategy $(\Xi^*, \Gamma^*, \omega^*) \in \mathcal{S}(t_0, T)$ is called a *causal feedback optimal strategy* of the new control problem if

$$J\big(t_0, \xi, \varsigma; (\Xi^*, \Gamma^*, \omega^*)[t_0, \xi, \varsigma]\big) \leqslant J\big(t_0, \xi, \varsigma; u\big),$$

for any $(\xi, \varsigma) \in C([-\delta, 0]; \mathbb{R}^n) \times L^2(-\delta, 0; \mathbb{R}^m)$ and any $u(\cdot) \in L^2_{\mathbb{F}}(t_0, T; \mathbb{R}^m)$.

The following result gives the closed-loop solvability for the new control problem.

Lemma A.1. *Let Assumptions* (A1)–(A2) *hold. Then, for any given $t_0 \in [0, T)$, the causal feedback optimal strategy $(\Xi^*, \Gamma^*, \omega^*)$ of the new control problem on $[t_0, T]$ with* (3.8) *and* (3.9)*, is given by* (3.14) *and for $t_0 < t < T$,*

$$\omega^*(t) = -\mathcal{R}(t)^{-1}\Big(D_1(t)^\top \mathcal{G}_1(t)\sigma(t) + \int_t^T \Big[\int_t^s \mathcal{B}(\theta, t)^\top \Pi(s, t, \theta)^\top \eta(s, t)d\theta + D_1(t)^\top \Upsilon(s, t)^\top \zeta(s, t)\Big] ds\Big), \quad \text{(A.22)}$$

where $\Pi(\cdot, \cdot, \cdot)$, $\Upsilon(\cdot, \cdot)$, $\mathcal{R}(\cdot)$, $\mathcal{G}_1(\cdot)$ and $\mathcal{G}_3(\cdot, \cdot, \cdot)$ are defined by (3.10),(3.12)*, respectively.*

Proof. First, given (η, ζ) satisfying (3.13), we define

$$\begin{aligned} &k(t) \equiv \big(D^\top \ltimes P \rtimes \tilde{\sigma}\big)(t) + \int_t^T B(s, t)^\top \eta(s, t)\, ds + \int_t^T D(s, t)^\top \zeta(s, t) ds, \\ &\omega^*(t) \equiv -\big(R(t) + (D^\top \ltimes P \rtimes D)(t)\big)^{-1} k(t), \ 0 < t < T. \end{aligned} \quad \text{(A.23)}$$

Under (A1)–(A2), by Theorem 5.4 in [29], and Theorem 3.3 and Theorem 3.5 in the current paper, we see that $(\Xi^*, \Gamma^*, \omega^*)$ is the optimal causal feedback strategy where (Ξ^*, Γ^*) is in (3.14). Based on this fact, it is sufficient to show that $\omega^*(\cdot)$ defined in (A.23) can be rewritten in the form of (A.22).

In fact, by (3.7) and (3.10), we obtain

$$B(s, t) = \int_t^s \Pi(s, t, \theta)\big(B_1(t)^\top, B_2(t+\delta)^\top, \tilde{F}(\theta, t)^\top B_3(\theta)^\top\big)^\top d\theta \mathbf{1}_{(0,s)}(t).$$

Recall that

$$\big(B^\top \ltimes P\big)(s, t) \equiv B(s, t)^\top P^{(1)}(s) + \int_t^T B(r, t)^\top P^{(2)}(r, s, t)\mathrm{d}r, \quad 0 < t < s < T.$$

Then, we deduce

$$\begin{aligned} (B^\top \ltimes P)(s, t) = &\int_t^s \big(B_1(t)^\top, B_2(t+\delta)^\top, \tilde{F}(\theta, t)^\top B_3(\theta)^\top\big)\Pi(s, t, \theta)^\top d\theta P^{(1)}(s) \\ &+ \int_t^T \int_t^r \big(B_1(t)^\top, B_2(t+\delta)^\top, \tilde{F}(\theta, t)^\top B_3(\theta)^\top\big)\Pi(r, t, \theta)^\top P^{(2)}(r, s, t)d\theta dr. \end{aligned} \quad \text{(A.24)}$$

Notice that $D(t, s) = \Upsilon(t, s)D_1(s)$ and

$$(D^\top \ltimes P \rtimes \tilde{\sigma})(t) = D_1(t)^\top \mathcal{G}_1(t)\sigma(t), \quad t \in [t_0, T].$$

Combining it with (A.23), we then have that

$$\begin{aligned} k(t) = D_1(t)^\top \mathcal{G}_1(t)\sigma(t) + &\int_t^T \Big[\int_t^s \big(B_1(t)^\top, B_2(t+\delta)^\top, \tilde{F}(\theta, t)^\top B_3(\theta)^\top\big)\Pi(s, t, \theta)^\top \eta(s, t)d\theta \\ &+ D_1(t)^\top \Upsilon(s, t)^\top \zeta(s, t)\Big] ds. \end{aligned} \quad \text{(A.25)}$$

This naturally implies the desired conclusion of Lemma A.1. ∎

The proof of Theorem 3.6:

Proof. The basic idea of the following arguments is to explicitly construct the desired five-tuple closed-loop strategy by the causal feedback strategy $(\Xi^*, \Gamma^*, \omega^*)$ in Lemma A.1. To this end, we divide the proof into six steps.

Step 1: Given $\Xi^*(\cdot), \Gamma^*(\cdot,\cdot)$ in (3.14), we decompose them as follows for later convenience:

$$\Xi^*(t) = \big[\Xi_1^*(t), \Xi_2^*(t), \Xi_3^*(t)\big], \Gamma^*(s,t) = \big[\Gamma_1^*(s,t), \Gamma_2^*(s,t), \Gamma_3^*(s,t)\big].$$

In this step, we prove that the following process is an optimal closed-loop outcome control of Problem (P) on $[t_0, T]$:

$$u^*(t) = K_1^*(t)x^*(t) + \int_{t_0}^{t} K_2^*(t,s)x^*(s)ds + K_3^*(t)x^*(t-\delta) + \int_{t_0}^{t} K_4^*(t,s)u^*(s)ds + v^*(t), \tag{A.26}$$

where

$$K_1^*(t) = \Xi_1^*(t) + \int_t^T \Big[\Gamma_1^*(s,t) + \Gamma_2^*(s,t)\mathbf{1}_{\big([t+\delta]\wedge T,\infty\big)}(s) + \int_t^s \Gamma_3^*(s,t)F(s,\theta)d\theta\Big]ds, \tag{A.27}$$

$$K_2^*(t,s) = \Xi_3^*(t)F(t,s) + \Gamma_2^*(s+\delta,t)\mathbf{1}_{\big[t-\delta,T-\delta\big]}(s) + \int_t^T \Gamma_3^*(\theta,t)F(\theta,s)d\theta, \tag{A.28}$$

$$K_3^*(t) = \Xi_2^*(t), \tag{A.29}$$

$$\begin{aligned} K_4^*(t,s) = &\Big\{\int_{s+\delta}^{T}\Big[\Gamma_1^*(r,t)+\Gamma_2^*(r,t)\mathbf{1}_{(s+2\delta,\infty)}(r)\mathbf{1}_{[0,T-\delta)}(t)+\int_{s+\delta}^{r}\Gamma_3^*(r,t)F(r,\theta)d\theta\Big]drB_2(s+\delta)\Big\} \\ &\times\mathbf{1}_{[t-\delta,T-\delta]}(s) + \int_t^T\int_{\theta'}^T\Big(\Gamma_1^*(r,t) + \Gamma_2^*(r,t)\mathbf{1}_{[t_0,T-\delta)}(\theta')\mathbf{1}_{[t_0,T-\delta)}(t)\mathbf{1}_{(\theta'+\delta,\infty)}(r) \\ &+ \int_{\theta'}^{r}\Gamma_3^*(r,t)F(r,\theta)d\theta\Big)B_3(\theta')\tilde{F}(\theta',s)drd\theta', \end{aligned} \tag{A.30}$$

$$\begin{aligned} v^*(t) = \omega^*(t)+&\int_t^{T\wedge(t_0+\delta)}\Gamma_2^*(s,t)\xi(s-t_0-\delta)ds + \int_{t-\delta}^{t_0}\Big\{\int_{s+\delta}^{T}\Big[\Gamma_1^*(r,t) + \Gamma_2^*(r,t)\mathbf{1}_{(s+2\delta,\infty)}(r) \\ &\times\mathbf{1}_{[0,T-\delta)}(t) + \int_{s+\delta}^{r}\Gamma_3^*(r,t)F(r,\theta)d\theta\Big]drB_2(s+\delta)\Big\}\varsigma(s-t_0)ds\mathbf{1}_{[t_0,t_0+\delta]}(t), \end{aligned} \tag{A.31}$$

and $\omega^*(\cdot)$ is defined by (A.22).

According to Proposition 3.1, given $(\Xi^*, \Gamma^*, \omega^*)$ in Lemma A.1, the following process is an optimal control of Problem (P):

$$u^*(t) = \Xi^*(t)X^*(t) + \int_t^T \Gamma^*(s,t)\Theta^*(s,t)ds + \omega^*(t), \;\; t_0 < t < T, \tag{A.32}$$

where $X^*(\cdot) \equiv \big[x^*(\cdot)^\top, y^*(\cdot)^\top, z^*(\cdot)^\top\big]^\top$, and

$$\begin{aligned} \Theta^*(s,t) = \varphi(s) &+ \int_{t_0}^{t}\Big[A(s,r)X^*(r) + B(s,r)u^*(r) + \tilde{b}(s,r)\Big]dr \\ &+ \int_{t_0}^{t}\Big[C(s,r)X^*(r) + D(s,r)u^*(r) + \tilde{\sigma}(s,r)\Big]dW(r), \quad t_0 < t < s < T. \end{aligned}$$

Our next idea is to rewrite (A.32) into (A.26). For later convenience, let $\Theta^*(\cdot,\cdot) = \big[\Theta_1^*(\cdot,\cdot)^\top, \Theta_2^*(\cdot,\cdot)^\top, \Theta_3^*(\cdot,\cdot)^\top\big]^\top$. It then yields

$$\begin{aligned} u^*(t) = &\,\Xi_1^*(t)x^*(t) + \Xi_2^*(t)x^*(t-\delta) + \Xi_3^*(t)\int_{t_0}^{t} F(t,s)x^*(s)ds \\ &+ \int_t^T \Big[\Gamma_1^*(s,t)\Theta_1^*(s,t) + \Gamma_2^*(s,t)\Theta_2^*(s,t) + \Gamma_3^*(s,t)\Theta_3^*(s,t)\Big]ds + w^*(t), \end{aligned} \tag{A.33}$$

where

$$\Theta_1^*(s,t)=\xi(0)+\int_{t_0}^{(t_0+\delta)\wedge s}B_2(r)\varsigma(r-t_0-\delta)dr+\int_{t_0}^{t}\Big[A_1(r)x^*(r)+A_2(r)y^*(r)+A_3(r)z^*(r)$$
$$+B_1(r)u^*(r)+B_2(r+\delta)\mathbf{1}_{[t_0,s-\delta)}(r)u^*(r)+\int_r^s B_3(\theta)\tilde F(\theta,r)d\theta u^*(r)+b(r)\Big]dr$$
$$+\int_{t_0}^{t}\Big[C_1(r)x^*(r)+C_2(r)y^*(r)+C_3(r)z^*(r)+D_1(r)u^*(r)+\sigma(r)\Big]dW(r),\ t_0<t<s<T, \tag{A.34}$$

and

$$\Theta_2^*(s,t)=\xi(s-\delta-t_0)\mathbf{1}_{[t_0,t_0+\delta]}(s)+\mathbf{1}_{(t_0+\delta,\infty)}(s)\Big\{\xi(0)+\int_{t_0}^{(t_0+\delta)\wedge(s-\delta)}B_2(r)\varsigma(r-t_0-\delta)dr\Big\}$$
$$+\int_{t_0}^{t\wedge(s-\delta)}\Big[A_1(r)x^*(r)+A_2(r)y^*(r)+A_3(r)z^*(r)+B_1(r)u^*(r)+B_2(r+\delta)u^*(r)\mathbf{1}_{[t_0,s-2\delta)}(r)$$
$$+\int_r^{s-\delta}B_3(\theta)\tilde F(\theta,r)d\theta u^*(r)+b(r)\Big]dr\mathbf{1}_{(t_0+\delta,\infty)}(s)+\int_{t_0}^{t\wedge(s-\delta)}\Big[C_1(r)x^*(r)+C_2(r)y^*(r)$$
$$+C_3(r)z^*(r)+D_1(r)u^*(r)+\sigma(r)\Big]dW(r)\mathbf{1}_{(t_0+\delta,\infty)}(s),\ t_0<t<s<T, \tag{A.35}$$

and

$$\Theta_3^*(s,t)=\int_{t_0}^{s}F(s,r)\Big(\xi(0)+\int_{t_0}^{t_0+\delta}B_2(\theta)\varsigma(\theta-t_0-\delta)\mathbf{1}_{[t_0,r)}(\theta)d\theta\Big)dr+\int_{t_0}^{t}\Big[\mathcal{E}(s,r)\big(A_1(r)x^*(r)+A_2(r)y^*(r)$$
$$+A_3(r)z^*(r)+B_1(r)u^*(r)\big)+\mathcal{E}(s,r+\delta)B_2(r+\delta)u^*(r)+\int_r^s\mathcal{E}(s,\theta)B_3(\theta)\tilde F(\theta,r)d\theta u^*(r)$$
$$+\mathcal{E}(s,r)b(r)\Big]dr+\int_{t_0}^{t}\mathcal{E}(s,r)\Big[C_1(r)x^*(r)+C_2(r)y^*(r)+C_3(r)z^*(r)$$
$$+D_1(r)u^*(r)+\sigma(r)\Big]dW(r),\ t_0<t<s<T.$$

As to Θ_1^*, Θ_2^*, Θ_3^*, by the Fubini theorem, we obtain

$$\Theta_1^*(s,t)=x^*(t)+\int_t^{s\wedge(t+\delta)}B_2(r)u^*(r-\delta)dr+\int_t^s B_3(r)\int_{t_0}^t\tilde F(r,\theta)u^*(\theta)d\theta dr, \tag{A.36}$$

and

$$\Theta_2^*(s,t)=\xi(s-\delta-t_0)\mathbf{1}_{[t_0,t_0+\delta]}(s)+\mathbf{1}_{(\delta+t_0,\infty)}(s)\Big\{x^*(t\wedge(s-\delta))$$
$$+\int_{t\wedge(s-\delta)}^{(t+\delta)\wedge(s-\delta)}B_2(r)u^*(r-\delta)dr+\int_{t\wedge(s-\delta)}^{s-\delta}B_3(r)\int_{t_0}^t\tilde F(r,\theta)u^*(\theta)d\theta dr\Big\}, \tag{A.37}$$

and

$$\Theta_3^*(s,t)=\int_{t_0}^s F(s,r)x^*(t\wedge r)dr$$
$$+\int_{t_0}^s F(s,r)\Big[\int_{t\wedge r}^{(t+\delta)\wedge r}B_2(\theta)u^*(\theta-\delta)d\theta+\int_{t_0}^{t\wedge r}\int_{t\wedge r}^{r}B_3(\theta)\tilde F(\theta,\theta')u^*(\theta')d\theta d\theta'\Big]dr. \tag{A.38}$$

By (A.33), (A.36)–(A.38), for $t\in[t_0,T]$, we derive

$$u^*(t)=\Big[\Xi_1^*(t)+\int_t^T\Gamma_1^*(s,t)ds+\int_{(t+\delta)\wedge T}^T\Gamma_2^*(s,t)ds+\int_t^T\int_r^T\Gamma_3^*(s,t)F(s,r)dsdr\Big]x^*(t)$$

$$
\begin{aligned}
&+\Xi_2^*(t)x^*(t-\delta)+\int_t^T \Gamma_2^*(s,t)\xi(s-t_0-\delta)\mathbf{1}_{[t_0,t_0+\delta]}(s)ds\\
&+\int_{t_0}^t \Big[\Xi_3^*(t)F(t,s)+\Gamma_2^*(s+\delta,t)\mathbf{1}_{[t-\delta,T-\delta]}(s)+\int_t^T \Gamma_3^*(r,t)F(r,s)dr\Big]x^*(s)ds\\
&+\int_t^{(t+\delta)\wedge T}\Big[\int_r^T \Gamma_1^*(s,t)B_2(r)ds+\int_{r+\delta}^T \Gamma_2^*(s,t)B_2(r)ds\mathbf{1}_{[t_0,T-\delta)}(t)\\
&+\int_r^T\int_r^{s\wedge(t+\delta)}\Gamma_3^*(s,t)F(s,\theta)B_2(r)d\theta ds+\int_t^T\int_{s\wedge(t+\delta)}^s \Gamma_3^*(s,t)F(s,\theta)B_2(r)d\theta ds\mathbf{1}_{[0,T-\delta)}(t)\Big]u^*(r-\delta)dr\\
&+\int_{t_0}^t\Big[\int_t^T\int_\theta^T \Gamma_1^*(s,t)B_3(\theta)\tilde F(\theta,r)dsd\theta+\int_{t+\delta}^T\int_t^{s-\delta}\Gamma_2^*(s,t)B_3(\theta)\tilde F(\theta,r)d\theta ds\mathbf{1}_{[t_0,T-\delta)}(t)\\
&+\int_t^T\int_t^s\int_t^{\theta'}\Gamma_3^*(s,t)F(s,\theta')B_3(\theta)\tilde F(\theta,r)d\theta d\theta' ds\Big]u^*(r)dr+w^*(t).
\end{aligned}
$$

Notice that

$$
\begin{aligned}
&\int_t^{(t+\delta)\wedge T}\Big[\int_r^T \Gamma_1^*(s,t)B_2(r)ds+\int_{r+\delta}^T \Gamma_2^*(s,t)B_2(r)ds\mathbf{1}_{[0,T-\delta)}(t)\\
&+\int_r^T\int_r^{s\wedge(t+\delta)}\Gamma_3^*(s,t)F(s,\theta)B_2(r)d\theta ds+\int_t^T\int_{s\wedge(t+\delta)}^s \Gamma_3^*(s,t)F(s,\theta)B_2(r)d\theta ds\mathbf{1}_{[0,T-\delta)}(t)\Big]u^*(r-\delta)dr\\
&=\int_t^{(t+\delta)\wedge T}\Big[\int_r^T \Gamma_1^*(s,t)ds+\int_{r+\delta}^T \Gamma_2^*(s,t)ds\mathbf{1}_{[t_0,T-\delta)}(t)+\int_r^T\int_r^s \Gamma_3^*(s,t)F(s,\theta)d\theta ds\Big]B_2(r)u^*(r-\delta)dr\\
&=\int_t^{(t+\delta)\wedge T}\int_r^T\Big[\Gamma_1^*(s,t)+\Gamma_2^*(s,t)\mathbf{1}_{(r+\delta,T]}(s)\mathbf{1}_{[t_0,T-\delta)}(t)+\int_r^s \Gamma_3^*(s,t)F(s,\theta)d\theta\Big]dsB_2(r)u^*(r-\delta)dr,
\end{aligned}
$$

and

$$
\begin{aligned}
&\int_{t_0}^t\Big[\int_t^T\int_\theta^T \Gamma_1^*(s,t)B_3(\theta)\tilde F(\theta,r)dsd\theta+\int_{t+\delta}^T\int_t^{s-\delta}\Gamma_2^*(s,t)B_3(\theta)\tilde F(\theta,r)d\theta ds\mathbf{1}_{[t_0,T-\delta)}(t)\\
&+\int_t^T\int_t^s\int_t^{\theta'}\Gamma_3^*(s,t)F(s,\theta')B_3(\theta)\tilde F(\theta,r)d\theta d\theta' ds\Big]u^*(r)dr\\
&=\int_{t_0}^t\int_t^T\Big(\int_\theta^T \Gamma_1^*(s,t)ds+\int_{\theta+\delta}^T \Gamma_2^*(s,t)ds\mathbf{1}_{[t_0,T-\delta)}(\theta)\mathbf{1}_{[t_0,T-\delta)}(t)+\int_\theta^T \Gamma_3^*(s,t)\int_\theta^s F(s,\theta')d\theta' ds\Big)B_3(\theta)\tilde F(\theta,r)d\theta u^*(r)dr\\
&=\int_{t_0}^t\int_t^T\int_\theta^T\Big(\Gamma_1^*(s,t)+\Gamma_2^*(s,t)\mathbf{1}_{[0,s-\delta)}(\theta)\mathbf{1}_{[t_0,T-\delta)}(t)+\Gamma_3^*(s,t)\int_\theta^s F(s,\theta')d\theta'\Big)dsB_3(\theta)\tilde F(\theta,r)d\theta u^*(r)dr.
\end{aligned}
$$

Then, the optimal closed-loop outcome control (A.32) of Problem (P) becomes (A.26), and $K_1^*(\cdot), K_2^*(\cdot,\cdot), K_3^*(\cdot), K_4^*(\cdot,\cdot), v^*(\cdot)$ are given by (A.27)–(A.31).

Step 2: In this step, we give furthermore explicit representation of K_1^* by means of $(P^{(1)},P^{(2)})$ and other given coefficients of the optimal control problem. Let $P=(P^{(1)},P^{(2)})$ be the solution to the Riccati–Volterra equation (3.11), and decompose them as $P^{(1)}(\cdot)=\big(P^{(1)}_{ij}(\cdot)\big)_{1\leqslant i,j\leqslant 3}$ and $P^{(2)}(\cdot)=\big(P^{(2)}_{ij}(\cdot)\big)_{1\leqslant i,j\leqslant 3}$. Then, by (3.14) and (3.10), for $i=1,2,3$, we have

$$
\begin{aligned}
&\Xi_i^*(t)=-\mathcal{R}(t)^{-1}D_1(t)^\top\mathcal{G}_1(t)C_i(t),\quad t_0<t<T,\\
&\Gamma_i^*(s,t)=-\mathcal{R}(t)^{-1}\int_t^T\mathcal{B}(\theta,t)^\top\Big(\Pi(s,t,\theta)^\top\mathbf{1}_{(t_0,s)}(\theta)\big[P^{(1)}_{1i}(s)^\top,P^{(1)}_{2i}(s)^\top,P^{(1)}_{3i}(s)^\top\big]^\top\\
&\qquad+\int_\theta^T\Pi(r,t,\theta)^\top\big[P^{(2)}_{1i}(r,s,t)^\top,P^{(2)}_{2i}(r,s,t)^\top,P^{(2)}_{3i}(r,s,t)^\top\big]^\top dr\Big)d\theta,\quad t_0<t<s<T.
\end{aligned}
\tag{A.39}
$$

By (A.27) and (A.39), we derive

$$
K_1^*(t)=-\mathcal{R}(t)^{-1}\Big\{D_1(t)^\top\mathcal{G}_1(t)C_1(t)+\int_t^T\int_t^s\mathcal{B}(\theta,t)^\top\Pi(s,t,\theta)^\top P^{(1)}(s)^\top(I,0,0)^\top d\theta ds
$$

$$+\int_{(t+\delta)\wedge T}^{T}\int_t^s \mathcal{B}(\theta,t)^\top\Pi(s,t,\theta)^\top P^{(1)}(s)^\top(0,I,0)^\top d\theta ds+\int_t^T\int_t^s \mathcal{B}(\theta,t)^\top\Pi(s,t,\theta)^\top P^{(1)}(s)^\top$$
$$\times(0,0,I)^\top\int_t^s F(s,\theta')d\theta' d\theta ds+\int_t^T\int_t^T\int_t^r \mathcal{B}(\theta,t)^\top\Pi(r,t,\theta)^\top P^{(2)}(s,r,t)^\top(I,0,0)^\top d\theta dr ds$$
$$+\int_{(t+\delta)\wedge T}^{T}\int_t^T\int_t^r \mathcal{B}(\theta,t)^\top\Pi(r,t,\theta)^\top P^{(2)}(s,r,t)^\top(0,I,0)^\top d\theta dr ds$$
$$+\int_t^T\int_t^T\int_t^r\mathcal{B}(\theta,t)^\top\Pi(r,t,\theta)^\top P^{(2)}(s,r,t)^\top(0,0,I)^\top\int_t^s F(s,\theta')d\theta' d\theta dr ds\Big\}.$$

It then follows from some basic calculations that

$$\begin{aligned}K_1^*(t)=&-\mathcal{R}(t)^{-1}\Big\{D_1(t)^\top\mathcal{G}_1(t)C_1(t)+\int_t^T\int_t^s \mathcal{B}(\theta,t)^\top\Pi(s,t,\theta)^\top\\&\times P^{(1)}(s)^\top\Big(I,I\mathbf{1}_{((t+\delta)\wedge T,\infty)}(s),\int_t^s F(s,\theta')^\top d\theta'\Big)^\top d\theta ds+\int_t^T\int_t^T\int_t^r\mathcal{B}(\theta,t)^\top\\&\times\Pi(r,t,\theta)^\top P^{(2)}(s,r,t)^\top\Big(I,I\mathbf{1}_{((t+\delta)\wedge T,\infty)}(s),\int_t^s F(s,\theta')^\top d\theta'\Big)^\top d\theta dr ds\Big\},\end{aligned}$$

which then implies

$$\begin{aligned}K_1^*(t)=&-\mathcal{R}(t)^{-1}\Big\{D_1(t)^\top\mathcal{G}_1(t)C_1(t)+\int_t^T\mathcal{B}(\theta,t)^\top\Big[\int_\theta^T\Pi(\alpha,t,\theta)^\top P^{(1)}(\alpha)^\top\Upsilon(\alpha,t)d\alpha\\&+\int_t^T\int_\theta^T\Pi(\alpha,t,\theta)^\top P^{(2)}(r,\alpha,t)^\top\Upsilon(r,t)d\alpha dr\Big]d\theta\Big\}.\end{aligned}$$

Step 3: In this step, we turn to calculate K_2^* and K_3^*. From (A.28) and (A.39), we deduce

$$\begin{aligned}K_2^*(t,s)=&-\mathcal{R}(t)^{-1}\Big\{D_1(t)^\top\mathcal{G}_1(t)C_3(t)F(t,s)+\int_t^{s+\delta}\mathcal{B}(\theta,t)^\top\Pi(s+\delta,t,\theta)^\top P^{(1)}(s+\delta)^\top\\&\times(0,I,0)^\top\mathbf{1}_{[t-\delta,T-\delta]}(s)d\theta+\int_t^T\int_t^r\mathcal{B}(\theta,t)^\top\Pi(r,t,\theta)^\top P^{(2)}(s+\delta,r,t)^\top(0,I,0)^\top d\theta dr\mathbf{1}_{[t-\delta,T-\delta]}(s)\\&+\int_t^T\int_t^\theta\mathcal{B}(\theta',t)^\top\Pi(\theta,t,\theta')^\top P^{(1)}(\theta)^\top(0,0,I)^\top F(\theta,s)d\theta' d\theta\\&+\int_t^T\int_t^T\int_t^r\mathcal{B}(\theta',t)^\top\Pi(r,t,\theta')^\top P^{(2)}(\theta,r,t)^\top(0,0,I)^\top F(\theta,s)d\theta' dr d\theta\Big\}.\end{aligned}$$

Applying Fubini theorem, we obtain

$$\begin{aligned}K_2^*(t,s)=&-\mathcal{R}(t)^{-1}\Big\{D_1(t)^\top\mathcal{G}_1(t)C_3(t)F(t,s)+\int_t^T\mathcal{B}(\theta,t)^\top\Big[\Big(\Pi(s+\delta,t,\theta)^\top P^{(1)}(s+\delta)^\top\mathbf{1}_{[t,s+\delta)}(\theta)\\&+\int_\theta^T\Pi(r,t,\theta)^\top P^{(2)}(s+\delta,r,t)^\top dr\Big)(0,I,0)^\top\mathbf{1}_{[t-\delta,T-\delta]}(s)+\int_\theta^T\Pi(\alpha,t,\theta)^\top P^{(1)}(\alpha)^\top(0,0,I)^\top F(\alpha,s)d\alpha\\&+\int_t^T\int_\theta^T\Pi(\alpha,t,\theta)^\top P^{(2)}(r,\alpha,t)^\top(0,0,I)^\top F(r,s)d\alpha dr\Big]d\theta\Big\}.\end{aligned}$$

Finally by (3.12), (A.29), we derive

$$K_3^*(t)=-\mathcal{R}(t)^{-1}D_1(t)^\top\mathcal{G}_1(t)C_2(t).$$

Step 4: In this step, we turn to look at the case of K_4^*. From (A.30), we have

$$K_4^*(t,s)=\Big\{\int_{s+\delta}^T\Big[\Gamma_1^*(r,t)+\Gamma_2^*(r,t)\mathbf{1}_{(s+2\delta,\infty)}(r)\mathbf{1}_{[t_0,T-\delta)}(t)\tag{A.40}$$

$$
\begin{aligned}
&+\int_{s+\delta}^{r}\Gamma_3^*(r,t)F(r,\theta)d\theta\Big]drB_2(s+\delta)\Big\}\mathbf{1}_{[t-\delta,T-\delta]}(s)+\int_t^T\int_{\theta'}^{T}\Big(\Gamma_1^*(r,t)+\Gamma_2^*(r,t)\\
&\times\mathbf{1}_{[t_0,r-\delta)}(\theta')\mathbf{1}_{[t_0,T-\delta)}(t)+\int_{\theta'}^{r}\Gamma_3^*(r,t)F(r,\theta)d\theta\Big)B_3(\theta')\tilde F(\theta',s)drd\theta'\\
&\equiv I_1(t,s+\delta)\mathbf{1}_{[t-\delta,T-\delta]}(s)+I_2(t,s),
\end{aligned}
$$

which implies that

$$
I_1(t,s)=\int_s^T\Big[\Gamma_1^*(r,t)+\Gamma_2^*(r,t)\mathbf{1}_{(s+\delta,\infty)}(r)\mathbf{1}_{[t_0,T-\delta)}(t)+\int_s^r\Gamma_3^*(r,t)F(r,\theta)d\theta\Big]drB_2(s).
$$

By (A.39), we get

$$
\begin{aligned}
I_1(t,s)=&-\mathcal{R}(t)^{-1}\Big\{\int_s^T\int_t^r\mathcal{B}(\theta,t)^\top\Pi(r,t,\theta)^\top P^{(1)}(r)^\top(I,0,0)^\top d\theta dr+\int_{s+\delta}^T\int_t^r\mathcal{B}(\theta,t)^\top\Pi(r,t,\theta)^\top\\
&\times P^{(1)}(r)^\top(0,I,0)^\top d\theta dr\mathbf{1}_{[0,T-\delta]}(t)+\int_s^T\int_s^r\int_t^r\mathcal{B}(\theta',t)^\top\Pi(r,t,\theta')^\top P^{(1)}(r)^\top(0,0,I)^\top F(r,\theta)d\theta'd\theta dr\\
&+\int_s^T\int_t^T\int_t^\alpha\mathcal{B}(\theta,t)^\top\Pi(\alpha,t,\theta)^\top P^{(2)}(r,\alpha,t)^\top(I,0,0)^\top d\theta d\alpha dr+\int_{s+\delta}^T\int_t^T\int_t^\alpha\mathcal{B}(\theta,t)^\top\Pi(\alpha,t,\theta)^\top P^{(2)}(r,\alpha,t)^\top\\
&\times(0,I,0)^\top d\theta d\alpha dr\mathbf{1}_{[0,T-\delta]}(t)+\int_s^T\int_s^r\int_t^T\int_t^\alpha\mathcal{B}(\theta',t)^\top\Pi(\alpha,t,\theta')^\top P^{(2)}(r,\alpha,t)^\top(0,0,I)^\top F(r,\theta)d\theta'd\alpha d\theta dr\Big\}B_2(s).
\end{aligned}
$$

It then follows from some calculations that

$$
\begin{aligned}
I_1(t,s)=&-\mathcal{R}(t)^{-1}\Big\{\int_t^T\mathcal{B}(\theta,t)^\top\Big[\int_{s\vee\theta}^T\Pi(r,t,\theta)^\top P^{(1)}(r)^\top\Big(I,\mathbf{1}_{(s+\delta,\infty)}(r)\\
&\times\mathbf{1}_{[0,T-\delta]}(t)I,\int_s^rF(r,\theta')^\top d\theta'\Big)^\top dr+\int_s^T\int_\theta^T\Pi(\alpha,t,\theta)^\top P^{(2)}(r,\alpha,t)^\top\\
&\times\Big(I,\mathbf{1}_{(s+\delta,\infty)}(r)\mathbf{1}_{[0,T-\delta]}(t)I,\int_s^rF(r,\theta')^\top d\theta'\Big)^\top d\alpha dr\Big]d\theta\Big\}B_2(s).
\end{aligned}
\tag{A.41}
$$

Similarly, by (A.39), we obtain

$$
\begin{aligned}
I_2(t,s)=&-\mathcal{R}(t)^{-1}\Big\{\int_t^T\mathcal{B}(\theta,t)^\top\Big\{\int_\theta^T\Pi(r,t,\theta)^\top P^{(1)}(r)^\top\Big(\int_t^r\tilde F(\theta',s)^\top B_3(\theta')^\top d\theta',\int_t^{r-\delta}\tilde F(\theta',s)^\top B_3(\theta')^\top d\theta',\\
&\int_t^r\int_t^\beta\tilde F(\theta',s)^\top B_3(\theta')^\top F(r,\beta)^\top d\theta'd\beta\Big)^\top dr+\int_t^T\int_\theta^T\Pi(\alpha,t,\theta)^\top P^{(2)}(r,\alpha,t)^\top\Big(\int_t^r\tilde F(\theta',s)^\top B_3(\theta')^\top d\theta',\\
&\int_t^{r-\delta}\tilde F(\theta',s)^\top B_3(\theta')^\top d\theta',\int_t^r\int_t^\beta\tilde F(\theta',s)^\top B_3(\theta')^\top F(r,\beta)^\top d\theta'd\beta\Big)^\top d\alpha dr\Big\}d\theta\Big\}.
\end{aligned}
\tag{A.42}
$$

Substituting (A.41) and (A.42) into (A.40), we deduce

$$
\begin{aligned}
K_4^*(t,s)=&-\mathcal{R}(t)^{-1}\Big\{\int_t^T\mathcal{B}(\theta,t)^\top\Big[\int_\theta^T\Pi(\alpha,t,\theta)^\top P^{(1)}(\alpha)^\top\Big\{\Big(I,\mathbf{1}_{(s+2\delta,\infty)}(\alpha)\mathbf{1}_{(0,T-\delta)}(t)I,\int_{s+\delta}^\alpha F(\alpha,\theta')^\top d\theta'\Big)^\top\\
&\times\mathbf{1}_{(s+\delta,T)}(\alpha)B_2(s+\delta)\mathbf{1}_{[t-\delta,T-\delta]}(s)+\Big(\int_t^\alpha\big(B_3(\theta')\tilde F(\theta',s)\big)^\top d\theta',\int_t^{\alpha-\delta}\big(B_3(\theta')\tilde F(\theta',s)\big)^\top d\theta',\\
&\int_t^\alpha\int_t^\beta\big(F(\alpha,\beta)B_3(\theta')\tilde F(\theta',s)\big)^\top d\theta'd\beta\Big)^\top\Big\}d\alpha+\int_t^T\int_\theta^T\Pi(\alpha,t,\theta)^\top P^{(2)}(r,\alpha,t)^\top\\
&\times\Big\{\Big(I,\mathbf{1}_{(s+2\delta,\infty)}(r)\mathbf{1}_{(0,T-\delta)}(t)I,\int_{s+\delta}^rF(r,\theta')^\top d\theta'\Big)^\top\mathbf{1}_{(s+\delta,T)}(r)B_2(s+\delta)\mathbf{1}_{[t-\delta,T-\delta]}(s)\\
&+\Big(\int_t^r\big(B_3(\theta')\tilde F(\theta',s)\big)^\top d\theta',\int_t^{r-\delta}\big(B_3(\theta')\tilde F(\theta',s)\big)^\top d\theta',\int_t^r\int_t^\beta\big(F(r,\beta)B_3(\theta')\tilde F(\theta',s)\big)^\top d\theta'd\beta\Big)^\top\Big\}d\alpha dr\Big]d\theta\Big\}.
\end{aligned}
$$

Step 5: In this step, we calculate v^*. Recalling (A.39), we have

$$
\begin{aligned}
&\int_{t-\delta}^{t_0}\Big\{\int_{s+\delta}^{T}\Big[\Gamma_1^*(r,t)+\Gamma_2^*(r,t)\mathbf{1}_{(s+2\delta,\infty)}(r)\mathbf{1}_{[0,T-\delta)}(t)\\
&\qquad+\int_{s+\delta}^{r}\Gamma_3^*(r,t)F(r,\theta)d\theta\Big]drB_2(s+\delta)\Big\}\varsigma(s-t_0)ds\mathbf{1}_{[t_0,t_0+\delta]}(t)\\
&=-\mathcal{R}(t)^{-1}\int_t^T\mathcal{B}(\theta,t)^\top\Big[\int_\theta^T\Pi(r,t,\theta)^\top P^{(1)}(r)^\top\int_{t-\delta}^{t_0}\mathbf{1}_{(s+\delta,\infty)}(r)\\
&\quad\times\Big(I,\mathbf{1}_{(s+2\delta,\infty)}(r)\mathbf{1}_{(0,T-\delta)}(t)I,\int_{s+\delta}^{r}F(r,\theta')^\top d\theta'\Big)^\top B_2(s+\delta)\varsigma(s-t_0)ds\mathbf{1}_{[t_0,t_0+\delta]}(t)dr\\
&\quad+\int_t^T\int_\theta^T\Pi(\alpha,t,\theta)^\top P^{(2)}(r,\alpha,t)^\top\int_{t-\delta}^{t_0}\mathbf{1}_{(s+\delta,\infty)}(r)\Big(I,\mathbf{1}_{(s+2\delta,\infty)}(r)\mathbf{1}_{(0,T-\delta)}(t)I,\\
&\quad\int_{s+\delta}^{r}F(r,\theta')^\top d\theta'\Big)^\top B_2(s+\delta)\varsigma(s-t_0)ds\mathbf{1}_{[t_0,t_0+\delta]}(t)d\alpha dr\Big]d\theta.
\end{aligned}
$$

Similarly, we obtain

$$
\begin{aligned}
&\int_t^T\Gamma_2^*(s,t)\xi(s-t_0-\delta)\mathbf{1}_{[t_0,t_0+\delta]}(s)ds\\
&=-\mathcal{R}(t)^{-1}\int_t^T\Big(\int_t^s\mathcal{B}(\theta,t)^\top\Pi(s,t,\theta)^\top d\theta P^{(1)}(s)^\top(0,I,0)^\top\\
&\quad+\int_t^T\int_t^r\mathcal{B}(\theta,t)^\top\Pi(r,t,\theta)^\top P^{(2)}(r,s,t)(0,I,0)^\top d\theta dr\Big)\xi(s-\delta-t_0)\mathbf{1}_{[t_0,t_0+\delta]}(s)ds.
\end{aligned}
\tag{A.43}
$$

Finally, by the Fubini theorem,

$$
\begin{aligned}
&\int_t^T\int_t^T\int_t^r\mathcal{B}(\theta,t)^\top\Pi(r,t,\theta)^\top P^{(2)}(r,s,t)(0,I,0)^\top d\theta dr\xi(s-\delta-t_0)\mathbf{1}_{[t_0,t_0+\delta]}(s)ds\\
&=\int_t^T\int_t^T\int_\theta^T\mathcal{B}(\theta,t)^\top\Pi(\alpha,t,\theta)^\top P^{(2)}(\alpha,r,t)(0,I,0)^\top d\alpha d\theta\xi(s-\delta-t_0)\mathbf{1}_{[t_0,t_0+\delta]}(s)ds,
\end{aligned}
$$

which and (A.22), (A.31), (A.43) imply that

$$
\begin{aligned}
v^*(t)&=-\mathcal{R}(t)^{-1}\Big\{D_1(t)^\top\mathcal{G}_1(t)\sigma(t)+\int_t^T D_1(t)^\top\Upsilon(\alpha,t)^\top\zeta(\alpha,t)d\alpha+\int_t^T\int_\theta^T\mathcal{B}(\theta,t)^\top\Pi(\alpha,t,\theta)^\top\eta(\alpha,t)d\alpha d\theta\\
&+\int_t^T\mathcal{B}(\theta,t)^\top\Big[\int_\theta^T\Pi(\alpha,t,\theta)^\top P^{(1)}(\alpha)^\top\Big((0,I,0)^\top\xi(\alpha-\delta-t_0)\mathbf{1}_{[t_0,t_0+\delta]}(\alpha)+\int_{t-\delta}^{t_0}\mathbf{1}_{(\theta'+\delta,\infty)}(\alpha)(I,\mathbf{1}_{(\theta'+2\delta,\infty)}(\alpha)\\
&\times\mathbf{1}_{(0,T-\delta)}(t)I,\int_{\theta'+\delta}^{\alpha}F(\alpha,\beta)^\top d\beta)^\top B_2(\theta'+\delta)\varsigma(\theta'-t_0)d\theta'\mathbf{1}_{[t_0,t_0+\delta]}(t)\Big)d\alpha+\int_t^T\int_\theta^T\Pi(\alpha,t,\theta)^\top P^{(2)}(r,\alpha,t)^\top\\
&\times\Big((0,I,0)^\top\xi(r-\delta-t_0)\mathbf{1}_{[t_0,t_0+\delta]}(r)+\int_{t-\delta}^{t_0}\mathbf{1}_{(\theta'+\delta,\infty)}(r)(I,\mathbf{1}_{(\theta'+2\delta,\infty)}(r)\mathbf{1}_{(0,T-\delta)}(t)I,\\
&\int_{\theta'+\delta}^{r}F(r,\beta)^\top d\beta)^\top B_2(\theta'+\delta)\varsigma(\theta'-t_0)d\theta'\mathbf{1}_{[t_0,t_0+\delta]}(t)\Big)d\alpha dr\Big]d\theta\Big\}.
\end{aligned}
$$

Step 6: In this step, we show that $(K_1^*(\cdot),K_2^*(\cdot,\cdot),K_3^*(\cdot),K_4^*(\cdot,\cdot),v^*(\cdot))$ is the optimal closed-loop strategy of Problem (P) on $[t_0,T]$ in terms of Definition 2.4.

In fact, by the optimality in Step 1, it is sufficient to prove that $(K_1^*(\cdot),K_2^*(\cdot,\cdot),K_3^*(\cdot),K_4^*(\cdot,\cdot),v^*(\cdot))\in\mathbb{L}$. Next we prove that $K_1(\cdot)\in L^2(t_0,T;\mathbb{R}^{n\times m})$. Since $(P^{(1)},P^{(2)})\in\Pi(0,T)$, we have

$$
\operatorname*{ess\,sup}_{t\in(0,T)}|P^{(1)}(t)|+\Big(\int_0^T\int_0^T\sup_{t\in[0,s_1\wedge s_2]}|P^{(2)}(s_1,s_2,t)|^2ds_1ds_2\Big)^{\frac{1}{2}}<\infty.
\tag{A.44}
$$

By the boundedness of $F(\cdot,\cdot)$, we obtain

$$\sup_{t\in(0,T)}\int_t^T|\Upsilon(s,t)^\top P^{(1)}(s)\Upsilon(s,t)|ds\leqslant M\operatorname*{ess\,sup}_{t\in(0,T)}|P^{(1)}(t)|\int_0^T\Big|1+\int_0^s|F(s,r)|dr\Big|^2ds<\infty,\tag{A.45}$$

and

$$\begin{aligned}&\sup_{t\in(0,T)}\int_t^T\int_t^T|\Upsilon(s_1,t)^\top P^{(2)}(s_1,s_2,t)\Upsilon(s_2,t)|ds_1ds_2\\&\leqslant M\Big(\int_0^T\int_0^T\sup_{t\in[0,s_1\wedge s_2]}|P^{(2)}(s_1,s_2,t)|^2ds_1ds_2\Big)^{\frac12}<\infty,\end{aligned}\tag{A.46}$$

where M is a generic constant. Thus (A.45) and (A.46) imply the boundedness of $\mathcal{G}_1(\cdot)$. Recall Theorem 3.3, $\mathcal{R}(\cdot)\geqslant\beta I$ for some constant $\beta>0$. Then, the boundedness of $R_1(\cdot)$, $R_2(\cdot)$, $D_1(\cdot)$ imply that $\mathcal{R}(\cdot)^{-1}$ is bounded. Notice that

$$\begin{aligned}&\int_{t_0}^T\Big|\int_t^T\int_t^T\mathcal{B}(\theta,t)^\top\int_\theta^T\Big(P^{(2)}(\alpha,r,t)\Pi(r,t,\theta)\Big)^\top dr\Upsilon(\alpha,t)d\alpha d\theta\Big|^2dt\\&\leqslant\int_{t_0}^T\Big|\int_t^T|\mathcal{B}(\theta,t)|\int_\theta^T\int_t^T\Big(\frac{1}{r-t}+1+\frac{1}{r-t}|\mathcal{E}(r,t)|+\frac{1}{r-t}|\mathcal{E}(r,t+\delta)|+|\mathcal{E}(r,\theta)|\Big)\\&\quad\times|P^{(2)}(\alpha,r,t)|\Big(1+|\mathcal{E}(\alpha,t)|\Big)d\alpha drd\theta\Big|^2dt.\end{aligned}\tag{A.47}$$

By the boundedness of $F(\cdot,\cdot)$, $B_1(\cdot)$, $B_2(\cdot)$ and $B_3(\cdot)$, we deduce

$$\begin{aligned}&\int_{t_0}^T\Big|\int_t^T|\mathcal{B}(\theta,t)|\int_\theta^T\int_t^T\frac{1}{r-t}|P^{(2)}(\alpha,r,t)||\mathcal{E}(\alpha,t)|d\alpha drd\theta\Big|^2dt\\&=\int_{t_0}^T\Big|\int_t^T\int_t^r|\mathcal{B}(\theta,t)|\int_t^T\frac{1}{r-t}|P^{(2)}(\alpha,r,t)||\mathcal{E}(\alpha,t)|d\alpha d\theta dr\Big|^2dt\\&\leqslant\int_{t_0}^T\int_t^T\Big(\int_t^r|\mathcal{B}(\theta,t)|d\theta\frac{1}{r-t}\Big)^2dr\int_t^T\Big(\int_t^T|P^{(2)}(\alpha,r,t)\mathcal{E}(\alpha,t)|d\alpha\Big)^2drdt\\&\leqslant M\int_0^T\int_0^T\sup_{t\in[0,\alpha\wedge r]}|P^{(2)}(\alpha,r,t)|^2d\alpha dr<\infty,\end{aligned}\tag{A.48}$$

and

$$\begin{aligned}&\int_{t_0}^T\Big|\int_t^T|\mathcal{B}(\theta,t)|\int_\theta^T\int_t^T\frac{1}{r-t}|\mathcal{E}(r,t)||P^{(2)}(\alpha,r,t)||\mathcal{E}(\alpha,t)|d\alpha drd\theta\Big|^2dt\\&=\int_{t_0}^T\Big|\int_t^T\int_t^r|\mathcal{B}(\theta,t)|\int_t^T\frac{1}{r-t}|\mathcal{E}(r,t)||P^{(2)}(\alpha,r,t)||\mathcal{E}(\alpha,t)|d\alpha d\theta dr\Big|^2dt\\&\leqslant M\int_0^T\int_0^T\sup_{t\in[0,\alpha\wedge r]}|P^{(2)}(\alpha,r,t)|^2d\alpha dr<\infty.\end{aligned}\tag{A.49}$$

Similar to (A.48) and (A.49), we can deal with other terms in (A.47) and prove that $K_1(\cdot)\in L^2(t_0,T;\mathbb{R}^{n\times m})$. Furthermore, we can verify that $(K_1^*(\cdot),K_2^*(\cdot,\cdot),K_3^*(\cdot),K_4^*(\cdot,\cdot),v^*(\cdot))\in\mathbb{L}$. ∎

Before proving Corollary 4.1 and Corollary 4.3, we firstly decompose the Riccati system (3.11) and the Type-II extended backward SVIE (3.13).

Let $P=(P^{(1)},P^{(2)})$ be the solution to the Riccati–Volterra equation (3.11) and decompose them as follows:

$$P^{(1)}(t)=\begin{bmatrix}P^{(1)}_{11}(t)&P^{(1)}_{12}(t)&P^{(1)}_{13}(t)\\P^{(1)}_{21}(t)&P^{(1)}_{22}(t)&P^{(1)}_{23}(t)\\P^{(1)}_{31}(t)&P^{(1)}_{32}(t)&P^{(1)}_{33}(t)\end{bmatrix},\quad P^{(2)}(s,t,r)=\begin{bmatrix}P^{(2)}_{11}(s,t,r)&P^{(2)}_{12}(s,t,r)&P^{(2)}_{13}(s,t,r)\\P^{(2)}_{21}(s,t,r)&P^{(2)}_{22}(s,t,r)&P^{(2)}_{23}(s,t,r)\\P^{(2)}_{31}(s,t,r)&P^{(2)}_{32}(s,t,r)&P^{(2)}_{33}(s,t,r)\end{bmatrix}.$$

Recall that (3.11) is equivalent to (A.8). Then, we decompose its coefficients as follows:

$$(P\rtimes A)(s,t)=\begin{bmatrix}\mathcal{T}_{PA}^{(11)}(s,t) & \mathcal{T}_{PA}^{(12)}(s,t) & \mathcal{T}_{PA}^{(13)}(s,t)\\ \mathcal{T}_{PA}^{(21)}(s,t) & \mathcal{T}_{PA}^{(22)}(s,t) & \mathcal{T}_{PA}^{(23)}(s,t)\\ \mathcal{T}_{PA}^{(31)}(s,t) & \mathcal{T}_{PA}^{(32)}(s,t) & \mathcal{T}_{PA}^{(33)}(s,t)\end{bmatrix},\quad (P\rtimes B)(s,t)=\begin{bmatrix}\mathcal{T}_{PB}^{(1)}(s,t)\\ \mathcal{T}_{PB}^{(2)}(s,t)\\ \mathcal{T}_{PB}^{(3)}(s,t)\end{bmatrix}.$$

In this case, we have

$$\begin{aligned}\mathcal{T}_{PA}^{(ij)}(s,t)&=\Big[P_{i1}^{(1)}(s)+P_{i2}^{(1)}(s)\mathbf{1}_{(t+\delta,\infty)}(s)+P_{i3}^{(1)}(s)\mathcal{E}(s,t)\Big]A_j(t)\\ &+\int_t^T\Big[P_{i1}^{(2)}(s,r,t)+P_{i2}^{(2)}(s,r,t)\mathbf{1}_{(t+\delta,\infty)}(r)+P_{i3}^{(2)}(s,r,t)\mathcal{E}(r,t)\Big]drA_j(t),\ 0<t<s<T, i,j=1,2,3,\end{aligned}\tag{A.50}$$

and

$$\begin{aligned}\mathcal{T}_{PB}^{(i)}(s,t)&=\Big\{P_{i1}^{(1)}(s)+P_{i2}^{(1)}(s)\mathbf{1}_{[0,s-\delta)}(t)+P_{i3}^{(1)}(s)\mathcal{E}(s,t)+\int_t^T\Big[P_{i1}^{(2)}(s,r,t)+P_{i2}^{(2)}(s,r,t)\mathbf{1}_{[0,r-\delta)}(t)\\ &\quad+P_{i3}^{(2)}(s,r,t)\mathcal{E}(r,t)\Big]dr\Big\}B_1(t)+\Big\{P_{i1}^{(1)}(s)\mathbf{1}_{[0,s-\delta)}(t)+P_{i2}^{(1)}(s)\mathbf{1}_{[0,s-2\delta)}(t)+P_{i3}^{(1)}(s)\mathcal{E}(s,t+\delta)\\ &\quad+\int_t^T\Big[P_{i1}^{(2)}(s,r,t)\mathbf{1}_{[0,r-\delta)}(t)+P_{i2}^{(2)}(s,r,t)\mathbf{1}_{[0,r-2\delta)}(t)+P_{i3}^{(2)}(s,r,t)\mathcal{E}(r,t+\delta)\Big]dr\Big\}B_2(t+\delta)\\ &\quad+\int_t^s\Big[P_{i1}^{(1)}(s)+P_{i2}^{(1)}(s)\mathbf{1}_{[0,s-\delta)}(\theta)+P_{i3}^{(1)}(s)\mathcal{E}(s,\theta)\Big]B_3(\theta)\tilde F(\theta,t)d\theta+\int_t^T\int_t^r\Big[P_{i1}^{(2)}(s,r,t)\\ &\quad+P_{i2}^{(2)}(s,r,t)\mathbf{1}_{[0,r-\delta)}(\theta)+P_{i3}^{(2)}(s,r,t)\mathcal{E}(r,\theta)\Big]B_3(\theta)\tilde F(\theta,t)d\theta dr, i=1,2,3.\end{aligned}\tag{A.51}$$

Hence, by (A.12)–(A.15), we derive the following decomposed coupled Riccati equation:

$$\begin{cases}P_{ii}^{(1)}(t)=Q_i(t)+C_i(t)^\top\mathcal{G}_1(t)C_i(t)-C_i(t)^\top\mathcal{G}_1(t)D_1(t)\mathcal{R}(t)^{-1}D_1(t)^\top\mathcal{G}_1(t)C_i(t),\\ \qquad\qquad 0<t<T,\ i=1,2,3,\\ P_{ij}^{(1)}(t)=C_i(t)^\top\mathcal{G}_1(t)C_j(t)-C_i(t)^\top\mathcal{G}_1(t)D_1(t)\mathcal{R}(t)^{-1}D_1(t)^\top\mathcal{G}_1(t)C_j(t),\\ \qquad\qquad 0<t<T,\quad i\neq j,\quad i,j=1,2,3,\\ \dot P_{ij}^{(2)}(s_1,s_2,t)=\mathcal{T}_{PB}^{(i)}(s_1,t)\mathcal{R}(t)^{-1}\mathcal{T}_{PB}^{(j)}(s_2,t)^\top,\quad 0<t<(s_1\wedge s_2)<T,\quad i,j=1,2,3,\\ P_{ij}^{(2)}(s,t,t)=P_{ji}^{(2)}(t,s,t)^\top=\mathcal{T}_{PA}^{(ij)}(s,t)-\mathcal{T}_{PB}^{(i)}(s,t)\mathcal{R}(t)^{-1}D_1(t)^\top\mathcal{G}_1(t)C_j(t),\quad 0<t<s<T,\end{cases}\tag{A.52}$$

where $\mathcal{R}(\cdot)$, $\mathcal{G}_1(\cdot)$, $\mathcal{T}_{PA}^{(ij)}(\cdot,\cdot)$ and $\mathcal{T}_{PB}^{(i)}(\cdot,\cdot)$ are defined by (3.12), (A.50), (A.51), and the decomposed $\mathcal{G}_1(\cdot)$ has a specific representation in the subsequent special cases, which is omitted here. Similarly one can decompose the Type-II extended backward SVIE (3.13). We will not go into the details.

The proof of Corollary 4.1:

Proof. In this special case, by (A.52) we get

$$P_{11}^{(1)}(t)=Q_1(t)+C_1(t)^\top\mathcal{G}_1(t)C_1(t)-C_1(t)^\top\mathcal{G}_1(t)D_1(t)\mathcal{R}(t)^{-1}D_1(t)^\top\mathcal{G}_1(t)C_1(t),\ 0<t<T,$$

and

$$P_{11}^{(2)}(s,t,t)=\mathcal{T}_{PA}^{(11)}(s,t)-\mathcal{T}_{PB}^{(1)}(s,t)\mathcal{R}(t)^{-1}D_1(t)^\top\mathcal{G}_1(t)C_1(t),\quad 0<t<s<T,\tag{A.53}$$

$$\dot P_{11}^{(2)}(s_1,s_2,t)=\mathcal{T}_{PB}^{(1)}(s_1,t)\mathcal{R}(t)^{-1}\mathcal{T}_{PB}^{(1)}(s_2,t)^\top,\ 0<t<(s_1\wedge s_2)<T,\tag{A.54}$$

$$P_{ij}^{(1)}(\cdot),P_{ij}^{(2)}(\cdot,\cdot,\cdot)=0,\quad (i,j)=(1,2),(1,3),(2,1),(2,2),(2,3),(3,1),(3,2),(3,3),\quad 0<t<T,$$

where $\mathcal{T}_{PB}^{(1)}(\cdot,\cdot)$ is given as follows:

$$\begin{aligned}\mathcal{T}_{PB}^{(1)}(s,t)&=\Big\{P_{11}^{(1)}(s)+\int_t^T P_{11}^{(2)}(s,r,t)dr\Big\}B_1(t)+\Big\{P_{11}^{(1)}(s)\mathbf{1}_{[0,s-\delta)}(t)+\int_t^T P_{11}^{(2)}(s,r,t)\mathbf{1}_{[0,r-\delta)}(t)dr\Big\}\\ &\quad\times B_2(t+\delta)+P_{11}^{(1)}(s)\int_t^s B_3(\theta)\tilde F(\theta,t)d\theta+\int_t^T P_{11}^{(2)}(s,r,t)\Big[\int_t^r B_3(\theta)\tilde F(\theta,t)d\theta\Big]dr, i=1,2,3.\end{aligned}$$

In addition,

$$\begin{aligned}
&\mathcal{G}_1(t)=\int_t^T P_{11}^{(1)}(s)ds+\int_t^T\int_t^T P_{11}^{(2)}(s_1,s_2,t)ds_1ds_2=\mathcal{P}_1(t,t,t)=\mathcal{S}_0(t),\\
&\mathcal{R}(t)=R_1(t)+D_1(t)^\top\mathcal{G}_1(t)D_1(t),\\
&\mathcal{T}_{PA}^{(11)}(s,t)=\Big[P_{11}^{(1)}(s)+\int_t^T P_{11}^{(2)}(s,r,t)dr\Big]A_1(t).
\end{aligned}$$

From (4.2) and (4.4), for $r\in[t-\delta,t]$, we have

$$\mathcal{S}_1(t,r-t)=B_2(r+\delta)^\top\mathcal{P}_1(t,t,r+\delta)+\int_t^T\tilde{F}(\theta',r)^\top B_3(\theta')^\top\mathcal{P}_1(t,t,\theta')d\theta'.$$

Notice that for $t\leqslant\theta'$, we have

$$\mathcal{P}_1(t,t,\theta')=\int_{\theta'}^T P_{11}^{(1)}(s)ds+\int_{\theta'}^T\int_t^T P_{11}^{(2)}(s,\alpha,t)d\alpha ds.$$

Therefore,

$$\begin{aligned}
\frac{\partial}{\partial t}\mathcal{P}_1(t,t,\theta')&=-\int_{\theta'}^T P_{11}^{(2)}(s,t,t)ds-\int_{\theta'}^T\int_t^T \dot{P}_{11}^{(2)}(s,\alpha,t)d\alpha ds\\
&=-\int_{\theta'}^T P_{11}^{(2)}(s,t,t)ds-\Big[\int_{\theta'}^T\mathcal{T}_{PB}^{(1)}(s,t)ds\Big]\mathcal{R}(t)^{-1}\Big[\int_t^T\mathcal{T}_{PB}^{(1)}(\alpha,t)^\top d\alpha\Big].
\end{aligned}$$

It then follows from some calculations that

$$\begin{aligned}
\frac{\partial}{\partial t}\mathcal{P}_1(t,t,\theta')&=-\int_{\theta'}^T P_{11}^{(1)}(s)dsA_1(t)-\int_{\theta'}^T\int_t^T P_{11}^{(2)}(s,r,t)drdsA_1(t)\\
&+\Big(\mathcal{P}_1(t,t,\theta')B_1(t)+\mathcal{P}_1(t,t+\delta,\theta')B_2(t+\delta)+\int_t^T\mathcal{P}_1(t,\theta,\theta')B_3(\theta)\tilde{F}(\theta,t)d\theta\Big)\\
&\times\mathcal{R}(t)^{-1}D_1(t)^\top\mathcal{G}_1(t)C_1(t)+\Big(\mathcal{P}_1(t,t,\theta')B_1(t)+\mathcal{P}_1(t,t+\delta,\theta')B_2(t+\delta)+\int_t^T\mathcal{P}_1(t,\theta,\theta')B_3(\theta)\tilde{F}(\theta,t)d\theta\Big)\\
&\times\mathcal{R}(t)^{-1}\Big(\mathcal{P}_1(t,t,t)B_1(t)+\mathcal{P}_1(t,t+\delta,t)B_2(t+\delta)+\int_t^T\mathcal{P}_1(t,\theta,t)B_3(\theta)\tilde{F}(\theta,t)d\theta\Big)^\top,\ \text{a.e. } t.
\end{aligned}$$

Then, we deduce

$$\begin{aligned}
\frac{\partial}{\partial t}\mathcal{S}_1(t,r-t)&=-B_2(r+\delta)^\top\int_{r+\delta}^T P_{11}^{(1)}(s)dsA_1(t)-B_2(r+\delta)^\top\int_{r+\delta}^T\int_t^T P_{11}^{(2)}(s,\tau,t)d\tau dsA_1(t)\\
&+B_2(r+\delta)^\top\Big(\mathcal{P}_1(t,t,r+\delta)B_1(t)+\mathcal{P}_1(t,t+\delta,r+\delta)B_2(t+\delta)+\int_t^T\mathcal{P}_1(t,\theta,r+\delta)B_3(\theta)\tilde{F}(\theta,t)d\theta\Big)\\
&\times\mathcal{R}(t)^{-1}D_1(t)^\top\mathcal{G}_1(t)C_1(t)+B_2(r+\delta)^\top\Big(\mathcal{P}_1(t,t,r+\delta)B_1(t)+\mathcal{P}_1(t,t+\delta,r+\delta)B_2(t+\delta)\\
&+\int_t^T\mathcal{P}_1(t,\theta,r+\delta)B_3(\theta)\tilde{F}(\theta,t)d\theta\Big)\mathcal{R}(t)^{-1}\Big(\mathcal{P}_1(t,t,t)B_1(t)+\mathcal{P}_1(t,t+\delta,t)B_2(t+\delta)\\
&+\int_t^T\mathcal{P}_1(t,\theta,t)B_3(\theta)\tilde{F}(\theta,t)d\theta\Big)^\top-\tilde{F}(t,r)^\top B_3(t)^\top\mathcal{P}_1(t,t,t)-\int_t^T\tilde{F}(\theta',r)^\top B_3(\theta')^\top\int_{\theta'}^T P_{11}^{(1)}(s)dsA_1(t)d\theta'\\
&-\int_t^T\tilde{F}(\theta',r)^\top B_3(\theta')^\top\int_{\theta'}^T\int_t^T P_{11}^{(2)}(s,\tau,t)d\tau dsd\theta'A_1(t)+\int_t^T\tilde{F}(\theta',r)^\top B_3(\theta')^\top\Big(\mathcal{P}_1(t,t,\theta')B_1(t)\\
&+\mathcal{P}_1(t,t+\delta,\theta')B_2(t+\delta)+\int_t^T\mathcal{P}_1(t,\theta,\theta')B_3(\theta)\tilde{F}(\theta,t)d\theta\Big)\mathcal{R}(t)^{-1}D_1(t)^\top\mathcal{G}_1(t)C_1(t)d\theta'\\
&+\int_t^T\tilde{F}(\theta',r)^\top B_3(\theta')^\top\Big(\mathcal{P}_1(t,t,\theta')B_1(t)+\mathcal{P}_1(t,t+\delta,\theta')B_2(t+\delta)+\int_t^T\mathcal{P}_1(t,\theta,\theta')B_3(\theta)\tilde{F}(\theta,t)d\theta\Big)
\end{aligned}$$

$$\times\mathcal{R}(t)^{-1}\Big(\mathcal{P}_1(t,t,t)B_1(t)+\mathcal{P}_1(t,t+\delta,t)B_2(t+\delta)+\int_t^T\mathcal{P}_1(t,\theta,t)B_3(\theta)\tilde{F}(\theta,t)d\theta\Big)^\top d\theta'.\ \text{a.e. } t.$$

It yields

$$\begin{aligned}&\frac{\partial}{\partial t}\mathcal{S}_1(t,r-t)+\tilde{F}(t,r)^\top B_3(t)^\top\mathcal{S}_0(t)+\mathcal{S}_1(t,r-t)A_1(t)-\big[\mathcal{S}_1(t,r-t)B_1(t)\\&\quad+\mathcal{S}_2(t,r-t,0)\big]\mathcal{R}(t)^{-1}\big[B_1(t)^\top\mathcal{S}_0(t)+\mathcal{S}_1(t,0)+D_1(t)^\top\mathcal{S}_0(t)C_1(t)\big]=0,\quad \text{a.e. } t\in[t_0,T],\end{aligned}$$

which implies that (4.7) holds. The proofs for (4.6) and (4.8) are similar.

Recalling (3.15)–(3.20), in this case the optimal closed-loop outcome control is as follows:

$$\begin{aligned}u^*(t)=&-\mathcal{R}(t)^{-1}\Big\{\Big(D_1(t)^\top\mathcal{S}_0(t)C_1(t)+B_1(t)^\top\mathcal{S}_0(t)+B_2(t+\delta)^\top\mathcal{P}_1(t,t+\delta,t)^\top\\&+\int_t^T\tilde{F}(\theta,t)^\top B_3(\theta)^\top\mathcal{P}_1(t,\theta,t)^\top d\theta\Big)x^*(t)+\int_{(t-\delta)\vee t_0}^{t\wedge(T-\delta)}\Big[B_1(t)^\top\mathcal{P}_1(t,t,r+\delta)^\top\\&+B_2(t+\delta)^\top\mathcal{P}_1(t,t+\delta,r+\delta)^\top+\int_t^T\tilde{F}(\theta,t)^\top B_3(\theta)^\top\mathcal{P}_1(t,\theta,r+\delta)^\top d\theta\Big]B_2(r+\delta)u^*(r)dr\\&+\int_{t_0}^t\int_t^T\Big[B_1(t)^\top\mathcal{P}_1(t,t,\theta')+B_2(t+\delta)^\top\mathcal{P}_1(t,t+\delta,\theta')^\top+\int_t^T\tilde{F}(\theta,t)^\top B_3(\theta)^\top\mathcal{P}_1(t,\theta,\theta')^\top d\theta\Big]\\&\times B_3(\theta')\tilde{F}(\theta',r)u^*(r)d\theta'dr\Big\}.\end{aligned}$$

By the definitions of (4.3)–(4.5), we deduce (4.9) and thus complete the proof of Corollary 4.1. ∎

The proof of Corollary 4.3:

Proof. In this case, by (A.52) we get

$$P_{i3}^{(j)}(t),P_{3i}^{(j)}(t),P_{3i}^{(j)}(t)=0,\ \ i=1,2,3,\ \ j=1,2,\ \ t\in(0,T),$$

and other $P_{ij}^{(1)}$ and $P_{ij}^{(2)}$ satisfy (A.52), $i,j=1,2$. By the definition (4.11) of $\mathcal{P}_3(\cdot,\cdot)$, we have

$$\begin{aligned}\frac{\partial\mathcal{P}_3(t,\theta)}{\partial t}=&-P_{21}^{(2)}(\theta,t,t)^\top-P_{22}^{(2)}(\theta,t+\delta,t)^\top+\int_t^T\Big[P_{11}^{(1)}(r)+P_{12}^{(1)}(r)\mathbf{1}_{[0,r-\delta)}(t)\\&+\int_t^T\Big(P_{11}^{(2)}(r,\alpha,t)+P_{12}^{(2)}(r,\alpha,t)\mathbf{1}_{[0,\alpha-\delta)}(t)\Big)d\alpha\Big]drB_1(R_1+D_1^\top\mathcal{P}_2(t)D_1)^{-1}B_1^\top\Big[P_{21}^{(1)}(\theta)^\top\\&+P_{22}^{(1)}(\theta)^\top\mathbf{1}_{[0,\theta-\delta)}(t)+\int_t^T\Big(P_{21}^{(2)}(\theta,\alpha,t)^\top+P_{22}^{(2)}(\theta,\alpha,t)^\top\mathbf{1}_{[0,\alpha-\delta)}(t)\Big)d\alpha\Big]\\&+\int_{t+\delta}^T\Big[P_{21}^{(1)}(r)+P_{22}^{(1)}(r)\mathbf{1}_{[0,r-\delta)}(t)+\int_t^T\Big(P_{21}^{(2)}(r,\alpha,t)+P_{22}^{(2)}(r,\alpha,t)\mathbf{1}_{[0,\alpha-\delta)}(t)\Big)d\alpha\Big]dr\\&\times B_1(R_1+D_1^\top\mathcal{P}_2(t)D_1)^{-1}B_1^\top\Big[P_{21}^{(1)}(\theta)^\top+P_{22}^{(1)}(\theta)^\top\mathbf{1}_{[0,\theta-\delta)}(t)+\int_t^T\Big(P_{21}^{(2)}(\theta,\alpha,t)^\top\\&+P_{22}^{(2)}(\theta,\alpha,t)^\top\mathbf{1}_{[0,\alpha-\delta)}(t)\Big)d\alpha\Big].\end{aligned}$$

Notice that

$$\begin{aligned}-P_{21}^{(2)}(\theta,t,t)^\top=&-A_1^\top\Big[P_{21}^{(1)}(\theta)^\top+P_{22}^{(2)}(\theta)^\top\mathbf{1}_{[0,\theta-\delta)}(t)\Big]-\int_t^T A_1^\top\Big[P_{21}^{(2)}(\theta,\alpha,t)^\top\\&+P_{22}^{(2)}(\theta,\alpha,t)^\top\mathbf{1}_{[0,\alpha-\delta)}(t)\Big]d\alpha+C_1^\top\mathcal{P}_2(t)D_1(R_1+D_1^\top\mathcal{P}_2(t)D_1)^{-1}B_1^\top\Big[P_{21}^{(1)}(\theta)^\top\\&+P_{22}^{(1)}(\theta)^\top\mathbf{1}_{[0,\theta-\delta)}(t)+\int_t^T\Big(P_{21}^{(2)}(\theta,\alpha,t)^\top+P_{22}^{(2)}(\theta,\alpha,t)^\top\mathbf{1}_{[0,\alpha-\delta)}(t)\Big)d\alpha\Big],\end{aligned}$$

and

$$
\begin{aligned}
&-P_{22}^{(2)}(\theta,t+\delta,t)^\top = -P_{22}^{(2)}(\theta,t+\delta,\theta)^\top + \int_t^\theta \dot{P}_{22}^{(2)}(t+\delta,\theta,s)ds \\
&= -\Big[P_{21}^{(1)}(t+\delta) + P_{22}^{(1)}(t+\delta)\mathbf{1}_{(\theta,\infty)}(t)\Big]A_2 - \int_\theta^T \Big(P_{21}^{(2)}(t+\delta,\alpha,\theta) + P_{22}^{(2)}(t+\delta,\alpha,\theta)\mathbf{1}_{(\theta+\delta,\infty)}(\alpha)\Big)A_2 d\alpha \\
&\quad + \Big\{P_{21}^{(1)}(t+\delta) + P_{22}^{(1)}(t+\delta)\mathbf{1}_{(\theta,\infty)}(t) + \int_\theta^T \Big(P_{21}^{(2)}(t+\delta,\alpha,\theta) + P_{22}^{(2)}(t+\delta,\alpha,\theta)\mathbf{1}_{(\theta+\delta,\infty)}(\alpha)\Big)d\alpha\Big\} \\
&\quad \times B_1(R_1 + D_1^\top \mathcal{P}_2(t)D_1)^{-1}D_1^\top \mathcal{P}_2(t)C_2 + \int_t^\theta \Big[P_{21}^{(1)}(t+\delta) + P_{22}^{(1)}(t+\delta)\mathbf{1}_{(s,\infty)}(t) \\
&\quad + \int_s^T \Big(P_{21}^{(2)}(t+\delta,\alpha,s) + P_{22}^{(2)}(t+\delta,\alpha,s)\mathbf{1}_{(s+\delta,\infty)}(\alpha)\Big)d\alpha\Big]B_1(R_1 + D_1^\top \mathcal{P}_2(t)D_1)^{-1}B_1^\top \Big[P_{21}^{(1)}(\theta)^\top \\
&\quad + P_{22}^{(1)}(\theta)^\top \mathbf{1}_{(t,\theta-\delta)}(s) + \int_s^T \Big(P_{21}^{(2)}(\theta,\alpha,s)^\top + P_{22}^{(2)}(\theta,\alpha,s)^\top \mathbf{1}_{(t,\alpha-\delta)}(s)\Big)d\alpha\Big]B_1 ds.
\end{aligned}
$$

Then, for $t \in (T-\delta, T]$, $\theta \in (t,T]$, with some calculations we deduce

$$
-\frac{\partial \mathcal{P}_3(t,\theta)}{\partial t} = A_1^\top \mathcal{P}_3(t,\theta) - \big(B_1^\top \mathcal{P}_2(t) + D_1^\top \mathcal{P}_2(t)C_1\big)^\top \big(R_1 + D_1^\top \mathcal{P}_2(t)D_1\big)^{-1}B_1^\top \mathcal{P}_3(t,\theta), \tag{A.55}
$$

and for $t \in [0, T-\delta]$, $\theta \in (t, t+\delta]$, we get

$$
\begin{aligned}
-\frac{\partial \mathcal{P}_3(t,\theta)}{\partial t} &= A_1^\top \mathcal{P}_3(t,\theta) - \big(B_1^\top \mathcal{P}_2(t) + D_1^\top \mathcal{P}_2(t)C_1\big)^\top \big(R_1 + D_1^\top \mathcal{P}_2(t)D_1\big)^{-1}B_1^\top \mathcal{P}_3(t,\theta) \\
&\quad + \mathcal{P}_3(\theta,t+\delta)^\top A_2 - (B_1^\top \mathcal{P}_3(\theta,t+\delta))^\top \big(R_1 + D_1^\top \mathcal{P}_2(t)D_1\big)^{-1}D_1\mathcal{P}_2(t)C_2 \\
&\quad - \int_t^\theta (B_1^\top \mathcal{P}_3(s,t+\delta))^\top \big(R_1 + D_1^\top \mathcal{P}_2(t)D_1\big)^{-1}B_1^\top \mathcal{P}_3(s,\theta)ds.
\end{aligned} \tag{A.56}
$$

For $\mathcal{P}_3(t,t)$, by (A.52) we have

$$
\begin{aligned}
\mathcal{P}_3(t,t) &= C_1^\top \mathcal{P}_2(t)C_2 - C_1^\top \mathcal{P}_2(t)D_1(R_1 + D_1^\top \mathcal{P}_2(t)D_1)^{-1}D_1^\top \mathcal{P}_2(t)C_2 \\
&\quad + \Big[A_2\mathcal{P}_2(t)^\top - C_2^\top \mathcal{P}_2(t)^\top D_1(R_1 + D_1^\top \mathcal{P}_2(t)D_1)^{-1}B_1^\top \mathcal{P}_2(t)\Big]^\top.
\end{aligned} \tag{A.57}
$$

Finally, (A.55)–(A.57) imply that the parts for $\mathcal{P}_3(\cdot,\cdot)$ in (4.12)–(4.13) hold, similarly we can also verify the parts for $\mathcal{P}_2(\cdot)$. In this case, (3.15)–(3.20) show that (4.14) is the optimal closed-loop outcome control. ∎