SIAM J. CONTROL OPTIM.
Vol. 61, No. 4, pp. 2595–2629

LINEAR-QUADRATIC OPTIMAL CONTROLS FOR STOCHASTIC VOLTERRA INTEGRAL EQUATIONS: CAUSAL STATE FEEDBACK AND PATH-DEPENDENT RICCATI EQUATIONS*

HANXIAO WANG†, JIONGMIN YONG‡, AND CHAO ZHOU§

Abstract. A linear-quadratic optimal control problem for a forward stochastic Volterra integral equation (FSVIE) is considered. Under the usual convexity conditions, open-loop optimal control exists, which can be characterized by the optimality system, a coupled system of an FSVIE and a type-II backward SVIE (BSVIE). To obtain a causal state feedback representation for the open-loop optimal control, a path-dependent Riccati equation for an operator-valued function is introduced, via which the optimality system can be decoupled. In the process of decoupling, a type-III BSVIE is introduced whose adapted solution can be used to represent the adapted M-solution of the corresponding type-II BSVIE. Under certain conditions, it is proved that the path-dependent Riccati equation admits a unique solution, which means that the decoupling field for the optimality system is found. Therefore, a causal state feedback representation of the open-loop optimal control is constructed. An additional interesting finding is that when the control only appears in the diffusion term, not in the drift term of the state system, the causal state feedback reduces to a Markovian state feedback.

Key words. linear-quadratic optimal control, stochastic Volterra integral equation, optimality system, coupled forward-backward stochastic Volterra integral equation, decoupling field, path-dependent Riccati equation, causal state feedback representation

MSC codes. 93E20, 49N10, 60H20, 45D05

DOI. 10.1137/22M1492696

1. Introduction. Let $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ be a complete filtered probability space, on which a one-dimensional standard Brownian motion W is defined, whose natural filtration augmented by all the $\mathbb{P}$-null sets in $\mathcal{F}$ is denoted by $\mathbb{F}\equiv\{\mathcal{F}_s\}_{s\ge0}$, and let $T>0$ be a fixed time horizon. For any $t\in[0,T)$, let $\mathbb{F}^t=\{\mathcal{F}^t_s\}_{s\ge0}$ with

$$\mathcal{F}^t_s=\begin{cases}\sigma\big(\{W(s)-W(t)\,\big|\,s\ge t\}\cup\{N\in\mathcal{F}\,\big|\,\mathbb{P}(N)=0\}\big), & s\in[t,T],\\ \sigma\big(\{N\in\mathcal{F}\,\big|\,\mathbb{P}(N)=0\}\big), & s\in[0,t).\end{cases}$$

Clearly, $\mathbb{F}^0=\mathbb{F}$. Next, let $\mathscr{X}_t=C([t,T];\mathbb{R}^n)=\big\{\mathbf{x}_t:[t,T]\to\mathbb{R}^n\,|\,\mathbf{x}_t(\cdot)\ \text{ is continuous}\big\}$, $t\in[0,T]$. For each $t\in[0,T]$, $\mathscr{X}_t$ is a Banach space (of deterministic continuous functions) under the norm $\|\mathbf{x}_t(\cdot)\|=\sup_{s\in[t,T]}|\mathbf{x}_t(s)|$. We will use $\mathbf{x}_t(\cdot)$ below to denote an element in $\mathscr{X}_t$ to emphasize the role played by t. For any $\mathbf{x}_0(\cdot)\in\mathscr{X}_0$, we

*Received by the editors April 26, 2022; accepted for publication (in revised form) April 14, 2023; published electronically August 18, 2023.
https://doi.org/10.1137/22M1492696
Funding: The first author is supported in part by NSFC grant 12201424 and Guangdong Basic and Applied Basic Research Foundation 2023A1515012104. The second author is supported in part by NSF grant DMS-2305475. The third author is supported by NSFC grant 11871364, as well as Singapore MOE AcRF grants R-146-000-271-112, A-8000453-00-00, A-0004273-00-00 and A-0004277-00-00.
†College of Mathematics and Statistics, Shenzhen University, Shenzhen 518060, China (hxwang@szu.edu.cn).
‡Department of Mathematics, University of Central Florida, Orlando, FL 32816 USA (jiongmin.yong@ucf.edu).
§Department of Mathematics and Risk Management Institute, National University of Singapore, Singapore 119076, Singapore (matzc@nus.edu.sg).

usually write $[\mathbf{x}_0]_t(\cdot)$ to represent $\mathbf{x}_0(s); s \in [t,T]$, the restriction of $\mathbf{x}_0(\cdot)$ on $[t,T]$, or simply as $\mathbf{x}_t(\cdot)$. Similarly, for any $\mathbf{x}_t(\cdot) \in \mathscr{X}_t$, we simply write $[\mathbf{x}_t]_0(\cdot)$ or $\mathbf{x}_0(\cdot)$ to represent $\mathbf{x}_t(s \vee t); s \in [0,T]$, the extension of $\mathbf{x}_t(\cdot)$ to $[0,T]$. Next, we introduce

$$\Lambda = \big\{(t, \mathbf{x}_t(\cdot)) \,|\, t \in [0,T),\, \mathbf{x}_t(\cdot) \in \mathscr{X}_t\big\}. \tag{1.1}$$

For any given *free pair* $(t, \mathbf{x}_t(\cdot)) \in \Lambda$, consider the following controlled linear forward stochastic Volterra integral equation (FSVIE) on $[t,T]$:

$$\begin{aligned} X(s) = \mathbf{x}_t(s) &+ \int_t^s \big[A(s,\tau)X(\tau) + B(s,\tau)u(\tau)\big]d\tau \\ &+ \int_t^s \big[C(s,\tau)X(\tau) + D(s,\tau)u(\tau)\big]dW(\tau), \qquad s \in [t,T], \end{aligned} \tag{1.2}$$

where $A, C : \Delta_*[0,T] \to \mathbb{R}^{n\times n}$, $B, D : \Delta_*[0,T] \to \mathbb{R}^{n\times m}$ are deterministic functions satisfying proper conditions, called the *coefficients* of the *state equation* (1.2). Here, $\Delta_*[0,T] \triangleq \{(s,r) \bigm| 0 \le r \le s \le T\}$ is the lower triangle domain. The process $u(\cdot)$ is called the *control process,* which belongs to the space

$$\mathscr{U}[t,T] = \Big\{ u : [t,T]\times\Omega \to \mathbb{R}^m \bigm| u(\cdot) \text{ is } \mathbb{F}^t\text{-progressively measurable}, \\ \mathbb{E}\int_t^T |u(\tau)|^2 d\tau < \infty \Big\},$$

and the corresponding solution $X(\cdot) \equiv X(\cdot\,; t, \mathbf{x}_t(\cdot), u(\cdot))$ of (1.2), which uniquely exists under some proper conditions on the coefficients, is called a *state process.* To measure the performance of the control $u(\cdot)$, we introduce the following cost functional:

$$\begin{aligned} J(t, \mathbf{x}_t(\cdot); u(\cdot)) = \frac{1}{2}\mathbb{E}_t\Big\{ &\int_t^T \big[\langle Q(\tau)X(\tau), X(\tau)\rangle + \langle R(\tau)u(\tau), u(\tau)\rangle\big] d\tau \\ &+ \langle GX(T), X(T)\rangle \Big\}, \end{aligned} \tag{1.3}$$

where $G \in \mathbb{S}^n$, the set of all $(n \times n)$ symmetric matrices; $Q : [0,T] \to \mathbb{S}^n$ and $R : [0,T] \to \mathbb{S}^m$ are deterministic functions. The problem that we are going to study can be stated as follows.

Problem (LQ-FSVIE). For any given free pair $(t, \mathbf{x}_t(\cdot)) \in \Lambda$, find a control $\bar u(\cdot) \in \mathscr{U}[t,T]$ such that

$$J(t, \mathbf{x}_t(\cdot); \bar u(\cdot)) \le J(t, \mathbf{x}_t(\cdot); u(\cdot)) \qquad \forall u(\cdot) \in \mathscr{U}[t,T]. \tag{1.4}$$

The above is called a *linear-quadratic* (LQ) *optimal control problem for FSVIEs.* Any $\bar u(\cdot) \in \mathscr{U}[t,T]$ satisfying (1.4) is called an *(open-loop) optimal control* of Problem (LQ-FSVIE) for the free pair $(t, \mathbf{x}_t(\cdot))$; the corresponding state process $\bar X(\cdot) \equiv X(\cdot\,; t, \mathbf{x}_t(\cdot), \bar u(\cdot))$ is called an *optimal state process*; and the function $V(\cdot\,,\cdot)$, defined by

$$V(t, \mathbf{x}_t(\cdot)) \triangleq \inf_{u(\cdot)\in\mathscr{U}[t,T]} J(t, \mathbf{x}_t(\cdot); u(\cdot)) \qquad \forall (t, \mathbf{x}_t(\cdot)) \in \Lambda, \tag{1.5}$$

is called the *value function* of Problem (LQ-FSVIE). We point out that the value function, $\mathbf{x}_t(\cdot) \mapsto V(t, \mathbf{x}_t(\cdot))$ is defined on the infinite dimensional Banach space $\mathscr{X}_t$ (of deterministic continuous functions on $[t,T]$). Moreover, by regarding Problem (LQ-FSVIE) as an optimization of the quadratic functional on $\mathscr{U}[t,T]$, with the parameter $(t, \mathbf{x}_t(\cdot))$, we expect that the value function has the following quadratic form:

$$V(t,\mathbf{x}_t(\cdot)) = \frac{1}{2}P(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big) \qquad \forall (t,\mathbf{x}_t(\cdot)) \in \Lambda. \tag{1.6}$$

Here, for any $t \in [0,T]$, $P(t)$ is a symmetric bilinear functional on $\mathscr{X}_t \times \mathscr{X}_t$.

When the coefficients $A(s,\tau)$, $B(s,\tau)$, $C(s,\tau)$, and $D(s,\tau)$ are independent of s and the free pair $(t,\mathbf{x}_t(\cdot)) \equiv (t,x)$, which is called an *initial pair,* for some $x \in \mathbb{R}^n$, Problem (LQ-FSVIE) reduces to a classical LQ optimal control for stochastic differential equations (SDEs), denoted by Problem (LQ-SDE). This has occupied a main part of the center stage for a long time in control theory. Since we prefer not to conduct a lengthy survey on the literature of Problem (LQ-SDE), let us just list some books [40, 4, 30, 31], where good surveys and tutorials along with extensive references (up to that time) can be found. It is well known by now that (see [40, Chapter 6], for example), Problem (LQ-SDE) can be solved by the following three steps in general: (i) By a variational method, the optimality system is derived, which is a coupled forward-backward SDE (FBSDE); (ii) the optimality system is decoupled by introducing the associated Riccati equation, which is solvable under certain conditions; and (iii) the optimal control is represented as a (Markovian) state feedback in terms of of the solution to the Riccati equation. This gives a very satisfactory solution to Problem (LQ-SDE) and provides a very good prototype of studying coupled FBSDEs as well. In recent years, FSVIEs have received more and more attention due to their applications in *rough volatility* models of mathematical finance; see, for example, Comte and Renault [6], Gatheral, Jaisson, and Rosenbaum [12], El Euch and Rosenbaum [9, 10], and Viens and Zhang [32]. In the control theory, the optimal control problem for general FSVIEs has been also widely studied, even before the above-mentioned literature of rough volatility appeared. Under the assumption that the control domain is convex, the maximum principle (MP) for FSVIEs was first established by Yong [39], in which the so-called *type-II backward stochastic Volterra integral equations* (BSVIEs) were introduced as the associated adjoint equations. See [3, 28, 37, 36, 14] for some further results on the MP for FSVIEs. More recently, the dynamical programming principle (DPP) for FSVIEs was first established by Viens and Zhang [32] by lifting the state space into the space of continuous functions. We point out that in [32] a functional Itô formula for FSVIEs was established, which serves as a fundamental tool in the current paper. On the other hand, for LQ problems of FSVIEs, namely, for Problem (LQ-FSVIE), some early stage of investigation can be found in Chen and Yong [5], Yong [39, section 5], and Wang [35], where the authors studied the corresponding MP under different assumptions, but the associated Riccati equation was not concerned at all. In Abi Jaber, Miller, and Pham [1, 2], the associated Riccati equation was derived, which, of course, brought some new insights into the LQ theory of FSVIEs. However, their problem was formulated only in the case of convolution form; that is, $\alpha(s,\tau) = \alpha(s-\tau)$ for $\alpha(\cdot\,,\cdot) = A(\cdot\,,\cdot), B(\cdot\,,\cdot), C(\cdot\,,\cdot), D(\cdot\,,\cdot)$, and one cannot directly extend the results in [1, 2] to the general LQ problems of FSVIEs, because of the limitation in their lifting methods.

We emphasize that in all the works mentioned above, the general question of *how to decouple the optimality system associated with Problem (LQ-FSVIE)* has not been touched. In other words, the crucial step (ii) in the standard path of solving Problem (LQ-SDE) mentioned above is completely missing for the case of FSVIEs. As a result, step (iii) for general FSVIEs does not have its foundation. In our opinion, this step (ii), which is essentially an *analytic* approach to the coupled system of linear FBSVIEs, is more important than solving Problem (LQ-FSVIE) itself, because on one hand, it links the Hamiltonian system and the (fully nonlinear) path-dependent HJB equation of controlled FSVIEs, and on the other hand, it provides some important prototypes for decoupling general coupled FBSVIEs. We refer the reader to Ma, Protter, and Yong [22], Delarue [7], Yong [38], Ma, et al. [23], and the books of Ma and Yong [24] and Zhang [41] for the related results in the SDE setting. From this point of view, we may also say that the main objective of the current paper is to explore the *decoupling method* of linear FBSVIEs (i.e., the famous *four-step scheme* of Ma, Protter, and Yong [22] for linear FBSVIEs), taking Problem (LQ-FSVIE) as a carrier, which needs a completely new creative method.

By a variational method and the duality principle of Yong [39], we can obtain the following *optimality system* associated with Problem (LQ-FSVIE) (see Theorem 3.3): Let $(\bar{X}(\cdot),\bar{u}(\cdot))$ be an open-loop optimal pair. Then

$$R(s)\bar{u}(s)+Y^0(s)=0,\qquad s\in[t,T], \tag{1.7}$$

with

$$\begin{cases}\bar{X}(s)=\mathbf{x}_t(s)+\displaystyle\int_t^s\big[A(s,\tau)\bar{X}(\tau)+B(s,\tau)\bar{u}(\tau)\big]d\tau\\ \qquad\qquad+\displaystyle\int_t^s\big[C(s,\tau)\bar{X}(\tau)+D(s,\tau)\bar{u}(\tau)\big]dW(\tau),\\ \eta(s)=G\bar{X}(T)-\displaystyle\int_s^T\zeta(\tau)dW(\tau),\\ Y(s)=Q(s)\bar{X}(s)+A(T,s)^\top G\bar{X}(T)+C(T,s)^\top\zeta(s)\\ \qquad\qquad+\displaystyle\int_s^T\big[A(\tau,s)^\top Y(\tau)+C(\tau,s)^\top Z(\tau,s)\big]d\tau-\int_s^T Z(s,\tau)dW(\tau),\\ Y^0(s)=B(T,s)^\top G\bar{X}(T)+D(T,s)^\top\zeta(s)\\ \qquad\qquad+\displaystyle\int_s^T\big[B(\tau,s)^\top Y(\tau)+D(\tau,s)^\top Z(\tau,s)\big]d\tau-\int_s^T Z^0(s,\tau)dW(\tau).\end{cases} \tag{1.8}$$

Basically, the above is the MP for Problem (LQ-FSVIE), that is, the step (i) mentioned above associated with Problem (LQ-FSVIE). Note that in the above (1.8), the second equation is a BSDE with unknown $(\eta(\cdot),\zeta(\cdot))$; the third is called a type-II BSVIE with unknown $(Y(\cdot),Z(\cdot\,,\cdot))$, whose main feature is that both $Z(s,\tau)$ and $Z(\tau,s)$ appear; the last equation is a type-I BSVIE with unknown $(Y^0(\cdot),Z^0(\cdot\,,\cdot))$, whose free term and drift term are known processes from the first three equations. Therefore, system (1.8) is a fully coupled FBSVIE, with the coupling being through (1.7).

It is not hard to see that the representation of the optimal control obtained from (1.7) (assuming $R(s)$ to be invertible for all $s\in[0,T]$) is not practically feasible. The reason is that in determining $Y^0(s)$, future information $\bar{X}(r);r\in[s,T]$ of the optimal state process $\bar{X}(\cdot)$ is involved. In the classical LQ theory (either for ODEs or SDEs), under proper conditions, the optimality system (which is a two-point boundary value problem for ODEs, or an FBSDE for SDEs) can be decoupled by the solution to a proper Riccati equation. Further, the state feedback representation of the open-loop optimal control can be obtained as a by-product. See [40, Chapter 6] for the standard LQ problems of ODEs and SDEs. The main tool used in the decoupling procedure for SDEs is the (classical) Itô formula (and the chain rule for ODEs). However, for FBSVIEs, the classical Itô formula is not applicable. By looking at the problem more deeply, one realizes that the decoupling technique essentially relies on the *flow* property of the equations, or some kind of *semigroup* property of the dynamic system.

Unfortunately, the controlled FSVIE in the optimality system does not satisfy the flow property in the standard sense. In fact, for $t \le r < s \le T$,

$$\bar{X}(s) \neq \bar{X}(r) + \int_r^s \big[A(s,\tau)\bar{X}(\tau) + B(s,\tau)\bar{u}(\tau)\big]d\tau + \int_r^s \big[C(s,\tau)\bar{X}(\tau) + D(s,\tau)\bar{u}(\tau)\big]dW(\tau), \tag{1.9}$$

due to which the usual decoupling method for SDEs (see Ma, Protter, and Yong [22], called the *four-step scheme*) cannot be applied here, and the corresponding Riccati-type equation, which plays the role of a *decoupling field* (see Delarue [7] and Ma et al. [23]), is completely unclear from this path. Indeed, this problem has been widely open for more than ten years (see [5, 39] for some early suggestions by the second author of this paper on the topic).

Recently, Viens and Zhang [32] and Wang, Yong, and Zhang [34] have developed a theory establishing some relations between decoupled type-I FBSVIEs and semilinear path-dependent PDEs, which are natural and significant extensions of the famous (path-dependent) *nonlinear Feynman–Kac formula* (see Pardoux and Peng [25], Peng and Wang [26], and Ekren et al. [8] for examples) to the FBSVIEs. These results provide some hope for our decoupling of the optimality system of Problem (LQ-FSVIE).

We now briefly describe the main clue of this paper. Let $(\bar{X}(\cdot), \bar{u}(\cdot))$ be an open-loop optimal pair. First, inspired by [32], we introduce the following auxiliary process $\bar{\mathcal{X}}(\cdot\,,\cdot)$ with two time variables:

$$\bar{\mathcal{X}}(s,r) = \mathbf{x}_t(s) + \int_t^r \big[A(s,\tau)\bar{X}(\tau) + B(s,\tau)\bar{u}(\tau)\big]d\tau + \int_t^r \big[C(s,\tau)\bar{X}(\tau) + D(s,\tau)\bar{u}(\tau)\big]dW(\tau), \qquad t \le r \le s \le T.$$

Then the flow property holds for the state process $\bar{X}(\cdot)$ in the following sense:

$$\bar{X}(s) = \bar{\mathcal{X}}(s,r) + \int_r^s \big[A(s,\tau)\bar{X}(\tau) + B(s,\tau)\bar{u}(\tau)\big]d\tau + \int_r^s \big[C(s,\tau)\bar{X}(\tau) + D(s,\tau)\bar{u}(\tau)\big]dW(\tau), \qquad t \le r \le s \le T.$$

It is worth pointing out that if r is the current time, then for $s \in [r,T]$, $s \mapsto \bar{\mathcal{X}}(s,r)$ only depends on the history $\{(\bar{X}(\tau), \bar{u}(\tau)) \,|\, \tau \in [t,r]\}$ of the optimal pair $(\bar{X}(\cdot), \bar{u}(\cdot))$ and no future information of the state and control is involved. Therefore, we say that $s \mapsto \bar{\mathcal{X}}(s,r)$ is *causal*. Next, we are trying to express the optimal control in the following manner:

$$\bar{u}(r) = \Theta(r)\bar{\mathcal{X}}(\cdot\,,r), \qquad r \in [t,T], \tag{1.10}$$

where $\Theta(r) : \mathscr{X}_r \to \mathbb{R}^m$ is a bounded linear operator which can be determined by the coefficients of the state equation and the weighting matrix functions in the cost functional. As $s \mapsto \bar{\mathcal{X}}(s,r)$ is causal, the above representation implies that the value $\bar{u}(r)$ of the optimal control $\bar{u}(\cdot)$ at current time r does not involve future information of the corresponding state process $\bar{X}(\cdot)$. Thus, we call the above (1.10) a *casual state feedback representation* of optimal control $\bar{u}(\cdot)$ (see Definition 2.5). Such a path is basically the analogue of steps (ii) and (iii) for solving classical LQ problems of ODEs and SDEs. The idea is pretty natural. But to achieve the goal, namely to determine the operator $\Theta(\cdot)$, is by no means trivial.

We now highlight the main contributions of this paper.

• We derive the (path-dependent) Riccati equation (4.10) for the bilinear operator-valued function $P(\cdot)$, through which the operator $\Theta(\cdot)$ can be determined. Note that if we mimicked the four-step scheme for FBSDEs (see [22, 24, 40]) trying to decouple the optimality system, we would encounter some difficulties that seem impossible to overcome. To get around this, we first make use of the above flow property and the functional Itô formula established in [32] to derive the path-dependent HJB equation for the value function $V(t, \mathbf{x}_t(\cdot))$ (see (4.5)), and from that we correctly identify the Riccati equation for the bilinear operator-valued function $P(\cdot)$, whose coefficients are path-dependent. In the process of deriving the (path-dependent) Riccati equation, a key point is that the value function at time $r \in [t, T]$ can be uniquely determined by the auxiliary process $\bar{\mathcal{X}}(\cdot, r)$ on $[r, T]$, without using the state process $\bar{X}(\cdot)|_{[0,r]}$. This is a new feature of our paper compared with [32, 34].

• We introduce a type-III BSVIE whose adapted solution can be used to represent the adapted M-solution of the type-II BSVIEs in the adjoint equation (see Proposition 3.5), which will play a crucial role in decoupling the optimality system via the solution of Riccati equation (4.10). By a type-III BSVIE, we mean a BSVIE that contains the diagonal value $Z(s, s)$ of $Z(\cdot, \cdot)$ in the drift. Such an equation was introduced by Wang and Yong [33] the first time and has been widely used in [16, 13, 17, 19, 18] while studying time-inconsistent optimal control problems. We coin the name of type-III BSVIEs (for the first time) here to distinguish this kind of equation from the other two types of BSVIEs. With such a relation, the optimality condition for Problem (LQ-FSVIE) can be characterized by a type-III BSVIE (see Theorem 3.6). From this, we are able to successfully decouple the optimality system (see Theorem 5.1) of Problem (LQ-FSVIE). The method is significantly different from that for Problem (LQ-SDE).

• We prove the existence and uniqueness of the *strongly regular solution* $P(\cdot)$ to the path-dependent Riccati equation (4.10) by an analytic method, under the following *standard condition*:

$$Q(s) \geq 0, \qquad R(s) \geq \lambda I_m, \qquad s \in [0, T]; \qquad G \geq 0, \tag{1.11}$$

where $\lambda > 0$ is a given constant (see Definition 4.1 and Theorem 6.1). It follows that the (fully nonlinear) path-dependent HJB equation (4.5) admits a unique classical solution, and the *decoupling field* of the optimality system really exists. Note that for any $t \in [0, T]$, $P(t)$ is a bilinear functional on $\mathscr{X}_t \times \mathscr{X}_t$, which is a Banach space depending on t (rather than a Hilbert space). This feature makes it different from the operator-valued Riccati equation derived from the LQ control problems for (stochastic) evolution equation (see [20, 21] for examples). Moreover, we see that the form of (4.10) is very similar to the classical stochastic Riccati equation (see (4.14)), except that the range of $P(\cdot)$ is not a Euclidean space. Needless to say, the form of (4.10) is more natural than the ones derived in [1, 2].

• An additional interesting finding is that when the drift term is not controlled, the causal state feedback representation of the optimal control will reduce to a *Markovian state feedback,* which means that the value $\bar{u}(s)$ of the optimal control $\bar{u}(\cdot)$ at current time s only depends on the current state $\bar{X}(s)$ (see Remark 5.3). Moreover, using the solution of the path-dependent Riccati equation, we can obtain a representation for $(\mathbb{E}_s[Y(\cdot)]\mathbf{1}_{[s,T]}(\cdot), Z(\cdot, s)\mathbf{1}_{[s,T]}(\cdot))$ in the dual space of $\mathscr{X}_s \equiv C([s, T]; \mathbb{R}^n)$ (see Theorem 5.5), by regarding it as a bounded linear functional on $\mathscr{X}_s$.

The rest of the paper is organized as follows. Section 2 collects some preliminary results. In section 3, the optimality system associated with Problem (LQ-FSVIE) is derived. We introduce the path-dependent Riccati equation in section 4 and establish the decoupling method for the optimality system in section 5. Finally, in section 6 the well-posedness of the Riccati equation is established.

2. Preliminaries. Throughout this paper, we let $T>0$ be a fixed time horizon and denote

$$\Delta_*[t,T]=\big\{(s,r)\,\big|\,t\le r\le s\le T\big\},\qquad \Delta^*[t,T]=\big\{(s,r)\,\big|\,t\le s\le r\le T\big\},$$

which are the lower and the upper triangle domains in $[t,T]^2$, respectively. For any Euclidean space $\mathbb{H}$ which could be $\mathbb{R}^{n\times m}$, or $\mathbb{S}^n$, and so on, we introduce the following spaces: For any $t\in[0,T)$ (with $\mathcal{B}([t,T])$ being the Borel σ-field of $[t,T]$),

$$\begin{aligned}
&L^\infty(t,T;\mathbb{H})=\big\{\varphi:[t,T]\to\mathbb{H}\,|\,\varphi(\cdot)\text{ is essentially bounded}\big\},\\
&L^2_{\mathcal{F}^t_T}(t,T;\mathbb{H})=\Big\{\varphi:[t,T]\times\Omega\to\mathbb{H}\,\Big|\,\varphi(\cdot)\text{ is }\mathcal{B}([t,T])\otimes\mathcal{F}^t_T\text{-measurable},\\
&\qquad\qquad\mathbb{E}\int_t^T|\varphi(\tau)|^2d\tau<\infty\Big\},\\
&L^2_{\mathbb{F}^t}(t,T;\mathbb{H})=\Big\{\varphi(\cdot)\in L^2_{\mathcal{F}^t_T}(t,T;\mathbb{H})\,\Big|\,\varphi(\cdot)\text{ is }\mathbb{F}^t\text{-progressively measurable on }[t,T]\Big\},\\
&L^2_{\mathbb{F}^t}(\Omega;C([t,T];\mathbb{H}))=\Big\{\varphi(\cdot)\in L^2_{\mathbb{F}^t}(t,T;\mathbb{H})\,\Big|\,\varphi(\cdot)\text{ has continuous paths},\\
&\qquad\qquad\mathbb{E}\Big[\sup_{s\in[t,T]}|\varphi(s)|^2\Big]<\infty\Big\},\\
&L^2_{\mathbb{F}^t}([t,T]^2;\mathbb{H})=\Big\{\varphi:[t,T]^2\times\Omega\to\mathbb{H}\,\Big|\,\varphi(s,\cdot)\in L^2_{\mathbb{F}^t}(t,T;\mathbb{H}),\text{ a.e. }s\in[t,T],\\
&\qquad\qquad\mathbb{E}\int_t^T\int_t^T|\varphi(s,\tau)|^2d\tau ds<\infty\Big\},\\
&L^2_{\mathbb{F}^t}(\Delta^*[t,T];\mathbb{H})=\Big\{\varphi:\Delta^*[t,T]\times\Omega\to\mathbb{H}\,\Big|\,\varphi(s,\cdot)\in L^2_{\mathbb{F}^t}(s,T;\mathbb{H}),\text{ a.e. }s\in[t,T],\\
&\qquad\qquad\operatorname{esssup}_{s\in[t,T]}\mathbb{E}\int_s^T|\varphi(s,\tau)|^2d\tau<\infty\Big\},\\
&L^2_{\mathbb{F}^t}(\Omega;C(\Delta_*[t,T];\mathbb{H}))=\Big\{\varphi:\Delta_*[t,T]\times\Omega\to\mathbb{H}\,\Big|\,\varphi(s,\cdot)\in L^2_{\mathbb{F}^t}(\Omega;C([t,s];\mathbb{H})),\\
&\qquad\qquad s\in[t,T],\text{ the map }(s,r)\mapsto\varphi(s,r)\text{ is jointly continuous},\\
&\qquad\qquad\mathbb{E}\Big[\sup_{(s,r)\in\Delta_*[t,T]}|\varphi(s,r)|^2\Big]<\infty\Big\}.
\end{aligned}$$

Similarly, one can define the spaces $L^\infty(\Delta^*[t,T];\mathbb{H})$ and $L^2_{\mathbb{F}^t}(\Omega;C([t,T]^2;\mathbb{H}))$.

2.1. Basic results for FSVIEs and BSVIEs. For the state equation (1.2) and the cost functional (1.3), we impose the following assumptions.

(H1) The coefficients $A,C:\Delta_*[0,T]\to\mathbb{R}^{n\times n}$ and $B,D:\Delta_*[0,T]\to\mathbb{R}^{n\times m}$ of the state equation (1.2) are (deterministic) bounded and partially differentiable with respect to the two variables, with bounded derivatives.

(H2) The weighting coefficients in the cost functional (1.3) satisfy

$$Q(\cdot)\in L^\infty(0,T;\mathbb{S}^n),\qquad R(\cdot)\in L^\infty(0,T;\mathbb{S}^m),\qquad G\in\mathbb{S}^n.$$

Applying the results of Ruan [27], we have the following result.

LEMMA 2.1. *Let* (H1) *hold. Then for any* $(t,\mathbf{x}_t(\cdot))\in\Lambda$ *and* $u(\cdot)\in\mathscr{U}[t,T]$, *state equation* (1.2) *admits a unique solution* $X(\cdot)\equiv X(\cdot\,;t,\mathbf{x}_t(\cdot),u(\cdot))\in L^2_{\mathbb{F}^t}(\Omega;C([t,T];\mathbb{R}^n))$. *Moreover, there exists a constant* $K>0$, *independent of* $(t,\mathbf{x}_t(\cdot))\in\Lambda$ *and* $u(\cdot)\in\mathscr{U}[t,T]$, *such that*

$$\mathbb{E}\left[\sup_{s\in[t,T]}|X(s)|^2\right]\le K\left[\sup_{s\in[t,T]}|\mathbf{x}_t(s)|^2+\mathbb{E}\int_t^T|u(s)|^2ds\right].\tag{2.1}$$

From Lemma 2.1, we see that Problem (LQ-FSVIE) is well formulated under the assumptions (H1) and (H2). The differentiability of $A(\cdot,\cdot)$, $B(\cdot,\cdot)$, $C(\cdot,\cdot)$, and $D(\cdot,\cdot)$ with respect to the first variable plays an important role in the proof of Lemma 2.1. The differentiability of $C(\cdot,\cdot)$ and $D(\cdot,\cdot)$ with respect to the second variable will be used in the proof of Proposition 3.5. Since we mainly focus on the form of the associated Riccati equation (or the decoupling field of the optimality system) in this paper, we prefer not to pursue the weakest possible conditions to simplify our presentation.

We now recall some fundamental results of the following type-II linear BSVIEs:

$$Y(s)=\psi(s)+\int_s^T\big[\mathcal{A}(s,\tau)Y(\tau)+\mathcal{C}(s,\tau)Z(\tau,s)\big]d\tau-\int_s^TZ(s,\tau)dW(\tau),\tag{2.2}$$

which can be found in Yong [39].

DEFINITION 2.2. *A pair of processes* $(Y(\cdot),Z(\cdot\,,\cdot))\in L^2_{\mathbb{F}^t}(t,T;\mathbb{R}^n)\times L^2_{\mathbb{F}^t}([t,T]^2;\mathbb{R}^n)$ *is called an* adapted M-solution *to BSVIE* (2.2) *if* (2.2) *is satisfied in the usual Itô sense for the Lebesgue measure almost every* $t\le s\le T$ *and, in addition, the following holds:*

$$Y(s)=\mathbb{E}_t[Y(s)]+\int_t^sZ(s,\tau)dW(\tau),\qquad t\le s\le T.\tag{2.3}$$

LEMMA 2.3. *Let* $\mathcal{A}(\cdot\,,\cdot),\mathcal{C}(\cdot\,,\cdot)\in L^\infty(\Delta^*[t,T];\mathbb{R}^{n\times n})$. *Then for any* $\psi(\cdot)\in L^2_{\mathcal{F}^t_T}(t,T;\mathbb{R}^n)$, *BSVIE* (2.2) *admits a unique adapted M-solution* $(Y(\cdot),Z(\cdot\,,\cdot))\in L^2_{\mathbb{F}^t}(t,T;\mathbb{R}^n)\times L^2_{\mathbb{F}^t}([t,T]^2;\mathbb{R}^n)$.

Let $\mathcal{L}(\mathscr{X}_t;\mathbb{R}^m)$ be the set of all bounded $\mathbb{R}^m$-valued linear functionals on $\mathscr{X}_t$, with the norm $\|\cdot\|_{\mathcal{L}}$ defined by

$$\|L\|_{\mathcal{L}}\triangleq\sup_{\|\mathbf{x}_t(\cdot)\|\le1}|L\mathbf{x}_t(\cdot)|\qquad\forall L\in\mathcal{L}(\mathscr{X}_t;\mathbb{R}^m).$$

Let $L^\infty(0,T;\mathcal{L}(\mathscr{X}_s;\mathbb{R}^m))$ be the set of all bounded $\mathcal{L}(\mathscr{X}_s;\mathbb{R}^m)$-valued functions defined on $[0,T]$. In other words, for any $L(\cdot)\in L^\infty(0,T;\mathcal{L}(\mathscr{X}_s;\mathbb{R}^m))$,

$$L(s)\in\mathcal{L}(\mathscr{X}_s;\mathbb{R}^m),\quad \text{a.e. } s\in[0,T],\quad \text{and}\quad \operatorname{esssup}_{s\in[0,T]}\|L(s)\|_{\mathcal{L}}<\infty.$$

For any $\Theta(\cdot)\in L^\infty(0,T;\mathcal{L}(\mathscr{X}_s;\mathbb{R}^m))$, we consider the closed-loop auxiliary system,

$$\begin{aligned}\mathcal{X}(s,r)&=\mathbf{x}_t(s)+\int_t^r\big[A(s,\tau)X(\tau)+B(s,\tau)\Theta(\tau)\mathcal{X}(\cdot\,,\tau)\big]d\tau\\&\quad+\int_t^r\big[C(s,\tau)X(\tau)+D(s,\tau)\Theta(\tau)\mathcal{X}(\cdot\,,\tau)\big]dW(\tau),\qquad(s,r)\in\Delta_*[t,T],\end{aligned}\tag{2.4}$$

and the closed-loop system,

$$
\begin{aligned}
X(s) = \mathbf{x}_t(s) &+ \int_t^s \big[A(s,\tau)X(\tau) + B(s,\tau)\Theta(\tau)\mathcal{X}(\cdot\,,\tau)\big]d\tau \\
&+ \int_t^s \big[C(s,\tau)X(\tau) + D(s,\tau)\Theta(\tau)\mathcal{X}(\cdot\,,\tau)\big]dW(\tau), \qquad s\in[t,T].
\end{aligned} \tag{2.5}
$$

We have the following well-posedness results of the above systems.

PROPOSITION 2.4. *Let* (H1) *hold. Then for any* $(t,\mathbf{x}_t(\cdot))\in\Lambda$ *and* $\Theta(\cdot)\in L^\infty(0,T;\mathcal{L}(\mathscr{X}_s;\mathbb{R}^m))$*, the auxiliary closed-loop system* (2.4) *and the closed-loop system* (2.5) *admit unique solutions* $\mathcal{X}(\cdot,\cdot)\in L^2_{\mathbb{F}^t}(\Omega;C(\Delta_*[t,T];\mathbb{R}^n))$ *and* $X(\cdot)\in L^2_{\mathbb{F}^t}(\Omega;C([t,T];\mathbb{R}^n))$*, respectively. Moreover, the relation* $\mathcal{X}(s,s)=X(s); s\in[t,T]$ *holds.*

Proof. Let us consider the following auxiliary systems with adapted solution $\tilde{\mathcal{X}}(\cdot,\cdot)$ on $[t,T]\times[t,T]$:

$$
\begin{aligned}
\tilde{\mathcal{X}}(s,r) = \mathbf{x}_t(s) &+ \int_t^r \big[A(s\vee\tau,\tau)\tilde{\mathcal{X}}(\tau,\tau) + B(s\vee\tau,\tau)\Theta(\tau)\tilde{\mathcal{X}}(\cdot\,,\tau)\big]d\tau \\
&+ \int_t^r \big[C(s\vee\tau,\tau)\tilde{\mathcal{X}}(\tau,\tau) + D(s\vee\tau,\tau)\Theta(\tau)\tilde{\mathcal{X}}(\cdot\,,\tau)\big]dW(\tau).
\end{aligned} \tag{2.6}
$$

Then to obtain the desired results, it is sufficient to show the above system admits a unique solution in $L^2_{\mathbb{F}^t}(\Omega;C([t,T]^2;\mathbb{R}^n))$.

For any $S\in(t,T]$ and $x(\cdot,\cdot)\in L^2_{\mathbb{F}^t}(\Omega;C([t,T]\times[t,S];\mathbb{R}^n))$, we can uniquely define the process $\mathcal{X}(\cdot,\cdot)$ on $[t,T]\times[t,S]$ as follows:

$$
\begin{aligned}
\mathcal{X}(s,r) = \mathbf{x}_t(s) &+ \int_t^r \big[A(s\vee\tau,\tau)x(\tau,\tau) + B(s\vee\tau,\tau)\Theta(\tau)x(\cdot\,,\tau)\big]d\tau \\
&+ \int_t^r \big[C(s\vee\tau,\tau)x(\tau,\tau) + D(s\vee\tau,\tau)\Theta(\tau)x(\cdot\,,\tau)\big]dW(\tau).
\end{aligned}
$$

Denote $\hat{\mathcal{X}}(s,r)=\mathcal{X}(s,r)-\mathbf{x}_t(s); (s,r)\in[t,T]\times[t,S]$. For any fixed $s',s''\in[t,T]$, by the B-D-G inequality, we have

$$
\mathbb{E}\left[\sup_{r\in[t,S]}|\hat{\mathcal{X}}(s',r)-\hat{\mathcal{X}}(s'',r)|^2\right] \le K(S-t)\mathbb{E}\left[\sup_{(s,r)\in[t,T]\times[t,S]}|x(s,r)|^2\right]|s'-s''|^2.
$$

Then following the proof of Kolmogorov's criterion theorem (see [11, Theorem 3.1], for example), for any given $\alpha\in(0,\frac{1}{2})$, there exists a square-integrable random variable $C_\alpha(\omega)>0$ such that

$$
\sup_{r\in[t,S]}|\hat{\mathcal{X}}(s',r)-\hat{\mathcal{X}}(s'',r)| \le C_\alpha(\omega)|s'-s''|^\alpha \qquad \forall s',s''\in\bigcup_n D_n, \text{ a.e. } \omega\in\Omega,
$$

where for any positive integer n, $D_n \triangleq \{t+k\frac{T-t}{2^n}\,|\,k=0,1,2,\dots,2^n\}$. For any $s\in[t,T]$, there exists a sequence $\{s_n\}\subseteq\bigcup_n D_n$ such that $s_n\to s$. Clearly, $\{\hat{\mathcal{X}}(s_n,\cdot)\}$ is a Cauchy sequence in $C([t,S];\mathbb{R}^n)$ almost surely. We denote the limit of $\{\hat{\mathcal{X}}(s_n,\cdot)\}$ as $n\to\infty$ by $\check{\mathcal{X}}(s,\cdot)$, which is independent of the choice of $\{s_n\}$. Since $\lim_{n\to\infty}\mathbb{E}[\sup_{r\in[t,S]}|\hat{\mathcal{X}}(s_n,r)-\hat{\mathcal{X}}(s,r)|^2]=0$, we know that $\check{\mathcal{X}}(s,\cdot)=\hat{\mathcal{X}}(s,\cdot)$ almost surely. We always consider the version $\check{\mathcal{X}}(s,\cdot)$ of $\hat{\mathcal{X}}(s,\cdot)$ and still denote it by $\hat{\mathcal{X}}(s,\cdot)$. Then from the fact $\lim_{n\to\infty}\sup_{r\in[t,S]}|\hat{\mathcal{X}}(s_n,r)-\hat{\mathcal{X}}(s,r)|=0$, we have

$$
\sup_{r\in[t,S]}|\hat{\mathcal{X}}(s',r)-\hat{\mathcal{X}}(s'',r)| \le C_\alpha(\omega)|s'-s''|^\alpha \qquad \forall s',s''\in[t,T], \text{ a.e. } \omega\in\Omega.
$$

Note that for any fixed $s \in [t,T]$, there exists an $\Omega^s \subseteq \Omega$ with $\mathbb{P}(\Omega^s)=1$ such that $r \mapsto \hat{\mathcal{X}}(s,r)(\omega)$ is continuous for any $\omega \in \Omega^s$. By the arguments employed in the proof of [34, Proposition 2.4], noting that $\bigcup_n D_n$ is countable and dense in $[t,T]$ and the definition of $\hat{\mathcal{X}}(s,\cdot)$ for $s \in [t,T]\backslash\big(\bigcup_n D_n\big)$, we can show that there exists an $\Omega_1 \subseteq \Omega$ with $\mathbb{P}(\Omega_1)=1$ such that for any $s \in [t,T]$ and $\omega \in \Omega_1$, $r \mapsto \hat{\mathcal{X}}(s,r)(\omega)$ is continuous. It follows that $\hat{\mathcal{X}}(s,r)$ is jointly continuous in (s,r) almost surely. By the Kolmogorov criterion theorem again (also see [32, Lemma 3.11]), we have

$$\begin{aligned}
&\mathbb{E}\left[\sup_{(s,r)\in[t,T]\times[t,S]} |\hat{\mathcal{X}}(s,r)|^2\right] \\
&\quad \le K\mathbb{E}\left[\sup_{r\in[t,S]} |\hat{\mathcal{X}}(t,r)|^2\right] + K(S-t)\mathbb{E}\left[\sup_{(s,r)\in[t,T]\times[t,S]} |x(s,r)|^2\right] \\
&\quad \le K(S-t)\mathbb{E}\left[\sup_{(s,r)\in[t,T]\times[t,S]} |x(s,r)|^2\right].
\end{aligned} \tag{2.7}$$

Then from $\mathcal{X}(s,r) = \hat{\mathcal{X}}(s,r) + \mathbf{x}_t(s)$, we get $\mathcal{X}(\cdot,\cdot) \in L^2_{\mathbb{F}^t}(\Omega; C([t,T]\times[t,S];\mathbb{R}^n))$. Thus, the map $\Gamma : L^2_{\mathbb{F}^t}(\Omega; C([t,T]\times[t,S];\mathbb{R}^n)) \to L^2_{\mathbb{F}^t}(\Omega; C([t,T]\times[t,S];\mathbb{R}^n))$, given by $\mathcal{X}(\cdot,\cdot) = \Gamma\big(x(\cdot,\cdot)\big)$, is well defined. For any $x_1(\cdot,\cdot),\ x_2(\cdot,\cdot) \in L^2_{\mathbb{F}^t}(\Omega; C([t,T]\times[t,S];\mathbb{R}^n))$, denote

$$\mathcal{X}_i(\cdot\cdot) = \Gamma(x_i(\cdot,\cdot)), \qquad \Delta\mathcal{X}(\cdot,\cdot) = \mathcal{X}_1(\cdot,\cdot) - \mathcal{X}_2(\cdot,\cdot), \qquad \Delta x(\cdot,\cdot) = x_1(\cdot,\cdot) - x_2(\cdot,\cdot).$$

Then

$$\begin{aligned}
\Delta\mathcal{X}(s,r) = &\int_t^r \big[A(s\vee\tau,\tau)\Delta x(\tau,\tau) + B(s\vee\tau,\tau)\Theta(\tau)\Delta x(\cdot,\tau)\big]d\tau \\
&+ \int_t^r \big[C(s\vee\tau,\tau)\Delta x(\tau,\tau) + D(s\vee\tau,\tau)\Theta(\tau)\Delta x(\cdot,\tau)\big]dW(\tau).
\end{aligned}$$

Similar to (2.7), we have

$$\mathbb{E}\left[\sup_{(s,r)\in[t,T]\times[t,S]} |\Delta\mathcal{X}(s,r)|^2\right] \le K(S-t)\mathbb{E}\left[\sup_{(s,r)\in[t,T]\times[t,S]} |\Delta x(s,r)|^2\right].$$

Thus, when $S-t>0$ is small, the map $\Gamma : L^2_{\mathbb{F}^t}(\Omega; C([t,T]\times[t,S];\mathbb{R}^n)) \to L^2_{\mathbb{F}^t}(\Omega; C([t,T]\times[t,S];\mathbb{R}^n))$ is a contraction. Hence, it admits a unique fixed point $\bar{\mathcal{X}}(\cdot,\cdot) \in L^2_{\mathbb{F}^t}(\Omega; C([t,T]\times[t,S];\mathbb{R}^n))$ which is the unique solution of (2.6) over $[t,T]\times[t,S]$. Next, we consider the following auxiliary systems on $[t,T]\times[S,T]$:

$$\begin{aligned}
\tilde{\mathcal{X}}(s,r) = \bar{\mathcal{X}}(s,S) &+ \int_S^r \big[A(s\vee\tau,\tau)\tilde{\mathcal{X}}(\tau,\tau) + B(s\vee\tau,\tau)\Theta(\tau)\tilde{\mathcal{X}}(\cdot,\tau)\big]d\tau \\
&+ \int_S^r \big[C(s\vee\tau,\tau)\tilde{\mathcal{X}}(\tau,\tau) + D(s\vee\tau,\tau)\Theta(\tau)\tilde{\mathcal{X}}(\cdot,\tau)\big]dW(\tau).
\end{aligned}$$

By continuing the above arguments, we can show that (2.6) admits a unique solution over $[t,T]\times[t,T]$. ∎

DEFINITION 2.5. *Any* $\bar{\Theta}(\cdot) \in L^\infty(0,T;\mathcal{L}(\mathscr{X}_s;\mathbb{R}^m))$ *is called an* optimal causal feedback operator *of Problem (LQ-FSVIE) if*

$$J(t,\mathbf{x}_t(\cdot);\bar{\Theta}(\cdot)\bar{\mathcal{X}}(\cdot,\cdot)) \le J(t,\mathbf{x}_t(\cdot);u(\cdot)) \quad \forall u(\cdot)\in\mathscr{U}[t,T] \quad \forall (t,\mathbf{x}_t)\in\Lambda, \tag{2.8}$$

where $(s,r)\mapsto\bar{\mathcal{X}}(s,r)\equiv\bar{\mathcal{X}}_r(s)$ *and* $s\mapsto\bar{X}(s)$ *are the unique solutions to* (2.4) *and* (2.5)*, with the operator-valued function* $\bar{\Theta}(\cdot)$*, respectively.*

Remark 2.6. It is clear that the outcome $\tau\mapsto\bar{u}(\tau)=\bar{\Theta}(\tau)\bar{\mathcal{X}}(\cdot\,,\tau)$ of the optimal causal state feedback operator $\bar{\Theta}(\cdot)$ is an open-loop optimal control. Note that the auxiliary process $\bar{\mathcal{X}}_r(\cdot)=\bar{\mathcal{X}}(\cdot\,,r)$ is uniquely determined by the portion $\bar{X}(l);l\in[t,r]$ of the state process $\bar{X}(\cdot)$. Thus, $\bar{u}(\cdot)$ has a causal state feedback representation. Namely, the value $\bar{u}(r)$ of the optimal control $\bar{u}(\cdot)$ at any time r does not involve future information of the corresponding state process $\bar{X}(\cdot)$.

Remark 2.7. In the proof of the well-posedness of the closed-loop auxiliary system (2.4), we use the property that the map $(\tau,\omega)\mapsto\Theta(\tau)\mathcal{X}(\cdot\,,\tau)$ is progressively measurable for any $\mathcal{X}(\cdot,\cdot)\in L^2_{\mathbb{F}}(\Omega;C([t,T]^2;\mathbb{R}^n))$, without imposing the specific measurable condition for the operator-valued function $\Theta(\cdot)$. We always assume this property holds for any $\Theta(\cdot)\in L^\infty(0,T;\mathcal{L}(\mathscr{X}_t;\mathbb{R}^m))$. It can be easily checked that the optimal causal feedback operator obtained in the paper satisfies such a property.

2.2. Bilinear operators. For any Banach space $\mathbb{X}$, let $\mathcal{L}^2(\mathbb{X})$ be the set of all bounded bilinear functionals on $\mathbb{X}\times\mathbb{X}$, with the norm $\|\cdot\|_{\mathcal{L}^2}$ defined by

$$\|P\|_{\mathcal{L}^2}\triangleq\sup_{\|\mathbf{x}\|,\|\mathbf{y}\|\le 1}|P(\mathbf{x},\mathbf{y})|\qquad\forall P\in\mathcal{L}^2(\mathbb{X}).$$

We denote

$$P(\mathbf{x},\mathbf{y})=\langle P\mathbf{x},\mathbf{y}\rangle=\mathbf{y}^\top P\mathbf{x}\qquad\forall\mathbf{x},\mathbf{y}\in\mathbb{X},$$

where $P\mathbf{x}\in\mathbb{X}^*$, the dual of $\mathbb{X}$, and $\langle\cdot\,,\cdot\rangle$ is the duality pairing between $\mathbb{X}^*$ and $\mathbb{X}$. A bilinear functional $P\in\mathcal{L}^2(\mathbb{X})$ is said to be *symmetric* if it satisfies

$$\mathbf{y}^\top P\mathbf{x}=P(\mathbf{x},\mathbf{y})=P(\mathbf{y},\mathbf{x})=\mathbf{x}^\top P\mathbf{y}\qquad\forall\mathbf{x},\mathbf{y}\in\mathbb{X}.$$

Let $\mathcal{S}(\mathbb{X})$ be the set of all symmetric bounded bilinear functionals on $\mathbb{X}$, and $\mathcal{S}_+(\mathbb{X})$ be the subset of $\mathcal{S}(\mathbb{X})$ consisting of all nonnegative bilinear functionals; that is, $P\in\mathcal{S}_+(\mathbb{X})$ if and only if $P\in\mathcal{S}(\mathbb{X})$ and

$$P(\mathbf{x},\mathbf{x})\ge 0\qquad\forall\mathbf{x}\in\mathbb{X}.$$

For any $P\in\mathcal{S}(\mathbb{X})$, define

$$\|P\|_{\mathcal{S}}\triangleq\sup_{\|\mathbf{x}\|\le 1}|P(\mathbf{x},\mathbf{x})|.$$

The following result shows that $\|\cdot\|_{\mathcal{S}}$ is an equivalent norm of $\|\cdot\|_{\mathcal{L}^2}$ on $\mathcal{S}(\mathbb{X})$.

LEMMA 2.8. *For any* $P\in\mathcal{S}(\mathbb{X})$*, it holds that*

$$\|P\|_{\mathcal{S}}\le\|P\|_{\mathcal{L}^2}\le 2\|P\|_{\mathcal{S}}.\tag{2.9}$$

Proof. Since the bilinear functional P satisfies $P(\mathbf{x},\mathbf{y})=P(\mathbf{y},\mathbf{x})$, we have

$$P(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y})-P(\mathbf{x}-\mathbf{y},\mathbf{x}-\mathbf{y})=2P(\mathbf{x},\mathbf{y})+2P(\mathbf{y},\mathbf{x})=4P(\mathbf{x},\mathbf{y})\qquad\forall\mathbf{x},\mathbf{y}\in\mathbb{X}.$$

Thus

$$\begin{aligned}|P(\mathbf{x},\mathbf{y})|&=\frac{|P(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y})-P(\mathbf{x}-\mathbf{y},\mathbf{x}-\mathbf{y})|}{4}\le\|P\|_{\mathcal{S}}\frac{\|\mathbf{x}+\mathbf{y}\|^2+\|\mathbf{x}-\mathbf{y}\|^2}{4}\\&\le\|P\|_{\mathcal{S}}[\|\mathbf{x}\|^2+\|\mathbf{y}\|^2]\qquad\forall\mathbf{x},\mathbf{y}\in\mathbb{X},\end{aligned}$$

which implies

$$\|P\|_{\mathcal{L}^2} = \sup_{\|\mathbf{x}\|,\|\mathbf{y}\|\le 1} |P(\mathbf{x},\mathbf{y})| \le \|P\|_{\mathcal{S}} \sup_{\|\mathbf{x}\|,\|\mathbf{y}\|\le 1} [\|\mathbf{x}|^2 + |\mathbf{y}\|^2] = 2\|P\|_{\mathcal{S}}.$$

Combining the above with the fact $\|P\|_{\mathcal{S}} \le \|P\|_{\mathcal{L}^2}$, we get the conclusion. ◻

Remark 2.9. Since $\mathbb{X}$ is merely a Banach space, it is not expected that $\|\cdot\|_{\mathcal{S}} = \|\cdot\|_{\mathcal{L}^2}$ holds in general. However, we still have the equivalence between $\|\cdot\|_{\mathcal{S}}$ and $\|\cdot\|_{\mathcal{L}^2}$ on $\mathcal{S}(\mathbb{X})$. This equivalence will play a crucial role in establishing the well-posedness of the Riccati equation associated with Problem (LQ-FSVIE). When there is no confusion, we often simply write $\|\cdot\|_{\mathcal{S}}$ (or $\|\cdot\|_{\mathcal{L}^2}$) as $\|\cdot\|$.

2.3. Functional Itô formula. Recall Λ from (1.1). As in Viens and Zhang [32], we introduce the following metric:

$$\mathbf{d}\big((t,\mathbf{x}_t(\cdot)),(t',\mathbf{x}'_{t'}(\cdot))\big) \triangleq |t-t'| + \sup_{s\in[0,T]} |\mathbf{x}_t(s\vee t) - \mathbf{x}'_{t'}(s\vee t')| \qquad \forall (t,\mathbf{x}_t(\cdot)),(t',\mathbf{x}'_{t'}(\cdot)) \in \Lambda.$$

It can be shown that Λ is a complete metric space under $\mathbf{d}$. Let $C^0(\Lambda)$ denote the set of all functions $v:\Lambda\to\mathbb{R}$ which are continuous under $\mathbf{d}$. For any $v(\cdot,\cdot)\in C^0(\Lambda)$ and given $(t,\mathbf{x}_t(\cdot))\in\Lambda$, $v(t,\mathbf{x}_t(\cdot))$ takes real values. We denote $v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot))$ to be the Fréchet derivative of $v(t,\mathbf{x}_t(\cdot))$ with respect to $\mathbf{x}_t(\cdot)$. Namely, $v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot)) : \mathscr{X}_t \to \mathbb{R}$ is the linear functional satisfying the following:

$$v(t,\mathbf{x}_t(\cdot)+\eta_t(\cdot)) - v(t,\mathbf{x}_t(\cdot)) = v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot))(\eta_t(\cdot)) + o(\|\eta_t(\cdot)\|) \qquad \forall \eta_t(\cdot)\in\mathscr{X}_t. \tag{2.10}$$

It is clear that the above Fréchet derivative can be calculated in the following way, which defines the Gâteaux derivative:

$$v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot))(\eta_t(\cdot)) = \lim_{\varepsilon\to 0} \frac{v(t,\mathbf{x}_t(\cdot)+\varepsilon\eta_t(\cdot)) - v(t,\mathbf{x}_t(\cdot))}{\varepsilon} \qquad \forall \eta_t(\cdot)\in\mathscr{X}_t. \tag{2.11}$$

Similarly, we define the second order derivative $v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot))$ as a bilinear functional on $\mathscr{X}_t\times\mathscr{X}_t$:

$$\begin{aligned} &v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot)+\eta_t(\cdot))(\eta'_t(\cdot)) - v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot))(\eta'_t(\cdot)) \\ &\quad = v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot))(\eta_t(\cdot),\eta'_t(\cdot)) + o(\|\eta_t\|) \qquad \forall \eta_t(\cdot),\eta'_t(\cdot)\in\mathscr{X}_t. \end{aligned} \tag{2.12}$$

To define the right t-partial derivative $v_t(t,\mathbf{x}_t(\cdot))$ of $v(\cdot,\cdot)$ at $(t,\mathbf{x}_t(\cdot))$, we need to "fix" $\mathbf{x}_t(\cdot)$ and define $v(t+\varepsilon,\mathbf{x}_t(\cdot))$. Naturally, we define

$$v(t+\varepsilon,\mathbf{x}_t(\cdot)) = v(t+\varepsilon,\mathbf{x}_{t+\varepsilon}(\cdot)) = v(t+\varepsilon,[\mathbf{x}_t]_{t+\varepsilon}(\cdot)) \qquad \forall \mathbf{x}_t(\cdot)\in\mathscr{X}_t, \tag{2.13}$$

where

$$\mathbf{x}_{t+\varepsilon}(s) = [\mathbf{x}_t]_{t+\varepsilon}(s) = \mathbf{x}_t(s), \qquad s\in[t+\varepsilon,T], \qquad \forall \mathbf{x}_t(\cdot)\in\mathscr{X}_t. \tag{2.14}$$

According to the above, we see that

$$\mathbf{x}_{t+\varepsilon}(\cdot) = [\mathbf{x}_t]_{t+\varepsilon}(\cdot) \in \mathscr{X}_{t+\varepsilon} \qquad \forall \mathbf{x}_t(\cdot)\in\mathscr{X}_t,$$

which can be regarded as the natural "projection" from $\mathscr{X}_t$ to $\mathscr{X}_{t+\varepsilon}$. Thus, (2.13) makes sense. Having such a natural restriction, we can define the right t-partial derivative $v_t(t,\mathbf{x}_t(\cdot))$ as follows:

$$v_t(t,\mathbf{x}_t(\cdot)) \triangleq \lim_{\varepsilon\to 0^+} \frac{v(t+\varepsilon,\mathbf{x}_t(\cdot)) - v(t,\mathbf{x}_t(\cdot))}{\varepsilon} \equiv \lim_{\varepsilon\to 0^+} \frac{v(t+\varepsilon,[\mathbf{x}_t]_{t+\varepsilon}(\cdot)) - v(t,\mathbf{x}_t(\cdot))}{\varepsilon}, \tag{2.15}$$

provided the limit exists. Let us introduce the following spaces.

DEFINITION 2.10. (i) *Let $C^{1,2}(\Lambda)$ be the set of all $v(\cdot,\cdot)\in C^0(\Lambda)$ such that $v_t(\cdot,\cdot), v_{\mathbf{x}}(\cdot,\cdot), v_{\mathbf{xx}}(\cdot,\cdot)$ exist on Λ.* (ii) *Let $C^{1,2}_+(\Lambda)$ denote the set of all $v(\cdot,\cdot)\in C^{1,2}(\Lambda)$ such that the following are satisfied:*

(a) *There exist constants $\kappa, K>0$ such that, for any $(t,\mathbf{x}_t(\cdot))\in\Lambda$,*

$$|v_t(t,\mathbf{x}_t(\cdot))| + \sup_{\substack{\eta_t(\cdot)\in\mathscr{X}_t\\ \|\eta_t(\cdot)\|\le 1}} |v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot))(\eta_t(\cdot))| + \sup_{\substack{\eta_t(\cdot),\eta_t'(\cdot)\in\mathscr{X}_t\\ \|\eta_t(\cdot)\|,\|\eta_t'(\cdot)\|\le 1}} |v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot))(\eta_t(\cdot),\eta_t'(\cdot))| \le K[1+\|\mathbf{x}_t(\cdot)\|^\kappa].$$

(b) *For any $\eta(\cdot),\eta'(\cdot)\in\mathscr{X}_0$, $v_t(t,\mathbf{x}_t(\cdot))$, $v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot))(\eta_t(\cdot))$, $v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot))(\eta_t(\cdot),\eta_t'(\cdot))$ are continuous in $(t,\mathbf{x}_t(\cdot))$, where the continuity in t always means right-continuity, $\eta_t(\cdot)=[\eta(\cdot)]_t(\cdot)$, and $\eta_t'(\cdot)=[\eta'(\cdot)]_t(\cdot)$.*

(c) *There exist $\kappa>0$ and a modulus of continuity $\rho(\cdot)$ such that*

$$\begin{aligned}&\big|[v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot)) - v_{\mathbf{xx}}(t,\mathbf{x}_t'(\cdot))](\eta_t(\cdot),\eta_t(\cdot))\big|\\ &\quad\le \big[1+\|\mathbf{x}_t(\cdot)\|^\kappa + \|\mathbf{x}_t'(\cdot)\|^\kappa\big]\|\eta_t(\cdot)\|^2\rho(\|\mathbf{x}_t(\cdot)-\mathbf{x}_t'(\cdot)\|)\\ &\quad\ \forall \mathbf{x}_t(\cdot),\mathbf{x}_t'(\cdot),\eta_t(\cdot)\in\mathscr{X}_t,\ t\in[0,T].\end{aligned}$$

The following is a version of the functional Itô formula for SVIEs found in Viens and Zhang [32].

PROPOSITION 2.11. *Let $b,\sigma:\Delta_*[0,T]\times\Omega\to\mathbb{R}^n$ be measurable such that $s\mapsto(b(s,\tau),\sigma(s,\tau))$ is differential with bounded derivatives, $\tau\mapsto(b(s,\tau),\sigma(s,\tau))$ is $\mathbb{F}^t$-progressively measurable on $[t,T]$, and*

$$\mathbb{E}\left[\int_t^T \sup_{s\in[\tau,T]}\Big(|b(s,\tau)|^2+|\sigma(s,\tau)|^2\Big)d\tau\right]<\infty.$$

Let

$$\mathcal{X}(s,r)=\mathbf{x}_t(s)+\int_t^r b(s,\tau)d\tau+\int_t^r\sigma(s,\tau)dW(\tau),\qquad t\le r\le s\le T, \tag{2.16}$$

and $v(\cdot,\cdot)\in C^{1,2}_+(\Lambda)$. Then the following functional Itô formula holds:

$$\begin{aligned}dv(r,\mathcal{X}(\cdot,r)) = \Big[&v_r(r,\mathcal{X}(\cdot,r))+\frac12 v_{\mathbf{xx}}(r,\mathcal{X}(\cdot,r))\big(\sigma(\cdot,r),\sigma(\cdot,r)\big)\\ &+v_{\mathbf{x}}(r,\mathcal{X}(\cdot,r))b(\cdot,r)\Big]dr + v_{\mathbf{x}}(r,\mathcal{X}(\cdot,r))\sigma(\cdot,r)dW(r),\qquad r\in[t,T].\end{aligned}$$

We denote by $\mathcal{S}^n_t\triangleq\mathcal{S}(\mathscr{X}_t)$ the set of all bounded symmetric bilinear functionals on $\mathscr{X}_t\times\mathscr{X}_t$, and we let $C([0,T];\mathcal{S}^n)$ be the set of all $P(\cdot)$ having the property that

$$P(t)\in\mathcal{S}^n_t\qquad\forall t\in[0,T],\qquad\text{and}\qquad\sup_{t\in[0,T]}\|P(t)\|<\infty,$$

and the map $t\mapsto P(t)$ is continuous in the following sense:

$$\lim_{t\to t_0}\big|P(t)\big([\mathbf{x}]_t(\cdot),[\mathbf{x}]_t(\cdot)\big)-P(t_0)\big([\mathbf{x}]_{t_0}(\cdot),[\mathbf{x}]_{t_0}(\cdot)\big)\big|=0\qquad\forall\mathbf{x}(\cdot)\in\mathscr{X}_0.$$

Note that for any $P(\cdot)\in C([0,T];\mathcal{S}^n)$, the map $(t,\mathbf{x}_t(\cdot))\mapsto v(t,\mathbf{x}_t(\cdot))$, defined by

$$v(t,\mathbf{x}_t(\cdot))\triangleq P(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big)\equiv\mathbf{x}_t(\cdot)^\top P(t)\mathbf{x}_t(\cdot)\qquad\forall(t,\mathbf{x}_t(\cdot))\in\Lambda,$$

is continuous on Λ under $\mathbf{d}$. Indeed,

$$\begin{aligned}|v(t,\tilde{\mathbf{x}}_t(\cdot))-v(t_0,\mathbf{x}_{t_0}(\cdot))|&=|P(t)\big(\tilde{\mathbf{x}}_t(\cdot),\tilde{\mathbf{x}}_t(\cdot)\big)-P(t_0)\big(\mathbf{x}_{t_0}(\cdot),\mathbf{x}_{t_0}(\cdot)\big)|\\ &\leq K\sup_{t\in[0,T]}\|P(t)\|\sup_{s\in[0,T]}|\tilde{\mathbf{x}}_t(s\vee t)-\mathbf{x}_{t_0}(s\vee t_0)|\\ &\quad+|P(t)(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot))-P(t_0)(\mathbf{x}_{t_0}(\cdot),\mathbf{x}_{t_0}(\cdot))|\to 0,\end{aligned}$$

as $(t,\tilde{\mathbf{x}}_t(\cdot))\to(t_0,\mathbf{x}_{t_0}(\cdot))$ under $\mathbf{d}$, where $K>0$ only depends on $\|\mathbf{x}_{t_0}(\cdot)\|$. Here, if $t<t_0$, we simply write $\mathbf{x}_t(\cdot)$ to represent $\mathbf{x}_{t_0}(s\vee t_0); s\in[t,T]$. Similar to the above, for any $\mathcal{X}(\cdot,\cdot),\mathcal{X}'(\cdot,\cdot)\in C(\Delta_*[0,T];\mathbb{R}^n)$ and $P(\cdot)\in C([0,T];\mathcal{S}^n)$, it can be shown that the map $t\mapsto P(t)\big(\mathcal{X}(\cdot,t),\mathcal{X}'(\cdot,t)\big)$ is continuous. Further, we let $C^1([0,T];\mathcal{S}^n)$ be the set of all $P(\cdot)\in C([0,T];\mathcal{S}^n)$ such that

$$v(t,\mathbf{x}_t(\cdot))=P(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big)\qquad\forall(t,\mathbf{x}_t(\cdot))\in\Lambda$$

admits

$$v_t(t,\mathbf{x}_t(\cdot))\equiv\dot{P}(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big)\qquad\forall(t,\mathbf{x}_t(\cdot))\in\Lambda,$$

for some $\dot{P}(\cdot)$, which is continuous. Similarly, we can define $C([a,b];\mathcal{S}^n)$ and $C^1([a,b];\mathcal{S}^n)$. We remark that for any $P(\cdot)\in C([a,b];\mathcal{S}^n)$ and $t\in[a,b]$, $P(t)$ is still a bounded symmetric bilinear functional on $\mathscr{X}_t\times\mathscr{X}_t=C([t,T];\mathbb{R}^n)\times C([t,T];\mathbb{R}^n)$. Note that for any $t\in[0,T]$, if the stochastic processes $X(\cdot)$ and $Y(\cdot)$ are in $\mathscr{X}_t$ almost surely, i.e., $X(\cdot)$ and $Y(\cdot)$ have continuous paths on $[t,T]$, then $Y(\cdot)^\top P(t)X(\cdot)\equiv P(t)\big(X(\cdot),Y(\cdot)\big)$ is well defined almost surely. As a direct consequence of Proposition 2.11, we have the following result.

PROPOSITION 2.12. *Let the assumptions of Proposition* 2.11 *hold. Let* $\mathcal{X}(\cdot,\cdot)$ *be defined by* (2.16)*. Then the following functional Itô formula holds:*

$$\begin{aligned}&d\big[P(r)\big([\mathbf{x}'_s(\cdot)]_r,\mathcal{X}(\cdot\,,r)\big)\big]=\big[\dot{P}(r)\big([\mathbf{x}'_s(\cdot)]_r,\mathcal{X}(\cdot\,,r)\big)+P(r)\big([\mathbf{x}'_s(\cdot)]_r,b(\cdot\,,r)\big)\big]dr\\ &\quad+P(r)\big([\mathbf{x}'_s(\cdot)]_r,\sigma(\cdot\,,r)\big)dW(r),\qquad r\in[s,T],\qquad\forall(s,\mathbf{x}'_s(\cdot))\in\Lambda,\\ &d\big[P(r)\big(\mathcal{X}(\cdot\,,r),\mathcal{X}(\cdot\,,r)\big)\big]=\big[\dot{P}(r)\big(\mathcal{X}(\cdot\,,r),\mathcal{X}(\cdot\,,r)\big)+2P(r)\big(\mathcal{X}(\cdot\,,r),b(\cdot\,,r)\big)\\ &\quad+P(r)\big(\sigma(\cdot\,,r),\sigma(\cdot\,,r)\big)\big]dr+2P(r)\big(\mathcal{X}(\cdot\,,r),\sigma(\cdot\,,r)\big)dW(r),\qquad r\in[t,T].\end{aligned}$$

3. Solvability of Problem (LQ-FSVIE) and optimality system. In this section, we shall study the cost functional (1.3) as a quadratic functional of the control $u(\cdot)$ on the Hilbert space $\mathscr{U}[t,T]$. A necessary condition and a sufficient condition for the existence of an open-loop optimal control will be derived by a standard variational method. This will be expressed by an optimality system which is a coupled system of an FSVIE and a type-II BSVIE. The innovation of this section is that we shall provide a new representation for the optimality system by establishing an interesting relation between linear type-II and type-III BSVIEs.

For any $u(\cdot)\in\mathscr{U}[t,T]$, consider the following FSVIE:

$$\begin{aligned}X^{0,u}(s)=&\int_t^s\big[A(s,\tau)X^{0,u}(\tau)+B(s,\tau)u(\tau)\big]d\tau\\ &+\int_t^s\big[C(s,\tau)X^{0,u}(\tau)+D(s,\tau)u(\tau)\big]dW(\tau),\qquad s\in[t,T].\end{aligned}$$

By Lemma 2.1, the above FSVIE admits a unique solution $X^{0,u}(\cdot)\in L^2_{\mathbb{F}^t}(\Omega;C([t,T];\mathbb{R}^n))$ satisfying

$$\mathbb{E}\left[\sup_{s\in[t,T]}|X^{0,u}(s)|^2\right]\leq K\mathbb{E}\int_t^T|u(s)|^2ds,$$

where the constant $K>0$ is independent of $u(\cdot)$. Thus we can define two bounded linear operators $\Gamma:\mathscr{U}[t,T]\to L^2_{\mathbb{F}^t}(\Omega;C([t,T];\mathbb{R}^n))$ and $\hat\Gamma:\mathscr{U}[t,T]\to L^2_{\mathcal{F}^t_T}(\Omega;\mathbb{R}^n)$ as follows:

$$[\Gamma u(\cdot)](\cdot)=X^{0,u}(\cdot),\qquad [\hat\Gamma u(\cdot)]=X^{0,u}(T)\qquad \forall u(\cdot)\in\mathscr{U}[t,T]. \tag{3.1}$$

For any $\mathbf{x}_t(\cdot)\in\mathscr{X}_t$, let $X^{\mathbf{x}_t,0}(\cdot)$ be the unique solution to the following linear uncontrolled FSVIE:

$$X^{\mathbf{x}_t,0}(s)=\mathbf{x}_t(s)+\int_t^s A(s,\tau)X^{\mathbf{x}_t,0}(\tau)d\tau+\int_t^s C(s,\tau)X^{\mathbf{x}_t,0}(\tau)dW(r),\qquad s\in[t,T].$$

Then the following linear operators $\Xi:C([t,T];\mathbb{R}^n)\to L^2_{\mathbb{F}^t}(\Omega;C([t,T];\mathbb{R}^n))$ and $\hat\Xi:C([t,T];\mathbb{R}^n)\to L^2_{\mathcal{F}^t_T}(\Omega;\mathbb{R}^n)$ can be also well defined:

$$[\Xi\mathbf{x}_t(\cdot)](\cdot)=X^{\mathbf{x}_t,0}(\cdot),\qquad [\hat\Xi\mathbf{x}_t(\cdot)]=X^{\mathbf{x}_t,0}(T)\qquad \forall\mathbf{x}_t(\cdot)\in\mathscr{X}_t. \tag{3.2}$$

With (3.1) and (3.2), the unique solution $X(\cdot)\equiv X(\cdot\,;t,\mathbf{x}_t(\cdot),u(\cdot))$ of state equation (1.2) corresponding to $(t,\mathbf{x}_t(\cdot))\in\Lambda$ and $u(\cdot)\in\mathscr{U}[t,T]$ can be represented by

$$X(\cdot)=[\Gamma u(\cdot)](\cdot)+[\Xi\mathbf{x}_t(\cdot)](\cdot),\qquad X(T)=[\hat\Gamma u(\cdot)]+[\hat\Xi\mathbf{x}_t(\cdot)]. \tag{3.3}$$

By substituting the above into (1.3), we obtain the following representation of the functional (1.3):

$$J(t,\mathbf{x}_t(\cdot);u(\cdot))=\langle\mathcal{M}_2u,u\rangle+2\langle\mathcal{M}_1\mathbf{x}_t,u\rangle+\frac{1}{2}\langle Q\Xi\mathbf{x}_t,\Xi\mathbf{x}_t\rangle+\frac{1}{2}\langle G\hat\Xi\mathbf{x}_t,\hat\Xi\mathbf{x}_t\rangle, \tag{3.4}$$

where

$$\mathcal{M}_2\triangleq\frac{\Gamma^*Q\Gamma+\hat\Gamma^*G\hat\Gamma+R}{2}\qquad\text{and}\qquad\mathcal{M}_1\triangleq\frac{\Gamma^*Q\Xi+\hat\Gamma^*G\hat\Xi}{2}.$$

Using the representation (3.4), we get the following abstract characterization for the open-loop optimal controls of Problem (LQ-FSVIE).

Proposition 3.1. *Let $(t,\mathbf{x}_t(\cdot))\in\Lambda$ be any given free pair and $\bar u(\cdot)\in\mathscr{U}[t,T]$. Then $\bar u(\cdot)$ is an open-loop optimal control of Problem* (LQ-FSVIE) *for $(t,\mathbf{x}_t(\cdot))$ if and only if*

$$\mathcal{M}_2\geq 0\qquad\text{and}\qquad\mathcal{M}_2\bar u+\mathcal{M}_1\mathbf{x}_t=0. \tag{3.5}$$

Proof. It is clear to see that $\bar u(\cdot)$ is an optimal control of Problem (LQ-FSVIE) if and only if

$$J(t,\mathbf{x}_t(\cdot);\bar u(\cdot)+\lambda u(\cdot))-J(t,\mathbf{x}_t(\cdot);\bar u(\cdot))\geq 0\qquad\forall u(\cdot)\in\mathscr{U}[t,T],\ \lambda\in\mathbb{R}. \tag{3.6}$$

For any $u(\cdot)\in\mathscr{U}[t,T]$ and $\lambda\in\mathbb{R}$, by (3.4) we have

$$J(t,\mathbf{x}_t(\cdot);\bar u(\cdot)+\lambda u(\cdot))-J(t,\mathbf{x}_t(\cdot);\bar u(\cdot))=\lambda^2\langle\mathcal{M}_2u,u\rangle+2\lambda\langle\mathcal{M}_2\bar u+\mathcal{M}_1\mathbf{x}_t,u\rangle. \tag{3.7}$$

Thus (3.6) holds if and only if (3.5) holds. The proof is thus complete. ∎

To solve Problem (LQ-FSVIE), we introduce the following assumption.

(H3) There exists a constant $\lambda > 0$ such that

$$\langle \mathcal{M}_2 u, u\rangle = J(t,0;u(\cdot)) \geq \lambda \mathbb{E}\int_t^T |u(s)|^2 ds \qquad \forall u(\cdot)\in \mathscr{U}[t,T]. \tag{3.8}$$

Combining this condition with Proposition 3.1, we can obtain the following result easily.

COROLLARY 3.2. *Let* (H1)–(H3) *hold. Then, for any free pair* $(t,\mathbf{x}_t(\cdot)) \in \Lambda$, *Problem* (LQ-FSVIE) *admits a unique optimal control* $\bar u(\cdot)$, *which is given by*

$$\bar u(\cdot) = -(\mathcal{M}_2^{-1}\mathcal{M}_1 \mathbf{x}_t)(\cdot). \tag{3.9}$$

Combining Proposition 3.1 and [39, Theorem 5.2], we have the following results.

THEOREM 3.3. *Suppose that the following convexity condition holds:*

$$\langle \mathcal{M}_2 u, u\rangle = J(t,0;u(\cdot)) \geq 0 \qquad \forall u(\cdot)\in \mathscr{U}[t,T]. \tag{3.10}$$

The control $\bar u(\cdot)\in\mathscr{U}[t,T]$ *is an optimal control of Problem* (LQ-FSVIE) *if and only if the optimality system* (1.7)–(1.8) *holds.*

Proof. From (3.7) we see that

$$\lim_{\lambda\to 0}\frac{J(t,\mathbf{x}_t(\cdot);\bar u(\cdot)+\lambda u(\cdot)) - J(t,\mathbf{x}_t(\cdot);\bar u(\cdot))}{\lambda} = 2\langle \mathcal{M}_2\bar u + \mathcal{M}_1\mathbf{x}_t, u\rangle \quad \forall u(\cdot)\in\mathscr{U}[t,T].$$

On the other hand, by [39, Theorem 5.2] we have

$$\lim_{\lambda\to 0}\frac{J(t,\mathbf{x}_t(\cdot);\bar u(\cdot)+\lambda u(\cdot)) - J(t,\mathbf{x}_t(\cdot);\bar u(\cdot))}{\lambda} = \mathbb{E}\int_t^T \langle R(s)\bar u(s) + Y^0(s), u(s)\rangle ds$$

for any $u(\cdot)\in\mathscr{U}[t,T]$, where $Y^0(\cdot)$ is uniquely determined by (1.8). Thus,

$$2(\mathcal{M}_2\bar u + \mathcal{M}_1\mathbf{x}_t)(\cdot) = (R\bar u + Y^0)(\cdot).$$

Then the desired results follow from Proposition 3.1. ∎

Remark 3.4. Let the controlled system (1.2) reduce to an SDE. In this case, if we denote

$$p(s) = \mathbb{E}_s\left(G\bar X(T) + \int_s^T Y(r)dr\right), \qquad q(s) = \zeta(s) + \int_s^T Z(r,s)dr, \qquad s\in[t,T],$$

then $(p(\cdot),q(\cdot))$ satisfies the BSDE

$$p(s) = G\bar X(T) + \int_s^T \big[A(r)^\top p(r) + C(r)^\top q(r) + Q(r)\bar X(r)\big]dr - \int_s^T q(r)dW(r)$$

on $[t,T]$ and the corresponding optimality condition (1.7) can be rewritten as

$$R(s)\bar u(s) + B(s)^\top p(s) + D(s)^\top q(s) = 0, \qquad s\in[t,T],$$

which recovers the corresponding results of Problem (LQ-SDE) (see [40, Chapter 6], for example).

Let us now return to BSVIEs. Denote

$$\begin{aligned} &\psi(s) = Q(s)\bar{X}(s) + A(T,s)^\top G\bar{X}(T) + C(T,s)^\top \zeta(s),\\ &\psi^0(s) = B(T,s)^\top G\bar{X}(T) + D(T,s)^\top \zeta(s), \qquad s\in[t,T]. \end{aligned} \tag{3.11}$$

Then (1.8) can be written as (assuming that $R(s)^{-1}$ exists)

$$\begin{cases} \bar{X}(s) = \mathbf{x}_t(s) + \int_t^s \big[A(s,\tau)\bar{X}(\tau) - B(s,\tau)R(\tau)^{-1}Y^0(\tau)\big]d\tau\\ \qquad\qquad + \int_t^s \big[C(s,\tau)\bar{X}(\tau) - D(s,\tau)R(\tau)^{-1}Y^0(\tau)\big]dW(\tau),\\ \eta(s) = G\bar{X}(T) - \int_s^T \zeta(\tau)dW(\tau),\\ Y(s) = \psi(s) + \int_s^T \big[A(\tau,s)^\top Y(\tau) + C(\tau,s)^\top Z(\tau,s)\big]d\tau - \int_s^T Z(s,\tau)dW(\tau),\\ Y^0(s) = \psi^0(s) + \int_s^T \big[B(\tau,s)^\top Y(\tau) + D(\tau,s)^\top Z(\tau,s)\big]d\tau - \int_s^T Z^0(s,\tau)dW(\tau). \end{cases} \tag{3.12}$$

Note that (3.12) is a coupled system of FBSVIEs. To the best of our knowledge, there is no general result on the solvability of coupled FBSVIEs on an arbitrary time horizon. In the rest of the paper, we are going to develop a decoupling method for (3.12), which can be regarded as the most important contribution of our paper. The key point is to find the so-called *decoupling field* for FBSVIE (3.12). As a preparation, we next provide a new representation for the optimality system (1.7)–(1.8) of Problem (LQ-FSVIE).

We introduce the following system of BSVIEs on $[t,T]$:

$$\begin{cases} Y^A(s) = \int_s^T A(\tau,s)^\top \Big[\mathbb{E}_\tau[\psi(\tau)] + Y^A(\tau) + Z^C(\tau,\tau)\Big]d\tau - \int_s^T Z^A(s,\tau)dW(\tau),\\ Y^B(s) = \int_s^T B(\tau,s)^\top \Big[\mathbb{E}_\tau[\psi(\tau)] + Y^A(\tau) + Z^C(\tau,\tau)\Big]d\tau - \int_s^T Z^B(s,\tau)dW(\tau),\\ Y^C(s) = \int_s^T C(\tau,s)^\top \Big[\mathbb{E}_\tau[\psi(\tau)] + Y^A(\tau) + Z^C(\tau,\tau)\Big]d\tau - \int_s^T Z^C(s,\tau)dW(\tau),\\ Y^D(s) = \int_s^T D(\tau,s)^\top \Big[\mathbb{E}_\tau[\psi(\tau)] + Y^A(\tau) + Z^C(\tau,\tau)\Big]d\tau - \int_s^T Z^D(s,\tau)dW(\tau). \end{cases} \tag{3.13}$$

Note that in the above, the diagonal value $Z^C(\tau,\tau)$ of the process $Z^C(\cdot\,,\cdot)$ appears, which makes such BSVIEs essentially different from type-I and type-II BSVIEs. Thus, we call the above *type-III BSVIEs* to distinguish them from the others.

BSVIEs with the diagonal values $Z(\tau,\tau)$ presented were initially introduced by Wang and Yong [33] while studying the time-inconsistent optimal control problems for SDEs. The well-posedness of such types of BSVIEs with general generators was studied by Hernández and Possamai [17]. We also highlight that Hamaguchi [14] gave a proper definition for the diagonal value $Z(\tau,\tau)$ of $Z(\cdot\,,\cdot)$. In [33, 16, 17], the diagonal value $Z(\tau,\tau)$ was due to the probabilistic representation of the equilibrium HJB equations. Here, we will use this type of BSVIEs to represent the solutions of adjoint equations in the optimality system. This is quite surprising.

Following Hamaguchi [13], we define the diagonal value $Z^\diamond(\tau,\tau)$ of $Z^\diamond(\cdot,\cdot)$ with $\diamond = C, D$ as follows:

$$Z^\diamond(\tau,\tau) := Z^\diamond(t,\tau) + \int_t^\tau \partial_s Z^\diamond(s,\tau)ds, \qquad \text{a.e., a.s.,} \tag{3.14}$$

where $\partial_s Z^\diamond(\cdot,\cdot)$ is the derivative of $Z^\diamond(\cdot,\cdot)$ with respect to the first variable. In other words, $(\partial_s Y^\diamond(\cdot,\cdot), \partial_s Z^\diamond(\cdot,\cdot))$ is the unique solution of the following (extended) type-I BSVIEs:

$$\partial_s Y^\diamond(s,r) = \int_r^T \diamond_s(\tau,s)^\top \Big[\mathbb{E}_\tau[\psi(\tau)] + Y^A(\tau) + Z^C(\tau,\tau)\Big]d\tau - \int_r^T \partial_s Z^\diamond(s,\tau)dW(\tau). \tag{3.15}$$

We call $(Y^*(\cdot), Z^*(\cdot,\cdot))$ with $* = A,B,C,D$ an *adapted solution* of (3.13) if it is satisfied in the usual Itô sense with the diagonal values $Z^C(\tau,\tau)$ and $Z^D(\tau,\tau)$ given by (3.14).

PROPOSITION 3.5. *Let* (H1) *hold. Then for any* $\psi(\cdot) \in L^2_{\mathcal{F}^t_T}([t,T];\mathbb{R}^n)$ *and* $\psi^0(\cdot) \in L^2_{\mathcal{F}^t_T}([t,T];\mathbb{R})$*, the third and the fourth BSVIEs in* (3.12) *admit unique adapted M-solution* $(Y(\cdot), Y^0(\cdot), Z(\cdot,\cdot), Z^0(\cdot,\cdot))$*, and the type-III BSVIE* (3.13) *also admits a unique adapted solution* $(Y^*(\cdot), Z^*(\cdot,\cdot))$ *with* $* = A,B,C,D$*. Moreover, the following representation holds:*

$$\begin{cases} Y(s) = \mathbb{E}_s[\psi(s)] + Y^A(s) + Z^C(s,s), \\ Y^0(s) = \mathbb{E}_s[\psi^0(s)] + Y^B(s) + Z^D(s,s), \end{cases} \qquad s \in [t,T]. \tag{3.16}$$

Proof. We first prove the uniqueness of the adapted solutions to (3.13). Suppose that $(Y^*(\cdot), Z^*(\cdot,\cdot))$ and $(\tilde{Y}^*(\cdot), \tilde{Z}^*(\cdot,\cdot))$ with $* = A,B,C,D$ are two adapted solutions of (3.13). Denote $(\hat{Y}^*(\cdot), \hat{Z}^*(\cdot,\cdot)) = (Y^*(\cdot) - \tilde{Y}^*(\cdot), Z^*(\cdot,\cdot) - \tilde{Z}^*(\cdot,\cdot))$. Then by (3.13) and (3.14), we have

$$\hat{Y}^*(s) = \int_s^T *(\tau,s)^\top \Big[\hat{Y}^A(\tau) + \hat{Z}^C(\tau,\tau)\Big]d\tau - \int_s^T \hat{Z}^*(s,\tau)dW(\tau), \tag{3.17}$$

and

$$\hat{Z}^C(\tau,\tau) = \hat{Z}^C(t,\tau) + \int_t^\tau \partial_s \hat{Z}^C(s,\tau)ds, \quad \text{a.e., a.s.,} \tag{3.18}$$

with

$$\partial_s \hat{Y}^C(s,r) = \int_r^T C_s(\tau,s)^\top \Big[\hat{Y}^A(\tau) + \hat{Z}^C(\tau,\tau)\Big]d\tau - \int_r^T \partial_s \hat{Z}^C(s,\tau)dW(\tau). \tag{3.19}$$

For any $S \in [t,T)$, by (3.18) we have

$$\begin{aligned} \mathbb{E}\int_S^T |\hat{Z}^C(\tau,\tau)|^2 d\tau &\le 2\mathbb{E}\int_S^T |\hat{Z}^C(t,\tau)|^2 d\tau + 2\mathbb{E}\int_S^T \left|\int_t^\tau \partial_s \hat{Z}^C(s,\tau)ds\right|^2 d\tau \\ &\le K\mathbb{E}\int_S^T |\hat{Z}^C(t,\tau)|^2 d\tau + K\mathbb{E}\int_S^T \int_t^\tau |\partial_s \hat{Z}^C(s,\tau)|^2 ds d\tau \\ &\le K\mathbb{E}\int_S^T |\hat{Z}^C(t,\tau)|^2 d\tau + K \sup_{s\in[t,T]} \mathbb{E}\int_{S\vee s}^T |\partial_s \hat{Z}^C(s,\tau)|^2 d\tau, \end{aligned}$$

where $K>0$ is a generic constant. Applying the standard results of BSVIEs to (3.17) and (3.19), taking the diagonal value $\hat Z^C(\tau,\tau)$ as a known process, we have

$$\mathbb{E}\int_S^T |\hat Z^C(t,\tau)|^2 d\tau + \sup_{s\in[t,T]}\mathbb{E}\int_{S\vee s}^T |\partial_s \hat Z^C(s,\tau)|^2 d\tau \le K(T-S)\mathbb{E}\int_S^T |\hat Z^C(\tau,\tau)|^2 d\tau.$$

It follows that

$$\mathbb{E}\int_S^T |\hat Z^C(\tau,\tau)|^2 d\tau \le K(T-S)\mathbb{E}\int_S^T |\hat Z^C(\tau,\tau)|^2 d\tau.$$

Taking $S=T-\frac{1}{2K}$, from the above we get $\hat Z^C(\tau,\tau)\equiv 0$ on $[S,T]$. By the standard results of BSVIEs, taking $\hat Z^C(\tau,\tau)\equiv 0$ in (3.17), we have $(\hat Y^*(\cdot),\hat Z^*(\cdot,\cdot))\equiv(0,0)$ with $*=A,B,C,D$ on $[S,T]$. Then from (3.17), $(\hat Y^*(\cdot),\hat Z^*(\cdot,\cdot))$ satisfies the following BSVIE on $[t,S]$:

$$\hat Y^*(s)=\int_s^S *(\tau,s)^\top\Big[\hat Y^A(\tau)+\hat Z^C(\tau,\tau)\Big]d\tau-\int_s^S \hat Z^*(s,\tau)dW(\tau).$$

Thus, by continuing the above arguments, we can get $(\hat Y^*(\cdot),\hat Z^*(\cdot,\cdot))\equiv(0,0)$ with $*=A,B,C,D$ on $[t,T]$, which implies the uniqueness of adapted solutions to (3.13).

Next we use the solutions of the third and the fourth BSVIEs in (3.12) to construct an adapted solution to (3.13), by which we can show that the representation (3.16) holds. Let $(Y(\cdot),Z(\cdot,\cdot))$ be the adapted M-solution to the third equation in (3.12). Using the fact

$$Y(\tau)=\mathbb{E}_s[Y(\tau)]+\int_s^\tau Z(\tau,r)dW(r),\qquad t\le s\le\tau\le T,$$

we get

$$\begin{aligned}\int_s^T *(\tau,s)^\top Y(\tau)d\tau &= \int_s^T *(\tau,s)^\top \mathbb{E}_s[Y(\tau)]d\tau+\int_s^T *(\tau,s)^\top\int_s^\tau Z(\tau,r)dW(r)d\tau\\ &=\mathbb{E}_s\int_s^T *(\tau,s)^\top Y(\tau)d\tau+\int_s^T\int_\tau^T *(r,s)^\top Z(r,\tau)drdW(\tau),\end{aligned}\tag{3.20}$$

with $*=A,B,C,D$. On the other hand, by the third equation in (3.12),

$$Y(\tau)=\mathbb{E}_\tau\left[\psi(\tau)+\int_\tau^T\big[A(r,\tau)^\top Y(r)+C(r,\tau)^\top Z(r,\tau)\big]dr\right],\qquad \tau\in[t,T].$$

Thus, we have

$$\begin{aligned}\int_s^T *(\tau,s)^\top Y(\tau)d\tau &= \int_s^T *(\tau,s)^\top\mathbb{E}_\tau[\psi(\tau)]d\tau+\int_s^T *(\tau,s)^\top\mathbb{E}_\tau\int_\tau^T A(r,\tau)^\top Y(r)drd\tau\\ &\quad+\int_s^T *(\tau,s)^\top\int_\tau^T C(r,\tau)^\top Z(r,\tau)drd\tau,\end{aligned}$$

in which we use the fact that $Z(r,\tau)$ is $\mathcal{F}_s^\tau$-measurable. Substituting the above into (3.20) yields that

$$\begin{aligned}\mathbb{E}_s\int_s^T *(r,s)^\top Y(r)dr &= \int_s^T *(\tau,s)^\top\bigg[\mathbb{E}_\tau[\psi(\tau)]+\mathbb{E}_\tau\int_\tau^T A(r,\tau)^\top Y(r)dr\\ &\quad+\int_\tau^T C(r,\tau)^\top Z(r,\tau)dr\bigg]d\tau-\int_s^T\int_\tau^T *(r,s)^\top Z(r,\tau)drdW(\tau).\end{aligned}$$

Thus, for $* = A, B, C, D$, if we denote

$$Y^*(s) = \mathbb{E}_s \int_s^T *(r,s)^\top Y(r)dr, \qquad Z^*(s,\tau) = \int_\tau^T *(r,s)^\top Z(r,\tau)dr, \tag{3.21}$$

then both (3.13) and (3.14) are satisfied. Thus, the process $(Y^*(\cdot), Z^*(\cdot\,,\cdot))$ defined by (3.21) is exactly the unique adapted solution to (3.13). Applying $\mathbb{E}_s[\,\cdot\,]$ on the last two equations in (3.12) yields that

$$\begin{aligned} Y(s) &= \mathbb{E}_s\left[\psi(s) + \int_s^T \Big(A(\tau,s)^\top Y(\tau) + C(\tau,s)^\top Z(\tau,s)\Big)d\tau\right] \\ &= \mathbb{E}_s[\psi(s)] + Y^A(s) + Z^C(s,s), \\ Y^0(s) &= \mathbb{E}_s\left[\psi^0(s) + \int_s^T \Big(B(\tau,s)^\top Y(\tau) + D(\tau,s)^\top Z(\tau,s)\Big)d\tau\right] \\ &= \mathbb{E}_s[\psi^0(s)] + Y^B(s) + Z^D(s,s), \end{aligned}$$

proving our conclusion. □

Proposition 3.5 gives an explicit relation between a type-II BSVIE and a type-III BSVIE. This relation will serve as a foundation in developing our decoupling approach for the optimality system associated Problem (LQ-FSVIE). In the proof of Proposition 3.5, we need to use the differentiability of the coefficients $C(\cdot\,,\cdot)$ and $D(\cdot\,,\cdot)$ with respect to the second variable.

Combining Theorem 3.3 and Proposition 3.5, we have the following new characterization of the optimal control.

THEOREM 3.6. *Let* (H1)–(H2) *hold. Suppose that the convexity condition* (3.10) *holds. Let* $\psi(\cdot)$ *and* $\psi^0(\cdot)$ *be defined by* (3.11). *Let* $(Y^A(\cdot), Y^B(\cdot), Y^C(\cdot), Y^D(\cdot), Z^A(\cdot\,,\cdot), Z^B(\cdot\,,\cdot), Z^C(\cdot\,,\cdot), Z^D(\cdot\,,\cdot))$ *be the unique adapted solution to BSVIE* (3.13). *Then the control* $\bar u(\cdot) \in \mathscr{U}[t,T]$ *is an open-loop optimal control of Problem* (LQ-FSVIE) *if and only if*

$$R(s)\bar u(s) + \mathbb{E}_s[\psi^0(s)] + Y^B(s) + Z^D(s,s) = 0, \qquad s \in [t,T], \text{ a.s.} \tag{3.22}$$

Remark 3.7. In the literature [33, 16, 13, 17, 19, 18] of time-inconsistent problems, the diagonal term $Z(s,s)$ was caused by the Voterra-type cost functionals, while in this paper, the appearance of $Z(s,s)$ was due to the Voterra-type state equation.

4. Derivation of the path-dependent Riccati equation. In this section, we will find the path-dependent Riccait equation for our Problem (LQ-FSVIE) via the HJB equation for the value function. The procedure is formal. However, the argument over verification is rigorous. Thus, once the well-posedness of the Riccati equation is established, our Problem (LQ-FSVIE) is solved. This also decouples the optimality system (3.12).

By [32, subsection 4.3], the path-dependent HJB equation associated with Problem (LQ-FSVIE) reads

$$\begin{cases} v_t(t,\mathbf{x}_t(\cdot)) + \inf\limits_{u\in\mathbb{R}^m} \mathcal{H}(t,\mathbf{x}_t(\cdot),u,v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot)),v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot))) = 0 \qquad \forall (t,\mathbf{x}_t(\cdot)) \in \Lambda, \\ v(T,\mathbf{x}(T)) = \dfrac{1}{2}\langle G\mathbf{x}(T),\mathbf{x}(T)\rangle \qquad \forall \mathbf{x}(T) \in \mathbb{R}^n, \end{cases} \tag{4.1}$$

where for any $(t, \mathbf{x}_t(\cdot), u) \in \Lambda \times \mathbb{R}^m$, the Hamiltonian $\mathcal{H}$ is defined by

$$\begin{aligned}
&\mathcal{H}\big(t, \mathbf{x}_t(\cdot), u, v_{\mathbf{x}}(t, \mathbf{x}_t(\cdot)), v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))\big) \\
&\quad \triangleq \frac{1}{2}\big[C(\cdot\,,t)\mathbf{x}_t(t) + D(\cdot\,,t)u\big]^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))\big[C(\cdot\,,t)\mathbf{x}_t(t) + D(\cdot\,,t)u\big] \\
&\qquad + v_{\mathbf{x}}(t, \mathbf{x}_t(\cdot))\big[A(\cdot\,,t)\mathbf{x}_t(t) + B(\cdot\,,t)u\big] + \frac{1}{2}\mathbf{x}_t(t)^\top Q(t)\mathbf{x}_t(t) + \frac{1}{2}u^\top R(t)u.
\end{aligned} \tag{4.2}$$

To understand each term in the above, let us denote

$$\begin{aligned}
A(s,t) &= \big(A_1(s,t), A_2(s,t), \cdots, A_n(s,t)\big), \\
B(s,t) &= \big(B_1(s,t), B_2(s,t), \cdots, B_m(s,t)\big), \\
C(s,t) &= \big(C_1(s,t), C_2(s,t), \cdots, C_n(s,t)\big), \\
D(s,t) &= \big(D_1(s,t), D_2(s,t), \cdots, D_m(s,t)\big),
\end{aligned}$$

where all $A_i(\cdot\,,\cdot), B_i(\cdot\,,\cdot), C_i(\cdot\,,\cdot), D_i(\cdot\,,\cdot)$ are $\mathbb{R}^n$-valued functions. Then, for any $v(\cdot\,,\cdot) \in C_+^{1,2}(\Lambda)$, we have the following:

$$\begin{aligned}
&\big[C(\cdot\,,t)\mathbf{x}_t(t) + D(\cdot\,,t)u\big]^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))\big[C(\cdot\,,t)\mathbf{x}_t(t) + D(\cdot\,,t)u\big] \\
&\quad = \mathbf{x}_t(t)^\top \big[C(\cdot\,,t)^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))C(\cdot\,,t)\big]\mathbf{x}_t(t) + \mathbf{x}_t(t)^\top \big[C(\cdot\,,t)^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))D(\cdot\,,t)\big]u \\
&\qquad + u^\top \big[D(\cdot\,,t)^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))C(\cdot\,,t)\big]\mathbf{x}_t(t) + u^\top \big[D(\cdot\,,t)^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))D(\cdot\,,t)\big]u,
\end{aligned}$$

with (for all $t \in [0,T]$)

$$\begin{aligned}
C(\cdot\,,t)^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))C(\cdot\,,t) &\equiv \Big(C_i(\cdot\,,t)^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))C_j(\cdot\,,t)\Big) \in \mathbb{S}^n, \\
D(\cdot\,,t)^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))D(\cdot\,,t) &\equiv \Big(D_i(\cdot\,,t)^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))D_j(\cdot\,,t)\Big) \in \mathbb{S}^m, \\
D(\cdot\,,t)^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))C(\cdot\,,t) &\equiv \Big(D_i(\cdot\,,t)^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))C_j(\cdot\,,t)\Big) \in \mathbb{R}^{m\times n}, \\
C(\cdot\,,t)^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))D(\cdot\,,t) &\equiv \Big(C_i(\cdot\,,t)^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))D_j(\cdot\,,t)\Big) \in \mathbb{R}^{n\times m},
\end{aligned}$$

and

$$v_{\mathbf{x}}(t, \mathbf{x}_t(\cdot))\big[A(\cdot\,,t)\mathbf{x}_t(t) + B(\cdot\,,t)u\big] = \big[v_{\mathbf{x}}(t, \mathbf{x}_t(\cdot))A(\cdot\,,t)\big]\mathbf{x}_t(t) + \big[v_{\mathbf{x}}(t, \mathbf{x}_t(\cdot))B(\cdot\,,t)\big]u,$$

with (for all $t \in [0,T]$)

$$\begin{aligned}
v_{\mathbf{x}}(t, \mathbf{x}_t(\cdot))A(\cdot\,,t) &= \big(v_{\mathbf{x}}(t, \mathbf{x}_t(\cdot))A_1(\cdot\,,t), \cdots, v_{\mathbf{x}}(t, \mathbf{x}_t(\cdot))A_n(\cdot\,,t)\big) \in \mathbb{R}^{1\times n}, \\
v_{\mathbf{x}}(t, \mathbf{x}_t(\cdot))B(\cdot\,,t) &= \big(v_{\mathbf{x}}(t, \mathbf{x}_t(\cdot))B_1(\cdot\,,t), \cdots, v_{\mathbf{x}}(t, \mathbf{x}_t(\cdot))B_m(\cdot\,,t)\big) \in \mathbb{R}^{1\times m}.
\end{aligned}$$

From the above, we see that the following condition makes sense:

$$\boldsymbol{R}(t, \mathbf{x}_t(\cdot)) \equiv D(\cdot\,,t)^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))D(\cdot\,,t) + R(t) \geq \lambda I_m \quad \text{for some } \lambda > 0. \tag{4.3}$$

If such a condition holds, then by denoting

$$\begin{aligned}
\boldsymbol{S}(t, \mathbf{x}_t(\cdot)) &= [D(\cdot\,,t)^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))C(\cdot\,,t)]\mathbf{x}_t(t) + [v_{\mathbf{x}}(t, \mathbf{x}_t(\cdot))B(\cdot\,,t)]^\top, \\
\boldsymbol{Q}(t, \mathbf{x}_t(\cdot)) &= [v_{\mathbf{x}}(t, \mathbf{x}_t(\cdot))A(\cdot\,,t)]\mathbf{x}_t(t) + \mathbf{x}_t(t)^\top [v_{\mathbf{x}}(t, \mathbf{x}_t(\cdot))A(\cdot\,,t)]^\top \\
&\quad + \mathbf{x}_t(t)^\top \big[C(\cdot\,,t)^\top v_{\mathbf{xx}}(t, \mathbf{x}_t(\cdot))C(\cdot\,,t) + Q(t)\big]\mathbf{x}_t(t),
\end{aligned}$$

we have

$$
\begin{aligned}
&2\mathcal{H}\big(t,\mathbf{x}_t(\cdot),u,v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot)),v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot))\big)\\
&\quad\geq \boldsymbol{Q}(t,\mathbf{x}_t(\cdot))-\boldsymbol{S}(t,\mathbf{x}_t(\cdot))^\top\boldsymbol{R}(t,\mathbf{x}_t(\cdot))^{-1}\boldsymbol{S}(t,\mathbf{x}_t(\cdot))\\
&\quad=\inf_{u\in\mathbb{R}^m}2\mathcal{H}\big(t,\mathbf{x}_t(\cdot),u,v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot)),v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot))\big)\\
&\quad\equiv 2\mathcal{H}\big(t,\mathbf{x}_t(\cdot),\bar{\Gamma}(t,\mathbf{x}_t(\cdot)),v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot)),v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot))\big),
\end{aligned}
$$

where

$$
\begin{aligned}
\bar{\Gamma}(t,\mathbf{x}_t(\cdot))&=-\boldsymbol{R}(t,\mathbf{x}_t(\cdot))^{-1}\boldsymbol{S}(t,\mathbf{x}_t(\cdot))=-\big[D(\cdot\,,t)^\top v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot))D(\cdot\,,t)+R(t)\big]^{-1}\\
&\quad\times\big\{[D(\cdot\,,t)^\top v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot))C(\cdot\,,t)]\mathbf{x}_t(t)+[v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot))B(\cdot\,,t)]^\top\big\}.
\end{aligned}
\tag{4.4}
$$

Substituting the above into (4.1), by some straightforward calculations, we get

$$
\begin{cases}
v_t(t,\mathbf{x}_t(\cdot))+\dfrac{1}{2}\mathbf{x}_t(t)^\top[C(\cdot\,,t)^\top v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot))C(\cdot\,,t)]\mathbf{x}_t(t)+\dfrac{1}{2}\mathbf{x}_t(t)^\top Q(t)\mathbf{x}_t(t)\\
\quad+[v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot))A(\cdot\,,t)]\mathbf{x}_t(t)-\dfrac{1}{2}\Big(D(\cdot\,,t)^\top v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot))C(\cdot\,,t)\mathbf{x}_t(t)\\
\quad+[v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot))B(\cdot\,,t)]^\top\Big)^\top\big[D(\cdot,t)^\top v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot))D(\cdot,t)+R(t)\big]^{-1}\\
\quad\times\Big(D(\cdot,t)^\top v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot))C(\cdot,t)\mathbf{x}_t(t)+[v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot))B(\cdot\,,t)]^\top\Big)=0\quad\forall(t,\mathbf{x}_t(\cdot))\in\Lambda,\\
v(T,\mathbf{x}(T))=\dfrac{1}{2}\mathbf{x}(T)^\top G\mathbf{x}(T)\qquad\forall\mathbf{x}(T)\in\mathbb{R}^n.
\end{cases}
\tag{4.5}
$$

In the rest of the paper, we shall prove that (4.5) admits a classical solution by introducing a new type of Riccati equation. With the solution of this Riccati equation, the form of the optimal strategy $\bar{\Gamma}(\cdot\,,\cdot)$, defined by (4.4), can be simplified, and some interesting phenomena will be found. More importantly, we will show that the solution of the derived Riccati equation is exactly the decoupling field for the optimality system (1.7)–(1.8).

Recall from subsection 2.3 the definition of the space $C([0,T];\mathcal{S}^n)$. From the definition (1.5) of the value function of Problem (LQ-FSVIE), we expect that

$$
v(t,\mathbf{x}_t(\cdot))=\frac{P(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big)}{2}\qquad\forall(t,\mathbf{x}_t(\cdot))\in\Lambda,
\tag{4.6}
$$

for some bilinear operator-valued function $P(\cdot)\in C([0,T];\mathcal{S}^n)$. We define

$$
\begin{aligned}
\dot{P}(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big)&\triangleq\lim_{\varepsilon\to0^+}\frac{P(t+\varepsilon)\big([\mathbf{x}_t]_{t+\varepsilon}(\cdot),[\mathbf{x}_t]_{t+\varepsilon}(\cdot)\big)-P(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big)}{\varepsilon}\\
&\qquad\forall(t,\mathbf{x}_t(\cdot))\in\Lambda,
\end{aligned}
\tag{4.7}
$$

provided the limit exists. Then, by (2.15) and (4.6) we have

$$
v_t(t,\mathbf{x}_t(\cdot))=\frac{\dot{P}(t)(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot))}{2}\qquad\forall(t,\mathbf{x}_t(\cdot))\in\Lambda.
\tag{4.8}
$$

Moreover, by (2.11)–(2.12) and (4.6), we get

$$
\begin{aligned}
&v_{\mathbf{x}}(t,\mathbf{x}_t(\cdot))(\eta_t(\cdot))=P(t)\big(\mathbf{x}_t(\cdot),\eta_t(\cdot)\big)\qquad\forall(t,\mathbf{x}_t(\cdot))\in\Lambda,\ \eta_t(\cdot)\in\mathscr{X}_t,\\
&v_{\mathbf{xx}}(t,\mathbf{x}_t(\cdot))(\eta_t(\cdot),\eta_t'(\cdot))=P(t)\big(\eta_t(\cdot),\eta_t'(\cdot)\big)\quad\forall(t,\mathbf{x}_t(\cdot))\in\Lambda,\ \eta_t(\cdot),\eta_t'(\cdot)\in\mathscr{X}_t.
\end{aligned}
\tag{4.9}
$$

With the representation (4.6), (4.8), and (4.9), from (4.5) we identify that $P(\cdot)$ should satisfy the following Riccati equation:

$$\begin{cases} \dot P(t) + P(t)A(\cdot\,,t)\delta_t + \delta_t^\top A(\cdot\,,t)^\top P(t) + \delta_t^\top C(\cdot\,,t)^\top P(t)C(\cdot\,,t)\delta_t + \delta_t^\top Q(t)\delta_t \\ \quad -\big[B(\cdot\,,t)^\top P(t) + D(\cdot\,,t)^\top P(t)C(\cdot\,,t)\delta_t\big]^\top \big[R(t) + D(\cdot\,,t)^\top P(t)D(\cdot\,,t)\big]^{-1} \\ \quad \times \big[B(\cdot\,,t)^\top P(t) + D(\cdot\,,t)^\top P(t)C(\cdot\,,t)\delta_t\big] = 0, \qquad t\in[0,T], \\ P(T) = G, \end{cases} \tag{4.10}$$

where δ_t is defined by

$$\delta_t \mathbf{x}_t(\cdot) = \mathbf{x}_t(t) \qquad \forall \mathbf{x}_t(\cdot)\in\mathscr{X}_t.$$

The above is understood as an equation for a symmetric bilinear operator-valued function defined on $[0,T]$. More precisely, for any $(t,\mathbf{x}_t(\cdot))\in\Lambda$, the term $\mathbf{x}_t(\cdot)^\top \dot P(t)\mathbf{x}_t(\cdot) \equiv \dot P(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big)$ is understood as in (4.7), and other terms are as follows:

$$\begin{aligned} &\mathbf{x}_t(\cdot)^\top[P(t)A(\cdot\,,t)\delta_t]\mathbf{x}_t(\cdot) = P(t)\big(A(\cdot\,,t)\mathbf{x}_t(t),\mathbf{x}_t(\cdot)\big) = P(t)\big(\mathbf{x}_t(\cdot),A(\cdot\,,t)\mathbf{x}_t(t)\big)\\ &\quad = \mathbf{x}_t(\cdot)^\top[\delta_t^\top A(\cdot\,,t)^\top P(t)]\mathbf{x}_t(\cdot)\in\mathbb{R},\\ &\mathbf{x}_t(\cdot)^\top[\delta_t^\top C(\cdot\,,t)^\top P(t)C(\cdot\,,t)\delta_t]\mathbf{x}_t(\cdot) = \mathbf{x}_t(t)^\top\Big(P(t)\big(C_i(\cdot\,,t),C_j(\cdot\,,t)\big)\Big)\mathbf{x}_t(t)\in\mathbb{R},\\ &\mathbf{x}_t(\cdot)^\top[\delta_t^\top Q(t)\delta_t]\mathbf{x}_t(\cdot) = \mathbf{x}_t(t)^\top Q(t)\mathbf{x}_t(t)\in\mathbb{R},\\ &D(\cdot\,,t)^\top P(t)D(\cdot,t) = \Big(P(t)\big(D_i(\cdot\,,t),D_j(\cdot\,,t)\big)\Big)\in\mathbb{S}^m,\\ &\big[B(\cdot\,,t)^\top P(t) + D(\cdot\,,t)^\top P(t)C(\cdot\,,t)\delta_t\big]\mathbf{x}_t(\cdot)\\ &\quad = \Big(P(t)\big(B_j(\cdot\,,t),\mathbf{x}_t(\cdot)\big) + P(t)\big(D_j(\cdot\,,t),C(\cdot\,,t)\mathbf{x}_t(t)\big)\Big)\in\mathbb{R}^m. \end{aligned}$$

Definition 4.1. *We call $P(\cdot)\in C([0,T];\mathcal{S}^n)$ a solution of Riccati equation* (4.10) *if the map $t\mapsto P(t)\big([\mathbf{x}]_t(\cdot),[\mathbf{x}]_t(\cdot)\big)$ is absolutely continuous for any $\mathbf{x}(\cdot)\in\mathscr{X}_0$, and it satisfies* (4.10) *for almost everywhere $t\in[0,T]$. Further, it is called a strongly regular solution, if, in addition, the following holds:*

$$R(t) + D(\cdot\,,t)^\top P(t)D(\cdot\,,t) \ge \lambda I_m, \qquad t\in[0,T], \qquad \text{for some } \lambda>0. \tag{4.11}$$

We emphasize that for any $t\in[0,T]$, the domain of $P(t)$ is $\mathscr{X}_t\times\mathscr{X}_t$, which is merely a Banach space rather than a Hilbert space, and as $t\in[0,T]$ varies, it changes. Thus, (4.10) is significantly different from the so-called *operator-valued Riccati equation* derived from the LQ control problem for (stochastic) evolution equations (see Li and Yong [20] and Lü [21] for examples). To emphasize the new features, we would like to call (4.10) a *path-dependent Riccati equation.*

Theorem 4.2. *Suppose that the path-dependent Riccati equation* (4.10) *admits a strongly regular solution $P(\cdot)\in C([0,T];\mathcal{S}^n)$. Then the function $v(\cdot,\cdot)$, defined by*

$$v(t,\mathbf{x}_t(\cdot)) = \frac{P(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big)}{2} \qquad \forall (t,\mathbf{x}_t(\cdot))\in\Lambda, \tag{4.12}$$

satisfies the path-dependent HJB equation (4.5) *for almost everywhere $t\in[0,T]$. Moreover, the minimizer $\bar\Gamma(\cdot,\cdot)$ of the Hamiltonian $\mathcal{H}$ can be represented as*

$$\begin{aligned} \bar\Gamma(t,\mathbf{x}_t(\cdot)) = -\big[D(\cdot\,,t)^\top P(t)D(\cdot\,,t) + R(t)\big]^{-1}\big[D(\cdot\,,t)^\top P(t)C(\cdot\,,t)\mathbf{x}_t(t)\\ + B(\cdot\,,t)^\top P(t)\mathbf{x}_t(\cdot)\big] \qquad \forall (t,\mathbf{x}_t(\cdot))\in\Lambda. \end{aligned} \tag{4.13}$$

Proof. By (4.10), with (4.8) and (4.9), we get that the function $v(\cdot,\cdot)$ defined by (4.12) satisfies (4.5) immediately. ∎

Remark 4.3. Recall that for any $t \in [0,T]$, the value function $v(t,\cdot)$ in Viens and Zhang [32] is defined on $\mathscr{X}_0 \equiv C([0,T];\mathbb{R}^n)$. The key of our derivation is that $v(t,\cdot)$, as well as $v_t(t,\cdot)$, $v_{\mathbf{x}}(t,\cdot)$, and $v_{\mathbf{xx}}(t,\cdot)$ can be well defined on the smaller space $\mathscr{X}_t \equiv C([t,T];\mathbb{R}^n)$ in the state-dependent setting, which means that the drift and diffusion terms of (1.2) only depend on the current value $X(\tau)$ of $X(l); l \in [t,\tau]$. Then, from the HJB equation (4.5), we correctly identify the Riccati equation associated with Problem (LQ-FSVIE).

Remark 4.4. If $\alpha(s,\tau) \equiv \alpha(\tau)$ for $\alpha(\cdot,\cdot) = A(\cdot,\cdot), B(\cdot,\cdot), C(\cdot,\cdot), D(\cdot,\cdot)$, and $\mathbf{x}_t(s) \equiv x; s \in [t,T]$ for some $x \in \mathbb{R}^n$, then state equation (1.2) becomes an SDE and Problem (LQ-FSVIE) reduces to a classical stochastic LQ optimal control problem. The corresponding Riccati equation reads (see [40, Chapter 6])

$$(4.14)\qquad \begin{cases} \dot{\Sigma}(t) + \Sigma(t)A(t) + A(t)^\top \Sigma(t) + C(t)^\top \Sigma(t)C(t) + Q(t) \\ \quad - [\Sigma(t)B(t) + C(t)^\top \Sigma(t)D(t)][R(t) + D(t)^\top \Sigma(t)D(t)]^{-1} \\ \quad \times [B(t)^\top \Sigma(t) + D(t)^\top \Sigma(t)C(t)] = 0, \qquad t \in [0,T], \\ \Sigma(T) = G. \end{cases}$$

By comparing the above with (4.10), we have

$$P(t)(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)) = \mathbf{x}_t(t)^\top \Sigma(t)\mathbf{x}_t(t) \qquad \forall (t,\mathbf{x}_t(\cdot)) \in \Lambda^0.$$

Here, Λ^0 is a subspace of Λ with the continuous function $\mathbf{x}_t(\cdot)$ being a constant. In the above sense, our result recovers the classical one for Problem (LQ-SDE).

5. Decoupling the optimality system. In this section, we shall show that the solution $P(\cdot)$ to Riccati equation (4.10) is exactly the so-called *decoupling field* for the optimality system (3.12), which is a coupled FBSVIE. Note that the solution of the path-dependent HJB equation associated with Problem (LQ-FSVIE) can be represented by (4.12). Thus, the connection between the Volterra-type stochastic Hamiltonian system (1.8) and the path-dependent HJB equation (4.5) associated with Problem (LQ-FSVIE) will be also established.

THEOREM 5.1. *Let* (H1)–(H3) *hold. Suppose that the path-dependent Riccati equation* (4.10) *admits a strongly regular solution* $P(\cdot) \in C([0,T];\mathcal{S}^n)$. *Then the solutions of the optimality system* (3.12) *admit the following representation:*

$$\begin{aligned} &\bar{X}(s) = \bar{\mathcal{X}}(s,s), \quad Y^0(s) = -R(s)\Theta(s)\bar{\mathcal{X}}(\cdot,s), \quad Y(s) = \big[C(\cdot,s)^\top P(s)C(\cdot,s) \\ &\quad + Q(s)\big]\bar{X}(s) + \big[A(\cdot,s)^\top P(s) + C(\cdot,s)^\top P(s)D(\cdot,s)\Theta(s)\big]\bar{\mathcal{X}}(\cdot,s), \qquad s \in [t,T], \end{aligned}$$

where

$$(5.1)\quad \Theta(s) \triangleq -[R(s)+D(\cdot,s)^\top P(s)D(\cdot,s)]^{-1}[B(\cdot,s)^\top P(s)+D(\cdot,s)^\top P(s)C(\cdot,s)\delta_s],$$

with $\bar{X}(\cdot)$ *being the unique solution to the closed-loop system*

$$(5.2)\quad \begin{aligned} \bar{X}(s) = \mathbf{x}_t(s) &+ \int_t^s \big[A(s,\tau)\bar{X}(\tau) + B(s,\tau)\Theta(\tau)\bar{\mathcal{X}}(\cdot,\tau)\big]d\tau \\ &+ \int_t^s \big[C(s,\tau)\bar{X}(\tau) + D(s,\tau)\Theta(\tau)\bar{\mathcal{X}}(\cdot,\tau)\big]dW(\tau), \qquad s \in [t,T], \end{aligned}$$

and $\bar{\mathcal{X}}(\cdot\,,\cdot)$ being the unique solution to the closed-loop auxiliary system

$$\begin{aligned}\bar{\mathcal{X}}(s,r) &= \mathbf{x}_t(s) + \int_t^r \big[A(s,\tau)\bar{\mathcal{X}}(\tau,\tau) + B(s,\tau)\Theta(\tau)\bar{\mathcal{X}}(\cdot\,,\tau)\big]d\tau \\ &\quad + \int_t^r \big[C(s,\tau)\bar{\mathcal{X}}(\tau,\tau) + D(s,\tau)\Theta(\tau)\bar{\mathcal{X}}(\cdot\,,\tau)\big]dW(\tau), \qquad (s,r)\in \Delta_*[t,T]. \end{aligned} \tag{5.3}$$

Moreover, the optimal control $\bar{u}(\cdot)$ admits the following causal state feedback representation:

$$\bar{u}(s) = \Theta(s)\bar{\mathcal{X}}(\cdot\,,s), \qquad s\in[t,T]. \tag{5.4}$$

Remark 5.2. By Theorem 5.1, the optimality system (3.12) is decoupled, and the optimal control $\bar{u}(\cdot)$ is represented as a causal state feedback (see (5.4)). Consequently, in principle, the optimal control is practically realizable.

Remark 5.3. Note that when $B(\cdot\,,\cdot)\equiv 0$, the optimal control $\bar{u}(\cdot)$ can be represented by

$$\begin{aligned}\bar{u}(s) &= \Theta(s)\bar{\mathcal{X}}(\cdot\,,s) = -[R(s)+D(\cdot\,,s)^\top P(s) D(\cdot\,,s)]^{-1} D(\cdot\,,s)^\top P(s) C(\cdot\,,s)\bar{X}(s) \\ &\triangleq \Theta^*(s)\bar{X}(s), \qquad s\in[t,T], \end{aligned} \tag{5.5}$$

with

$$\begin{aligned}\bar{X}(s) &= \mathbf{x}_t(s) + \int_t^s A(s,\tau)\bar{X}(\tau)d\tau \\ &\quad + \int_t^s \big[C(s,\tau)\bar{X}(\tau) + D(s,\tau)\Theta^*(\tau)\bar{X}(\tau)\big]dW(\tau), \qquad s\in[t,T]. \end{aligned} \tag{5.6}$$

It is particularly noteworthy that though the state equation (5.6) is a non-Markovian system, the optimal control $\bar{u}(\cdot)$, defined by (5.5), can be uniquely determined by the current value of the state process $\bar{X}(\cdot)$. Thus, when the drift term of (1.2) does not contain controls, the causal feedback representation of the optimal control, obtained by Theorem 5.1, reduces to a *state feedback* (also called a *Markovian feedback*), which is really unexpected.

Proof of Theorem 5.1. By Proposition 2.4, the systems (5.2) and (5.3) admit unique solution $\bar{X}(\cdot)\in L^2_{\mathbb{F}^t}(\Omega;C([t,T];\mathbb{R}^n))$ and $\bar{\mathcal{X}}(\cdot\,,\cdot)\in L^2_{\mathbb{F}^t}(\Omega;C(\Delta_*[t,T];\mathbb{R}^n))$, respectively. Moreover, we have $\bar{X}(s)=\bar{\mathcal{X}}(s,s); s\in[t,T]$. Let

$$\eta(s) = G\bar{\mathcal{X}}(T,T) - \int_s^T \zeta(r)dW(r), \qquad s\in[t,T], \tag{5.7}$$

and denote

$$\begin{aligned}\psi(s) &= Q(s)\bar{\mathcal{X}}(s,s) + A(T,s)^\top G\bar{\mathcal{X}}(T,T) + C(T,s)^\top\zeta(s), \\ \psi^0(s) &= B(T,s)^\top G\bar{\mathcal{X}}(T,T) + D(T,s)^\top\zeta(s), \qquad s\in[t,T]. \end{aligned} \tag{5.8}$$

Then for any $(s,\mathbf{x}_s(\cdot))\in\Lambda$, by the functional Itô formula (see Proposition 2.12), we have

$$\begin{aligned}
&d\big\{[\mathbf{x}_s]_r(\cdot)^\top P(r)\bar{\mathcal{X}}(\cdot,r)\big\} \equiv d\big\{P(r)\big(\bar{\mathcal{X}}(\cdot,r),[\mathbf{x}_s]_r(\cdot)\big)\big\}\\
&\quad=\big\{[\mathbf{x}_s]_r(\cdot)^\top \dot{P}(r)\bar{\mathcal{X}}(\cdot,r)+[\mathbf{x}_s]_r(\cdot)^\top P(r)[A(\cdot,r)\bar{\mathcal{X}}(r,r)+B(\cdot,r)\Theta(r)\bar{\mathcal{X}}(\cdot,r)]\big\}dr\\
&\qquad+[\mathbf{x}_s]_r(\cdot)^\top P(r)\big[C(\cdot,r)\bar{\mathcal{X}}(r,r)+D(\cdot,r)\Theta(r)\bar{\mathcal{X}}(\cdot,r)\big]dW(r)\\
&\quad=-\big\{\mathbf{x}_s(r)^\top A(\cdot,r)^\top P(r)\bar{\mathcal{X}}(\cdot,r)+\mathbf{x}_s(r)^\top C(\cdot,r)^\top P(r)C(\cdot,r)\bar{\mathcal{X}}(r,r)\\
&\qquad+\mathbf{x}_s(r)^\top Q(r)\bar{\mathcal{X}}(r,r)+\mathbf{x}_s(r)^\top C(\cdot,r)^\top P(r)D(\cdot,r)\Theta(r)\bar{\mathcal{X}}(\cdot,r)\big\}dr\\
&\qquad+[\mathbf{x}_s]_r(\cdot)^\top P(r)\big[C(\cdot,r)\bar{\mathcal{X}}(r,r)+D(\cdot,r)\Theta(r)\bar{\mathcal{X}}(\cdot,r)\big]dW(r),\qquad r\in[s,T].
\end{aligned}$$

Combining the above with (5.7), we have

$$\begin{aligned}
&[\mathbf{x}_s]_r(\cdot)^\top P(r)\bar{\mathcal{X}}(\cdot,r)-\mathbf{x}_s(T)^\top\mathbb{E}_r[G\bar{\mathcal{X}}(T,T)]\\
&\quad=\int_r^T\Big\{\mathbf{x}_s(\tau)^\top A(\cdot,\tau)^\top P(\tau)\bar{\mathcal{X}}(\cdot,\tau)+\mathbf{x}_s(\tau)^\top C(\cdot,\tau)^\top P(\tau)C(\cdot,\tau)\bar{\mathcal{X}}(\tau,\tau)\\
&\qquad+\mathbf{x}_s(\tau)^\top Q(\tau)\bar{\mathcal{X}}(\tau,\tau)+\mathbf{x}_s(\tau)^\top C(\cdot,\tau)^\top P(\tau)D(\cdot,\tau)\Theta(\tau)\bar{\mathcal{X}}(\cdot,\tau)\Big\}d\tau\\
&\qquad-\int_r^T\Big\{[\mathbf{x}_s]_\tau(\cdot)^\top P(\tau)[C(\cdot,\tau)\bar{\mathcal{X}}(\tau,\tau)+D(\cdot,\tau)\Theta(\tau)\bar{\mathcal{X}}(\cdot,\tau)]\\
&\qquad-\mathbf{x}_s(T)^\top\zeta(\tau)\Big\}dW(\tau).
\end{aligned}\tag{5.9}$$

For $*=A,B,C,D$, denote

$$\begin{aligned}
&Y^*(s)=*(\cdot,s)^\top P(s)\bar{\mathcal{X}}(\cdot,s)-*(T,s)^\top\mathbb{E}_s[G\bar{\mathcal{X}}(T,T)],\\
&Z^*(s,\tau)=[*(\cdot,s)]_\tau^\top P(\tau)\big[C(\cdot,\tau)\bar{\mathcal{X}}(\tau,\tau)+D(\cdot,\tau)\Theta(\tau)\bar{\mathcal{X}}(\cdot,\tau)\big]\\
&\qquad\qquad-*(T,s)^\top\zeta(\tau).
\end{aligned}\tag{5.10}$$

Taking $\mathbf{x}_s(\cdot)=*(\cdot,s)$ and $r=s$ in (5.9), we have

$$\begin{aligned}
Y^*(s)&=\int_s^T\Big\{*(\tau,s)^\top Y^A(\tau)+*(\tau,s)^\top\mathbb{E}_\tau[A(T,\tau)^\top G\bar{\mathcal{X}}(T,T)]+*(\tau,s)^\top Z^C(\tau,\tau)\\
&\qquad+*(\tau,s)^\top C(T,\tau)^\top\zeta(\tau)+*(\tau,s)^\top Q(\tau)\bar{\mathcal{X}}(\tau,\tau)\Big\}d\tau-\int_s^T Z^*(s,\tau)dW(\tau)\\
&=\int_s^T *(\tau,s)^\top\Big[\mathbb{E}_\tau[\psi(\tau)]+Y^A(\tau)+Z^C(\tau,\tau)\Big]d\tau-\int_s^T Z^*(s,\tau)dW(\tau),
\end{aligned}\tag{5.11}$$

where $\psi(\cdot)$ is defined by (5.8). Thus, by Proposition 3.5, the process $(Y^*(\cdot),Z^*(\cdot,\cdot))$ (with $*=A,B,C,D$) defined by (5.10) is the unique adapted solution to the type-III BSVIE (3.13). Let $(Y(\cdot),Z(\cdot,\cdot),Y^0(\cdot),Z^0(\cdot,\cdot))$ be the unique adapted solution to the following type-II BSVIE:

$$\begin{cases}
Y(s)=\psi(s)+\displaystyle\int_s^T\big[A(\tau,s)^\top Y(\tau)+C(\tau,s)^\top Z(\tau,s)\big]d\tau-\int_s^T Z(s,\tau)dW(\tau),\\
Y^0(s)=\psi^0(s)+\displaystyle\int_s^T\big[B(\tau,s)^\top Y(\tau)+D(\tau,s)^\top Z(\tau,s)\big]d\tau-\int_s^T Z^0(s,\tau)dW(\tau).
\end{cases}$$

Then by Proposition 3.5 again, we have

$$Y(s)=\mathbb{E}_s[\psi(s)]+Y^A(s)+Z^C(s,s),\qquad Y^0(s)=\mathbb{E}_s[\psi^0(s)]+Y^B(s)+Z^D(s,s).$$

Substituting (5.10) into the above, we get

$$\begin{aligned}Y(s) &= Q(s)\bar{\mathcal{X}}(s,s) + A(T,s)^\top G\mathbb{E}_s[\bar{\mathcal{X}}(T,T)] + C(T,s)^\top\zeta(s) + A(\cdot\,,s)^\top P(s)\bar{\mathcal{X}}(\cdot\,,s)\\ &\quad - A(T,s)^\top\mathbb{E}_s[G\bar{\mathcal{X}}(T,T)] + C(\cdot\,,s)^\top P(s)[C(\cdot\,,s)\bar{\mathcal{X}}(s,s) + D(\cdot\,,s)\Theta(s)\bar{\mathcal{X}}(\cdot\,,s)]\\ &\quad - C(T,s)^\top\zeta(s)\\ &= A(\cdot\,,s)^\top P(s)\bar{\mathcal{X}}(\cdot\,,s) + C(\cdot\,,s)^\top P(s)C(\cdot\,,s)\bar{\mathcal{X}}(s,s) + Q(s)\bar{\mathcal{X}}(s,s)\\ &\quad + C(\cdot\,,s)^\top P(s)D(\cdot\,,s)\Theta(s)\bar{\mathcal{X}}(\cdot\,,s), \qquad s\in[t,T],\end{aligned}$$

and

$$\begin{aligned}Y^0(s) &= B(T,s)^\top G\mathbb{E}_s[\bar{\mathcal{X}}(T,T)] + D(T,s)^\top\zeta(s) + B(\cdot\,,s)^\top P(s)\bar{\mathcal{X}}(\cdot\,,s)\\ &\quad - B(T,s)^\top\mathbb{E}_s[G\bar{\mathcal{X}}(T,T)] + D(\cdot\,,s)^\top P(s)[C(\cdot\,,s)\bar{\mathcal{X}}(s,s)\\ &\quad + D(\cdot\,,s)\Theta(s)\bar{\mathcal{X}}(\cdot\,,s)] - D(T,s)^\top\zeta(s)\\ &= B(\cdot\,,s)^\top P(s)\bar{\mathcal{X}}(\cdot\,,s) + D(\cdot\,,s)^\top P(s)[C(\cdot\,,s)\bar{\mathcal{X}}(s,s) + D(\cdot\,,s)\Theta(s)\bar{\mathcal{X}}(\cdot\,,s)]\\ &= -R(s)\Theta(s)\bar{\mathcal{X}}(\cdot\,,s), \qquad s\in[t,T].\end{aligned}$$

Then the desired results can be obtained easily. ∎

Remark 5.4. Note that the functional $(r,\mathbf{x}_r(\cdot))\mapsto\dot{P}(r)(\mathbf{x}_r(\cdot),\mathbf{x}_r(\cdot))$ is not continuous if $Q(\cdot)$ or $R(\cdot)$ is not a continuous function. However, the map $r\mapsto P(r)(\mathbf{x}_r(\cdot),\mathbf{x}_r(\cdot))$ is absolutely continuous, and from (4.10) the maps $r\mapsto\dot{P}(r)\big([\mathbf{x}]_r(\cdot),\bar{\mathcal{X}}(\cdot\,,r)\big)$ and $r\mapsto\dot{P}(r)\big(\bar{\mathcal{X}}(\cdot\,,r),\bar{\mathcal{X}}(\cdot\,,r)\big)$ are progressively measurable and square integrable on $[t,T]$. Thus, the functional Itô formula still holds true here. In what follows, we will always apply the functional Itô formula to this case without the explanation.

Let $\mathscr{X}_s^*$ be the dual space of $\mathscr{X}_s\equiv C([s,T];\mathbb{R}^n)$ that is, the space consisting of all the bounded linear functionals on $\mathscr{X}_s$. Clearly, for a.s. $\omega\in\Omega$,

$$\{\mathbb{E}_s[Y(\cdot)]\mathbf{1}_{[s,T]}(\cdot)\}(\omega)\in\mathscr{X}_s^* \qquad\text{and}\qquad \{Z(\cdot,s)\mathbf{1}_{[s,T]}(\cdot)\}(\omega)\in\mathscr{X}_s^*,$$

by letting

$$\begin{aligned}\langle\mathbb{E}_s[Y(\cdot)]\mathbf{1}_{[s,T]}(\cdot),\mathbf{x}_s(\cdot)\rangle(\omega) &\triangleq \int_s^T\mathbf{x}_s(r)^\top\mathbb{E}_s[Y(r)](\omega)dr = \left[\mathbb{E}_s\int_s^T\mathbf{x}_s(r)^\top Y(r)dr\right](\omega),\\ \langle Z(\cdot,s)\mathbf{1}_{[s,T]}(\cdot),\mathbf{x}_s(\cdot)\rangle(\omega) &\triangleq \left[\int_s^T\mathbf{x}_s(r)^\top Z(r,s)dr\right](\omega) \qquad \forall\mathbf{x}_s(\cdot)\in\mathscr{X}_s.\end{aligned}$$

By the similar arguments to that employed in the proof of Theorem 5.1, we get the following representation for $(\mathbb{E}_s[Y(\cdot)]\mathbf{1}_{[s,T]}(\cdot),\, Z(\cdot,s)\mathbf{1}_{[s,T]}(\cdot))$ in the space $\mathscr{X}_s^*$.

THEOREM 5.5. *For any $s\in[t,T]$, the following equalities hold:*

$$\begin{aligned}&\mathbb{E}_s[Y(\cdot)]\mathbf{1}_{[s,T]}(\cdot) = P(s)\bar{\mathcal{X}}(\cdot,s) - \delta_T^\top\mathbb{E}_s[G\bar{\mathcal{X}}(T,T)],\\ &Z(\cdot,s)\mathbf{1}_{[s,T]}(\cdot) = P(s)C(\cdot,s)\bar{\mathcal{X}}(s,s) + P(s)D(\cdot,s)\Theta(s)\bar{\mathcal{X}}(\cdot,s) - \delta_T^\top\zeta(s),\end{aligned}$$

in the space $\mathscr{X}_s^*$*; that is for any* $\mathbf{x}_s(\cdot)\in\mathscr{X}_s$ *and a.s.* $\omega\in\Omega$*,*

$$\begin{aligned}\left[\mathbb{E}_s\int_s^T \mathbf{x}_s(r)^\top Y(r)dr\right](\omega)&=\big\{\mathbf{x}_s(\cdot)^\top P(s)\bar{\mathcal{X}}(\cdot,s)-\mathbf{x}_s(T)^\top\mathbb{E}_s[G\bar{\mathcal{X}}(T,T)]\big\}(\omega),\\ \left[\int_s^T \mathbf{x}_s(r)^\top Z(r,s)dr\right](\omega)&=\big\{\mathbf{x}_s(\cdot)^\top P(s)C(\cdot,s)\bar{\mathcal{X}}(s,s)-\mathbf{x}_s(T)^\top\zeta(s)\\ &\qquad+\mathbf{x}_s(\cdot)^\top P(s)D(\cdot,s)\Theta(s)\bar{\mathcal{X}}(\cdot,s)\big\}(\omega).\end{aligned}$$

Remark 5.6. If we let $G=0$, then

$$\begin{aligned}&\mathbb{E}_s[Y(\cdot)]\mathbf{1}_{[s,T]}(\cdot)=P(s)\bar{\mathcal{X}}(\cdot,s),\\ &Z(\cdot,s)\mathbf{1}_{[s,T]}(\cdot)=P(s)C(\cdot,s)\bar{\mathcal{X}}(s,s)+P(s)D(\cdot,s)\Theta(s)\bar{\mathcal{X}}(\cdot,s),\end{aligned}$$

which is very similar to the results of classical stochastic LQ control problems (see [40, Chapter 6], for example). Thus, it should be more natural to regard the solution $(Y(\cdot),Z(\cdot,\cdot))$ as a bounded linear functional on $\mathscr{X}_s; s\in[0,T]$.

6. Well-posedness of the path-dependent Riccati equation. In this section, we shall establish the well-posedness of path-dependent Riccati equation (4.10) which is rewritten here, for convenience:

(6.1)
$$\begin{cases}\dot P(t)+P(t)A(\cdot,t)\delta_t+\delta_t^\top A(\cdot,t)^\top P(t)+\delta_t^\top C(\cdot,t)^\top P(t)C(\cdot,t)\delta_t+\delta_t^\top Q(t)\delta_t\\ \qquad-\big[D(\cdot,t)^\top P(t)C(\cdot,t)\delta_t+B(\cdot,t)^\top P(t)\big]^\top\big[D(\cdot,t)^\top P(t)D(\cdot,t)+R(t)\big]^{-1}\\ \qquad\times\big[D(\cdot,t)^\top P(t)C(\cdot,t)\delta_t+B(\cdot,t)^\top P(t)\big]=0,\qquad t\in[0,T],\\ P(T)=G.\end{cases}$$

We assume the following condition.

(H4) The weighting matrices $Q(\cdot)$, $R(\cdot)$, and G satisfy

$$Q(t)\ge 0,\qquad R(t)\ge\lambda I_m,\qquad t\in[0,T];\qquad G\ge 0,\tag{6.2}$$

where $\lambda>0$ is a given constant.

THEOREM 6.1. *Let* (H1)–(H2) *and* (H4) *hold. Then the path-dependent Riccati equation* (6.1) *admits a unique strongly regular solution* $P(\cdot)\in C([0,T];\mathcal{S}^n)$*. Moreover,*

$$P(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big)\ge 0\qquad\forall(t,\mathbf{x}_t(\cdot))\in\Lambda.$$

Remark 6.2. For the corresponding results of Theorem 6.1 in the SDE setting, we refer the reader to [40, Chapter 6], in which (H4) was called a *standard condition.* It is known that (H4) implies the uniformly convexity condition (H3). An interesting question is whether Riccati equation (6.1) has the well-posedness under (H3), as shown by Sun, Li, and Yong [29] in the SDE setting. We shall explore that in the near future. We highlight that in the recent work [15] by Hamaguchi and Wang, they provide another method of constructing the optimal causal feedback strategy under the uniform convexity condition.

To establish the well-posedness of (6.1), we introduce the following *path-dependent Lyapunov equation*:

$$
\begin{cases} \dot{P}(t) + P(t)A(\cdot\,,t)\delta_t + \delta_t^\top A(\cdot\,,t)^\top P(t) + \delta_t^\top C(\cdot\,,t)^\top P(t) C(\cdot\,,t)\delta_t \\ \quad + \delta_t^\top Q(t)\delta_t - \big[P(t)B(\cdot,t) + \delta_t^\top C(\cdot,t)^\top P(t)D(\cdot,t)\big]\Psi(t) \\ \quad - \Psi(t)^\top \big[B(\cdot,t)^\top P(t) + D(\cdot,t)^\top P(t) C(\cdot,t)\delta_t\big] \\ \quad + \Psi(t)^\top \big[R(t) + D(\cdot\,,t)^\top P(t) D(\cdot\,,t)\big]\Psi(t) = 0, \qquad t \in [0,T], \\ P(T) = G, \end{cases} \tag{6.3}
$$

where $\Psi(\cdot) \in L^\infty(0,T;\mathcal{L}(\mathscr{X}_t;\mathbb{R}^m))$. As pointed out in Remark 2.7, we assume that the map $(t,\omega) \mapsto \Psi(t)\mathcal{X}(\cdot,t)$ is progressively measurable for any $\mathcal{X}(\cdot,\cdot) \in L^2_{\mathbb{F}}(\Omega;C(\Delta_*[0,T];\mathbb{R}^n))$. For any $\mathbf{x}(\cdot) \in \mathscr{X}_0$, the map $t \mapsto \Psi(t)[\mathbf{x}]_t(\cdot)$ is in $L^\infty(0,T;\mathbb{R}^m)$.

We call $P(\cdot) \in C([0,T];\mathcal{S}^n)$ a solution of (6.3) if the map $t \mapsto P(t)\big([\mathbf{x}]_t(\cdot),[\mathbf{x}]_t(\cdot)\big)$ is absolutely continuous for any $\mathbf{x}(\cdot) \in \mathscr{X}_0$, and it satisfies (6.3) for almost everywhere $t \in [0,T]$. Clearly, noting that $P(\cdot)$ is symmetric, (6.3) is equivalent to that for any $\mathbf{x}(\cdot),\mathbf{x}'(\cdot) \in \mathscr{X}_0$, the following holds on $[0,T]$:

$$
\begin{aligned}
&[\mathbf{x}]_t(\cdot)^\top P(t)[\mathbf{x}']_t(\cdot) = \mathbf{x}(T)^\top G\mathbf{x}'(T) + \int_t^T \Big\{[\mathbf{x}]_s(\cdot)^\top P(s)A(\cdot,s)\mathbf{x}'(s) \\
&\quad + \mathbf{x}(s)^\top A(\cdot\,,s)^\top P(s)[\mathbf{x}']_s(\cdot) + \mathbf{x}(s)^\top C(\cdot\,,s)^\top P(s)C(\cdot\,,s)\mathbf{x}'(s) \\
&\quad + \mathbf{x}(s)^\top Q(s)\mathbf{x}'(s) - [\mathbf{x}]_s(\cdot)^\top P(s)B(\cdot\,,s)\Psi(s)[\mathbf{x}']_s(\cdot) - \mathbf{x}(s)^\top C(\cdot\,,s)^\top \\
&\quad \times P(s)D(\cdot\,,s)\Psi(s)[\mathbf{x}']_s(\cdot) - \{\Psi(s)[\mathbf{x}]_s(\cdot)\}^\top B(\cdot\,,s)^\top P(s)[\mathbf{x}']_s(\cdot) \\
&\quad - \{\Psi(s)[\mathbf{x}]_s(\cdot)\}^\top D(\cdot\,,s)^\top P(s)C(\cdot\,,s)\mathbf{x}'(s) + \{\Psi(s)[\mathbf{x}]_s(\cdot)\}^\top R(s) \\
&\quad \times \{\Psi(s)[\mathbf{x}']_s(\cdot)\} + \{\Psi(s)[\mathbf{x}]_s(\cdot)\}^\top D(\cdot\,,s)^\top P(s)D(\cdot\,,s)\{\Psi(s)[\mathbf{x}']_s(\cdot)\}\Big\}ds.
\end{aligned}
$$

LEMMA 6.3. *Let* (H1)–(H2) *hold. Then the path-dependent Lyapunov equation* (6.3) *admits a unique solution* $P(\cdot) \in C([0,T];\mathcal{S}^n)$. *If, in addition,* (H4) *holds, then we have*

$$
P(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big) \geq 0 \qquad \forall (t,\mathbf{x}_t(\cdot)) \in \Lambda. \tag{6.4}
$$

Proof. We first prove the existence and the uniqueness of the solution to Lyapunov equation (6.3). For any $\varepsilon \in (0,T]$ and $\bar{P}(\cdot) \in C([T-\varepsilon,T];\mathcal{S}^n)$, denote the operator-valued function $P(\cdot)$ on $[T-\varepsilon,T]$ by

$$
\begin{aligned}
&[\mathbf{x}]_t(\cdot)^\top P(t)[\mathbf{x}']_t(\cdot) = \mathbf{x}(T)^\top G\mathbf{x}'(T) + \int_t^T \Big\{[\mathbf{x}]_s(\cdot)^\top \bar{P}(s)A(\cdot,s)\mathbf{x}'(s) \\
&\quad + \mathbf{x}(s)^\top A(\cdot\,,s)^\top \bar{P}(s)[\mathbf{x}']_s(\cdot) + \mathbf{x}(s)^\top C(\cdot\,,s)^\top \bar{P}(s)C(\cdot\,,s)\mathbf{x}'(s) \\
&\quad + \mathbf{x}(s)^\top Q(s)\mathbf{x}'(s) - [\mathbf{x}]_s(\cdot)^\top \bar{P}(s)B(\cdot\,,s)\Psi(s)[\mathbf{x}']_s(\cdot) - \mathbf{x}(s)^\top C(\cdot\,,s)^\top \\
&\quad \times \bar{P}(s)D(\cdot\,,s)\Psi(s)[\mathbf{x}']_s(\cdot) - \{\Psi(s)[\mathbf{x}]_s(\cdot)\}^\top B(\cdot\,,s)^\top \bar{P}(s)[\mathbf{x}']_s(\cdot) \\
&\quad - \{\Psi(s)[\mathbf{x}]_s(\cdot)\}^\top D(\cdot\,,s)^\top \bar{P}(s)C(\cdot\,,s)\mathbf{x}'(s) + \{\Psi(s)[\mathbf{x}]_s(\cdot)\}^\top R(s) \\
&\quad \times \{\Psi(s)[\mathbf{x}']_s(\cdot)\} + \{\Psi(s)[\mathbf{x}]_s(\cdot)\}^\top D(\cdot\,,s)^\top \bar{P}(s)D(\cdot\,,s)\{\Psi(s)[\mathbf{x}']_s(\cdot)\}\Big\}ds.
\end{aligned} \tag{6.5}
$$

From the fact that $\bar{P}(\cdot)$ is symmetric, we see that $P(\cdot) \in C([T-\varepsilon,T];\mathcal{S}^n)$. Thus, the map $\varGamma : C([T-\varepsilon,T];\mathcal{S}^n) \to C([T-\varepsilon,T];\mathcal{S}^n)$, given by

$$
P(\cdot) = \varGamma\big(\bar{P}(\cdot)\big),
$$

is well defined. For any $\bar{P}_1(\cdot)$, $\bar{P}_2(\cdot) \in C([T-\varepsilon,T];\mathcal{S}^n)$, denote

$$
P_i(\cdot) = \varGamma(\bar{P}_i(\cdot)), \qquad \varDelta P(\cdot) = P_1(\cdot) - P_2(\cdot), \qquad \varDelta\bar{P}(\cdot) = \bar{P}_1(\cdot) - \bar{P}_2(\cdot).
$$

Then from (6.5), taking $\mathbf{x}'(\cdot)=\mathbf{x}(\cdot)$, we have

$$
\begin{aligned}
&[\mathbf{x}]_t(\cdot)^\top \mathit{\Delta} P(t)[\mathbf{x}]_t(\cdot)=\int_t^T \Big\{2[\mathbf{x}]_s(\cdot)^\top \mathit{\Delta}\bar P(s)A(\cdot\,,s)\mathbf{x}(s)-2[\mathbf{x}]_s(\cdot)^\top \mathit{\Delta}\bar P(s)B(\cdot\,,s)\\
&\quad\times\Psi(s)[\mathbf{x}]_s(\cdot)+\mathbf{x}(s)^\top C(\cdot\,,s)^\top \mathit{\Delta}\bar P(s)C(\cdot\,,s)\mathbf{x}(s)-2\mathbf{x}(s)^\top C(\cdot\,,s)^\top \mathit{\Delta}\bar P(s)\\
&\quad\times D(\cdot\,,s)\Psi(s)[\mathbf{x}]_s(\cdot)+\{\Psi(s)[\mathbf{x}]_s(\cdot)\}^\top D(\cdot\,,s)^\top \mathit{\Delta}\bar P(s)D(\cdot\,,s)\{\Psi(s)[\mathbf{x}]_s(\cdot)\}\Big\}ds.
\end{aligned}
$$

Thus, by Lemma 2.8, we get

$$
\sup_{s\in[T-\varepsilon,T]}\|\mathit{\Delta} P(s)\|\le K\varepsilon \sup_{s\in[T-\varepsilon,T]}\|\mathit{\Delta}\bar P(s)\|,
$$

where $K>0$, only depending on the norms of the coefficients, is a fixed constant. Then when $\varepsilon>0$ is small, by the fixed point theorem, the Lyapunov equation (6.3) admits a unique solution $\hat P(\cdot)\in C([T-\varepsilon,T];\mathcal{S}^n)$ over $[T-\varepsilon,T]$. Next, on $[0,T-\varepsilon]$, we consider the following equation:

$$
\begin{aligned}
&[\mathbf{x}]_t(\cdot)^\top P(t)[\mathbf{x}']_t(\cdot)=[\mathbf{x}]_{T-\varepsilon}(\cdot)^\top \hat P(T-\varepsilon)[\mathbf{x}']_{T-\varepsilon}(\cdot)\\
&\quad+\int_t^{T-\varepsilon}\Big\{[\mathbf{x}]_s(\cdot)^\top P(s)A(\cdot,s)\mathbf{x}'(s)+\mathbf{x}(s)^\top A(\cdot\,,s)^\top P(s)[\mathbf{x}']_s(\cdot)+\mathbf{x}(s)^\top C(\cdot\,,s)^\top\\
&\quad\times P(s)C(\cdot\,,s)\mathbf{x}'(s)+\mathbf{x}(s)^\top Q(s)\mathbf{x}'(s)-[\mathbf{x}]_s(\cdot)^\top P(s)B(\cdot\,,s)\Psi(s)[\mathbf{x}']_s(\cdot)\\
&\quad-\mathbf{x}(s)^\top C(\cdot\,,s)^\top P(s)D(\cdot\,,s)\Psi(s)[\mathbf{x}']_s(\cdot)-\{\Psi(s)[\mathbf{x}]_s(\cdot)\}^\top B(\cdot\,,s)^\top P(s)[\mathbf{x}']_s(\cdot)\\
&\quad-\{\Psi(s)[\mathbf{x}]_s(\cdot)\}^\top D(\cdot\,,s)^\top P(s)C(\cdot\,,s)\mathbf{x}'(s)+\{\Psi(s)[\mathbf{x}]_s(\cdot)\}^\top R(s)\{\Psi(s)[\mathbf{x}']_s(\cdot)\}\\
&\quad+\{\Psi(s)[\mathbf{x}]_s(\cdot)\}^\top D(\cdot\,,s)^\top P(s)D(\cdot,s)\{\Psi(s)[\mathbf{x}']_s(\cdot)\}\Big\}ds.
\end{aligned}
$$

Note that in the above, the integral is from t to $T-\varepsilon$, and for any $\mathbf{x}(\cdot),\mathbf{x}'(\cdot)\in\mathscr{X}_0$, the map $s\mapsto[\mathbf{x}]_s(\cdot)^\top P(s)[\mathbf{x}']_s(\cdot)$ is only considered on $[0,T-\varepsilon]$. Moreover, we see that for any $t\in[0,T-\varepsilon]$, $P(t)$ is still defined on $\mathscr{X}_t\times\mathscr{X}_t\equiv C([t,T];\mathbb{R}^n)\times C([t,T];\mathbb{R}^n)$. Thus, using the same arguments as above, we can prove that the above equation admits a unique solution in $C([T-2\varepsilon,T-\varepsilon];\mathcal{S}^n)$. Then we can complete the proof by induction.

We now prove (6.4). Let $\mathcal{X}(\cdot\,,\cdot)$ be the unique solution to the following SVIE:

$$
\begin{aligned}
\mathcal{X}(s,r)=\mathbf{x}_t(s)&+\int_t^r[A(s,\tau)\mathcal{X}(\tau,\tau)-B(s,\tau)\Psi(\tau)\mathcal{X}(\cdot\,,\tau)]d\tau\\
&+\int_t^r[C(s,\tau)\mathcal{X}(\tau,\tau)-D(s,\tau)\Psi(\tau)\mathcal{X}(\cdot\,,\tau)]dW(\tau),\qquad (s,r)\in\Delta_*[t,T].
\end{aligned}
$$

Then by the functional Itô formula (see Proposition 2.12), we have

$$
\begin{aligned}
P(s)\big(\mathcal{X}(\cdot\,,s),\mathcal{X}(\cdot,s)\big)=\mathbb{E}_s\bigg\{&\int_s^T\big[\langle Q(r)\mathcal{X}(r,r),\mathcal{X}(r,r)\rangle\\
&+\langle R(r)\Psi(r)\mathcal{X}(\cdot\,,r),\Psi(r)\mathcal{X}(\cdot\,,r)\rangle\big]dr+\langle G\mathcal{X}(T,T),\mathcal{X}(T,T)\rangle\bigg\},\qquad s\in[t,T].
\end{aligned}
$$

In particular, taking $s=t$, we have (noting $\mathcal{X}(\cdot,t)=\mathbf{x}_t(\cdot)$)

$$P(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big)=\mathbb{E}_t\Bigg\{\int_t^T\big[\langle Q(r)\mathcal{X}(r,r),\mathcal{X}(r,r)\rangle+\langle R(r)\Psi(r)\mathcal{X}(\cdot,r),\Psi(r)\mathcal{X}(\cdot,r)\rangle\big]dr+\langle G\mathcal{X}(T,T),\mathcal{X}(T,T)\rangle\Bigg\}\geq 0\qquad\forall(t,\mathbf{x}_t(\cdot))\in\Lambda,$$

where the last inequality is due to the facts $G\geq 0$, $Q(\cdot)\geq 0$, and $R(\cdot)\geq 0$. ∎

Proof of Theorem 6.1. The uniqueness of the solution to Riccati equation (6.1) can be obtained by a standard method. We now prove the existence of a solution to Riccati equation (6.1) by a iterative method. For any $t\in[0,T]$, we denote

$$\Psi(t)=[D(\cdot,t)^\top P(t)D(\cdot,t)+R(t)]^{-1}[D(\cdot,t)^\top P(t)C(\cdot,t)\delta_t+B(\cdot,t)^\top P(t)].$$

Then from the fact

$$[D(\cdot,t)^\top P(t)D(\cdot,t)+R(t)]\Psi(t)=[D(\cdot,t)^\top P(t)C(\cdot,t)\delta_t+B(\cdot,t)^\top P(t)],\quad t\in[0,T],$$

it is easily checked that (6.1) is equivalent to the following:

$$\begin{cases}\dot P(t)+P(t)A(\cdot,t)\delta_t+\delta_t^\top A(\cdot,t)^\top P(t)-P(t)B(\cdot,t)\Psi(t)-\Psi(t)^\top B(\cdot,t)^\top P(t)\\ \quad+\delta_t^\top C(\cdot,t)^\top P(t)C(\cdot,t)\delta_t-\delta_t^\top C(\cdot,t)^\top P(t)D(\cdot,t)\Psi(t)\\ \quad-\Psi(t)^\top D(\cdot,t)^\top P(t)C(\cdot,t)\delta_t+\Psi(t)^\top D(\cdot,t)^\top P(t)D(\cdot,t)\Psi(t)\\ \quad+\delta_t^\top Q(t)\delta_t+\Psi(t)^\top R(t)\Psi(t)=0,\\ P(T)=G.\end{cases}$$

By Lemma 6.3, with $\Psi(\cdot)\equiv 0$, the following equation admits a unique solution $P_0(\cdot)$:

$$\begin{cases}\dot P_0(t)+P_0(t)A(\cdot,t)\delta_t+\delta_t^\top A(\cdot,t)^\top P_0(t)+\delta_t^\top C(\cdot,t)^\top P_0(t)C(\cdot,t)\delta_t+\delta_t^\top Q(t)\delta_t=0,\\ P_0(T)=G.\end{cases}$$

Moreover, $P_0(\cdot)$ satisfies

$$P_0(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big)\geq 0\qquad\forall(t,\mathbf{x}_t(\cdot))\in\Lambda,$$

which, together with (H4), implies

$$R(t)+D(\cdot,t)^\top P_0(t)D(\cdot,t)\geq\lambda I_m,\qquad t\in[0,T].$$

For $i=0,1,2,\ldots$, define

$$\Psi_i(t)=[D(\cdot,t)^\top P_i(t)D(\cdot,t)+R(t)]^{-1}[D(\cdot,t)^\top P_i(t)C(\cdot,t)\delta_t+B(\cdot,t)^\top P_i(t)],\tag{6.6}$$

with $P_i(\cdot)$ being the unique solution to the following Lyapunov equation:

$$\begin{cases}\dot P_{i+1}(t)+P_{i+1}(t)A(\cdot,t)\delta_t+\delta_t^\top A(\cdot,t)^\top P_{i+1}(t)-P_{i+1}(t)B(\cdot,t)\Psi_i(t)\\ \quad-\Psi_i(t)^\top B(\cdot,t)^\top P_{i+1}(t)+\delta_t^\top C(\cdot,t)^\top P_{i+1}(t)C(\cdot,t)\delta_t\\ \quad-\delta_t^\top C(\cdot,t)^\top P_{i+1}(t)D(\cdot,t)\Psi_i(t)-\Psi_i(t)^\top D(\cdot,t)^\top P_{i+1}(t)C(\cdot,t)\delta_t\\ \quad+\Psi_i(t)^\top D(\cdot,t)^\top P_{i+1}(t)D(\cdot,t)\Psi_i(t)+\delta_t^\top Q(t)\delta_t+\Psi_i(t)^\top R(t)\Psi_i(t)=0,\\ P_{i+1}(T)=G,\end{cases}\tag{6.7}$$

which is equivalent to

$$
\begin{aligned}
&[\mathbf{x}]_t(\cdot)^\top P_{i+1}(t)[\mathbf{x}']_t(\cdot) = \mathbf{x}(T)^\top G\mathbf{x}'(T) + \int_t^T \Big\{[\mathbf{x}]_s(\cdot)^\top P_{i+1}(s)A(\cdot,s)\mathbf{x}'(s) \\
&\quad + \mathbf{x}(s)^\top A(\cdot\,,s)^\top P_{i+1}(s)[\mathbf{x}']_s(\cdot) + \mathbf{x}(s)^\top C(\cdot\,,s)^\top P_{i+1}(s)C(\cdot\,,s)\mathbf{x}'(s) \\
&\quad + \mathbf{x}(s)^\top Q(s)\mathbf{x}'(s) - [\mathbf{x}]_s(\cdot)^\top P_{i+1}(s)B(\cdot\,,s)\Psi_i(s)[\mathbf{x}']_s(\cdot) \\
&\quad - \mathbf{x}(s)^\top C(\cdot\,,s)^\top P_{i+1}(s)D(\cdot\,,s)\Psi_i(s)[\mathbf{x}']_s(\cdot) - \{\Psi_i(s)[\mathbf{x}]_s(\cdot)\}^\top B(\cdot\,,s)^\top \\
&\quad \times P_{i+1}(s)[\mathbf{x}']_s(\cdot) - \{\Psi_i(s)[\mathbf{x}]_s(\cdot)\}^\top D(\cdot\,,s)^\top P_{i+1}(s)C(\cdot\,,s)\mathbf{x}'(s) \\
&\quad + \{\Psi_i(s)[\mathbf{x}]_s(\cdot)\}^\top R(s)\{\Psi_i(s)[\mathbf{x}']_s(\cdot)\} + \{\Psi_i(s)[\mathbf{x}]_s(\cdot)\}^\top D(\cdot\,,s)^\top \\
&\quad \times P_{i+1}(s)D(\cdot,s)\{\Psi_i(s)[\mathbf{x}']_s(\cdot)\}\Big\}ds, \qquad t\in[0,T] \quad \forall \mathbf{x}(\cdot),\mathbf{x}'(\cdot)\in\mathscr{X}_0.
\end{aligned}
\tag{6.8}
$$

It is easily checked that the above operator-valued function $\Psi_i(\cdot)$ satisfies the assumption imposed for $\Psi(\cdot)$ in the Lyapunov equation (6.3). Then by Lemma 6.3 again, we have

$$P_i(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big)\geq 0 \qquad \forall (t,\mathbf{x}_t(\cdot))\in\Lambda, \quad i\geq 0.$$

It follows that

$$R(t) + D(\cdot,t)^\top P_i(t)D(\cdot,t) \geq \lambda I_m, \qquad t\in[0,T], \quad i\geq 0. \tag{6.9}$$

For $i\geq 1$, denote

$$\mathit{\Delta}_i(\cdot) = P_i(\cdot) - P_{i+1}(\cdot), \qquad \Pi_i(\cdot) = \Psi_i(\cdot) - \Psi_{i-1}(\cdot).$$

Then by (6.7), we get

$$
\begin{cases}
\dot{\mathit{\Delta}}_i(t) + \mathit{\Delta}_i(t)A(\cdot\,,t)\delta_t + \delta_t^\top A(\cdot\,,t)^\top \mathit{\Delta}_i(t) - \mathit{\Delta}_i(t)B(\cdot,t)\Psi_i(t) \\
\quad - \Psi_i(t)^\top B(\cdot,t)^\top \mathit{\Delta}_i(t) + P_i(t)B(\cdot,t)\Pi_i(t) + \Pi_i(t)^\top B(\cdot,t)^\top P_i(t) \\
\quad + \delta_t^\top C(\cdot,t)^\top \mathit{\Delta}_i(t)C(\cdot,t)\delta_t - \delta_t^\top C(\cdot,t)^\top \mathit{\Delta}_i(t)D(\cdot,t)\Psi_i(t) \\
\quad - \Psi_i(t)^\top D(\cdot,t)^\top \mathit{\Delta}_i(t)C(\cdot,t)\delta_t + \delta_t^\top C(\cdot,t)^\top P_i(t)D(\cdot,t)\Pi_i(t) \\
\quad + \Pi_i(t)^\top D(\cdot,t)^\top P_i(t)C(\cdot,t)\delta_t + \Psi_i(t)^\top D(\cdot\,,t)^\top \mathit{\Delta}_i(t)D(\cdot,t)\Psi_i(t) \\
\quad + \Psi_{i-1}(t)^\top D(\cdot\,,t)^\top P_i(t)D(\cdot\,,t)\Psi_{i-1}(t) - \Psi_i(t)^\top D(\cdot\,,t)^\top P_i(t)D(\cdot\,,t)\Psi_i(t) \\
\quad + \Psi_{i-1}(t)^\top R(t)\Psi_{i-1}(t) - \Psi_i(t)^\top R(t)\Psi_i(t) = 0, \\
\mathit{\Delta}_i(T) = 0.
\end{cases}
\tag{6.10}
$$

By some straightforward calculations, we have

$$
\begin{aligned}
&P_i(t)B(\cdot,t)\Pi_i(t) + \Pi_i(t)^\top B(\cdot,t)^\top P_i(t) + \delta_t^\top C(\cdot,t)^\top P_i(t)D(\cdot,t)\Pi_i(t) \\
&\qquad + \Pi_i(t)^\top D(\cdot,t)^\top P_i(t)C(\cdot,t)\delta_t + \Psi_{i-1}(t)^\top D(\cdot\,,t)^\top P_i(t)D(\cdot\,,t)\Psi_{i-1}(t) \\
&\qquad - \Psi_i(t)^\top D(\cdot\,,t)^\top P_i(t)D(\cdot\,,t)\Psi_i(t) + \Psi_{i-1}(t)^\top R(t)\Psi_{i-1}(t) - \Psi_i(t)^\top R(t)\Psi_i(t) \\
&\quad = P_i(t)B(\cdot,t)\Pi_i(t) + \Pi_i(t)^\top B(\cdot,t)^\top P_i(t) + \delta_t^\top C(\cdot,t)^\top P_i(t)D(\cdot,t)\Pi_i(t) \\
&\qquad + \Pi_i(t)^\top D(\cdot,t)^\top P_i(t)C(\cdot,t)\delta_t + \Pi_i(t)^\top [D(\cdot\,,t)^\top P_i(t)D(\cdot\,,t) + R(t)]\Pi_i(t) \\
&\qquad - \Pi_i(t)^\top [D(\cdot\,,t)^\top P_i(t)D(\cdot\,,t) + R(t)]\Psi_i(t) \\
&\qquad - \Psi_i(t)^\top [D(\cdot\,,t)^\top P_i(t)D(\cdot\,,t) + R(t)]\Pi_i(t) \\
&\quad = \Pi_i(t)^\top [D(\cdot\,,t)^\top P_i(t)D(\cdot\,,t) + R(t)]\Pi_i(t) \geq 0.
\end{aligned}
$$

Then by Lemma 6.3 again, (6.10) admits a unique solution $\varDelta_i(\cdot)$ satisfying $\varDelta_i(\cdot)\geq 0$. It follows that

$$P_i(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big)\geq P_{i+1}(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big)\geq 0 \qquad \forall (t,\mathbf{x}_t(\cdot))\in\Lambda, \quad i\geq 0. \tag{6.11}$$

Thus, for any $(t,\mathbf{x}_t(\cdot))\in\Lambda$, $\{P_i(t)(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot))\}_{i\geq 0}$ is a decreasing sequence, which is convergent as $i\to\infty$. Denote

$$P(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big)\triangleq\lim_{i\to\infty}P_i(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot)\big) \qquad \forall (t,\mathbf{x}_t(\cdot))\in\Lambda. \tag{6.12}$$

Note that

$$\begin{aligned}P_i(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}'_t(\cdot)\big)=&\ \frac{P_i(t)\big(\mathbf{x}_t(\cdot)+\mathbf{x}'_t(\cdot),\mathbf{x}_t(\cdot)+\mathbf{x}'_t(\cdot)\big)}{4}\\&-\frac{P_i(t)\big(\mathbf{x}_t(\cdot)-\mathbf{x}'_t(\cdot),\mathbf{x}_t(\cdot)-\mathbf{x}'_t(\cdot)\big)}{4}.\end{aligned}$$

Thus, the limit of $P_i(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}'_t(\cdot)\big)$ also exists as $i\to\infty$, and then for any $t\in[0,T]$, we can extend the domain of $P(t)$ to $\mathscr{X}_t\times\mathscr{X}_t$ by

$$P(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}'_t(\cdot)\big)=\lim_{i\to\infty}P_i(t)\big(\mathbf{x}_t(\cdot),\mathbf{x}'_t(\cdot)\big) \qquad \forall \mathbf{x}_t(\cdot),\mathbf{x}'_t(\cdot)\in\mathscr{X}_t.$$

Combining the above with (6.6) and (6.9) yields that

$$\lim_{i\to\infty}\Psi_i(t)=[D(\cdot\,,t)^\top P(t)D(\cdot\,,t)+R(t)]^{-1}[D(\cdot\,,t)^\top P(t)C(\cdot\,,t)\delta_t+B(\cdot\,,t)^\top P(t)]$$

for any $t\in[0,T]$. Moreover, by Lemma 2.8 and (6.11), we have

$$\sup_{i\geq 0}\sup_{t\in[0,T]}\|P_i(t)\|_{\mathcal{L}^2}\leq 2\sup_{i\geq 0}\sup_{t\in[0,T]}\|P_i(t)\|_{\mathcal{S}}\leq 2\sup_{t\in[0,T]}\|P_0(t)\|_{\mathcal{S}}<\infty,$$

and then

$$\sup_{i\geq 0}\sup_{t\in[0,T]}\|\Psi_i(t)\|\leq K\sup_{t\in[0,T]}\|P_0(t)\|_{\mathcal{S}}<\infty.$$

Taking $i\to\infty$ in (6.8), we see that $P(\cdot)$ satisfies (6.6), and hence (6.1). Moreover, by (6.9) we get that $P(\cdot)$ is a strongly regular solution. ∎

By Theorem 6.1, we show that under (H1)–(H2) and (H4), the decoupling field of the optimality system associated with Problem (LQ-FSVIE) really exists. Note that if, in addition, $Q(\cdot)$ and $R(\cdot)$ are continuous functions, then $P(\cdot)\in C^1([0,T];\mathcal{S}_n)$. Combining Theorems 4.2 and 6.1, we get the following result, in which the uniqueness of the classical solution to (4.5) is due to the uniqueness of the value function of Problem (LQ-FSVIE).

COROLLARY 6.4. *Let* (H1)–(H2) *and* (H4) *hold. In addition, we assume that* $Q(\cdot)$ *and* $R(\cdot)$ *are continuous functions. Then* $v(t,\mathbf{x}_t(\cdot))\triangleq\frac{P(t)(\mathbf{x}_t(\cdot),\mathbf{x}_t(\cdot))}{2}\in C_+^{1,2}(\Lambda)$*, and it is the unique classical solution to the path-dependent HJB equation* (4.5).

7. Conclusion. The most important contribution of this paper is that we develop a decoupling method (Theorems 5.1 and 5.5) for the optimality system associated with Problem (LQ-FSVIE), which is a linear coupled system of an FSVIE and a type-II BSVIE. As a result, the open-loop optimal control of Problem (LQ-FSVIE) can be represented as a causal feedback of the state process, and then an interesting phenomenon is found (Remark 5.3). The key techniques developed in the paper are establishing a link between the type-II and type-III BSVIEs (Proposition 3.5), deriving the associated path-dependent Riccati equation (Theorem 4.2), and proving the solvability of this new type of Riccati equation (Theorem 6.1). Such a Riccati-equation approach should have a big impact on the study of LQ control (or game) problems for SVIEs. It is also expected that by modifying this approach, one can decouple some general coupled FBSVIEs and solve some general control problems evolved by SVIEs. We will report the related results in our future publications.

Acknowledgments. The first author would like to thank Professor Jianfeng Zhang (of USC) for some useful suggestions. The authors would like to thank the anonymous referees and the associate editor for their constructive comments, which led to this improved version of the paper.

REFERENCES

[1] E. Abi Jaber, E. Miller, and H. Pham, *Linear-quadratic control for a class of stochastic Volterra equations: Solvability and approximation*, Ann. Appl. Probab., 31 (2021), pp. 2244–2274.

[2] E. Abi Jaber, E. Miller, and H. Pham, *Integral operator Riccati equations arising in stochastic Volterra control problems*, SIAM J. Control Optim., 59 (2021), pp. 1581–1603, https://doi.org/10.1137/19M1298287.

[3] N. Agram and B. Øksendal, *Malliavin calculus and optimal control of stchastic Volterra equations*, J. Optim. Theory Appl., 167 (2015), pp. 1070–1094.

[4] A. Bensoussan, *Estimation and Control of Dynamical Systems*, Springer, Heidelberg, 2018.

[5] S. Chen and J. Yong, *A linear quadratic optimal control problem for stochastic Volterra integral equation*, in Control Theory and Related Topics—in Memory of Professor Xunjing Li, Fudan University, China, pp. 44–66.

[6] F. Comte and E. Renault, *Long memory in continuous-time stochastic volatility models*, Math. Finance, 8 (1998), pp. 291–323.

[7] F. Delarue, *On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case*, Stochastic Process Appl., 99 (2002), pp. 209–286.

[8] I. Ekren, C. Keller, N. Touzi, and J. Zhang, *On viscosity solutions of path dependent PDEs*, Ann. Probab., 42 (2014), pp. 204–236.

[9] O. El Euch and M. Rosenbaum, *Perfect hedging in rough Heston models*, Ann. Appl. Probab., 28 (2018), pp. 3813–3856.

[10] O. El Euch and M. Rosenbaum, *The characteristic function of rough Heston models*, Math. Finance, 29 (2019), pp. 3–38.

[11] P. Friz and M. Hairer, *A Course on Rough Paths: With an Introduction to Regularity Structures*, Springer, New York, 2014.

[12] J. Gatheral, T. Jaisson, and M. Rosenbaum, *Volatility is rough*, Quant. Finance, 18 (2018), pp. 933–949.

[13] Y. Hamaguchi, *Extended backward stochastic Volterra integral equations and their applications to time-inconsistent stochastic recursive control problems*, Math. Control Relat. Fields, 11 (2021), pp. 433–478.

[14] Y. Hamaguchi, *Infinite horizon backward stochastic Volterra integral equations and discounted control problems*, ESAIM Control Optim. Calc. Var., 27 (2021), 101.

[15] Y. Hamaguchi and T. Wang, *Linear-Quadratic Stochastic Volterra Controls* II*: Optimal Strategies and Riccati-Volterra Equations*, preprint, arXiv:2204.10239, 2022.

[16] C. Hernández and D. Possamai, *Me, myself and* I*: A general theory of non-Markovian time-inconsistent stochastic control for sophisticated agents*, Ann. Appl. Probab., 33 (2023), pp. 1196–1258.

[17] C. Hernández and D. Possamai, *A unified approach to well-posedness of type-I backward stochastic Volterra integral equations*, Electron. J. Probab., 26 (2021), pp. 1–35.

[18] C. Hernández, *On Quadratic Multidimensional Type*-I *BSVIEs, Infinite Families of BSDEs and Their Applications*, preprint, arXiv:2111.11131, 2021.

[19] Q. Lei and C. S. Pun, *Nonlocal fully nonlinear parabolic differential equations arising in time-inconsistent problems*, J. Differential Equations, 358 (2023), pp. 339–385.

[20] X. Li and J. Yong, *Optimal Control Theory for Infinite Dimensional Systems*, Birkhäuser Boston, Boston, MA, 1995.
[21] Q. Lü, *Well-posedness of stochastic Riccati equations and closed-loop solvability for stochastic linear quadratic optimal control problems*, J. Differential Equations, 267 (2019), pp. 180–227.
[22] J. Ma, P. Protter, and J. Yong, *Solving forward-backward stochastic differential equations explicitly—a four step scheme*, Probab. Theory Relat. Fields, 98 (1994), pp. 339–359.
[23] J. Ma, Z. Wu, D. Zhang, and J. Zhang, *On well-posedness of forward-backward SDEs—a unified approach*, Ann. Appl. Probab., 25 (2015), pp. 2168–2214.
[24] J. Ma and J. Yong, *Forward-Backward Stochastic Differential Equations and Their Applications*, Lecture Notes in Math. 1702, Springer-Verlag, Berlin, 1999.
[25] E. Pardoux and S. Peng, *Backward stochastic differential equations and quasilinear parabolic partial differential equations*, in Stochastic Partial Differential Equations and Their Applications, Springer, Berlin, 1992, pp. 200–217.
[26] S. Peng and F. Wang, *BSDE, path-dependent PDE and nonlinear Feynman-Kac formula*, Sci. China Math., 59 (2016), pp. 19–36.
[27] J. Ruan, *Numerical Methods for High Dimensional Path Dependent PDEs Driven by Stochastic Volterra Integral Equations*, Ph.D. dissertation, University of Southern California, 2020.
[28] Y. Shi, T. Wang, and J. Yong, *Optimal control problems of forward-backward stochastic Volterra integral equations*, Math. Control Relat. Fields, 5 (2015), pp. 613–649.
[29] 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 (2016), pp. 2274–2308, https://doi.org/10.1137/15M103532X.
[30] J. Sun and J. Yong, *Stochastic Linear-Quadratic Optimal Control Theory: Open-Loop and Closed-Loop Solutions*, SpringerBriefs in Math., Springer, Cham, 2019.
[31] J. Sun and J. Yong, *Stochastic Linear-Quadratic Optimal Control Theory: Differential Games and Mean-Field Problems*, SpringerBriefs in Math., Springer, Cham, 2020.
[32] F. Viens and J. Zhang, *A martingale approach for fractional Brownian motions and related path dependent PDEs*, Ann. Appl. Probab., 29 (2019), pp. 3489–3540.
[33] H. Wang and J. Yong, *Time-inconsistent stochastic optimal control problems and backward stochastic Volterra integral equations*, ESAIM Control Optim. Calc. Var., 27 (2021), 22.
[34] H. Wang, J. Yong, and J. Zhang, *Path dependent Feynman-Kac formula for forward backward stochastic Volterra integral equations*, Ann. Inst. Henri Poincaré Probab. Stat., 58 (2022), pp. 603–638.
[35] T. Wang, *Linear quadratic control problems of stochastic Volterra integral equations*, ESAIM Control Optim. Calc. Var., 24 (2018), pp. 1849–1879.
[36] T. Wang, *Necessary conditions of Pontraygin's type for general controlled stochastic Volterra integral equations*, ESAIM Control Optim. Calc. Var., 26 (2020), 165.
[37] T. Wang and H. Zhang, *Optimal control problems of forward-backward stochastic Volterra integral equations with closed control regions*, SIAM J. Control Optim., 55 (2017), pp. 2574–2602, https://doi.org/10.1137/16M1059801.
[38] J. Yong, *Linear forward-backward stochastic differential equations with random coefficients*, Probab. Theory Relat. Fields, 135 (2006), pp. 53–83.
[39] J. Yong, *Well-posedness and regularity of backward stochastic Volterra integral equations*, Probab. Theory Related Fields, 142 (2008), pp. 21–77.
[40] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York, 1999.
[41] J. Zhang, *Backward Stochastic Differential Equations: From Linear to Fully Nonlinear Theory*, Springer, New York, 2017.